EQUISUM MATRICES AND THEIR PERMANENCE

BY

PHILIP J. DAVIS (*Division of Applied Mathematics, Brown University, Providence, RI*)

AND

IGOR NAJFELD (*Dept. of Mathematics and Statistics, University of Vermont, Burlington, VT*)

Abstract. In a square equisum matrix, all row and column sums are equal. In a rectangular equisum matrix, the common row sum is a rational multiple of the common column sum. This paper explores properties of equisum matrices, in particular, the preservation of the equisum condition under a variety of linear, nonlinear and pattern-maintaining transformations. A principal tool employed is a representation via the Fourier matrix or the circulant projectors associated with it.

1. Introduction. Patterns and their persistence under transformations constitute one of the major themes of mathematics. In this paper we take as the basic pattern to be studied matrices whose rows or columns have equal sums. The approach we take is through the discrete Fourier transform or the circulant projectors associated with it. The projector representation tends to be simpler than the Fourier one but the latter generalizes easily and the proofs of most theorems are vastly simplified.

Matrices with equal row and column sums are fairly common, e.g., permutation matrices, circulant matrices, and, with further nonlinear restrictions, they appear in the theory of stochastic or doubly stochastic matrices. They may be found in the equations of chemical kinetics, economics, graph theory, game theory, neural network modeling [6], cerebrovascular modeling [9], and in the collection of mathematical problems [13]. The equisum condition appears as a simple but important tool in a variety of nonlinear problems in matrix theory, such as the scaling problem [8] and the estimation problem of the Perron root [3].

Bold letters will denote matrices (upper case) and vectors (lower case). Unless otherwise indicated, all matrices are in $\mathbf{C}^{m \times n}$ and are labeled by two subscripts, e.g., $\mathbf{A}_{mn}$ or $\mathbf{A}_{m \times n}$ to avoid ambiguity. Square matrices are labeled by a single subscript, e.g., $\mathbf{A}_n$, which is often omitted when it is clearly understood from the context or when a statement holds for any $n \geq 1$. We shall use standard notation for transpose, $\mathbf{A}^\top$, conjugate, $\overline{\mathbf{A}}$, and conjugate transpose, $\mathbf{A}^*$. The notation $\mathbf{A}^{-1}$ designates the inverse if

Received April 2, 1998.

2000 *Mathematics Subject Classification.* Primary 15A57, 15A21.

$\mathbf{A}$ is square and nonsingular, or the Moore-Penrose generalized inverse otherwise. For an arbitrary rectangular matrix, $\mathbf{A}^{-1}$ always exists (see, e.g., [10], [7]). The notation $s(\mathbf{A}) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}$ stands for the sum of all entries of $\mathbf{A}$. By $\mathbf{A} = \text{circ}(\mathbf{a})$, or $\mathbf{A} = \text{semicirc}(\mathbf{a})$, we mean a circulant, or a semicirculant, matrix generated by a vector $\mathbf{a} = [a_1, \dots, a_n]$ (see, e.g., [2]). For example,

$$\mathbf{A} = \text{semicirc}(a_1, a_2, a_3) = \begin{bmatrix} a_1 & a_2 & a_3 \\ 0 & a_1 & a_2 \\ 0 & 0 & a_1 \end{bmatrix}.$$

The notation $\text{diag}(\mathbf{A}, \mathbf{B}, \dots)$ designates the block diagonal matrix whose square blocks are, reading from northwest to southeast, $\mathbf{A}, \mathbf{B}, \dots$, and $\text{semicirc}(\mathbf{A}, \mathbf{B}, \dots)$ stands for a block semicirculant matrix.

Three elementary square matrices are the zero matrix $\mathbf{0}_n$, the identity matrix $\mathbf{I}_n$, and the Fourier matrix $\mathbf{F}_n$ defined in terms of a scalar $\omega = e^{2\pi i/n}$ by

$$\mathbf{F}_n^* = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \omega^{(n-1)(n-1)} \end{bmatrix}. \quad (1)$$

The Fourier matrix $\mathbf{F}$ is unitary and satisfies the following properties (see, e.g., [2]):

$$\begin{aligned} \mathbf{F} &= \mathbf{F}^\top, \\ \mathbf{F}^{-1} &= \mathbf{F}^* = \overline{\mathbf{F}}. \end{aligned} \quad (2)$$

2. The basic projector matrices. Take the first column of $\mathbf{F}_n$, denoted by $\mathbf{f}_n = n^{-1/2}[1, \dots, 1]^*$, and form the product with its conjugate, namely

$$\mathbf{J}_n = \mathbf{f}_n \mathbf{f}_n^* = \frac{1}{n} \text{circ}(1, \dots, 1). \quad (3)$$

The matrix $\mathbf{J}_n$ is a projector onto the one-dimensional space spanned by $\mathbf{f}_n$ and satisfies the following properties:

$$\mathbf{J} = \mathbf{J}^\top = \mathbf{J}^* = \mathbf{J}^{-1} = \mathbf{J}^2. \quad (4)$$

In the rectangular case we write

$$\mathbf{J}_{mn} = \mathbf{f}_m \mathbf{f}_n^* = \frac{1}{\sqrt{mn}} \begin{bmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{bmatrix}_{m \times n}, \quad (5)$$

and the matrix $\mathbf{J}_{mn}$ satisfies the following properties:

$$\begin{aligned} \mathbf{J}_{mn} &= \mathbf{J}_{nm}^\top = \mathbf{J}_{nm}^{-1}, \\ \mathbf{J}_n &= \mathbf{J}_{nm} \mathbf{J}_{mn} = \mathbf{J}_{mn}^\top \mathbf{J}_{mn}, \end{aligned} \quad (6)$$

as well as

$$\mathbf{J}_{mn} \mathbf{J}_n = \mathbf{J}_m \mathbf{J}_{mn} = \mathbf{J}_{mn}. \quad (7)$$

The second basic matrix, $\mathbf{Q}_n$, is defined by

$$\mathbf{Q}_n = \mathbf{I}_n - \mathbf{J}_n = \frac{1}{n} \text{circ}(n-1, -1, \dots, -1). \quad (8)$$

This matrix is always square and satisfies the same relations (4) as does $\mathbf{J}$, namely

$$\mathbf{Q} = \mathbf{Q}^\top = \mathbf{Q}^* = \mathbf{Q}^{-1} = \mathbf{Q}^2. \quad (9)$$

The matrix $\mathbf{Q}$ is a projector onto a subspace orthogonal to $\mathbf{f}$. The square matrices $\mathbf{J}$ and $\mathbf{Q}$ commute (circulants) and the orthogonality relations extend to the rectangular case,

$$\mathbf{J}_{mn} \mathbf{Q}_n = \mathbf{Q}_m \mathbf{J}_{mn} = \mathbf{0}_{mn}. \quad (10)$$

Two other related matrices are

$$\begin{aligned} \mathbf{U}_n &= \text{diag}(1, \mathbf{0}_{n-1}), \\ \mathbf{V}_n &= \mathbf{I}_n - \mathbf{U}_n = \text{diag}(0, \mathbf{I}_{n-1}). \end{aligned} \quad (11)$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ satisfy the same relations as do $\mathbf{J}$ and $\mathbf{Q}$, namely

$$\begin{aligned} \mathbf{U} &= \mathbf{U}^\top = \mathbf{U}^* = \mathbf{U}^{-1} = \mathbf{U}^2, \\ \mathbf{V} &= \mathbf{V}^\top = \mathbf{V}^* = \mathbf{V}^{-1} = \mathbf{V}^2, \end{aligned} \quad (12)$$

as well as the orthogonality relations

$$\mathbf{UV} = \mathbf{VU} = \mathbf{0}. \quad (13)$$

Much of what follows depends on

THEOREM 1. The matrices $\mathbf{J}$ and $\mathbf{Q}$ are diagonalized by the Fourier matrix

$$\begin{aligned} \mathbf{J} &= \mathbf{F}^* \mathbf{U} \mathbf{F} = \mathbf{F} \mathbf{U} \mathbf{F}^*, \\ \mathbf{Q} &= \mathbf{F}^* \mathbf{V} \mathbf{F} = \mathbf{F} \mathbf{V} \mathbf{F}^*. \end{aligned} \quad (14)$$

Proof. Both of these matrices are circulants, and all circulants are diagonalized by $\mathbf{F}$ [2]. The rank($\mathbf{J}$) = 1, by Eq. (3), and the only nonzero eigenvalue, corresponding to the eigenvector $\mathbf{f}$, has the value of $\mathbf{f}^* \mathbf{f} = 1$. The proof for $\mathbf{Q}$ is similar and follows from Eqs. (8) and (14) (the Fourier transform of $\mathbf{J}$).

THEOREM 2. Fourier matrices provide a singular value decomposition of $\mathbf{J}_{mn}$,

$$\mathbf{J}_{mn} = \mathbf{F}_m^* \mathbf{U}_{mn} \mathbf{F}_n \quad (15)$$

with

$$\mathbf{U}_{mn} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}_{m \times n}. \quad (16)$$

Proof. To verify, we first postmultiply the matrix $\mathbf{J}_{mn}$ by $\mathbf{F}_n^*$ and use the property that for any integer $0 < p < n$ the finite geometric series $\sum_{k=0}^{n-1} \omega^{pk}$ sums to zero. The result is the matrix $\mathbf{J}_{mn} \mathbf{F}_n^* = [\mathbf{f}_m | \mathbf{0}_{m \times n-1}]$. Then, premultiplying by matrix $\mathbf{F}_m$ and using the same property, we get the matrix $\mathbf{U}_{mn}$.

