On square roots and norms of matrices with symmetry properties.

R. Brunnock, M. C. Lettington and K. M. Schmidt

(On the occasion of the 500th anniversary of Albrecht Dürer's work Melencolia I)

Abstract

The present work concerns the algebra of semi-magic square matrices. These can be decomposed into matrices of specific rotational symmetry types, where the square of a matrix of pure type always has a particular type. We examine the converse problem of categorising the square roots of such matrices, observing that roots of either type occur, but only one type is generated by the functional calculus for matrices. Some explicit construction methods are given. Moreover, we take an observation by N. J. Higham as a motivation for determining bounds on the operator p-norms of semi-magic square matrices.

1 Introduction

Traditionally a square matrix $M = (m_{ij})_{i,j\in\{1,\dots,n\}}$, which satisfies the constant row and column sum symmetry condition

$$\text{(s1)} \qquad \sum_{j=1}^{n} m_{ij} = c \quad (1 \le i \le n), \qquad \sum_{i=1}^{n} m_{ij} = c \quad (1 \le j \le n),$$

for some constant number c, is called a *semi-magic square* with weight c/n (*SMS* or *type S matrix* for short). If all the entries of M are non-negative, then we can write $M = (1/c)H$ with H a doubly-stochastic matrix.

If, in addition to condition $(s1)$, M also satisfies the associated pairwise symmetry condition

$$\text{(s2)} \qquad m_{ij} + m_{(n+1-i)(n+1-j)} = 2c/n \qquad (1 \le i, j \le n),$$

then M is called an *associated magic square* [1] with weight c/n, (*AMS* or *type A matrix* for short). In contrast, if M satisfies condition (s1) and the balanced pairwise symmetry condition

$$\text{(s3)} \qquad m_{ij} = m_{(n+1-i)(n+1-j)}, \quad (1 \le i, j \le n),$$

then we say that M is a *balanced magic square* with weight c/n (*BMS* or *type B matrix* for short). Hence all type A and type B matrices are of type S; a matrix

2010 *Mathematics Subject Classification*: 15A30; 15A60; 15A03; 16S50
Key words and phrases: Matrix Rings; Matrix Norms, Matrix Square Roots; Magic Squares.

which is simultaneously type A and type B must have every entry the same (see Lemma 2.3).

In [13] the multiplicative properties of these matrix types in the 3×3 case are considered. For general n, the result is as follows (see [8], Lemma 3.1).

Lemma 1.1. *Let M and N be $n \times n$ type* S *matrices with respective weights z and w. Then it follows that*

(1) MN is type S *with weight nzw.*

(2) If M and N are both type A *or both type* B*, then MN is type* B*.*

(3) If M is type A *and N is type* B*, then MN and NM are type* A*.*

(4) If M is invertible, then M^{-1} is type S *with weight $1/n^2z$.*

(5) If M is type B *and invertible, then M^{-1} is type* B*.*

(6) If M is type A *and invertible, then M^{-1} is type* A*.*

Remark. It follows from the above lemma that if M is type A then M^r is type A for all positive odd r and type B for all positive even r. Similarly, if M is type B then M^r is type B for all positive r. (Clearly, if M is also non-singular, then this holds for all $r \in \mathbb{Z}$). It is also clear from Lemma 1.1 (2) that a type A matrix cannot have any square root of type A or B.

A natural consequence of these relations is that for a given n, the set of all $n \times n$ type A and type B matrices generates an algebra which is contained in the algebra of type S matrices $\mathcal{R}(\mathrm{S})$. In fact, they generate the algebra $\mathcal{R}(\mathrm{S})$, as Lemma 2.1 shows. The algebra of all type $n \times n$ type B matrices, $\mathcal{R}(\mathrm{B})$, is a subalgebra of $\mathcal{R}(\mathrm{S})$. The subset of $\mathcal{R}(\mathrm{B})$ containing all matrices N for which there exists at least one $M \in \mathcal{R}(\mathrm{S})$ with $M^2 = N$ can be partitioned into two sets; those which have at least one type A square root matrix and those which have only type B square root matrices. Conceptually this partition can be thought of as being similar to the partition of $\mathbb{R}$ into the positive and negative real numbers.

The motivation for the present paper partially stems from considering the reciprocal statement which asks, given an $n \times n$ matrix M, if M^2 is type B, then must M necessarily be of type A or type B? Our results show that the answer is not as straightforward as one might hope and depends on a number of factors.

We begin by outlining fundamental results on representations of these matrices and give examples of general construction methods including one that always gives type A matrices M, with M^2 symmetric and type B. We show that every type S matrix M with weight w has a natural representation as a sum of a matrix with weight 0 (called its *kernel matrix*) and a multiple of a universal matrix E. Furthermore the kernel matrix can always be decomposed uniquely into its type A and type B parts, both of which also have weight 0.

We briefly consider matrix equations in the form of non-commutative polynomial functions of type S matrices. Solving such equations can be extremely complicated, but it turns out that particular types of equations involving type A and type B matrices are soluble; their solutions highlight the implicit duality underlying these matrix types.

We then focus on the question of matrix square roots. We outline the functional calculus for matrices which extends standard functions, such as the square root function, to matrices in general (see [6], [5], [4]). We find that if M is a square root matrix of a type B matrix N obtained using the functional calculus, then M is of type B; this implies that the type A square roots of type B matrices (which exist) do not arise in this way.

Furthermore we give conditions on the characteristic polynomial of the square root matrix M which ensure that M is at least type S.

In the final section of this paper we study the operator p-norms of type S matrices. The motivation for this line of enquiry stems from the observation by N. J. Higham that the operator p-norm of a type S matrix whose entries all have the same sign is invariant with regard to the choice of p. However, if this hypothesis is not satisfied, the norms depend on p in a significant way, and we provide upper and lower bounds and discuss their sharpness.

2 Matrix Definitions and Representations

We begin by translating the symmetry conditions (s1), (s2), (s3) of the previous section into matrix algebraic relations.

Let E be the $n \times n$ matrix with all entries equal to 1, and let J be the $n \times n$ matrix with the anti-diagonal entries equal to 1 and 0 everywhere else. When $n = 3$ this gives

$$E = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then the $n \times n$ matrix M is

(S1) type S with weight w if

$$ME = nwE = M^T E,$$

(S2) type A with weight w if M satisfies $(S1)$ and

$$M + JMJ = 2wE,$$

(S3) type B with weight w if M satisfies $(S1)$ and

$$JMJ = M.$$

The condition (S1) says that the rows and columns of M sum to nw, and so condition $(S1)$ is equivalent to condition $(s1)$ with $nw = c$. Similarly the product JMJ^{-1} has the original (i, j) entry of M in row $n+1-i$, column $n+1-j$. Hence conditions $(S2)$ and $(S3)$ are equivalent to conditions $(s2)$ and $(s3)$, respectively, again with $nw = c$.

Examples of a type B matrix with $n = 3, w = 4$, a traditional type A matrix with $n = 4, w = 17/2$, and a type A matrix with $n = 4, w = 41/2$ are

$$\begin{pmatrix} 2 & 4 & 6 \\ 4 & 4 & 4 \\ 6 & 4 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 8 & 27 & 21 & 26 \\ 40 & 7 & 13 & 22 \\ 19 & 28 & 34 & 1 \\ 15 & 20 & 14 & 33 \end{pmatrix}.$$

The second example is known as Dürer's magic square, as it appears in an engraving entitled *Melencolia I* by Albrecht Dürer. The date of the engraving is 1514 and appears in the middle of the bottom row of the matrix.

Lemma 2.1. *Any $n \times n$ matrix of type* S *is the sum of a type* A *matrix with weight* 0 *and a type* B *matrix (which has the same weight as the original matrix).*

Proof. Let M be a type S matrix; then

$$M = \frac{1}{2}(M - JMJ) + \frac{1}{2}(M + JMJ),$$

and by a straightforward calculation, $M - JMJ$ satisfies (S2) with $w = 0$, and $M + JMJ$ satisfies (S3). Both are type S because JMJ is. $\square$

Type S matrices have the following *natural representation.*

Lemma 2.2. *Let M be an $n \times n$ matrix of type* S *with weight w. Then M can be written in the form $M = L + wE$, where L is a type* S *matrix with weight* 0*. If M is of type* A *or* B*, L has the same type. E generates an ideal in $R(\mathrm{S})$. Moreover,*

$$M^r = L^r + n^{r-1} w^r E \qquad (r \in \mathbb{N}). \tag{2.1}$$

We call L the *kernel matrix* of M.

Proof. Clearly $L := M - wE$ has the stated properties. From $LE = EL = 0_n$ it follows that the product of M with E (in either order) always gives a multiple of E.

To prove the last statement by induction, we note that

$$M^{r+1} = (L + wE)(L^r + n^{r-1} w^n E) = L^{r+1} + n^r w^{r+1} E,$$

as $E^2 = nE$. $\square$

Lemma 2.3. *Let M be an $n \times n$ type* S *matrix with weight w.*

(a) If M is both type A and type B, then $M = wE$.

Figure 1: Albrecht Dürer's engraving Melencolia I of 1514.

(b) The decomposition

$$M = L_A + L_B + wE,$$

where L_A, L_B are kernel matrices of type A *and type* B, *resp., is unique.*

Proof. (a) From (S2) and (S3), we have $M = JMJ$ and also $M = -JMJ + 2wE$, so $2M = 2wE$.

(b) If $L_A + L_B + wE = M = L'_A + L'_B + wE$, then $L_A - L'_A = L'_B - L_B$, and from (a) it follows that both are $0E = 0_n$. ☐

The natural representation for Dürer's square is

$$\frac{1}{2}\begin{pmatrix} 15 & -11 & -13 & 9 \\ -7 & 3 & 5 & -1 \\ 1 & -5 & -3 & 7 \\ -9 & 13 & 11 & -15 \end{pmatrix} + \frac{17}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{2.2}$$

The following lemma means that in order to find a square root matrix of a type S matrix, we need only look for a square root matrix of its kernel matrix.