The transformation UAU annihilates all but one of the entries of a square matrix A , namely

$$UAU = \text{diag}(a_{11}, \mathbf{0}_{n-1}), \quad (17)$$

while the transformation VAV is complementary, i.e.,

$$VAV = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & a_{nn} \end{bmatrix}. \quad (18)$$

3. Equisum matrices. Suppose that each row of a matrix $A \in C^{n \times n}$ sums to the same scalar σ . We may call this an equal row sum matrix and write it as $A \in ERS(n, \sigma)$. Similarly, $ECS(n, \sigma)$ denotes the set of all square matrices such that each column sums to σ . Examples are null column sum matrices, $ECS(n, 0)$ and generalized row stochastic matrices $ERS(n, 1)$. A matrix $A \in C^{n \times n}$ is said to be an *equisum* matrix, denoted by $A \in ES(n, \sigma)$, if each row and each column sums to σ . Examples are nullsum matrices, $ES(n, 0)$ and generalized doubly stochastic matrices, $ES(n, 1)$.

In these definitions, nothing is said about the reality or the nonnegativity of the matrix elements. It should be noted that the above conditions impose linear constraints and might therefore be treated by a general theory of linear structured matrices (see, e.g., [11]). However, this paper takes a different route.

We start with a few basic observations: if r_i and c_j denote the i th row and the j th column of an arbitrary square matrix A , then each row of the matrix JA is equal to the average row $r = n^{-1} \sum_{i=1}^n r_i$, while each column of the matrix AJ is equal to the average column $c = n^{-1} \sum_{j=1}^n c_j$. All entries in the matrix $JAJ = (s(A)/n)J$ are the same and equal to the arithmetic average $s(A)/n^2$. This means that if $A \in ERS(n, \sigma)$, then $AJ = \sigma J$. Similarly, if $A \in ECS(n, \sigma)$, then $JA = \sigma J$. Finally, a matrix $A \in ES(n, \sigma)$ iff it satisfies $AJ = JA = \sigma J$.

Next, the action of the matrix Q results in the following: each column of the matrix QA and each row of the matrix AQ sums to zero, while in the matrix QAQ each row and each column sums to zero. This follows from Eq. (10).

THEOREM 3. Let $A \in C^{n \times n}$. Then the following three representations in terms of projector matrices J and Q hold:

1. $A \in ERS(n, \sigma)$ iff it has the form $A = \sigma J + RQ$;
2. $A \in ECS(n, \sigma)$ iff it has the form $A = \sigma J + QR$;
3. $A \in ES(n, \sigma)$ iff it has the form $A = \sigma J + QRQ$;

The matrix $R \in C^{n \times n}$ is not unique.

Proof. It is sufficient to prove the third statement only. If $A = \sigma J + QRQ$, then $AJ = \sigma J^2 + QRQJ = \sigma J + \mathbf{0} = \sigma J$, and similarly for JA . Conversely, if $A \in ES(n, \sigma)$, then $A - \sigma J \in ES(n, 0)$. Such a matrix can always be represented as $A - \sigma J = QRQ$, for some $R \in C^{n \times n}$, and the third statement follows.

None of the factorizations above, $\mathbf{RQ}$, $\mathbf{QR}$, and $\mathbf{QRQ}$, is unique. We may add to $\mathbf{R}$ in, say, the $\mathbf{RQ}$ case, any matrix of the form $\mathbf{XJ}$, for $\mathbf{X} \in \mathbb{C}^{n \times n}$, since by (10) $\mathbf{XJQ} = \mathbf{0}$.

We now write these representations in Fourier language and recall that $\mathbf{f}$ stands for the first column of the Fourier matrix.

THEOREM 4. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$. Then the following three factorizations in terms of the Fourier matrix $\mathbf{F}$ hold:

1. $\mathbf{A} \in \mathbf{ERS}(n, \sigma)$ iff it has the form $\mathbf{A} = [\sigma \mathbf{f} \mid \mathbf{R}_{n \times n-1}] \mathbf{F}$
(or $\mathbf{A} = [\sigma \mathbf{f} \mid \mathbf{R}_{n \times n-1}] \mathbf{F}^*$);
2. $\mathbf{A} \in \mathbf{ECS}(n, \sigma)$ iff it has the form $\mathbf{A} = \mathbf{F} \begin{bmatrix} \sigma \mathbf{f}^* \\ \mathbf{R}_{n-1 \times n} \end{bmatrix}$ (or $\mathbf{A} = \mathbf{F}^* \begin{bmatrix} \sigma \mathbf{f}^* \\ \mathbf{R}_{n-1 \times n} \end{bmatrix}$);
3. $\mathbf{A} \in \mathbf{ES}(n, \sigma)$ iff it has the form

$$\mathbf{A} = \mathbf{F}^* \text{diag}(\sigma, \mathbf{R}) \mathbf{F} = \mathbf{F}^* \begin{bmatrix} \sigma & \mathbf{0}_{1 \times n-1} \\ \mathbf{0}_{n-1 \times 1} & \mathbf{R}_{n-1 \times n-1} \end{bmatrix} \mathbf{F}. \quad (19)$$

In each factorization the matrix $\mathbf{R}$ is unique.

Proof. We shall prove only the third factorization. If $\mathbf{A} \in \mathbf{ES}(n, \sigma)$ then, by Theorem 3, there is a matrix $\mathbf{M} \in \mathbb{C}^{n \times n}$ so that one can write $\mathbf{A} = \sigma \mathbf{J} + \mathbf{QMQ}$. By using the Fourier decompositions of matrices $\mathbf{J}$ and $\mathbf{Q}$ (see Eq. (14)), one gets

$$\mathbf{A} = \sigma \mathbf{F}^* \mathbf{U} \mathbf{F} + (\mathbf{F}^* \mathbf{V} \mathbf{F}) \mathbf{M} (\mathbf{F}^* \mathbf{V} \mathbf{F}) = \mathbf{F}^* (\sigma \mathbf{U} + \mathbf{V} \mathbf{F} \mathbf{M} \mathbf{F}^* \mathbf{V}) \mathbf{F}. \quad (20)$$

Since the matrix $\mathbf{F} \mathbf{M} \mathbf{F}^*$ is simply “another” $\mathbf{M}$ we may write $\mathbf{A} = \mathbf{F}^* (\sigma \mathbf{U} + \mathbf{V} \mathbf{M} \mathbf{V}) \mathbf{F}$. Now, by Eq. (11), the matrix $\sigma \mathbf{U} = \text{diag}(\sigma, \mathbf{0}_{n-1})$ and, by Eq. (18), the matrix $\mathbf{V} \mathbf{M} \mathbf{V} = \text{diag}(\mathbf{0}, \mathbf{R})$, with $\mathbf{R} \in \mathbb{C}^{n-1 \times n-1}$. This means that $\mathbf{A}$ takes on the form shown in Eq. (19).

For uniqueness, let $\mathbf{A} = \mathbf{F}^* \text{diag}(\sigma, \mathbf{R}_1) \mathbf{F} = \mathbf{F}^* \text{diag}(\sigma, \mathbf{R}_2) \mathbf{F}$. Then multiplying on the left by $\mathbf{F}$ and on the right by $\mathbf{F}^*$, we obtain $\text{diag}(\sigma, \mathbf{R}_1) = \text{diag}(\sigma, \mathbf{R}_2)$, whence $\mathbf{R}_1 = \mathbf{R}_2$.

The projector representation, $\mathbf{A} = \sigma \mathbf{J} + \mathbf{QRQ}$, is of course simpler, not requiring the Fourier matrix, but the Fourier representation (19), which is a block (or a partial) diagonalization, has its own virtues. For example, let $\mathbf{A} = \mathbf{F}^* \text{diag}(\sigma, \mathbf{R}) \mathbf{F}$. If k is a positive integer then, as we shall see later, $\mathbf{A}^k = \mathbf{F}^* \text{diag}(\sigma^k, \mathbf{R}^k) \mathbf{F}$ and the asymptotic properties of $\mathbf{A}^k$ can be read off from those of σ and $\mathbf{R}$. Next, suppose we have a linear system of equations $\mathbf{Ax} = \mathbf{b}$ with a nonsingular equisum matrix. Then the Fourier representation exhibits a conservation property for the solution, namely $\sum_{i=1}^n x_i = \sigma^{-1} \sum_{i=1}^n b_i$. If $\sigma = 0$, it is still possible to have a solution provided $\sum_{i=1}^n b_i = 0$. As we shall see, Fourier representation easily generalizes to rectangular and block matrices while the proofs of most theorems are considerably simplified.

To avoid needless repetitions of results, we shall work from now on with matrices in $\mathbf{ES}(n, \sigma)$ only, since all the results easily carry over to matrices in $\mathbf{ECS}(n, \sigma)$ and $\mathbf{ERS}(n, \sigma)$.

We now list several useful elementary observations: if $\mathbf{A} \in \mathbf{ES}(n, \sigma)$, then $\mathbf{A}, \mathbf{A}^\top \in \mathbf{ES}(n, \sigma)$, while $\overline{\mathbf{A}}, \mathbf{A}^* \in \mathbf{ES}(n, \bar{\sigma})$. At least one of the eigenvalues of $\mathbf{A}$ is equal to σ , while the corresponding left and right eigenvector is $\mathbf{f}$. Also, $|\sigma|$ is one of its singular values. A matrix $\mathbf{A}$ is diagonalizable iff the corresponding $\mathbf{R}$ in representation (19) is

diagonalizable. Suppose that $\mathbf{R}$ is diagonalized by $\mathbf{T}$, i.e., $\mathbf{F} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$. Then, except for the eigenvector $\mathbf{f}$ corresponding to the eigenvalue σ , the other $n - 1$ eigenvectors of $\mathbf{A}$ are obtained from $\mathbf{T}$ as follows:

$$\mathbf{A} = \mathbf{F}^* \begin{bmatrix} \sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \mathbf{F} = \mathbf{F}^* \begin{bmatrix} \sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} \end{bmatrix} \mathbf{F} = \mathbf{F}^* \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix} \begin{bmatrix} \sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{bmatrix} \mathbf{F}. \quad (21)$$

In other words, the eigenvectors of $\mathbf{A}$ are obtained from the Fourier transform

$$\mathbf{F}^* \text{diag}(1, \mathbf{T}). \quad (22)$$

The Fourier representation makes it very simple to derive several closure results: if $\mathbf{A} \in \mathbf{ES}(n, \alpha)$ and $\mathbf{B} \in \mathbf{ES}(n, \beta)$, and a, b are arbitrary constants, then

$$\begin{aligned} a\mathbf{A} + b\mathbf{B} &\in \mathbf{ES}(n, a\alpha + b\beta), \\ \mathbf{A}\mathbf{B}, \mathbf{B}\mathbf{A} &\in \mathbf{ES}(n, \alpha\beta), \\ [\mathbf{A}, \mathbf{B}] &\in \mathbf{ES}(n, 0). \end{aligned} \quad (23)$$

This shows that the space of all equisum matrices, denoted by $\mathbf{ES}(n)$, is closed under linear combinations, the standard and the commutator products. This closure carries over to the Kronecker product and sum, that is, if $\mathbf{A} \in \mathbf{ES}(m, \alpha)$ and $\mathbf{B} \in \mathbf{ES}(n, \beta)$, then

$$\begin{aligned} \mathbf{A} \otimes \mathbf{B} &\in \mathbf{ES}(mn, \alpha\beta), \\ \mathbf{A} \oplus \mathbf{B} &\in \mathbf{ES}(mn, \alpha + \beta), \end{aligned} \quad (24)$$

and both results are also blockwise in $\mathbf{ES}(n)$. This can be extended as follow: if $\mathbf{A}_k \in \mathbf{ES}(n_k, \alpha_k)$, $k = 1, \dots, l$, are l square equisum matrices, then their Kronecker product and sum are also equisum,

$$\begin{aligned} \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \dots \otimes \mathbf{A}_r &\in \mathbf{ES} \left(\prod_{k=1}^l n_k, \prod_{k=1}^l \alpha_k \right), \\ \mathbf{A}_1 \oplus \mathbf{A}_2 \oplus \dots \oplus \mathbf{A}_r &\in \mathbf{ES} \left(\prod_{k=1}^l n_k, \sum_{k=1}^l \alpha_k \right). \end{aligned} \quad (25)$$

Here, e.g., $\mathbf{A}_1 \oplus \mathbf{A}_2 \oplus \mathbf{A}_3 = \mathbf{A}_{n_1} \otimes \mathbf{I}_{n_2} \otimes \mathbf{I}_{n_3} + \mathbf{I}_{n_1} \otimes \mathbf{A}_{n_2} \otimes \mathbf{I}_{n_3} + \mathbf{I}_{n_1} \otimes \mathbf{I}_{n_2} \otimes \mathbf{A}_{n_3}$.

The closure property leads immediately to the question whether the limit of a given sequence of matrices $\mathbf{A}_k \in \mathbf{ES}(n, \alpha_k)$, such that $\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A}$, is also equisum.

THEOREM 5. If $\mathbf{A}_k \in \mathbf{ES}(n, \alpha_k)$ is a convergent sequence of equisum matrices, then the limit matrix is equisum satisfying

$$\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A} \in \mathbf{ES}(n, \lim_{k \rightarrow \infty} \alpha_k) = \mathbf{ES}(n, \alpha). \quad (26)$$

Proof. Convergence $\mathbf{A}_k \rightarrow \mathbf{A}$ means that any fixed finite linear combination of entries of $\mathbf{A}_k$ converges to the same linear combination of entries of $\mathbf{A}$; therefore,

$$\begin{aligned} \mathbf{A} &= \lim_{k \rightarrow \infty} (\mathbf{F} \text{diag}(\alpha_k, \mathbf{R}_k) \mathbf{F}^*) = \mathbf{F} \lim_{k \rightarrow \infty} (\text{diag}(\alpha_k, \mathbf{R}_k)) \mathbf{F}^* \\ &= \mathbf{F} \text{diag}(\lim_{k \rightarrow \infty} \alpha_k, \lim_{k \rightarrow \infty} \mathbf{R}_k) \mathbf{F}^* = \mathbf{F} \text{diag}(\alpha, \mathbf{R}) \mathbf{F}^* \in \mathbf{ES}(n, \alpha). \end{aligned} \quad (27)$$

THEOREM 6. If an integer $k \neq 0$, then $\mathbf{A}^k \in \mathbf{ES}(n, \sigma^k)$. If $f(z) = \sum_{k=0}^{\infty} c_k z^k$ designates any power series whose radius of convergence is greater than the spectral radius of $\mathbf{A}$, then $f(\mathbf{A}) \in \mathbf{ES}(n, f(\sigma))$.

Proof. The first part of the theorem follows from the representation (19),

$$\mathbf{A}^k = (\mathbf{F}^* \text{diag}(\sigma, \mathbf{R}) \mathbf{F})^k = \mathbf{F}^* \text{diag}(\sigma^k, \mathbf{R}^k) \mathbf{F} \in \mathbf{ES}(n, \sigma^k). \quad (28)$$

This is used to prove the second part,

$$\begin{aligned} f(\mathbf{A}) &= \sum_{k=0}^{\infty} c_k \mathbf{A}^k = \sum_{k=0}^{\infty} c_k (\mathbf{F}^* \text{diag}(\sigma, \mathbf{R}) \mathbf{F})^k \\ &= \mathbf{F}^* \sum_{k=0}^{\infty} c_k \text{diag}(\sigma^k, \mathbf{R}^k) \mathbf{F} = \mathbf{F}^* \text{diag} \left(\sum_{k=0}^{\infty} c_k \sigma^k, \sum_{k=0}^{\infty} c_k \mathbf{R}^k \right) \mathbf{F} \\ &= \mathbf{F}^* \text{diag}(f(\sigma), f(\mathbf{R})) \mathbf{F} \in \mathbf{ES}(n, f(\sigma)). \end{aligned} \quad (29)$$

In connection with Theorem 6, we note an exponential relation between matrices $\mathbf{J}$ and $\mathbf{Q}$,

$$\mathbf{Q} = \lim_{t \rightarrow \infty} e^{-t\mathbf{J}} = \lim_{t \rightarrow \infty} (\mathbf{I} + (e^{-t} - 1)\mathbf{J}). \quad (30)$$

The equisum pattern inheritance will now be explored in some depth. Starting with the Fourier representation (19), as a function of $\mathbf{R}$, one easily verifies the following:

- If $\mathbf{R}$ is a diagonal matrix, then $\mathbf{A}$ is a circulant (and equisum). However, if $\mathbf{R}$ is a block diagonal matrix with each block being an equisum, it does not follow that $\mathbf{A}$ is a circulant or even a block circulant, although certain symmetries persist.
- If $\mathbf{R}$ is normal, or hermitian with σ real, then so is $\mathbf{A}$.
- If $\mathbf{R}$ is centrohermitian (hermitian symmetry with respect to the “central position” of the matrix, in other words $\mathbf{K}\mathbf{R} = \overline{\mathbf{R}}\mathbf{K}$, where $\mathbf{K}$ is the counter-identity matrix with ones on the counter-diagonal, zeros elsewhere, see, e.g. [5]) and σ is real, then $\mathbf{A}$ is real but not quite centrohermitian. Its lower $(n-1) \times (n-1)$ principal submatrix is centrohermitian and the first row and column exhibit reflection symmetry.
- If $\mathbf{R}$ is unitary and $|\sigma| = 1$, then $\mathbf{A}$ is unitary.
- If $\mathbf{R}$ is a real M -matrix (i.e., $\mathbf{R} = \rho\mathbf{I} - \mathbf{P}$ such that $\mathbf{P} \geq \mathbf{0}$ and $\rho \geq \max |\lambda(\mathbf{P})|$) and $\sigma = 0$, then $e^{-t\mathbf{A}} \geq \mathbf{0}$ (see, e.g., [1]) and is also doubly stochastic for all $t \geq 0$.
- $\text{tr}(\mathbf{A}) = \sigma + \text{tr}(\mathbf{R})$.
- $\det(\mathbf{A}) = \sigma \det(\mathbf{R})$.

From the projector representation one can show the following:

- If $\mathbf{R}$ is skewhermitian, then the Cayley transform $\mathbf{A} = (\mathbf{I} - t\mathbf{Q}\mathbf{R}\mathbf{Q})/(\mathbf{I} + t\mathbf{Q}\mathbf{R}\mathbf{Q})$ is both unitary and a generalized doubly stochastic matrix for all t .
- If $\mathbf{R}$ is skewhermitian, then $\mathbf{A} = e^{t\mathbf{Q}\mathbf{R}\mathbf{Q}}$ is both unitary and a generalized doubly stochastic matrix for all t .
- If $\mathbf{R}$ is real, then $\mathbf{A} = e^{it\mathbf{Q}\mathbf{R}\mathbf{Q}}$ is a generalized doubly stochastic matrix and a circle matrix (i.e., it satisfies $\mathbf{A}^{-1} = \overline{\mathbf{A}}$).

- If $\mathbf{R}$ is real and the entries satisfy $\mu^- \leq (\mathbf{Q}\mathbf{R}\mathbf{Q})_{ij} \leq \mu^+$, then $\mathbf{A} = \mathbf{J} + t\mathbf{Q}\mathbf{R}\mathbf{Q}$ is doubly stochastic for

$$t < \min \left(\frac{1}{n|\mu^-|}, \frac{n-1}{n\mu^+} \right). \quad (31)$$

This list is by no means complete; its main advantage is in efficient computer generation of matrices with a given set of patterns or properties. We now turn to the other matrix functions that preserve the equisum pattern.