Lemma 2.4. *Let M and N be $n \times n$ type* S *matrices with respective weights w and nw^2, and respective kernel matrices L_M and L_N. Then M is a square root matrix of N if and only if L_M is a square root matrix of L_N.*

Proof. To see the sufficiency of this condition let us assume that $L_M^2 = L_N$, where L_M and L_N both have weight 0. Then by Lemma 2.2,

$$M = (L_M + wE)^2 = L_M^2 + nw^2E = L_N + nw^2E = N.$$

To prove the necessity of the condition we assume $M^2 = N$, which gives the identity $L_M^2 + nw^2E = L_N + nw^2E$. After cancellation we obtain $L_M^2 = L_N$ and hence the result. □

We now show that every type S matrix with weight w can be decomposed into the sum of a type S symmetric matrix with weight w and a type S antisymmetric kernel matrix with weight 0. Moreover, by subtracting wE, M_S can be decomposed into the sum of a kernel symmetric matrix of weight 0 and the matrix wE. This allows us to split the kernel matrix L into two kernel matrices, one of which is symmetric and the other antisymmetric.

Lemma 2.5. *Let M be an $n \times n$ type* S *matrix with weight w and kernel matrix L, so that $M = L + wE$. Denote by M_s and M_a, respectively L_s and L_a, the symmetric and antisymmetric parts of M and L, so that*

$$M_s = \frac{1}{2}\left(M + M^T\right), \qquad M_a = \frac{1}{2}\left(M - M^T\right),$$
$$L_s = \frac{1}{2}\left(L + L^T\right), \qquad L_a = \frac{1}{2}\left(L - L^T\right).$$

Then $M_s = L_s + wE$ is type S *with weight w and $M_a = L_a$ is type* S *with weight 0.*

Proof. Denote by $r_j^{(s)}$ and $c_j^{(s)}$ the jth row and column sums of the symmetric matrix M_s, and by $r_j^{(a)}$ and $c_j^{(a)}$ the jth row and column sums of the antisymmetric matrix M_a. By properties of (anti)symmetric matrices we have

$$c_j^{(s)} = r_j^{(s)}, \qquad c_j^{(a)} = -r_j^{(a)},$$

and using $M = M_s + M_a$ gives

$$nw = r_j^{(s)} + r_j^{(a)} = c_j^{(s)} + c_j^{(a)} = r_j^{(s)} - r_j^{(a)}.$$

Hence $r_j^{(a)} = -r_j^{(a)}$, from which we deduce that $r_j^{(a)} = 0$ and $r_j^{(s)} = nw$. Similarly we find that $c_j^{(s)} = nw$ and $c_j^{(a)} = 0$. □

The full decomposition of Dürer's square into type A symmetric and antisymmetric kernels and weight matrix is

$$\frac{1}{2}\begin{pmatrix} 15 & -9 & -6 & 0 \\ -9 & 3 & 0 & 6 \\ -6 & 0 & -3 & 9 \\ 0 & 6 & 9 & -15 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & -2 & -7 & 9 \\ 2 & 0 & 5 & -7 \\ 7 & -5 & 0 & -2 \\ -9 & 7 & 2 & 0 \end{pmatrix} + \frac{17}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

We conclude this section with an observation on the centre of mass of semi-magic square matrices. For a matrix $M = (m_{ij})_{i,j\in\{1,\dots,n\}}$ whose entries do not sum up to 0, the centre of mass is defined as

$$\left(\frac{\sum_{i,j=1}^n i\, m_{ij}}{\sum_{i,j=1}^n m_{ij}}, \frac{\sum_{i,j=1}^n j\, m_{ij}}{\sum_{i,j=1}^n m_{ij}}\right).$$

Theorem 2.6. *The centre of mass of any $n \times n$ type S matrix is* $\left(\frac{n+1}{2}, \frac{n+1}{2}\right)$.

Proof. Let $M = (m_{ij})_{i,j\in\{1,\dots,n\}}$ be a type S matrix. If M has non-zero weight w, we find

$$\sum_{i,j=1}^n m_{ij} = n^2 w$$

and

$$\sum_{i,j=1}^n i\, m_{ij} = \sum_{i=1}^n i \sum_{j=1}^n m_{ij} = \frac{n+1}{2}\, n^2 w$$

and similarly

$$\sum_{i,j=1}^n j\, m_{ij} = \frac{n+1}{2}\, n^2 w.$$

The matrix E has centre of mass as above; if the matrix M has weight 0, we can consider its centre of mass the limit of those of $M + wE$ as $w \to 0$, which will again be as stated above. $\square$

3 Methods of Solution and Construction

It is obvious from Lemma 1.1 (2) that a type B matrix may have a type A or a type B square root matrix. In fact, it may have both.

Lemma 3.1. *There are $n \times n$ type* B *matrices which simultaneously have a square root matrix of type A and a square root matrix of type B.*

We illustrate this lemma with examples of the simultaneous case when $n = 3$ and also when $n = 4$ with Dürer's square. By Lemma 2.2 we need only give examples of kernel matrices.

As a 3×3 example, we have

$$\begin{pmatrix} -2 & 1 & 1 \\ 3 & 0 & -3 \\ -1 & -1 & 2 \end{pmatrix}^2 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}^2 = \begin{pmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{pmatrix},$$

and as a 4×4 example, with Dürer's square

$$\begin{pmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{pmatrix}^2 = \begin{pmatrix} 15 & 8 & 5 & 6 \\ 5 & 10 & 11 & 8 \\ 8 & 11 & 10 & 5 \\ 6 & 5 & 8 & 15 \end{pmatrix}^2 = \begin{pmatrix} 341 & 285 & 261 & 269 \\ 261 & 301 & 309 & 285 \\ 285 & 309 & 301 & 261 \\ 269 & 261 & 285 & 341 \end{pmatrix}.$$

The first example also shows that the square root matrix of a symmetric type B matrix need not be symmetric.

We now consider some general forms and construction methods for square root matrices of a type B matrix.

The case $\mathbf{n = 3}$. The smallest type A matrix that can be constructed such that not all of its entries are equal is of size 3×3. Solving via the symmetry conditions we find that a 3×3 type A matrix M has the general form

$$M = \begin{pmatrix} a & b & -a-b \\ -2a-b & 0 & 2a+b \\ a+b & -b & -a \end{pmatrix} + w \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \tag{3.1}$$

so M^2 is type B and symmetric and of the form

$$M^2 = b(2a+b) \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} + 3w^2 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Hence any 3×3 type B kernel matrix which has a type A square root must be of the form

$$N' = \begin{pmatrix} -2B & B & B \\ B & -2B & B \\ B & B & -2B \end{pmatrix},$$

with $B = b(2a+b)$, so that $a = (B - b^2)/(2b)$. For any given real number B we can therefore choose b to be any other non-zero real number and then solve for a, so that N' has an infinite number of type A square root matrices. The choice $b = 0$ gives type A square roots of 0_n.

In a similar fashion we use the symmetry conditions for a 3×3 type B matrix to obtain the general form

$$N = \begin{pmatrix} a & b & -a-b \\ b & -2b & b \\ -a-b & b & a \end{pmatrix} + w \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \tag{3.2}$$

where we find that both N and

$$N^2 = \begin{pmatrix} 2\left(a^2+ba+b^2\right) & -3b^2 & -2a^2-2ba+b^2 \\ -3b^2 & 6b^2 & -3b^2 \\ -2a^2-2ba+b^2 & -3b^2 & 2\left(a^2+ba+b^2\right) \end{pmatrix} + 3w^2 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

are type B symmetric. Therefore any 3×3 type B kernel matrix which has a type B square root must be of the form

$$N' = \begin{pmatrix} A & B & -A-B \\ B & -2B & B \\ -A-B & B & A \end{pmatrix},$$

so that, given values of A and B, our square root matrix parameters are

$$b = \pm\sqrt{\frac{-B}{3}}, \qquad \text{and} \qquad a = \frac{1}{2}\left(b \pm \sqrt{2A+B}\right).$$

Hence N' has four type B square root matrices, which are all real if $B \leq 0$ and $2A \geq |B|$.

The case $\mathbf{n = 4}$. Proceeding as in the case $n = 3$ we use the symmetry conditions to simplify the solution space of all 4×4 type A and B kernel matrices. For M type A, the general algebraic form is given by

$$M = \begin{pmatrix} a & b & c & -a-b-c \\ d & -a-b-d & -a-c-d & 2a+b+c+d \\ -2a-b-c-d & a+c+d & a+b+d & -d \\ a+b+c & -c & -b & -a \end{pmatrix}, \tag{3.3}$$

where setting $a = 15/2$, $b = -11/2$, $c = -13/2$ and $d = -7/2$ yields Dürer's square. For N type B, the general form is given by

$$N = \begin{pmatrix} a & b & c & -a-b-c \\ d & e & -b-c-e & b+c-d \\ b+c-d & -b-c-e & e & d \\ -a-b-c & c & b & a \end{pmatrix}, \tag{3.4}$$

and the special choice $b = d$ ensures that N is symmetric. In the 3×3 case the squares of both general forms were symmetric and for simplicity we continue with this restriction and look for square roots of symmetric type B matrices. We find that any 4×4 type B symmetric kernel matrix N' which has a type A square root must be of the form

$$N' = \begin{pmatrix} -A-B-\sqrt{\mathrm{AB}} & A & B & \sqrt{\mathrm{AB}} \\ A & -A-B+\sqrt{\mathrm{AB}} & -\sqrt{\mathrm{AB}} & B \\ B & -\sqrt{\mathrm{AB}} & -A-B+\sqrt{\mathrm{AB}} & A \\ \sqrt{\mathrm{AB}} & B & A & -A-B-\sqrt{\mathrm{AB}} \end{pmatrix}.$$

For given values of A and B the corresponding type A square root matrix M' is given by fM (M as in (3.3) and f a scalar, possibly imaginary) where a takes an arbitrary value and setting

$$b = -a \pm \sqrt{B}, \quad c = -a \pm \sqrt{A}, \quad d = \frac{-2a^2 - 2ab - 2ac - b^2 - bc}{2a+b+c}, \quad f = \sqrt{\frac{2a+b+c}{2(b+c)}}.$$

(assuming that $2a + b + c, b + c \neq 0$) yields $(M')^2 = N'$. Hence N' has as infinite number of type A square root matrices.