THEOREM 7. If $\mathbf{A} \in \mathbf{ES}(n, \alpha)$, then its polar decomposition $\mathbf{A} = \mathbf{W}\mathbf{S}$ satisfies

$$\mathbf{S} \in \mathbf{ES}(n, |\alpha|), \quad \mathbf{W} \in \mathbf{ES}(n, \text{sign}(\alpha)). \quad (32)$$

Proof. The matrix $\mathbf{H} = \mathbf{A}^* \mathbf{A} \in \mathbf{ES}(n, |\alpha|^2)$, by Eq. (19), is hermitian nonnegative definite. Hence, it has a unique hermitian nonnegative square root, which we denote by $\mathbf{S}$. By Theorem 6, $\mathbf{S} = \sqrt{\mathbf{H}} \in \mathbf{ES}(n, |\alpha|)$ and if $\mathbf{A}$ is nonsingular, then so is $\mathbf{S}$ and its inverse $\mathbf{S}^{-1} \in \mathbf{ES}(n, |\alpha|^{-1})$. In this case we can uniquely determine the unitary factor by

$$\mathbf{W} = \mathbf{A}\mathbf{S}^{-1} \in \mathbf{ES}(n, \alpha/|\alpha|) = \mathbf{ES}(n, \text{sign}(\alpha)). \quad (33)$$

If $\mathbf{A}$ is singular, we shall work with its singular value decomposition, denoted by $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{Y}^*$ (see, e.g., [7]). Regardless of singularity, the hermitian factor $\mathbf{S} = \mathbf{Y}\mathbf{D}\mathbf{Y}^*$ is unique and, as we have just shown, equisum. If we write $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{Y}^* = \mathbf{W}\mathbf{Y}\mathbf{D}\mathbf{Y}^*$ and assume that $r = \text{rank}(\mathbf{D}) < n$, then the first r columns (or rows) of $\mathbf{W}\mathbf{Y}$ and $\mathbf{X}$ must be the same. The remaining $n - r$ columns (or rows) of $\mathbf{W}\mathbf{Y}$ can be chosen arbitrarily except that they must form an orthonormal system. Nothing prevents us from choosing them to be equal to those of $\mathbf{X}$; in other words, we set

$$\mathbf{W} = \mathbf{X}\mathbf{Y}^*. \quad (34)$$

To show that $\mathbf{W} \in \mathbf{ES}(n, \theta)$ write its Fourier transform as

$$\mathbf{F}\mathbf{W}\mathbf{F}^* = \begin{bmatrix} \theta & \mathbf{r} \\ \mathbf{c} & \mathbf{W}' \end{bmatrix}. \quad (35)$$

Since both $\mathbf{A}$ and $\mathbf{S}$ are equisum, then the Fourier transform of the equation $\mathbf{A} = \mathbf{W}\mathbf{S}$ implies that

$$\begin{bmatrix} \alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{A}' \end{bmatrix} = \begin{bmatrix} \theta & \mathbf{r} \\ \mathbf{c} & \mathbf{W}' \end{bmatrix} \begin{bmatrix} |\alpha| & \mathbf{0} \\ \mathbf{0} & \mathbf{S}' \end{bmatrix}. \quad (36)$$

This equation can be satisfied only if $\theta = \text{sign}(\alpha)$. When $\alpha = 0$ we can set $\theta = e^{ix}$, for some $x \in (-\pi, \pi]$. It follows that $|\theta| = 1$, which in turn implies that $\mathbf{c} = \mathbf{0}$ and $\mathbf{r} = \mathbf{0}$ since otherwise the first row and column of $\mathbf{W}$ could not be of unit length.

REMARK. A similar proof holds for the other polar decomposition, $\mathbf{A} = \mathbf{S}\mathbf{W}$.

Another matrix function that maintains the equisum property is $\text{sign}(\mathbf{A})$. If $\mathbf{A} = \mathbf{T}\mathbf{J}\mathbf{T}^{-1}$ is the Jordan canonical form of $\mathbf{A}$ and $\Lambda = \text{diag}(\mathbf{J})$, then

$$\text{sign}(\mathbf{A}) = \mathbf{T} \text{sign}(\text{Re}(\Lambda)) \mathbf{T}^{-1}. \quad (37)$$

In order to avoid ambiguity when $\text{Re}(\lambda_i) = 0$, this function is defined only for matrices having no eigenvalues on the imaginary axis. With this assumption, three important

properties immediately follow from Eq. (37): a) $[\mathbf{A}, \text{sign}(\mathbf{A})] = \mathbf{0}$, b) $(\text{sign}(\mathbf{A}))^2 = \mathbf{I}$, and c) $\text{sign}(\mathbf{MAM}^{-1}) = \mathbf{M} \text{sign}(\mathbf{A}) \mathbf{M}^{-1}$ for any nonsingular matrix $\mathbf{M}$.

THEOREM 8. If $\mathbf{A} \in \mathbf{ES}(n, \alpha)$, then $\text{sign}(\mathbf{A}) \in \mathbf{ES}(n, \text{sign}(\alpha))$.

Proof. From the Fourier representation of $\mathbf{A}$ and $\text{sign}(\mathbf{A}) = \text{sign}(\mathbf{F}^* \text{diag}(\alpha, \mathbf{R}) \mathbf{F})$, we may, by property c), pull out both Fourier matrices, namely $\text{sign}(\mathbf{A}) = \mathbf{F}^* \text{sign}(\text{diag}(\alpha, \mathbf{R})) \mathbf{F} = \mathbf{F}^* \text{diag}(\text{sign}(\alpha), \text{sign}(\mathbf{R})) \mathbf{F}$ and the last matrix is in $\mathbf{ES}(n, \text{sign}(\alpha))$.

REMARK. It is well known that the iteration $\mathbf{A}_{k+1} = (\mathbf{A}_k + \mathbf{A}_k^{-1})/2$, with the initial condition $\mathbf{A}_0 = \mathbf{A}$, converges to $\text{sign}(\mathbf{A})$ provided that no eigenvalue is on the imaginary axis. This restriction can be somewhat relaxed in the case of nullsum matrices: suppose $\sigma = 0$ is the only zero eigenvalue of $\mathbf{A}$ and define $\text{sign}(\sigma) = 0$. Then, if we carry out the iteration above using the generalized inverse, the following convergence will occur:

$$\mathbf{A}_k \rightarrow \mathbf{F}^* \text{diag}(0, \text{sign}(\mathbf{R})) \mathbf{F} = \text{sign}(\mathbf{A}) \in \mathbf{ES}(n, 0). \quad (38)$$

If we square the limit we get $(\text{sign}(\mathbf{A}))^2 = \mathbf{F}^* \text{diag}(0, (\text{sign}(\mathbf{R}))^2) \mathbf{F} = \mathbf{Q}$, a variation of property b) above.

We have seen that if $\mathbf{A} \in \mathbf{ES}(n, \sigma)$ and if $\mathbf{T} \in \mathbf{ES}(n, \tau)$ is nonsingular, then $\mathbf{TAT}^{-1} \in \mathbf{ES}(n, \sigma)$. The converse is more involved.

THEOREM 9. If $\mathbf{A} \in \mathbf{ES}(n, \sigma)$ and $\mathbf{B} \in \mathbf{ES}(n, \sigma)$ are similar, i.e., $\mathbf{B} = \mathbf{TAT}^{-1}$, then $\mathbf{T} \in \mathbf{ES}(n, \tau)$, provided σ is a distinct eigenvalue of $\mathbf{A}$.

Proof. Writing the similarity equation in terms of Fourier representations for both matrices, i.e., $\mathbf{F}^* \text{diag}(\sigma, \mathbf{B}) \mathbf{FT} = \mathbf{TF}^* \text{diag}(\sigma, \mathbf{A}') \mathbf{F}$, it follows that

$$\text{diag}(\sigma, \mathbf{B}')(\mathbf{FTF}^*) = (\mathbf{FTF}^*) \text{diag}(\sigma, \mathbf{A}'). \quad (39)$$

Denote the key submatrices of the Fourier transform of $\mathbf{T}$ by

$$\mathbf{FTF}^* = \begin{bmatrix} \tau & \mathbf{r} \\ \mathbf{c} & \mathbf{T}' \end{bmatrix}. \quad (40)$$

Equation (39) then splits into four equations, namely

$$\begin{aligned} \sigma\tau &= \tau\sigma, \\ \sigma\mathbf{r} &= \mathbf{r}\mathbf{A}', \\ \sigma\mathbf{c} &= \mathbf{B}'\mathbf{c}, \\ \mathbf{B}'\mathbf{T}' &= \mathbf{T}'\mathbf{A}'. \end{aligned} \quad (41)$$

The first equation is satisfied for any τ . The second and the third equations are the left and the right eigenvector equations, corresponding to the same eigenvalue σ , associated with two matrices $\mathbf{A}'$ and $\mathbf{B}'$ known, by assumption, to share the same set of $n - 1$ eigenvalues. Thus, if σ is an algebraically distinct eigenvalue of matrix $\mathbf{A}$ (as well as of $\mathbf{B}$) it cannot be an eigenvalue of either $\mathbf{A}'$ or $\mathbf{B}'$ and the only way these two equations can be satisfied is if $\mathbf{r} = \mathbf{0}$ and $\mathbf{c} = \mathbf{0}$. By assumption on nonsingularity of $\mathbf{T}$ we may set $\tau = 1$. Finally, there will be at least one nontrivial solution of the fourth equation and for each solution the corresponding matrix $\mathbf{T} \in \mathbf{ES}(n, 1)$.

Suppose now that $\mathbf{A}(t) \in \mathbf{ES}(n, \alpha(t))$ is a sufficiently smooth matrix function for all $t \in [a, b]$ in a given interval, which implies that $\alpha(t)$ inherits the smoothness. Then

there is a number of important linear operators $\mathbf{L}: \mathbf{C}^{n \times n} \rightarrow \mathbf{C}^{n \times n}$ satisfying $\mathbf{L}\{\mathbf{A}(t)\} \in \mathbf{ES}(n, \mathbf{L}(\alpha(t)))$. For example, derivatives behave according to

$$\frac{d^k \mathbf{A}(t)}{dt^k} \in \mathbf{ES} \left(n, \frac{d^k \alpha(t)}{dt} \right), \quad k = 1, 2, \dots, \quad (42)$$

and if $\mathbf{B}(t) \in \mathbf{ES}(n, \beta(t))$ and is integrable, then the convolution integral satisfies

$$\mathbf{C}(t) = \int_a^b \mathbf{A}(\tau) \mathbf{B}(t - \tau) d\tau \in \mathbf{ES} \left(n, \int_a^b \alpha(\tau) \beta(t - \tau) d\tau \right). \quad (43)$$

This relation follows from Eq. (2.3).

Consider now a linear system of ordinary differential equations

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{A}(t)\mathbf{X}(t), \quad (44)$$

when $\mathbf{A}(t) \in \mathbf{ES}(n, \alpha(t))$ and is also Lipschitz continuous on some domain $t \in D_{t_0}$ that includes an initial point t_0 . The fundamental matrix of solutions for system (44) satisfies $\mathbf{X}(t_0) = \mathbf{I}$, which is an equisum matrix, and the question is, does it remain so for other values of $t \in D_{t_0}$.

THEOREM 10. The fundamental matrix of solutions generated by $\mathbf{A}(t) \in \mathbf{ES}(n, \alpha(t))$ satisfies

$$\mathbf{X}(t) \in \mathbf{ES} \left(n, \exp \left(\int_{t_0}^t \alpha(\tau) d\tau \right) \right). \quad (45)$$

Proof. We shall use the standard notation for the product integral representation (see, e.g., [4]): denote a partition by $P_k: t_0 = \tau_0 < \tau_1 < \dots < \tau_k = t$, the i th subinterval length by $\delta_i = \tau_i - \tau_{i-1}$, and the i th interior point by $\hat{\tau}_i \in (\tau_{i-1}, \tau_i)$. One then forms the ordered product (index i increases from right to left) of k matrices,

$$\mathbf{X}(t) = \lim_{k \rightarrow \infty} \prod_{i=1}^k e^{\delta_i \mathbf{A}(\hat{\tau}_i)}, \quad (46)$$

and then passes to the limit so that the maximum partition interval $\mu(P_k) \rightarrow 0$ as $k \rightarrow \infty$. This infinite matrix product converges uniformly to the fundamental matrix of solutions [4]. Since $\mathbf{A}(\hat{\tau}_i) \in \mathbf{ES}(n, \alpha(\hat{\tau}_i))$ then, by Theorem 6,

$$e^{\delta_i \mathbf{A}(\hat{\tau}_i)} \in \mathbf{ES}(n, e^{\delta_i \alpha(\hat{\tau}_i)}) \quad (47)$$

and, by the closure property of products in Eq. (23), each finite product satisfies

$$\prod_{i=1}^k e^{\delta_i \mathbf{A}(\hat{\tau}_i)} \in \mathbf{ES} \left(n, \exp \left(\sum_{i=1}^k \delta_i \alpha(\hat{\tau}_i) \right) \right). \quad (48)$$

The left-hand side of Eq. (48) defines a sequence of matrices known to converge uniformly to the limit. It follows by Theorem 5 that this limit satisfies

$$\lim_{k \rightarrow \infty} \prod_{i=1}^k e^{\delta_i \mathbf{A}(\hat{\tau}_i)} \in \mathbf{ES} \left(n, \exp \left(\lim_{k \rightarrow \infty} \sum_{i=1}^k \delta_i \alpha(\hat{\tau}_i) \right) \right) = \mathbf{ES} \left(n, \exp \left(\int_{t_0}^t \alpha(\tau) d\tau \right) \right). \quad (49)$$

REMARK. A proof based on the Peano-Baker-Hausdorff series is a bit more difficult since it involves both products and sums of equisum matrices.

REMARK. It is worth mentioning that if $\mathbf{A}(t) \in \mathbf{ERS}(n, \alpha(t))$, then $\mathbf{X}(t) \in \mathbf{ERS}(n, \exp(\int_{t_0}^t \alpha(\tau) d\tau))$, and similarly for matrices in $\mathbf{ECS}(n)$.

Equation (45), which relates the equisum property of the matrices $\mathbf{A}(t)$ and its fundamental matrix of solutions $\mathbf{X}(t)$, should be compared to the Abel-Liouville identity, which states that the $\det(\mathbf{X}(t)) = \exp(\int_{t_0}^t \text{Tr}(\mathbf{A}(\tau)) d\tau)$. The common appearance of the exponential of an indefinite integral is due to linearity of the differential system (44), but the different integrands convey different meanings.

4. Rectangular equisum matrices. By a rectangular equisum matrix we mean a matrix $\mathbf{A} \in \mathbf{C}^{m \times n}$ such that all its row sums are equal to σ_r and all its column sums equal to σ_c . The set of all equisum matrices is denoted by $\mathbf{ES}(m, n, \sigma)$, where we define an “average sum” by

$$\sigma = \frac{s(\mathbf{A})}{\sqrt{mn}}. \quad (50)$$

The five numbers $m, n, \sigma_r, \sigma_c, \sigma$ are not independent.

THEOREM 11. If $\mathbf{A} \in \mathbf{ES}(m, n, \sigma)$, then

$$s(\mathbf{A}) = m\sigma_r = n\sigma_c \quad (51)$$

and

$$\sigma = \sigma_r \sqrt{\frac{m}{n}} = \sigma_c \sqrt{\frac{n}{m}}. \quad (52)$$

Proof. The two equalities in Eq. (51) come from writing the $s(\mathbf{A})$ first in terms of row sums and then in terms of column sums,

$$s(\mathbf{A}) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \right) = m\sigma_r = \sum_{j=1}^n \left(\sum_{i=1}^m a_{ij} \right) = n\sigma_c. \quad (53)$$

Equation (52) follows from dividing Eq. (51) by $\sqrt{mn}$ and using definition (50). On the other hand, definition (50) comes from writing equisum conditions in terms of the matrix $\mathbf{J}_{mn}$ (see Eq. (5)) as follows:

$$\begin{aligned} \mathbf{A}\mathbf{J}_{nm} &= \frac{m\sigma_r}{\sqrt{mn}} \mathbf{J}_m = \sigma_r \sqrt{\frac{m}{n}} \mathbf{J}_m = \frac{s(\mathbf{A})}{\sqrt{mn}} \mathbf{J}_m = \sigma \mathbf{J}_m, \\ \mathbf{J}_{nm}\mathbf{A} &= \frac{n\sigma_c}{\sqrt{mn}} \mathbf{J}_n = \sigma_c \sqrt{\frac{n}{m}} \mathbf{J}_n = \frac{s(\mathbf{A})}{\sqrt{mn}} \mathbf{J}_n = \sigma \mathbf{J}_n. \end{aligned} \quad (54)$$

Theorem 11 shows that it is impossible to have a rectangular equisum matrix whose row and column sums have different signs. Therefore, σ can also be interpreted as the signed geometric average of σ_r and σ_c , i.e., $\sigma = \text{sign}(s(\mathbf{A}))\sqrt{\sigma_r\sigma_c}$. We now extend projector representation to rectangular matrices.

THEOREM 12. Let $\mathbf{A} \in \mathbf{C}^{m \times n}$. Then the following three representations in terms of projector matrices $\mathbf{J}_{mn}$, $\mathbf{Q}_m$, and $\mathbf{Q}_n$ hold:

1. $\mathbf{A} \in \mathbf{ERS}(m, n, \sigma)$ iff it has the form $\mathbf{A} = \sigma \mathbf{J}_{mn} + \mathbf{R}_{mn} \mathbf{Q}_n$;
2. $\mathbf{A} \in \mathbf{ECS}(m, n, \sigma)$ iff it has the form $\mathbf{A} = \sigma \mathbf{J}_{mn} + \mathbf{Q}_m \mathbf{R}_{mn}$;
3. $\mathbf{A} \in \mathbf{ES}(m, n, \sigma)$ iff it has the form $\mathbf{A} = \sigma \mathbf{J}_{mn} + \mathbf{Q}_m \mathbf{R}_{mn} \mathbf{Q}_n$.

The matrix $\mathbf{R}_{mn}$ is not unique in any of these representations.

Proof. As before, it is sufficient to prove only the third statement. If $\mathbf{A} = \sigma \mathbf{J}_{mn} + \mathbf{Q}_m \mathbf{R}_{mn} \mathbf{Q}_n$, then $\mathbf{A} \mathbf{J}_{nm} = \sigma \mathbf{J}_{mn} \mathbf{J}_{nm} + \mathbf{Q}_m \mathbf{R}_{mn} \mathbf{Q}_n \mathbf{J}_{nm}$. The last term is annihilated, by the orthogonality property (10), and the first term simplifies, by Eq. (6), to $\mathbf{A} \mathbf{J}_{nm} = \sigma \mathbf{J}_m$. This is the first of Eqs. (54). The second one is obtained in a similar way by premultiplying with $\mathbf{J}_{nm}$. The proof that any rectangular matrix, whose rows and columns sum to zero, must have the form $\mathbf{Q}_m \mathbf{R}_{mn} \mathbf{Q}_n$ comes from Eq. (10).

The Fourier representation of rectangular equisum matrices is given by the following:

THEOREM 13. Let $\mathbf{A} \in \mathbf{C}^{m \times n}$. Then the following three factorizations in terms of the Fourier matrix hold:

1. $\mathbf{A} \in \mathbf{ERS}(m, n, \sigma)$ iff it has the form $\mathbf{A} = [\sigma \mathbf{f}_m \mid \mathbf{R}_{m \times n-1}] \mathbf{F}_n$;
2. $\mathbf{A} \in \mathbf{ECS}(m, n, \sigma)$ iff it has the form $\mathbf{A} = \mathbf{F}_m [\mathbf{R}_{m-1 \times n}^{\sigma \mathbf{f}_n^*}]$;
3. $\mathbf{A} \in \mathbf{ES}(m, n, \sigma)$ iff it has the form $\mathbf{A} = \mathbf{F}_m^* \text{diag}(\sigma, \mathbf{R}_{m-1 \times n-1}) \mathbf{F}_n$.

In each factorization the associated matrix $\mathbf{R}$ is unique. Also, in each factorization the matrix $\mathbf{F}$ can be replaced with $\mathbf{F}^*$.