Employing similar methods we find that if N is a type B square root of the symmetric type B matrix N', then N' must be of the form

$$N' = \begin{pmatrix} A & B & B & -A-2B \\ B & A & -A-2B & B \\ B & -A-2B & A & B \\ -A-2B & B & B & A \end{pmatrix}.$$

For given values of A and B the corresponding type B square root matrix M' is given by (3.4) with b an arbitrary value, $d = b$ and $e = -a - b - c$ with

$$c = -b \pm \sqrt{-B}, \quad \text{and} \qquad a = -\frac{1}{2}\left(\pm\sqrt{-B} \pm \sqrt{2A + 3B - 4bc}\right).$$

Therefore there also exist an infinite number of symmetric type B square root matrices of N'.

For larger values of n similar general solutions can be obtained but the formulae become more complicated. To conclude this section we therefore focus on construction methods which work for all $n \times n$ matrices. Specifically we give one method for construction of symmetric type B square roots and one for construction of type A square roots of symmetric type B matrices. The type B construction is a generalisation of the above case.

Lemma 3.2. *Let N' and N both be type* B *matrices with weight zero defined by*

$$N = (a-b)I + (c-b)J + bE, \qquad N' = (A-B)I + (C-B)J + BE,$$

where N and N' are related by

$$\begin{aligned} C &= -A-(n-2)B & \qquad b &= -\sqrt{-B/n} \\ a &= \frac{1}{2}(-(n-2)b + \sqrt{A-C}) & \qquad c &= -a-(n-2)b. \end{aligned}$$

Then $N^2 = N'$.

Proof. Writing $N = (a-b)I + (c-b)J + bE$, we find that

$$N^2 = \left((a-b)^2 + (c-b)^2\right) I + 2(a-b)(c-b)J + (b^2n + 2b(a-2b+c))E,$$

and using the identities

(i) $(a-b)^2 + (c-b)^2 = A - B$,

(ii) $2(a-b)(c-b) = -A-(n-1)B$,

(iii) $b^2n + 2b(a-2b+c) = B$,

we deduce that $N^2 = (A-B)I + (-A-(n-1)B)J + BE = N'$. □

Remark. With a bit more work it can be shown that N is a primary square root matrix of N', as defined in Section 5.

We now describe a construction of matrices from permutations of the canonical basis vectors. Let σ be a permutation on the integers $1, 2, 3, \dots n$, so that

$$\sigma : (1, 2, 3, \dots, n) \mapsto (\sigma(1), \sigma(2), \sigma(3), \dots, \sigma(n)).$$

Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ be the unit vectors $(1, 0, \cdots)^T, (0, 1, 0, \cdots)^T, \cdots, (0, \cdots, 0, 1)^T$, so that $I = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$.

We consider the permutation σ acting on the rows of I and define the permutation matrix $P_\sigma = (\mathbf{e}_{\sigma(1)}, \cdots, \mathbf{e}_{\sigma(n)})^T$. A permutation of the n rows $\mathbf{m}_1, \cdots, \mathbf{m}_n$ of an $n \times n$ matrix M can be accomplished by the product $P_\sigma M$, whose rows are $\mathbf{m}_{\sigma(1)}, \cdots, \mathbf{m}_{\sigma(n)}$. Similarly MP_σ has columns $\mathbf{k}_{\tau(1)}, \cdots, \mathbf{k}_{\tau(n)}$ where $\mathbf{k}_1, \cdots, \mathbf{k}_n$ are the columns of M, and τ is the permutation inverse to σ. With E and J defined as previously, J is the matrix with rows $\mathbf{e}_n, \mathbf{e}_{n-1}, \cdots, \mathbf{e}_1$, and we define K to be the matrix with rows $\mathbf{e}_2, \cdots, \mathbf{e}_n, \mathbf{e}_1$; e.g. when $n = 3$,

$$K = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

The matrices J and K (under multiplication) generate the dihedral group D_{2n}. We call any matrix that can be expressed as a linear combination of products of powers of J, K and E *diagonally expressible.* In fact we can form the diagonally expressible type A and type B basis matrices as follows.

Definition. For $n = 2m+1$, let K be the $n \times n$ permutation matrix as above; then, for $r \in \mathbb{Z}$, K^r gives a cyclic permutation of n elements. Let

$$A_r = K^{2r-1} - K^{-(2r-1)} \qquad \text{and} \qquad B_r = K^{2r} + K^{-2r}. \tag{3.5}$$

Then $JA_rJ = -A_r$ and $JB_rJ = B_r$, so that A_r, JA_r are both type A with weight 0 and B_r, JB_r are both type B with weight $w = 2/n$. Furthermore $A_r^T = -A_r = A_{1-r}$, $B_r^T = B_r = B_{-r}$, where $A_{m+1} = 0_n$ and $B_0 = 2I$, which together imply that A_r is an anti-symmetric type A kernel matrix and B_r is a symmetric type B matrix. This means that A_r^2 and B_r^2 are both symmetric type B matrices. They obey the identities

$$A_rA_s = B_{r+s-1} - B_{r-s}, \qquad A_rB_s = A_{r+s} + A_{r-s}, \qquad B_rB_s = B_{r+s} + B_{r-s}. \tag{3.6}$$

Using the above notation, we have the following lemma.

Lemma 3.3. *Let*

$$M = \sum_{r=1}^{[n/2]} (\alpha_r I + \beta_r J)A_r + \gamma E, \quad \text{and} \quad N = \sum_{r=0}^{[n/2]} (\alpha_r I + \beta_r J)B_r + \gamma E.$$

Then M is a type A *matrix with weight γ and anti-symmetric kernel matrix, and N is symmetric and type* B. *Moreover, M^r and N^r are symmetric type* B *for all positive even r and M^r is type A with anti-symmetric kernel for all positive odd r.*

The following is an example of such a type A square M and its inverse matrix M^{-1}, as stated in [8].

For natural number m, let $n = 2m+1$, and let a three-parameter family of $n \times n$ type A matrices be defined as

$$M(z, y, x) = (zI - yJ) \sum_{r=1}^{m} (m+1-r)A_r + (m(z+y)+x)E. \tag{3.7}$$

Then the (type B) square of the type A matrix $M(z, y, x)$ is given by

$$\begin{aligned} M^2(z,y,x) &= n(m(z+y)+x)^2 E \\ &\quad + (z^2 - y^2)\left(\frac{m(m+1)}{2} I + \frac{m(m+1)}{6} E - \sum_{q=1}^{m-1} \frac{(m+1-q)(m-q)}{2} B_q\right), \end{aligned}$$

and the inverse matrix has the simple structure

$$M^{-1}(z,y,x) = \frac{(zI - yJ)}{n(z^2-y^2)}A_0 + \frac{E}{n^2(m(z+y)+x)}.$$

4 Solving Matrix Equations in $\mathcal{R}(\mathrm{S})$

Theorem 4.1. *Let $r \in \mathbb{N}$, let M_i $(i \in \{1,\dots,r\})$ be $n\times n$ type* S *matrices and $M_i = L_i + w_iE$ their natural representations. Moreover, let f be an element of the non-commutative polynomial algebra (free algebra) in r variables with vanishing constant term, i.e. $f(M_1,\dots,M_r)$ a linear combination of finite products of $M_1,\dots,M_r$ with at least one factor. Then*

$$f(M_1,\dots,M_r) = f(L_1,\dots,L_r) + f(w_1E,\dots,w_rE). \tag{4.1}$$

In consequence, $M_1,\dots,M_r$ and k are a solution of the equation

$$f(M_1,\dots,M_r) = kE \tag{4.2}$$

if and only if

$$f(L_1,\dots,L_r) = 0_n \qquad \textit{and} \qquad f(w_1E,\dots,w_rE) = kE. \tag{4.3}$$

Proof. For any $j,k \in \{1,\dots,r\}$ we have

$$M_jM_k = (L_j + w_jE)(L_k + w_kE) = L_jL_k + w_jw_kE^2,$$

since $L_mE = 0_n = EL_m$ $(m \in \{1,\dots,r\})$. Thus any product of matrices M_j splits into the sum of the corresponding products of matrices L_j and w_jE; this gives (4.1).

Hence, if $L_1,\dots,L_r$, $w_1,\dots,w_r$ satisfy the equations (4.3), then $M_j := L_j + w_jE$ $(j \in \{1,\dots,r\})$ satisfy (4.2). Conversely, if

$$f(L_1,\dots,L_r) + f(w_1E,\dots,w_rE) = kE,$$

then

$$f(L_1,\dots,L_r) = kE - f(w_1E,\dots,w_rE) = uE$$

for some number u, since f applied to multiples of E will give a multiple of E. As products of matrices L_j have weight 0 and the left-hand side is a sum of such products, it follows that $u = 0$, and hence (4.3). □

Example 1. In the notation of (3.7), consider the Pell type $n\times n$ matrix equation

$$M(a,0,w_a - ma)^2 - \lambda M(b,0,w_b - mb)^2 = M(c,0,w_c - mc)^2. \tag{4.4}$$

Using the natural representation

$$M(z,0,w_z - mz) = L_z + w_zE \qquad (z \in \{a,b,c\}),$$

this is equivalent to

$$(L_a + w_a E)^2 - \lambda(L_b + w_b E)^2 = (L_c + w_c E)^2,$$

which, by Theorem 4.1, reduces to $L_a^2 - \lambda L_b^2 = L_c^2$ and $nw_a^2 E - \lambda n w_b^2 E = n w_c^2 E$. As here $L_z = zL_1$ by (3.7), this gives the uncoupled pair of equations

$$c^2 = a^2 - \lambda b^2, \qquad w_c^2 = w_a^2 - \lambda w_b^2,$$

so that any pair (a, b, c) and (w_a, w_b, w_c) of solution triples to Pell's equation give a solution to (4.4).