Proof. The reader will have no difficulty in modifying the proof in Theorem 4 to the rectangular case.

If $\mathbf{A} \in \mathbf{ES}(m, n, \sigma)$ then $\mathbf{A}^\top \in \mathbf{ES}(n, m, \sigma)$, while $\mathbf{A}^* \in \mathbf{ES}(n, m, \bar{\sigma})$. One of the singular values of $\mathbf{A}$ is $|\sigma|$ with the corresponding singular vectors being $\text{sign}(\sigma) \mathbf{f}_m$ and $\mathbf{f}_n$. From the Fourier representation one can immediately obtain several other results. If $\mathbf{A} \in \mathbf{ES}(m, n, \alpha)$, $\mathbf{B} \in \mathbf{ES}(m, n, \beta)$, and a, b are arbitrary constants, then

$$\begin{aligned} \alpha \mathbf{A} + \beta \mathbf{B} &\in \mathbf{ES}(m, n, a\alpha + b\beta), \\ \mathbf{A} \mathbf{B}^* &\in \mathbf{ES}(m, \alpha \bar{\beta}), \\ \mathbf{B}^* \mathbf{A} &\in \mathbf{ES}(n, \alpha \bar{\beta}). \end{aligned} \tag{55}$$

Two immediate consequences are: $\sqrt{\mathbf{A}^* \mathbf{A}} \in \mathbf{ES}(n, |\sigma|)$ and $\sqrt{\mathbf{A} \mathbf{A}^*} \in \mathbf{ES}(m, |\sigma|)$. These results can be used to show (see, e.g., [2]) that the generalized inverse

$$\mathbf{A}^{-1} \in \mathbf{ES}(n, m, \sigma^{-1}). \tag{56}$$

If $\mathbf{A}_k \in \mathbf{ES}(m_k, n_k, \alpha_k)$, $k = 1, \dots, l$, are l rectangular equisum matrices, then their Kronecker product is also equisum:

$$\mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \dots \otimes \mathbf{A}_l \in \mathbf{ES} \left(\prod_{k=1}^l m_k, \prod_{k=1}^l n_k, \prod_{k=1}^l \alpha_k \right). \tag{57}$$

The definition of the Kronecker sum may be extended to the rectangular matrices, namely

$$\mathbf{A}_{mn} \oplus \mathbf{B}_{pq} = \mathbf{A}_{mn} \otimes \mathbf{I}(\mathbf{B}_{pq}) + \mathbf{I}(\mathbf{A}_{mn}) \otimes \mathbf{B}_{pq}. \tag{58}$$

Here $I(A_{mn})$ denotes the rectangular matrix of the same size as A_{mn} with ones on the main diagonal and zeros elsewhere. It acts as a dimension equalizer for the ordinary summation so that the result is in $C^{mp \times nq}$. When $A_{mn} \in ES(m, n, \alpha)$ and $B_{pq} \in ES(p, q, \beta)$, their Kronecker sum exhibits block partition with various equisum properties. This motivates the next section.

5. Square matrices with equisum blocks. When the previous theory is extended to partitioned matrices with equisum blocks, some new features appear. By a nontrivial k -partition of an integer n we mean a set of k integers $n_1, \dots, n_k$ such that $n_j > 0$ and $n = \sum_{j=1}^k n_j$. A matrix $A \in C^{n \times n}$ is said to be symmetrically partitioned (see, e.g., [2]) if its (i, j) th block $A_{ij} \in C^{n_i \times n_j}$, in other words,

$$A = \begin{bmatrix} A_{11} & \dots & A_{1k} \\ \dots & \dots & \dots \\ A_{k1} & \dots & A_{kk} \end{bmatrix}. \quad (59)$$

If, in addition, each block satisfies the equisum condition, namely $A_{ij} \in ES(n_i, n_j, \alpha_{ij})$, $i, j = 1, \dots, k$, then we shall say that $A \in ESB(n_1, \dots, n_k, \square A)$. The square matrix $\square A \in C^{k \times k}$, which we shall call the *sumblock* of A , has entries $(\square A)_{ij} = \alpha_{ij}$. Note that such an A need not be in $ES(n, \alpha)$. This definition can be extended to matrices whose blocks satisfy $A_{ij} \in ERS(n_i, n_j, \alpha_{ij})$ (or $ECS(n_i, n_j, \alpha_{ij})$) and all the results of this section could then be carried over with obvious modifications.

THEOREM 14. If $A \in ESB(n_1, \dots, n_k, \square A)$ then it is unitarily similar to the block diagonal matrix

$$A = T^* \text{diag}(\square A, R)T. \quad (60)$$

The matrix T , which depends only on a given k -partition of n , is given by

$$T = P \text{diag}(F_{n_1}, \dots, F_{n_k}), \quad (61)$$

with P being a permutation matrix depending only on $n_1, \dots, n_k$. The matrix $r \in C^{n-k \times n-k}$ is unique.

Proof. By Theorem 13, each block has a unique Fourier factorization so that matrix A can be written as

$$A = \begin{bmatrix} F_{n_1}^* \begin{bmatrix} \alpha_{11} & 0 \\ 0 & R_{11} \end{bmatrix} F_{n_1} & \dots & F_{n_1}^* \begin{bmatrix} \alpha_{1k} & 0 \\ 0 & R_{1k} \end{bmatrix} F_{n_k} \\ \dots & \dots & \dots \\ F_{n_k}^* \begin{bmatrix} \alpha_{k1} & 0 \\ 0 & R_{k1} \end{bmatrix} F_{n_1} & \dots & F_{n_k}^* \begin{bmatrix} \alpha_{kk} & 0 \\ 0 & R_{kk} \end{bmatrix} F_{n_k} \end{bmatrix}. \quad (62)$$

Factoring out the Fourier matrices on both sides of each block results in the following:

$$A = \text{diag}(F_{n_1}, \dots, F_{n_k})^* \begin{bmatrix} \begin{bmatrix} \alpha_{11} & 0 \\ 0 & R_{11} \end{bmatrix} & \dots & \begin{bmatrix} \alpha_{1k} & 0 \\ 0 & R_{1k} \end{bmatrix} \\ \dots & \dots & \dots \\ \begin{bmatrix} \alpha_{k1} & 0 \\ 0 & R_{k1} \end{bmatrix} & \dots & \begin{bmatrix} \alpha_{kk} & 0 \\ 0 & R_{kk} \end{bmatrix} \end{bmatrix} \text{diag}(F_{n_1}, \dots, F_{n_k}). \quad (63)$$

We now construct the permutation matrix $\mathbf{P}$ defined by

$$\mathbf{P} = \begin{bmatrix} e_{n_1} & z_{n_2} & \cdots & z_{n_k} \\ z_{n_1} & e_{n_2} & \cdots & z_{n_k} \\ \cdots & \cdots & \cdots & \cdots \\ z_{n_1} & z_{n_2} & \cdots & e_{n_k} \\ \mathbf{L}_{n_1} & \mathbf{Z}_{n_1 n_2} & \cdots & \mathbf{Z}_{n_1 n_k} \\ \mathbf{Z}_{n_2 n_1} & \mathbf{L}_{n_2} & \cdots & \mathbf{Z}_{n_2 n_k} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{Z}_{n_k n_1} & \mathbf{Z}_{n_k n_2} & \cdots & \mathbf{L}_{n_k} \end{bmatrix}. \quad (64)$$

Here, $e_{n_i} = [1, 0, \dots, 0]$ and $z_{n_i} = [0, 0, \dots, 0]$ are row vectors of length n_i , while the rectangular matrices $\mathbf{L}_{n_i}$ and $\mathbf{Z}_{n_i n_j}$ are subscripted in a nonstandard manner, i.e., $\mathbf{L}_{n_i} = [\mathbf{0}_{n_i-1 \times 1} \mid \mathbf{I}_{n_i-1}]$ and $\mathbf{Z}_{n_i n_j} = \mathbf{0}_{n_i-1 \times n_j}$. They both have $n_i - 1$ rows. The matrix $\mathbf{L}_{n_i}$ has n_i and matrix $\mathbf{Z}_{n_i n_j}$ has n_j columns. A short check will persuade the reader that $\mathbf{T} = \mathbf{P} \operatorname{diag}(\mathbf{F}_{n_1}, \dots, \mathbf{F}_{n_k})$ is unitary (as a product of unitary matrices) and that it does perform a block diagonalization of $\mathbf{A}$. The proof of uniqueness of $\mathbf{R}$ follows the same steps as in Theorem 13 and is omitted.

REMARK. The representation in terms of $\mathbf{J}_{n_i n_j}$ and $\mathbf{Q}_{n_i}$ matrices, as in Theorem 12, does not appear very useful except for easy computer generation.

The representation (60) shows that the space $\mathbf{ESB}(n_1, \dots, n_k)$ is closed under linear combinations, and ordinary and commutator products; in other words, if $\mathbf{A} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A})$, $\mathbf{B} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{B})$, while a and b are arbitrary constants, then

$$\begin{aligned} \mathbf{A}^* &\in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A}^*), \\ \overline{\mathbf{A}} &\in \mathbf{ESB}(n_1, \dots, n_k, \square \overline{\mathbf{A}}), \\ a\mathbf{A} + b\mathbf{B} &\in \mathbf{ESB}(n_1, \dots, n_k, a\square \mathbf{A} + b\square \mathbf{B}), \\ \mathbf{AB} &\in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A} \square \mathbf{B}), \\ [\mathbf{A}, \mathbf{B}] &\in \mathbf{ESB}(n_1, \dots, n_k, \square [\mathbf{A}, \mathbf{B}]). \end{aligned} \quad (65)$$

While the Kronecker product and sum of two such matrices inherit the structure to some extent (as shown by (60)), the inheritance mechanism is more complex since it involves other permutation matrices.

We shall now extend Theorem 6 as follows:

THEOREM 15. If $\mathbf{A} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A})$ then $\mathbf{A}^p \in \mathbf{ESB}(n_1, \dots, n_k, (\square \mathbf{A})^p)$ for any integer $p \neq 0$. If $f(z)$ is any analytic function whose convergence radius is greater than the spectral radius of $\mathbf{A}$, then

$$f(\mathbf{A}) \in \mathbf{ESB}(n_1, \dots, n_k, f(\square \mathbf{A})). \quad (66)$$

Proof. By representation (60),

$$\mathbf{A}^p = (\mathbf{T}^* \operatorname{diag}(\square \mathbf{A}, \mathbf{R}) \mathbf{T})^p = \mathbf{T}^* \operatorname{diag}((\square \mathbf{A})^p, \mathbf{R}^p) \mathbf{T}. \quad (67)$$

This result is then used in the following:

$$\begin{aligned}
 f(\mathbf{A}) &= \sum_{j=0}^{\infty} c_j \mathbf{A}^j = \sum_{k=0}^{\infty} c_j \mathbf{T}^* \text{diag}((\square \mathbf{A})^j, \mathbf{R}^j) \mathbf{T} \\
 &= \mathbf{T}^* \left(\sum_{j=0}^{\infty} c_j \text{diag}((\square \mathbf{A})^j, \mathbf{R}^j) \right) \mathbf{T} = \mathbf{T}^* \text{diag} \left(\sum_{j=0}^{\infty} c_j (\square \mathbf{A})^j, \sum_{j=0}^{\infty} c_j \mathbf{R}^j \right) \mathbf{T} \quad (68) \\
 &= \mathbf{T}^* \text{diag}(f(\square \mathbf{A}), f(\mathbf{R})) \mathbf{T} \in \mathbf{ESB}(n_1, \dots, n_k, f(\square \mathbf{A})).
 \end{aligned}$$

THEOREM 16. If $\mathbf{A} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A})$ is nonsingular with the nonsingular sum-block, then both factors in its polar decomposition $\mathbf{A} = \mathbf{W}\mathbf{S}$ inherit the equisum block pattern: the unitary factor $\mathbf{W} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{W})$ while the symmetric positive definite factor $\mathbf{S} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{S})$. Furthermore, the sumblocks $\square \mathbf{W}$ and $\square \mathbf{S}$ are the factors in the polar decomposition of sumblock $\square \mathbf{A}$,

$$\square \mathbf{A} = \square \mathbf{W} \square \mathbf{S}. \quad (69)$$

Proof. We start with the definition of $\mathbf{S} = \sqrt{\mathbf{H}} = \sqrt{\mathbf{A}^* \mathbf{A}}$. From the representation (60),

$$\begin{aligned}
 \mathbf{H} &= \mathbf{A}^* \mathbf{A} = \mathbf{T}^* \text{diag}((\square \mathbf{A})^*, \mathbf{R}^*) \mathbf{T} \mathbf{T}^* \text{diag}(\square \mathbf{A}, \mathbf{R}) \mathbf{T} \\
 &= \mathbf{T}^* \text{diag}((\square \mathbf{A})^* \square \mathbf{A}, \mathbf{R}^* \mathbf{R}) \mathbf{T} = \mathbf{T}^* \text{diag}(\square \mathbf{H}, \mathbf{G}) \mathbf{T}, \quad (70)
 \end{aligned}$$

which shows that $\mathbf{H} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{H})$. By Theorem 15, its unique hermitian positive definite square root, denoted by $\mathbf{S}$, satisfies

$$\mathbf{S} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{S}), \quad (71)$$

where we set $\square \mathbf{S} = \sqrt{\square \mathbf{H}}$. By assumption, $\mathbf{A}$ is nonsingular and so is $\mathbf{S}$. In that case we can multiply $\mathbf{A} = \mathbf{W}\mathbf{S}$ on the right by the inverse $\mathbf{S}^{-1} \in \mathbf{ESB}(n_1, \dots, n_k, (\square \mathbf{S})^{-1})$ and obtain the following:

$$\mathbf{W} = \mathbf{A}\mathbf{S}^{-1} \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A}(\square \mathbf{S})^{-1}). \quad (72)$$

REMARK. There are three cases involving singularity: 1) $\mathbf{A}$ is nonsingular, $\square \mathbf{A}$ is singular; 2) $\mathbf{A}$ is singular, $\square \mathbf{A}$ is nonsingular; 3) $\mathbf{A}$ and $\square \mathbf{A}$ are singular. Each case can be proved by means of a singular value decomposition of either $\mathbf{A}$ or $\square \mathbf{A}$ using the steps in the proof of Theorem 7.

REMARK. A similar proof holds for the other polar decomposition in the reverse order $\mathbf{A} = \mathbf{S}\mathbf{W}$.

We know from Theorem 6 that if $\mathbf{A} \in \mathbf{ES}(n, \alpha)$, then $f(\mathbf{A}) \in \mathbf{ES}(n, f(\sigma))$. If $\mathbf{B} \in \mathbf{ES}(n, \beta)$, then $f(\mathbf{A} + \mathbf{B}) \in \mathbf{ES}(n, f(\alpha + \beta))$ provided that the radius of convergence of f is sufficiently large. We shall now show that each directional, or Fréchet, derivative of $f(\mathbf{A})$ in the direction $\mathbf{B}$ is an equisum matrix.

THEOREM 17. If $\mathbf{A} \in \mathbf{ES}(n, \alpha)$, $\mathbf{B} \in \mathbf{ES}(n, \beta)$, and $f(z)$ is any power series whose convergence radius is greater than the sum of the spectral radii of $\mathbf{A}$ and $\mathbf{B}$, then each

directional derivative ($\mathbf{B}$) is an equisum matrix, namely

$$\mathbf{D}_B^{(j)}(f(\mathbf{A})) \in \mathbf{ES} \left(n, f^{(j)}(\alpha) \frac{\beta^j}{j!} \right), \quad j = 1, 2, \dots \quad (73)$$

Proof. The proof is based on the directional block representation of directional derivatives of $f(\mathbf{A})$ (see, [12]), which states that if $\mathbf{C}$ is a block semicirculant matrix,

$$\mathbf{C} = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{B} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{A} \end{bmatrix} = \text{semicirc}(\mathbf{A}, \mathbf{B}, \mathbf{0}, \dots, \mathbf{0}), \quad (74)$$

having the matrix $\mathbf{B}$ appearing precisely k times on the first block “superdiagonal”, then

$$f(\mathbf{C}) = \text{semicirc} \left(f(\mathbf{A}), \mathbf{D}_B(f(\mathbf{A})), \frac{\mathbf{D}_B^{(2)}(f(\mathbf{A}))}{2!}, \dots, \frac{\mathbf{D}_B^{(k)}(f(\mathbf{A}))}{k!} \right). \quad (75)$$

Since both $\mathbf{A}$ and $\mathbf{B}$ are equisum, it follows that the matrix $\mathbf{C} \in \mathbf{ESB}(n, \dots, n, \square \mathbf{C})$, with n appearing $k+1$ times. The sumblock $\square \mathbf{C}$ is a semicirculant matrix of dimension $k+1$, with the representer given by $\square \mathbf{C} = \text{semicirc}(\alpha, \beta, 0, \dots, 0)$.

By Theorem 15, $f(\mathbf{C}) \in \mathbf{ESB}(n, \dots, n, f(\square \mathbf{C}))$ and it follows that each directional derivative $\mathbf{D}_B^{(j)}(f(\mathbf{A}))$, $j = 1, \dots, k$ is an equisum matrix. To find out its sum, we invoke the theorem stating that if $\square \mathbf{C}$ is semicirculant, then so is $f(\square \mathbf{C})$ and we only need determine its representer. It is not difficult to show (e.g., induction) that the representer of $(\square \mathbf{C})^j$ is given by

$$(\square \mathbf{C})^j = \text{semicirc} \left(\alpha^j, \frac{j}{1!} \alpha^{j-1} \beta, \dots, \frac{j(j-1) \cdots (j-k+1)}{k!} \alpha^{j-k} \beta^k \right). \quad (76)$$

Hence,

$$f(\square \mathbf{C}) = \text{semicirc} \left[f(\alpha), f^{(1)}(\alpha) \beta, \dots, \frac{f^{(k)}(\alpha) \beta^k}{k!} \right] \quad (77)$$

and this proves Eq. (73).

We now extend Theorem 10 to a linear system of ordinary differential equations

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{A}(t)\mathbf{X}(t), \quad (78)$$

given that $\mathbf{A}(t) \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{A}(t))$ and is also Lipschitz continuous on some interval that includes an initial point t_0 .

THEOREM 18. The fundamental matrix of solutions for the system (78) satisfies

$$\mathbf{X}(t) \in \mathbf{ESB}(n_1, \dots, n_k, \square \mathbf{X}(t)), \quad (79)$$

and the sumblock of the fundamental matrix $\square \mathbf{X}(t)$ satisfies the differential equation

$$\frac{d}{dt} \square \mathbf{X} = \square \mathbf{A}(t) \square \mathbf{X}, \quad \square \mathbf{X}(t_0) = \mathbf{I}. \quad (80)$$

Proof. Writing the Fourier representation (60) as $\mathbf{A}(t) = \mathbf{T}^* \text{diag}(\square \mathbf{A}(t), \mathbf{R}(t)) \mathbf{T}$, where the matrix $\mathbf{T}$ is independent of t , we see that the Fourier transform of the solution $\mathbf{Y}(t) = \mathbf{T} \mathbf{X}(t) \mathbf{T}^*$ satisfies $d\mathbf{Y}(t)/dt = \text{diag}(\square \mathbf{A}(t), \mathbf{R}(t)) \mathbf{Y}(t)$, $\mathbf{Y}(t_0) = \mathbf{I}$. The system is block diagonalized, the first block being Eq. (80).