It is interesting to note that as the kernel matrices commute in this instance, we can write

$$(L_a - \sqrt{\lambda} L_b)(L_a + \sqrt{\lambda} L_b) + n w_c^2 E = L_c^2 + n w_c^2 E = M(c, 0, w_c - mc)^2,$$

which factorises as

$$(L_a - \sqrt{\lambda} L_b + w_c E)(L_a + \sqrt{\lambda} L_b + w_c E) = L_c^2 + n w_c^2 E.$$

If $(L_a + \sqrt{\lambda} L_b + w_c E)^{-1}$ exists then

$$(L_a - \sqrt{\lambda} L_b + w_c E) = (L_a + \sqrt{\lambda} L_b + w_c E)^{-1}(L_c^2 + n w_c^2 E),$$

and using properties of matrix p-norms we have

$$\|(L_a - \sqrt{\lambda} L_b + w_c E)\|_p \leq \|(L_a + \sqrt{\lambda} L_b + w_c E)^{-1}\|_p \|(L_c^2 + n w_c^2 E)\|_p.$$

If all entries of a type S matrix are non-negative or non-positive, then its p-norm is equal to $n|w|$ (see Proposition 7.1 below). Hence if we choose $w_c > 0$ large enough to ensure that all entries of $(L_c^2 + n w_c^2 E)$ are non-negative, then we have the bound

$$\|(L_a - \sqrt{\lambda} L_b + w_c E)\|_p \leq n w_c \|(L_a + \sqrt{\lambda} L_b + w_c E)^{-1}\|_p.$$

Example 2. In this example we modify the definitions of $M(z, y, x)$ to obtain the $n \times n$ matrix $N(n, a)$ defined as

$$N(n, a) = (nI - J) \sum_{r=1}^{\left[\frac{n}{2}\right]} \left(na + \left[\frac{n}{2}\right] + 1 - r \right) A_r + \frac{2n(n+1)a + (n^2 - 1)}{2} E;$$

e.g. when $n = 7$ and $a = 2$,

$$N(7,2) = \begin{pmatrix} 119 & 270 & 15 & 264 & 9 & 258 & 17 \\ 32 & 120 & 271 & 16 & 265 & 24 & 224 \\ 225 & 33 & 121 & 272 & 31 & 231 & 39 \\ 40 & 226 & 34 & 136 & 238 & 46 & 232 \\ 233 & 41 & 241 & 0 & 151 & 239 & 47 \\ 48 & 248 & 7 & 256 & 1 & 152 & 240 \\ 255 & 14 & 263 & 8 & 257 & 2 & 153 \end{pmatrix}.$$

For fixed values of n, and working with the structure of the characteristic polynomials, it turns out that we can obtain general solutions for equations in $N(n, a)$. For example, when $n = 5$ the characteristic polynomial of $N(5, a)$ is given by $(150a - x + 60)\times$

$$\left(600(120 + 1200a + 4200a^2 + 6000a^3 + 3000a^4 + x^2 + 6ax^2 + 10a^2x^2) + x^4\right),$$

and the equation

$$N(5, a)^5 + bN(5, a)^3 + cN(5, a) = kE$$

has the general solution given by

$$a \in \mathbb{Z}, \qquad b = 600(1 + 6a + 10a^2),$$

$$c = 600(120+3000a+24000a^2+90000a^3+165000a^4+120000a^5-3ab-15a^2b-20a^3b),$$

and

$$k = 1800(85680 + 1068000a + 5316000a^2 + 13200000a^3 + 16335000a^4 + 8047500a^5$$
$$+26b + 198ab + 510a^2b + 445a^3b),$$

where the kernel matrix $L(5, a)$satisfies

$$L(5, a)^5 + bL(5, a)^3 + cL(5, a) = 0_n.$$

5 Existence of Type S Square Roots

In this section we shall use the concept of functional calculus for square matrices, which can be defined as follows.

Let N be an $n \times n$ matrix and Y its Jordan normal form, so

$$N = ZYZ^{-1},$$

where Z is a regular $n \times n$ matrix and Y is a block-diagonal matrix. More precisely, let $\lambda_1, \lambda_2, \ldots, \lambda_p$ be the eigenvalues of N, repeated according to *geometric* multiplicity; then there are $m_1, m_2, \ldots, m_p \in \mathbb{N}$ with $\sum_{i=1}^{p} m_i = n$ such that $Y = \mathrm{diag}(Y_1, Y_2, \ldots, Y_p)$, where the ith Jordan block takes the form

$$Y_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{m_i \times m_i}.$$

Now let f be a function which is defined on the spectrum $\sigma(N) = \{\lambda_i \mid 1 \le i \le p\}$ and holomorphic in an open neighbourhood of those points λ_i for which $m_i > 1$; we shall call such a function *suitable*.

Then $f(N)$ is defined as $f(N) = Zf(Y)Z^{-1}$, where

$$f(Y) = \mathrm{diag}(f(Y_1), f(Y_2), \dots, f(Y_p))$$

and

$$f(Y_i) = \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{f^{(m_i-1)}(\lambda_i)}{(m_i-1)!} \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix}$$

(see [6] page 3).

If f is a polynomial, then this definition coincides with the evaluation of the polynomial in the matrix algebra. Conversely, given the matrix N, a suitable function f could clearly be replaced with an interpolating polynomial which, along with the derivatives up to order $m_i - 1$, coincides with f at the points λ_i, $1 \le i \le p$.

Theorem 5.1. *Let N be an $n \times n$ type* S *matrix with weight w, and f a suitable function. Then $f(N)$ is type* S *with weight $\frac{1}{n} f(nw)$.*

Proof. A matrix satisfies condition (S1) if and only if the vector $v = (1, 1, \dots, 1)^T$ is an eigenvector of the matrix and of its transpose, with eigenvalue $\lambda = nw$. Hence we have $Nv = \lambda v$, so using the Jordan normal form for N, $Yy = \lambda y$, where $y = Z^{-1}v$. Thus y is an eigenvector of Y with eigenvalue λ, and it follows from the structure of Y that y has non-zero entries only in the top slot of each Jordan block for eigenvalue λ.

From the structure of $f(Y)$, it is then clear that y is an eigenvector of $f(Y)$ with eigenvalue $f(\lambda)$. Consequently,

$$f(N)v = Zf(Y)Z^{-1}v = Zf(Y)y = f(\lambda)Zy = f(\lambda)v,$$

so v is an eigenvector of $f(N)$ with eigenvalue $f(\lambda)$.

We also have $N^T v = \lambda v$, so considering that $N^T = (Z^{-1})^T Y^T Z^T$, we find that $Y^T x = \lambda x$, where $x = Z^T v$. From the structure of Y^T is follows that x has non-zero entries only in the bottom slots of the Jordan blocks for eigenvalue λ.

Hence, the structure of $f(Y)^T$ shows that x is an eigenvector of $f(Y)^T$ with eigenvalue $f(\lambda)$, so

$$f(N)^T v = (Z^{-1})^T f(Y)^T Z^T v = (Z^{-1})^T f(Y)^T x = f(\lambda)(Z^{-1})^T x = f(\lambda)v.$$

Therefore v is an eigenvector of both $f(N)$ and $f(N)^T$ with eigenvalue $f(\lambda)$, which shows that $f(N)$ is type S with weight $\frac{1}{n} f(nw)$. □

Corollary. *Let N be an $n \times n$ type* S *matrix with weight nw^2 which has the property that if* 0 *is an eigenvalue, then its algebraic and geometric multiplicities coincide. Then there are $2^{p'}$ distinct type* S *matrices M of weight w such that $M^2 = N$; here p' is the number of non-zero eigenvalues of N, repeated according to geometric multiplicity.*

Proof. The condition on the algebraic and geometric multiplicities of 0 ensures that the function f, defined to be 0 at 0 and either one of the two square roots in a neighbourhood of each non-zero eigenvalue of N, is suitable. Hence we can apply the preceding Theorem to obtain a square root matrix $M = f(N)$ with the required properties. As there are two possible choices for f in each of the p' neighbourhoods, there are $2^{p'}$ such square root matrices. $\square$

Remark. We call the square-root matrices obtained by the functional calculus *primary square roots*. There may be other, *non-primary square roots* of N which cannot be expressed as a function of N; this is the case if two or more Jordan blocks have the same eigenvalue. For example,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.1}$$

and

$$\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{5.2}$$

are both factorisations of type S matrices into non-type S square roots.