6. Least squares characterizations. For $\mathbf{A} \in \mathbb{C}^{n \times n}$, we designate the operator $\mathbf{P}_\theta(\mathbf{A}) = \theta \mathbf{J} + \mathbf{Q} \mathbf{A} \mathbf{Q}$. This operator is a projection since $\mathbf{P}_\theta^2(\mathbf{A}) = \theta \mathbf{J} + \mathbf{Q}(\theta \mathbf{J} + \mathbf{Q} \mathbf{A} \mathbf{Q}) \mathbf{Q} = \theta \mathbf{J} + \mathbf{Q} \mathbf{A} \mathbf{Q} = \mathbf{P}_\theta(\mathbf{A})$. Therefore, one would expect it to have certain properties of best approximation. These will now be developed. We employ the Euclidean (or Frobenius) norm, defined by

$$\|\mathbf{A}\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 = \text{tr}(\mathbf{A} \mathbf{A}^*) = \text{tr}(\mathbf{A}^* \mathbf{A}), \quad (81)$$

which is unitarily invariant, i.e., $\|\mathbf{X} \mathbf{A}\|_F = \|\mathbf{A} \mathbf{X}\|_F = \|\mathbf{A}\|_F$ for any unitary matrix $\mathbf{X}$.

Let $\mathbf{E}$ denote the difference between a given matrix $\mathbf{A}$ and an equisum matrix $\mathbf{B} = \mathbf{F}^* \text{diag}(\beta, \mathbf{R}) \mathbf{F}$. Since $\mathbf{F}, \mathbf{F}^*$ are unitary, we have the following equality:

$$\|\mathbf{E}\|_F = \|\mathbf{A} - \mathbf{B}\|_F = \|\mathbf{A} - \mathbf{F}^* \text{diag}(\beta, \mathbf{R}) \mathbf{F}\|_F = \|\mathbf{F} \mathbf{A} \mathbf{F}^* - \text{diag}(\beta, \mathbf{R})\|_F. \quad (82)$$

By the definition in (81), $\|\mathbf{E}\|_F$ will be minimized if we select

$$\beta = (\mathbf{F} \mathbf{A} \mathbf{F}^*)_{1,1} = \frac{s(\mathbf{A})}{n} \quad (83)$$

and $\mathbf{R}$ to be the $(n-1) \times (n-1)$ principal submatrix of $\mathbf{F} \mathbf{A} \mathbf{F}^*$ sitting in the lower right corner. This proves the following:

THEOREM 19. For any $\mathbf{A} \in \mathbb{C}^{n \times n}$, the matrix $\mathbf{B} = \beta \mathbf{J} + \mathbf{Q} \mathbf{A} \mathbf{Q}$ is the best approximation to $\mathbf{A}$, in the sense of the Frobenius norm, by matrices in $\mathbf{ES}(n)$ provided β is given by Eq. (83). The Fourier representation is given by

$$\mathbf{B} = \mathbf{F}^* (\text{diag}(1, (n-1)\mathbf{J}_{n-1}) \bullet (\mathbf{F} \mathbf{A} \mathbf{F}^*)) \mathbf{F}. \quad (84)$$

Here, the matrix $\text{diag}(1, (n-1)\mathbf{J}_{n-1})$ is a “stencil” matrix consisting of zeros and ones only. It acts as a selector of entries in the matrix $\mathbf{F} \mathbf{A} \mathbf{F}^*$ by using the Schur-Hadamard product denoted by $\bullet$. The quality of this approximation may be judged by the following theorem.

THEOREM 20. The difference matrix $\mathbf{E}$ between an arbitrary matrix $\mathbf{A}$ and its best equisum approximation $\mathbf{B}$ has at most two nonzero eigenvalues, i.e., $\lambda(\mathbf{E}) = (\varepsilon, -\varepsilon, 0, \dots, 0)$. The value of ε is specified in the proof by Eqs. (87) and (89).

Proof. The difference matrix $\mathbf{E} = \mathbf{A} - \mathbf{B} = \mathbf{A} - \mathbf{F}^* \text{diag}(\beta, \mathbf{R}) \mathbf{F}$ and its Fourier transform $\mathbf{F} \mathbf{E} \mathbf{F}^* = \mathbf{F} \mathbf{A} \mathbf{F}^* - \text{diag}(\beta, \mathbf{R})$ have the same characteristic equation. In view of Eq. (84), the Fourier transform of $\mathbf{E}$ is a rather sparse matrix,

$$\mathbf{F} \mathbf{E} \mathbf{F}^* = \begin{bmatrix} 0 & r_1 & \dots & r_{n-1} \\ c_1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ c_{n-1} & 0 & \dots & 0 \end{bmatrix}. \quad (85)$$

The constants $r_1, \dots, r_{n-1}, c_1, \dots, c_{n-1}$ are in general complex, but if a matrix $\mathbf{A}$ is real, then so is the difference matrix $\mathbf{E}$ (Theorem 19). In either case, it follows from Eq. (85) that

$$\|\mathbf{E}\|_F^2 = \sum_{k=1}^{n-1} |c_k|^2 |r_k|^2. \quad (86)$$

Furthermore, if we denote by

$$p = \sum_{k=1}^{n-1} c_k r_k, \quad (87)$$

then the characteristic equation of $\mathbf{F}\mathbf{E}\mathbf{F}^*$ can be derived by a recurrence (e.g., start the expansion of the determinant from the lower right corner) and contains only two terms:

$$\lambda^n - p\lambda^{n-2} = 0. \quad (88)$$

It follows that there are at most two nonzero eigenvalues, namely

$$\varepsilon = \pm\sqrt{p}. \quad (89)$$

REMARK. The meaning of the coefficients in Eq. (85) can be inferred from the following analysis: denote by $\mathbf{e}_r$ the vector of differences between row sums and β in Eq. (83). Then the first row in the matrix (85) is the Fourier transform of $\mathbf{e}_r$ scaled by $n^{-1/2}$, and similarly for $\mathbf{e}_c$, the vector of differences between column sums and β .

We now extend the best approximation of a square matrix $\mathbf{A}$ by a partitioned matrix $\mathbf{B} \in \mathbf{ESB}(n_1, \dots, n_k, \square\mathbf{B})$.

THEOREM 21. Given a square matrix $\mathbf{A}$, and a k -partition of $n, n_1, \dots, n_k$, the matrix

$$\mathbf{B} = \mathbf{T}^*(\text{diag}(k\mathbf{J}_k, (n-k)\mathbf{J}_{n-k}) \bullet (\mathbf{T}\mathbf{A}\mathbf{T}^*))\mathbf{T} \quad (90)$$

is the best approximation to $\mathbf{A}$ in the Frobenius norm by matrices in $\mathbf{EBS}(n_1, \dots, n_k)$. The matrix $\mathbf{T}$ is the same as in Theorem 14 and the sumblock of $\mathbf{B}$ is given by

$$(\square\mathbf{B})_{ij} = \frac{s(\mathbf{A}_{ij})}{\sqrt{n_i n_j}}. \quad (91)$$

We omit the proof of this theorem since it repeats the arguments used in Theorem 14 and Theorem 19.

7. Conclusion. Possible further investigations suggest themselves: solution of nonlinear matrix equations, e.g., matrix Riccati equation, with equisum coefficients, or the use of equisum (block) matrices for preconditioning processes. Similarly, consider equisum iterations of the form $\mathbf{X}_{k+1} = \sigma_k \mathbf{J} + \mathbf{Q}(\mathbf{S} \bullet \mathbf{X}_k) \mathbf{Q}$, where $\mathbf{S}$ is a selector matrix of 0's and 1's, $\sigma_k \rightarrow \sigma$, and which converges to an equisum matrix with 0's coinciding with those of $\mathbf{S}$.

Instead of constructing matrix $\mathbf{J}$ as the dyadic product of the first column of $\mathbf{F}$ with itself, one can use any other column to generate a projector matrix $\mathbf{J}$ (as a dyadic product) and the corresponding orthogonal matrix $\mathbf{Q}$. The entire theory replicates itself to a considerable extent, as the results of our computation with the second column of $\mathbf{F}$

show. This raises the question of which unitary matrices other than the Fourier matrices generate useful patterns invariant under linear and nonlinear transformations.

This paper has given a brief introduction of the extent to which the equisum property persists through both linear and nonlinear transformations. It seems to the writers that the general phenomenon of pattern permanence in the matrix context is both significant and wide open for further study.

REFERENCES

- [1] Richard Bellman, *Introduction to Matrix Analysis*, McGraw-Hill Book Co., 1970
- [2] Philip J. Davis, *Circulant Matrices*, Chelsea, 1994
- [3] Emeric Deutsch and Helmut Wielandt, *Nested Bounds for the Perron Root of a Nonnegative Matrix*, *Linear Algebra Appl.* **52/53**, 235–251 (1983)
- [4] John D. Dollard and Charles N. Friedman, *Product Integration*, Addison-Wesley Publ. Co., 1979
- [5] John A. Eisele and Robert M. Mason, *Applied Matrix and Tensor Analysis*, Wiley-Interscience, 1970
- [6] Frank C. Hoppensteadt, *An Introduction to the Mathematics of Neurons*, second edition, Cambridge University Press, 1997
- [7] Roger Horn and Charles Johnson, *Matrix Analysis*, Cambridge University Press, 1990
- [8] Leonid Kachian, *Diagonal Matrix Scaling is NP-Hard*, *Linear Algebra Appl.* **234**, 173–179 (1996)
- [9] Z. M. Kadas, W. D. Lakin, J. Yu, and P. L. Penar, *A mathematical model of the intracranial system including autoregulation*, *Neurological Research* **19**, August, 441–450 (1997)
- [10] Peter Lancaster and Miron Tismenetsky, *The Theory of Matrices*, second edition, *Computer Science and Applied Mathematics*, Academic Press, Orlando, FL, 1985
- [11] Jan R. Magnus, *Linear Structures*, Oxford University Press, 1988
- [12] Igor Najfeld and Timothy F. Havel, *Derivatives of the Matrix Exponential and Their Computation*, *Adv. Appl. Math.* **16**, 321–375 (1995)
- [13] Kenneth S. Williams and Kenneth Hardy, *The Red Book of Mathematical Problems*, Dover Publ., 1996