Theorem 5.2. *Let N be an $n \times n$ type* B *matrix with weight nw^2 and M a primary square root matrix of N. Then M is type* B *with weight w.*

Corollary. *Let N be an $n \times n$ type* B *matrix with weight nw^2 and M a type A square root matrix of N. Then M is a non-primary square root matrix of N.*

Proof. By Theorem 5.1, M is type S with weight w. A type S matrix is type B if and only if it commutes with J. Let $p(x)$ be the interpolating polynomial for the given primary square root of N, so that

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0,$$

and

$$p(N) = a_n N^n + a_{n-1}N^{n-1} + \ldots + a_1 N + a_0 I = \sqrt{N} = M.$$

Then

$$\begin{aligned} Jp(N) &= a_n JN^n + a_{n-1}JN^{n-1} + \ldots + a_1 JN + a_0 J = JM, \\ p(N)J &= a_n N^n J + a_{n-1}N^{n-1}J + \ldots + a_1 NJ + a_0 J = MJ \end{aligned}$$

as $JN = NJ$, it follows that $JN^r = N^r J$ for $r \in 1, \ldots, n$. Hence $JM = MJ$ and so M is type B. $\square$

Theorem 5.3. *Let N be an $n \times n$ type* B *matrix with weight nw^2, and let M be any $n \times n$ matrix such that $M^2 = N$. Let $\chi(x) = \sum_{j=0}^{n} \lambda_j x^j$ be the characteristic polynomial of M, and assume that*

$$\sum_{i=0}^{[(n-1)/2]} \lambda_{2i+1} n^{2i-1} w^{2i} \neq 0. \tag{5.3}$$

Then M is of type S *with weight w, and its type* A *and type* B *kernel matrices L_A, L_B anticommute. Moreover, for any natural number r*

$$M^r = \sum_{k=0}^{r} \sigma_{k,r-k} L_A^k L_B^{r-k} + n^{r-1} w^r E. \tag{5.4}$$

Remark. Here $\sigma_{k,l}$ $(k, l \in \mathbb{N}_0)$ are the (-1)-binomial coefficients arising in the binomial formula for anticommuting terms and defined by the recurrence

$$\sigma_{k,l} = \sigma_{k-1,l} + (-1)^k \sigma_{k,l-1}; \qquad \sigma_{0,1} = 1 = \sigma_{1,0}$$

(cf. sequence A051159 in [9]).

By Lemma 1.1 (2), (3), splitting the sum in (5.4) into the sums with odd and even k gives the type A and type B kernel matrices for M^r, resp. (cf. Lemma 2.3).

Note that the statement of the theorem is false in general if condition (5.3) is not satisfied, as seen in (5.1), (5.2).

Proof. By the Cayley-Hamilton theorem, $\chi(M) = 0_n$, so

$$-\sum_{i=0}^{[(n-1)/2]} \lambda_{2i+1} M^{2i+1} = \sum_{i=0}^{[n/2]} \lambda_{2i} M^{2i} = \sum_{i=0}^{[n/2]} \lambda_{2i} N^i,$$

a type B matrix with weight $u = \sum_{i=0}^{[n/2]} \lambda_{2i} w^{2i} n^{2i-1}$ (see (2.1)). Defining

$$U = \sum_{i=0}^{[n/2]} \lambda_{2i} M^{2i}, \qquad V = -\sum_{i=0}^{[(n-1)/2]} \lambda_{2i+1} M^{2i}, \tag{5.5}$$

we see that U has weight u and V has weight v, which is the negative of (5.3). Then $MV = U$, so $MVE = UE$ giving

$$ME = \frac{nu}{nv} E = \frac{u}{v} E,$$

and similarly, from $VM = U$,

$$EM = \frac{nu}{nv} E = \frac{u}{v} E.$$

Therefore M satisfies (S1) and thus is type S with weight $w = u/(nv)$.

By Lemma 2.3, M has a unique decomposition into a type A, a type B kernel matrix and a multiple of E, $M = L_A + L_B + wE$, and upon squaring we find

$$N = M^2 = L_A^2 + L_B^2 + L_A L_B + L_B L_A + n w^2 E;$$

as N is type B, as are L_A^2 and L_B^2, it follows from the uniqueness (Lemma 2.3) that the type A part $L_A L_B + L_B L_A$ vanishes, so

$$L_A L_B = -L_B L_A.$$

The representation (5.4) now follows from (2.1) by rearranging the factors in the power $L^r = (L_A + L_B)^r$ after expansion. $\square$

Remark. Since U and V, defined in (5.5), are type B, we have

$$MV = U = JUJ = JMVJ = JMJJVJ = JMJV,$$

so $(M - JMJ)V = 0_n$. Similarly we obtain $V(M - JMJ) = 0_n$. Thus $M - JMJ$ is a two-sided zero divisor of V; clearly it is the trivial zero divisor 0_n is M is type B.

Corollary. *In the situation of Theorem 5.3, let U, V be the type* B *matrices defined in (5.5), and L_A the type* A *part of the kernel matrix of M. Then*

$$VL_A = 0_n = L_A V.$$

If M is type A*, then U is a multiple of E.*

Proof. Writing $M = L_A + L_B + wE$ as in Lemma 2.3 (b), we have $U = MV = L_A V + L_B V + nwvE$, so comparing the (unique) type A parts of the kernel matrix of both sides, we conclude that $0_n = L_A V$. The other identity follows in the same way from $U = VM$.

If M is type A, then so is VM (by Lemma 1.1 (3)), and the last statement follows by Lemma 2.3. $\square$

To illustrate this corollary, consider the type A matrix

$$M(7,1,0) = (7I - J)\sum_{r=1}^{3}(4-r)A_r + 24E = \begin{pmatrix} 21 & 46 & 15 & 40 & 9 & 34 & 3 \\ 4 & 22 & 47 & 16 & 41 & 10 & 28 \\ 29 & 5 & 23 & 48 & 17 & 35 & 11 \\ 12 & 30 & 6 & 24 & 42 & 18 & 36 \\ 37 & 13 & 31 & 0 & 25 & 43 & 19 \\ 20 & 38 & 7 & 32 & 1 & 26 & 44 \\ 45 & 14 & 39 & 8 & 33 & 2 & 27 \end{pmatrix};$$

then

$$\sum_{k=0}^{3} \frac{7^{6-2k}}{7-2k}\binom{6-k}{2k} M(7,1,0)^{2k+1} = 168\sum_{k=0}^{3} \frac{7^{6-2k}}{7-2k}\binom{6-k}{2k} M(7,1,0)^{2k}$$

$$= 633317860933632E = 7 \times 8 \times 168^4(4^7 - 3^7)E,$$

where the antisymmetric type A part $M_a(7,1,0)$ and the symmetric type B part $M_s(7,1,0)$ of $M(7,1,0)$ are given by

$$M_a(7,1,0) = -J\sum_{r=1}^{3}(4-r)A_r, \qquad M_s(7,1,0) = 7\sum_{r=1}^{3}(4-r)A_r + 24E.$$

We conclude this section with a further observation on the characteristic polynomial of type S matrices.

Theorem 5.4. *Let N be an $n \times n$ type* S *matrix with weight w, kernel matrix L, and characteristic polynomial $\chi(x) = \sum_{j=0}^{n} \lambda_j x^j$. Then $\chi(wn) = 0$ and*

$$\chi(L) - \frac{\lambda_0}{n} E = 0_n.$$

Moreover, in terms of the type A *and type* B *parts of L^j,*

$$L_{A,j} = \sum_{k=0}^{[(j-1)/2]} \sigma_{2k+1,j-2k-1} L_A^{2k+1} L_B^{j-2k-1}, \qquad L_{B,j} = \sum_{k=0}^{[j/2]} \sigma_{2k,j-2k} L_A^{2k} L_B^{j-2k} \qquad (j \in \mathbb{N})$$

(cf. Theorem 5.3) and setting $L_{B,0} := I - \frac{1}{n} E$, we have

$$\sum_{j=1}^{n} \lambda_j L_{A,j} = 0_n, \qquad \sum_{j=0}^{n} \lambda_j L_{B,j} = 0_n.$$

Proof. By the Cayley-Hamilton theorem and Lemma 2.2 we have

$$0_n = \sum_{j=0}^{n} \lambda_j N^j = \lambda_0 I + \sum_{j=1}^{n} \lambda_j (L^j + n^{j-1} w^j E) = \lambda_0 \left(I - \frac{1}{n} E \right) + \sum_{j=1}^{n} \lambda_j L^j + \sum_{j=0}^{n} \lambda_j w^j n^{j-1} E.$$

Hence

$$\lambda_0 (I - \frac{1}{n} E) + \sum_{j=0}^{n} \lambda_j L^j = - \sum_{j=0}^{n} \lambda_j w^j n^{j-1} E,$$

and as the weight of the matrix on the left-hand side is 0, it follows that

$$\chi(L) - \frac{\lambda_0}{n} E = E \sum_{j=0}^{n} \lambda_j w^j n^{j-1} = 0_n.$$

Further,

$$\frac{\lambda_0}{n} E = \chi(L) = \lambda_0 I + \sum_{j=1}^{n} \lambda_j L^j = \lambda_0 I + \sum_{j=1}^{n} \lambda_j L_{A,j} + \sum_{j=1}^{n} \lambda_j L_{B,j},$$

and hence

$$\sum_{j=1}^{n} \lambda_j L_{A,j} = - \sum_{j=1}^{n} \lambda_j L_{B,j} - \lambda_0 I + \frac{\lambda_0}{n} E = - \sum_{j=0}^{n} \lambda_j L_{B,j};$$

the last statement now follows by Lemma 2.3. □

6 Dimensionality

The dimension of the vector space of magic squares has long been of interest and it was shown in 1959 by L. J. Ratcliff that the vector space of $n \times n$ semi-magic squares (type S matrices) is of dimension $n^2 - 2n + 2$ [3], [10]. If in addition to the $2n$ linear constraints on the rows and columns, we require that the two principal diagonals also sum to the row and column constant, then the total number of linear

	Type A	Type B	**Totals**
$n{=}2m{+}1$ odd	$\frac{(n-1)^2}{2}$	$\frac{(n-1)^2}{2}$	$(n-1)^2$
$n{=}2m$ even	$\frac{n(n-2)}{2}$	$\frac{n^2-2n+2}{2}$	$(n-1)^2$

Table 1: Vector space dimensions for kernel type A and type B, $n \times n$ matrices.

constraints is $2n+2$ and the corresponding dimension of the vector space is $n^2 - 2n$. This is the usual definition of a magic square, which we state here for completeness.

The dimension of the vector spaces of $n \times n$ kernel type A and type B matrices is given above for n odd or n even. One can obtain these dimensions diagrammatically and in the following two tables we give such examples for the dimension of the vector space of $n \times n$ type A matrices. In both tables, once the grey cells are chosen, then the white cells are determined, and in the second table the final step is to solve the equations relating to the four remaining cells a, b, c, d.

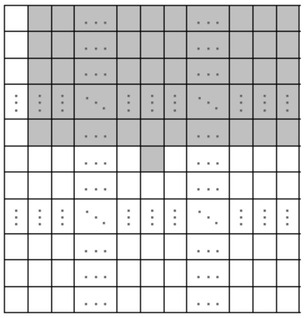

Table 2: Vector space dimension of $n \times n$ (n odd) type A matrices $= (n^2 - 2n + 3)/2$.

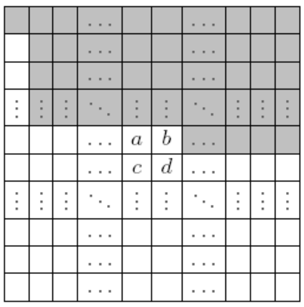

Table 3: Vector space dimension of $n \times n$ (n even) type A matrices $= (n^2 - 2n + 2)/2$.

It is easy to see that the bounds in Table 1 are exact, for if we omit one of the grey cells in either table then we are able to construct a matrix that is not of type A (unless we stipulate the kernel matrix condition of weight zero). For the kernel matrices the vector space dimension can be represented by removing the central grey cell in the first table and the top left grey cell in the second table. Hence the dimension for the kernel matrices is one less than that of the matrices with non-zero weight, as expected.

We can use similar arguments to obtain the dimensions of the vector spaces of kernel type A and type B matrices, with the extra condition that the matrix be either symmetric or anti-symmetric. For kernel type A symmetric/anti-symmetric we find respectively, that the second/leading principal diagonal must have each entry 0, and for kernel type B that both principal diagonals must have each entry 0. In each case, choosing the possible entries for the first row, determines the first column, nth row, and nth column. Similarly, the second row determines the second column and the $(n-1)$th row and column. Hence the counting argument proceeds in a concentric fashion ending up in the centre of the matrix. The resulting dimensions are given below in Table 4.

	Type A Sym	Type A A-sym	Type B Sym	Type B A-sym	**Totals**
$n{=}2m{+}1$ odd	$\frac{(n-1)^2}{4}$	$\frac{(n-1)^2}{4}$	$\frac{(n-1)(n+1)}{4}$	$\frac{(n-1)(n-3)}{4}$	$(n-1)^2$
$n{=}2m$ even	$\frac{n(n-2)}{4}$	$\frac{n(n-2)}{4}$	$\frac{n^2}{4}$	$\frac{(n-2)^2}{4}$	$(n-1)^2$

Table 4: Vector space dimensions for kernel type A and type B, symmetric and anti-symmetric $n \times n$ matrices.

In both Table 1 and Table 4, it can be seen that each row sums to $(n-1)^2$, which is the dimension of the vector space of kernel type S matrices. This confirms the results of Lemmas 2.1, 2.3 and 2.5, which together imply that every type S kernel matrix has a unique representation as a linear combination of the four types of matrix considered in Table 4.

The dimension of the set of all kernel $n \times n$ matrices is always one less than the set of weighted $n \times n$ matrices. A nice way of viewing this is to consider the group $\mathcal{M}$ of all non-singular weighted $n \times n$ matrices, so that for each matrix $M \in \mathcal{M}$ we define the mapping ϕ such that

$$\phi : M \to MI_L, \qquad \text{where} \qquad I_L = \left(I - \frac{1}{n}E\right). \tag{6.1}$$

Then $\phi : I \to I_L$, and $I_L I_L = I_L$. Using Lemma 2.2 we can write $M = L + wE$ so that

$$\phi(M) = (L + wE)\left(I - \frac{1}{n}E\right) = L + wE - \frac{1}{n}LE - \frac{w}{n}E^2 = L. \tag{6.2}$$

Hence ϕ maps M onto its kernel matrix and similarly it can be seen that $\phi(L) = L$.

As M is non-singular, there exists $M^{-1} \in \mathcal{M}$ with kernel matrix $L^{\sim 1}$, where $LL^{\sim 1} = L^{\sim 1}L = I_L$. If we define $\mathcal{N}$ to be the image of the set $\mathcal{M}$ under ϕ, then $\mathcal{N}$ is the set of all kernel matrices L that have a pseudo-inverse kernel matrix $L^{\sim 1}$, which satisfy $LL^{\sim 1} = I_L$. Under the normal definitions of matrix multiplication we see that $\mathcal{N}$ forms a group with identity I_L, as the inverse of $L_1 L_2^{\sim 1}$ is given by $L_2 L_1^{\sim 1}$, which is in $\mathcal{N}$ as $M_1 M_2^{-1}, M_2 M_1^{-1}$ are both in $\mathcal{M}$.

7 Matrix p-Norms

The operator p-norm of an $n \times n$ matrix $A \in \mathbb{R}^{n \times n}$ is defined by

$$\|A\|_p = \max_{\substack{x \in \mathbb{R}^n \\ \|x\|_p \leq 1}} \|Ax\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}, \tag{7.1}$$

where $\|x\|_p = \left(\sum_{j=1}^n |x_j|^p\right)^{\frac{1}{p}}$ $(x \in \mathbb{R}^n)$ is the usual vector norm defined for $p \geq 1$. For $p = \infty$, the vector norm is defined as $\|x\|_\infty = \max_{j \in \{1,\ldots,n\}} |x_j|$ $(x \in \mathbb{R}^n)$. It is well known that $\|A\|_1$ is the maximum column sum and $\|A\|_\infty$ the maximum row sum of the matrix of absolute values of A. For $p = 2$, there is the relationship $\|A\|_2 = \rho(A^T A)^{1/2}$, where ρ denotes the spectral radius, i.e. the maximal absolute value of an eigenvalue. For other values of p, no simple expressions of $\|A\|_p$ in terms of the entries of A are known.

An interesting observation due to N. J. Higham concerning the matrix p-norm of a semi-magic square is given in the following Proposition 7.1 (see page 115, Problem 6.4 of [4], and also [12]).

Proposition 7.1. *Let $M = (m_{ij})_{i,j \in \{1,\ldots,n\}}$ be a type* S *matrix with weight w, and let $p \in [1, \infty]$.*

(a) Then $\|M\|_p \geq n|w|$.

(b) If either all entries of M are non-negative or all entries of M are non-positive, then $\|M\|_p = n|w|$.

Proof. (a) Using the test vector $1_n := (1, 1, \ldots, 1)^T \in \mathbb{R}^n$, we obtain from (7.1) that

$$\|M\|_p = \max_{x \neq 0} \frac{\|Mx\|_p}{\|x\|_p} \geq \frac{\|M 1_n\|_p}{\|1_n\|_p} = \frac{\|n w 1_n\|_p}{\|1_n\|_p} = n|w|.$$

(b) By the Riesz-Thorin interpolation inequality (see Theorem IX.17 in [11]),

$$\|M\|_{p_t} \leq \|M\|_{p_0}^{1-t} \|M\|_{p_1}^{t}, \tag{7.2}$$

where $1 \leq p_0, p_1 \leq \infty$, $t \in [0, 1]$ and

$$\frac{1}{p_t} = \frac{t}{p_1} + \frac{1-t}{p_0}.$$

Specifically for $p_0 = \infty$, $p_1 = 1$ and $t = 1/p$, this gives

$$\|M\|_p \leq \|M\|_1^{\frac{1}{p}} \|M\|_\infty^{1-\frac{1}{p}}. \tag{7.3}$$

If the entries of M are non-negative or non-positive throughout, then $\|M\|_1$ and $\|M\|_\infty$ are both equal to the (constant) column and row sum $n|w|$, and it follows that $\|M\|_p \leq n|w|$.

□

The last result on the p-norm of non-negative type S matrices can be used to obtain estimates for the p-norm of kernel (i.e., weight 0) matrices L. The following is a simple upper bound on the basis of convexity.

Corollary. *Let $L = (l_{ij})_{i,j\in\{1,\dots,n\}}$ be a type S matrix with weight* 0 *and set*

$$w_+ = \max_{i,j\in\{1,\dots,n\}} l_{ij} \geq 0, \qquad w_- = \max_{i,j\in\{1,\dots,n\}} (-l_{ij}) \geq 0, \qquad w_0 = \min\{w_+, w_-\}; \tag{7.4}$$

furthermore, let $\alpha \geq 1$ be such that

$$\{w_+, w_-\} = \{w_0, \alpha w_0\}. \tag{7.5}$$

Then

$$\|L\|_p \leq n w_0 \left(\frac{\alpha}{1+\alpha} \, (1 + \alpha^{p-1}) \right)^{\frac{1}{p}} \tag{7.6}$$

for all $p \in [1, \infty)$.

Proof. For the matrix-valued function $L(x) = L - xE$ $(x \in \mathbb{R})$, Proposition 7.1 implies that

$$\|L(x)\|_p = n|x| \qquad (x \in \mathbb{R} \setminus (-w_-, w_+)).$$

For each fixed $u \in \mathbb{R}^n \setminus \{0\}$,

$$\|L(x)\, u\|_p^p = \sum_{j=1}^n |(Lu)_j - xn\overline{u}|^p \qquad (x \in \mathbb{R})$$

where $\overline{u} = \frac{1}{n}\sum_{j=1}^n u_j$, is a sum of convex functions, hence convex, and bounded above by $(n|x|)^p \|u\|_p^p$ on $\mathbb{R} \setminus (-w_-, w_+)$. Hence

$$\|L(x)\|_p \leq n \left(\frac{w_- w_+}{w_+ + w_-} \, (w_-^{p-1} + w_+^{p-1}) + x \, \frac{w_+^p - w_-^p}{w_+ + w_-} \right)^{\frac{1}{p}}.$$

Specifically for $L = L(0)$, this gives (7.6). □

To put this upper bound into perspective, we note that for *any* $n \times n$ matrix L, the trivial bound

$$\|L\|_p \leq n w_0 \alpha \tag{7.7}$$

holds, as can be seen by estimating each matrix entry by its absolute value and applying Hölder's inequality in

$$\|Lu\|_p^p = \sum_{j=1}^n \left| \sum_{k=1}^n l_{jk} u_k \right|^p.$$

In the special case $\alpha = 1$, in particular if L is a type A kernel matrix, (7.7) coincides with (7.6); moreover, the trivial bond is sharp for type A kernel matrices if $4 \mid n$.

Indeed, consider the $\frac{n}{2} \times \frac{n}{2}$ matrix

$$A = \begin{pmatrix} 1 & -1 & \cdots & 1 & -1 \\ -1 & 1 & \cdots & -1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & -1 & \cdots & 1 & -1 \\ -1 & 1 & \cdots & -1 & 1 \end{pmatrix};$$

then

$$L = \begin{pmatrix} A_n & -A_n \\ A_n & -A_n \end{pmatrix}$$

is a type A $n \times n$ matrix with weight 0 and $\|L\|_p \geq 2nw_0$, as can be seen using the test vector $u = (1, -1, \dots, 1, -1; -1, 1, \dots, -1, 1)$, where the semicolon separates the first $\frac{n}{2}$ entries from the rest.

Lemma 7.2. *Let $L = (l_{ij})_{i,j \in \{1,\dots,n\}}$ be a type* S *matrix with weight* 0 *and w_0 as in (7.4). Then*

$$\|L\|_1, \|L\|_\infty \leq (2n-2)w_0, \qquad \|L\|_2 \leq nw_0.$$

These inequalities are sharp.

Remark. The fact that $\|L\|_1$ and $\|L\|_\infty$ have the same sharp upper estimate in Theorem 7.2 does not mean that $\|L\|_1 = \|L\|_\infty$ in all cases; for example, the kernel matrix of Dürer's square (2.2) has $w_0 = 15/2$, $\|L\|_1 = 16$, $\|L\|_\infty = 24$. This example also shows that the above estimates are not equal to the norm in general.

Nevertheless, there is the following general symmetry between the norms for conjugate exponents (see [4] eq. (6.21)).

Lemma 7.3. *Let M be an $n \times n$ matrix and $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then $\|M\|_p = \|M^T\|_q$.*

Proof of Lemma 7.2. We use the matrix I_L defined in (6.1). By a straightforward calculation,

$$\|I_L\|_1 = \|I_L\|_\infty = 2 - 2/n.$$

Now by (6.2) we have $MI_L = L$, where we take $M = L \pm w_0 E$ (choosing the sign such that M has all non-negative or non-positive entries), and hence, by Proposition 7.1(b), we obtain

$$\|L\|_p = \|MI_L\|_p \leq \|M\|_p \|I_L\|_p \leq \left(2 - \frac{2}{n}\right) nw_0 = (2n-2)\, w_0 \tag{7.8}$$

for $p \in \{1, \infty\}$. For $p = 2$, we note that I_L is a symmetric projector (i.e. idempotent matrix) with characteristic equation $0 = \lambda(\lambda - 1)^{n-1}$, so that the spectral radius of $I_L^T I_L$ is just 1, and so $\|I_L\|_2 = 1$. Estimating as in (7.8), we find $\|L\|_2 \leq \|M\| = nw_0$. These estimates are sharp, as can be seen by taking $L = I_L$ and noting that $w_0 = 1/n$ in this case. $\square$

Comparing the upper estimates in the Corollary to Proposition 7.1 and Lemma 7.2, we see that in the case $p = 2$, the latter bound is always better. For $p = 1$ (and correspondingly $p = \infty$ in view of Lemma 7.3), however, it will depend on n and α which of the two upper bounds is smaller.

Taking the smaller upper bounds for $p \in \{1, 2\}$, using the interpolation inequality (7.2) with $p_0 = 2$, $p_1 = 1$, $t = \frac{2}{p} - 1$ for $p \in (1, 2)$ and the symmetry Lemma 7.3, we arrive at the following bound.

Theorem 7.4. *Let $L = (l_{ij})_{i,j\in\{1,\dots,n\}}$ be a type* S *matrix with weight* 0*, $p \in [1, \infty]$, and w_0, α as in (7.4), (7.5). Then, reading* 0 *for $\frac{1}{\infty}$,*

$$\|L\|_p \le n w_0 \left(\min\left\{ 2 - \frac{2}{n}, \frac{2\alpha}{1+\alpha} \right\} \right)^{\left|\frac{2}{p}-1\right|}.$$

The following theorem gives a lower bound for the p-norm.

Theorem 7.5. *Let $L = (l_{ij})_{i,j\in\{1,\dots,n\}}$ be a type* S *matrix with weight* 0*, w_0 as in (7.4), and $p \in [1, \infty]$. Then, reading* 0 *for $\frac{1}{\infty}$,*

$$\|L\|_p \ge 2w_0 \max\{n^{-\frac{1}{p}}, n^{\frac{1}{p}-1}\};$$

in the cases $p \in \{1, \infty\}$ this inequality is sharp.

Proof. Let $i, j \in \{1, \dots, n\}$ be such that $w_0 = -l_{ij}$, and consider the test vector $v = (1, 1, \dots, 1, -1, 1, \dots, 1)^T$, where the -1 is in the jth position. Then, using the fact that the ith row of L adds up to 0, we find that

$$(Lv)_i = \sum_{k\ne j} l_{ik} - l_{ij} = -2l_{ij} = 2w_0,$$

so $\|Lv\|_p \ge 2w_0$; and as $\|v\|_p = n^{\frac{1}{p}}$, this implies $\|L\|_p \ge 2n^{-1/p} w_0$. This is our lower estimate if $p \ge 2$; hence the case $p < 2$ follows by Lemma 7.3, noting that L^T is another type S matrix with weight 0.

If $p \in \{1, \infty\}$, the lower bound is sharp; indeed, the matrix

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

has $w_0 = 1$ and $\|L\|_1 = \|L\|_\infty = 2 = 2w_0$. □

There is a widening gap between the upper and lower bounds of Theorems 7.4 and 7.5 as the dimension n of the kernel matrix L grows. For example, for the kernel matrices of traditional magic squares, we have $w_0 = \frac{n^2-1}{2}$, so the upper bound grows like n^3 while the lower bound has growth between $n^{3/2}$ and n^2, depending on the value of p.

While we know that both estimates are best-possible for $p \in \{1, \infty\}$, one could hope for some improvement for the other values of p. We expect that there is some loss in the first inequality of (7.8), since the vector 1_n, which is a maximising eigenvector for $\|M\|_p$ by the proof of Proposition 7.1, is in the null space of I_L. Furthermore, Riesz-Thorin interpolation between $p = 2$ and $p \in \{1, \infty\}$ is very likely to overestimate the p-norm of I_L.

A possible approach to finding the p-norm of a kernel matrix L (or specifically the matrix I_L) is to look for critical vectors of the ratio

$$r(u) = \frac{\|Lu\|_p^p}{\|u\|_p^p},$$

which will include the globally maximising vector $\hat{u}$ such that $\|L\|_p = r(\hat{u})^{\frac{1}{p}}$. This will involve differentiation of the p-th power of the absolute value function; specifically,

$$\frac{d}{dx}|x|^p = p\, x^{(p-1)} \qquad (x \in \mathbb{R} \setminus \{0\}),$$

where $x^{(p)} := |x|^{p-1}x$ is the *signed p-th power* of x.

For any $u \in \mathbb{R}^n \setminus \{0\}$, the partial derivative of $r(u)$ with respect to u_j is

$$\begin{aligned}
\partial_j r(u) &= \partial_j \frac{\sum_{k=1}^n |\sum_{m=1}^n l_{km} u_m|^p}{\sum_{k=1}^n |u_k|^p} \\
&= \|u\|_p^{-2p} \left(\left(\sum_{k=1}^n p \left(\sum_{m=1}^n l_{km} u_m \right)^{(p-1)} l_{kj} \right) \|u\|_p^p - \|Lu\|_p^p \, p \, u_j^{(p-1)} \right) \\
&= p\|u\|_p^{-p} \left(\sum_{k=1}^n \left(\sum_{m=1}^n l_{km} u_m \right)^{(p-1)} l_{kj} - r(u)\, u_j^{(p-1)} \right).
\end{aligned}$$

Hence, if u is a (non-null) critical vector for r, then

$$0 = \sum_{k=1}^n l_{kj} \left(\sum_{m=1}^n l_{km} u_m \right)^{(p-1)} - r(u)\, u_j^{(p-1)} \qquad (j \in \{1, \ldots, n\}). \tag{7.9}$$

Summing the equations (7.9) over all j and using the fact that $\sum_{j=1}^n l_{kj} = 0$, we also find

$$\sum_{j=1}^n u_j^{(p-1)} = 0 \tag{7.10}$$

for any critical vector u such that $r(u) > 0$.

Specifically for the matrix I_L, a special case of a type S matrix with weight 0 which has $l_{km} = \delta_{km} - \frac{1}{n}$, the critical equations (7.9) take the form

$$0 = (u_j - \overline{u})^{(p-1)} - r(u)\, u_j^{(p-1)} - \frac{1}{n} \sum_{k=1}^n (u_k - \overline{u})^{(p-1)} \qquad (j \in \{1, \ldots, n\}), \tag{7.11}$$

where $\overline{u} := \frac{1}{n} \sum_{k=1}^n u_k$. Clearly, if $\overline{u} = 0$, then $I_L u = 0$ and therefore

$$r(u) = \frac{\|I_L u\|_p^p}{\|u\|_p^p} = 0, \tag{7.12}$$

and the ratio is homogeneous, so we can assume w.l.o.g. that $\overline{u} = -1$ in the following. Setting $c := \frac{1}{n}\sum_{k=1}^{n}(u_k+1)^{(p-1)}$ and $v_j := u_j^{(p-1)}$, we can then rewrite (7.11) in the form

$$\left(v_j^{(\frac{1}{p-1})}+1\right)^{(p-1)} = r(u)\, v_j + c \qquad (j \in \{1,\dots,n\}),$$

which shows that all v_j are abscissae of points where the straight line of slope $r(u)$ and intercept c crosses the graph of the function

$$g_p(x) = (x^{(\frac{1}{p-1})}+1)^{(p-1)} = \begin{cases} (x^{\frac{1}{p-1}}+1)^{p-1} & \text{if } x \geq 0 \\ (-(-x)^{\frac{1}{p-1}}+1)^{p-1} & \text{if } x \in [-1,0] \\ -((-x)^{\frac{1}{p-1}}-1)^{p-1} & \text{if } x \leq -1. \end{cases}$$

By considering the first and second derivatives of g_p, we can see the following properties, assuming $p > 2$ without loss of generality (see Lemma 7.3). The function g_p is strictly increasing throughout, with derivative $g_p'(x) \in (0,1)$ if $x < -1$. The function g_p is strictly concave in $(-\infty,-1)\cup(0,\infty)$ and strictly convex in $(-1,0)$. Moreover, $g_p(-1) = 0 = g_p'(-1)$; $g_p(0) = 1$ and $g_p'(0) = \infty$.

Further, we note that $\|I_L\|_p \geq 1$, since any vector orthogonal to 1_n is invariant under multiplication with I_L, so we are only interested in the case $r(u) > 1$.

Now we observe that a straight line of slope $r(u) > 1$ can intersect the graph of g_p in at most one point to the left of -1; that such a straight line which intersects the graph to the left of -1 does not intersect it at all in $[-1,0]$; and that any straight line intersects the graph of g_p in no more than two points to the right of 0, due to strict concavity.

From the fact that $\sum_{k=1}^{n} u_k = -n$ and from (7.10) we conclude that there must be some $v_j < -1$ and some $v_j > 0$. Hence we obtain the following theorem.

Theorem 7.6. *Let $p > 2$ and $u \in \mathbb{R}^n \setminus \{0\}$ be a critical vector for the ratio (7.12) with $r(u) > 1$. Then the set $\{u_j \mid j \in \{1,\dots,n\}\}$ has at least two and at most three elements.*

Remark. More precisely, if, in the situation of Theorem 7.6, additionally $\overline{u} = -1$, then $\{u_j \mid j \in \{1,\dots,n\}\}$ has exactly one negative (in fact, less than -1) and either one or two positive elements. (7.12) with $r(u) > 1$ and $\overline{u} := \frac{1}{n}\sum_{k=1}^{n} u_k = -1$.

Theorem 7.6 shows that the search for a maximising vector for $\|I_L\|_p$ only needs to consider critical vectors with either two or three different values. Moreover, due to the invariance of I_L under simultaneous identical permutations of its rows and columns, any permutation of the entries of a critical vector will give another critical vector with the same $r(u)$, so it only matters how often the two or three values are repeated in the vector.

In the two-value case, the situation is quite clear and can be summarised as follows.

Theorem 7.7. *If $p > 2$ and $u \in \mathbb{R}^n \setminus \{0\}$ is a critical vector for the ratio (7.12) with two different values, repeated k and $n-k$ times, where $k \in \{1, \dots, n-1\}$, then*

$$r(u)^{\frac{1}{p}} = f_p\left(\frac{k}{n}\right),$$

where

$$f_p(x) = ((1-x)^{q-1} + x^{q-1})^{\frac{1}{q}}((1-x)^{p-1} + x^{p-1})^{\frac{1}{p}} \qquad (x \in [0,1]),$$

$\frac{1}{p} + \frac{1}{q} = 1.$

Clearly $f_p(1-x) = f_p(x)$ $(x \in [0,1])$ and $f_p(\frac{1}{2}) = 1$, so it is sufficient to consider $k \in \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$. Theorem 7.7 gives the following lower bound for the p-norm of I_L.

Corollary. *For any $p \in (1, \infty)$ and $n \in \mathbb{N}$, $n \geq 2$,*

$$\|I_L\|_p \geq \max_{k \in \{1,\dots,n\}} f_p\left(\frac{k}{n}\right). \tag{7.13}$$

Proof of Theorem 7.7. Assume u has the two values a, repeated k times, and $-b$, repeated $n-k$ times, where $a, b > 0$. Then $\overline{u} = \frac{k}{n}\,a - \frac{n-k}{n}\,b$, and consequently I_L has the two values

$$\frac{n-k}{n}\,(a+b), \qquad -\frac{k}{n}\,(a+b),$$

repeated k and $n-k$ times, respectively.

From (7.10) we obtain that $k\,a^{p-1} = (n-k)\,b^{p-1}$, so

$$a = c\,(n-k)^{\frac{1}{p-1}}, \qquad b = c\,k^{\frac{1}{p-1}}$$

with some $c > 0$. Then

$$\|u\|_p^p = k\,a^p + (n-k)\,b^p = c^p(k\,(n-k)^{\frac{p}{p-1}} + (n-k)\,k^{\frac{p}{p-1}})$$

and

$$\|I_L\,u\|_p^p = (a+b)^p\,(\frac{k(n-k)}{n})^p\,(\frac{1}{k^{p-1}} + \frac{1}{(n-k)^{p-1}});$$

hence

$$\begin{aligned} r(u) &= \frac{(a+b)^{p-1}}{c^{p-1}}\,\frac{k^{p-1}(n-k)^{p-1}}{n^p}\left(\frac{1}{k^{p-1}} + \frac{1}{(n-k)^{p-1}}\right) \\ &= \frac{1}{n^p}\,\frac{(a+b)^{p-1}}{c^{p-1}}\,((n-k)^{p-1} + k^{p-1}) \\ &= \frac{1}{n^p}((n-k)^{q-1} + k^{q-1})^{p-1})\,((n-k)^{p-1} + k^{p-1}), \end{aligned}$$

where we have used

$$\frac{a+b}{c} = (n-k)^{\frac{1}{p-1}} + k^{\frac{1}{p-1}} = (n-k)^{q-1} + k^{q-1}$$

in the last step. $\square$

A lower bound for $\|I_L\|_p$, such as given in Corollary 7, is not helpful in (7.8), where an upper bound is required. Thus, critical vectors u with three different values will also need to be considered, as they may conceivably give rise to a higher ratio $r(u)^{\frac{1}{p}}$ than the maximum value on the right-hand side of (7.13).

Unfortunately, critical vectors with three different values are much harder to analyse than the two-value vectors of Theorem 7.7, where equation (7.10) can be used essentially to linearise the critical equations (7.11). Here we have a system of three non-linear equations which apparently does not allow explicit solution and has a variable number of solutions. Through analysis of this equation system and numerical experimentation, we discovered that in some situations critical vectors with three different values exist and may have a ratio $r(u)$ greater than that of any two-value critical vector with the same repeat count for the negative value; for example, when we take $p = 3$ and the vector $u \in \mathbb{R}^{27}$ which has entries 4.918 (repeated 6 times), 7.888 and -3.22 (repeated 20 times), then $r(u) = 1.15617$, while $f_3(\frac{7}{27})^3 = 1.15573$. However, in dimension $n = 27$, the maximum ratio, for $p = 3$, for two-value critical vectors will be $f_3(\frac{2}{27})^3 = 1.31476$. We do not know of any example of a three-value critical vector which would raise the norm $\|I_L\|_p$ above the maximum in (7.13).

This leads to the following **conjecture**:

$\|I_L\|_p = \max_{k\in\{1,\dots,n\}} f_p(\frac{k}{n})$ *for all* $p \in [1,\infty]$ *and* $n \in \mathbb{N}$, $n \geq 2$.

Acknowledgement

The authors would like to thank Prof N J Higham for his support and guidance in this problem, and Prof H Kalf for useful discussions.

References

[1] W. S. Andrews, *Magic Squares and Cubes*, Dover, 1960.

[2] J. Chernick, *Solution of the general magic square*, Amer. Math. Monthly, 45, 1938, 172-175.

[3] A. V. Den Essen, *Magic squares and linear algebra*, Amer. Math. Monthly, 97(1), 1990, 60-62.

[4] N. J. Higham, *Accuracy and Stability of Numerical Algorithms* (2nd Ed.), SIAM, 2002.

[5] N. J. Higham, *Estimating the matrix p-norm*, Numer. Math. 62, 1992, 539-555.

[6] N. J. Higham, *Functions of Matrices, Theory and Computation*, SIAM, 2008.

[7] M. C. Lettington, *Topics Related to the Theory of Numbers*, PhD Thesis, Cardiff University, 2008.

[8] M. C. Lettington, *Fleck's congruence, associated magic squares and a zeta identity*, Funct. Approx. Comment Math. (2) 45, 2011, 165-205.

[9] The On-Line Encyclopaedia of Integer Sequences, URL: http://oeis.org/A051159

[10] L. J. Ratcliff Jr, *The dimension of the magic square vector space*, Amer. Math. Monthly, (9) 66, 1959, 793-795.

[11] M. Reed and B. Simon, *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness*, Academic Press, San Diego 1975

[12] J. Stoer and C. Witzgall, *Transformations by diagonal matrices in a normed space*, Numer. Math. 4, 1962, 158-177.

[13] A. C. Thompson, *Odd magic powers*, Amer. Math. Monthly, (4) 101, 1994, 339-342.

School of Mathematics
Cardiff University
Senghennydd Road
Cardiff CF24 4AG
UK
email: LettingtonMC@cardiff.ac.uk; SchmidtKM@cardiff.ac.uk