The Electronic Journal of Linear Algebra.
A publication of the International Linear Algebra Society.
Volume 3, pp. 75-92, May 1998.
ISSN 1081-3810. http://math.technion.ac.il/iic/ela

ELA

PARTIAL CONVERGENCE AND SIGN CONVERGENCE OF MATRIX POWERS VIA EIGEN C–ADS*

FRANK UHLIG†

Dedicated to Hans Schneider, 70 + ϵ

Abstract. This paper investigates the convergence over the reals and in the sign $+, -, 0$ sense of individual elements of matrix powers A^k as $k \to \infty$ in dependence on the eigenvalues and left and right (generalized) eigenvector structure of a given matrix $A \in \mathbb{R}^{nn}$ theoretically and computationally.

Key words. Matrix powers, partial convergence, sign convergence, c–ad, dyad, matrix eigenstructure

AMS subject classifications. 40A99, 15A99, 15A18, 15A24, 15A48

1. Introduction. The convergence of matrix powers A^k has generally been treated in a global, norm–wise sense. Standard theoretical knowledge relates the spectral radius $r(A)$ of $A \in \mathbb{R}^{nn}$ to the convergence of the matrix powers A^k as follows.

If $r(A) > 1$, then $||A^k|| \to \infty$ and A^k diverges as $k \to \infty$.

$r(A) < 1$ if and only if $A^k \to O_{nn}$ as $k \to \infty$.

This global approach says little about the convergence behaviour of individual elements of A^k over $\mathbb{R}$ or in the $+, -, 0$ sign sense. For example for triangular matrices A nearly half of the entries of A^k never change with k. Thus many results may be discovered for individual matrix entry convergence of matrix powers in either of the two mentioned convergence senses.

After an introductory example we shall define some elementary tools and deduce a theoretical characterization of partial matrix power convergence over the reals, as well as for sign convergence of the entries of matrix powers. These characterizations involve the *eigen c–ads* of A, which we can prove to be unique for A.

Our introductory example involves the upper triangular matrix

$$A = \begin{bmatrix} 2 & 1 & 1 & -1 \\ 0 & -0.9 & -2.9 & 2.9 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

*Received by the editors on 2 November 1997. Final manuscript accepted on 21 April 1997. Handling editor: Daniel Hershkowitz.

†Department of Mathematics, Auburn University, Auburn, AL 36849–5310, USA (uhligfd@mail.auburn.edu).

A right column eigenvector matrix X for A is

$$X = \begin{bmatrix} 1 & \frac{1}{2.9} & -1 & -1 \\ -2.9 & 0 & 2.9 & 2.9 \\ 3.9 & 0 & 0 & 0 \\ 0 & 0 & 1.9 & 0 \end{bmatrix}, \quad \text{with} \quad Y = X^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{3.9} & 0 \\ 2.9 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{1.9} \\ 0 & \frac{1}{2.9} & \frac{1}{3.9} & -\frac{1}{1.9} \end{bmatrix}$$

being the corresponding left row eigenvector matrix for the identical decreasing ordering of the eigenvalues of A so that $YX = I_4$.
Since $YAX = D =$ diag$(3,\ 2,\ 1,\ -0.9)$ is diagonal, we can readily obtain the following explicit formula for the matrix powers of A,

$$A^k = X\ D^k\ Y =$$

$$(1) \qquad = \begin{bmatrix} 2^k & \frac{1}{2.9}(2^k - (-0.9)^k) & \frac{1}{3.9}(3^k - (-0.9)^k) & \frac{1}{1.9}(-1 + (-0.9)^k) \\ 0 & (-0.9)^k & \frac{2.9}{3.9}(-3^k + (-0.9)^k) & \frac{2.9}{1.9}(1 - (-0.9)^k) \\ 0 & 0 & 3^k & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus as $k \to \infty$,

$$A^k \to \begin{bmatrix} \infty & \infty & \infty & -\frac{1}{1.9} \\ 0 & \pm 0 & -\infty & \frac{2.9}{1.9} \\ 0 & 0 & \infty & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

i.e., all subdiagonal entries, the (2,2) entry and all entries in the last column of A^k converge as elements in $\mathbb{R}$, while in the $+, -, 0$ sign sense only the subdiagonal entries and those of the last column of A^k converge. Clearly the spectral radius $r(A) = 3$ makes A^k norm–wise divergent. We note that numerically the last column of A^k should become stationary at around $\begin{bmatrix} -0.52632 \\ 1.5263 \\ 0 \\ 1 \end{bmatrix}$. Matlab for example computes matrix powers of diagonalizable matrices A by evaluating $A^k = X{*}D^k{*}X^{-1}$, where X is a right eigenvector matrix for A and D is the diagonal eigenvalue matrix; see, e.g., [7, p. 294]. The theoretical stationary value of 1.5263... for the (2,4) entry is reached in Matlab at approximately A^{100}, when the value for the (1,4) entry has unfortunately grown to -10^{13}. In fact this (1,4) entry seems to approach its theoretical steady state value of –0.52632... from below for exponents up to around $k = 45$, after which the (1,4) entry increases slowly at first and diverges rapidly after $k = 64$.
This failed numerical convergence example involving such a simple matrix A seems to indicate that computationally the problem of the convergence of matrix powers is not too easily understood; see, e.g., [5] on floating point matrix power convergence or divergence for matrices A with $r(A) < 1$ and [3] for some of the history of our subject, as well as [8] for the related problem of evaluating matrix exponentials

Our theoretical analysis of the problem below shall be built on the Jordan Normal Form of real matrices and on dyads and their generalizations, called *c–ads* by us, built up from eigenvalues, left and right eigenvectors of A and their generalizations.

2. c–adic Eigenvector Expansion of Matrices. If $A \in \mathbb{R}^{nn}$ is real diagonalizable, i.e., if $A = XDY$ with $XY = YX = I_n$ and $D = \operatorname{diag}(\Lambda_1, ..., \Lambda_n) \in \mathbb{R}^{nn}$, then we may expand A^k dyadically for any exponent k (see, e.g., [4, problem 11.1.3, p. 545]) as follows,

$$\begin{aligned} A^k &= \begin{bmatrix} \vdots & & \vdots \\ x_1 & \dots & x_n \\ \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \Lambda_1^k & & \\ & \ddots & \\ & & \Lambda_n^k \end{bmatrix} \begin{bmatrix} \dots & y_1 & \dots \\ & \vdots & \\ \dots & y_n & \dots \end{bmatrix} \\ &= \sum_{i=1}^n \begin{bmatrix} \vdots \\ x_i \\ \vdots \end{bmatrix} \Lambda_i^k \left[\; \dots \quad y_i \quad \dots \; \right] \quad = \sum_{i=1}^n \Lambda_i^k \; x_i \; y_i \; . \end{aligned} \tag{2}$$

This formula gives a dyadic expansion of A^k in terms of the eigenvalue powers Λ_i^k and the right times left eigenvector dyads $x_i \; y_i$ of A for $i = 1, ..., n$. Looking again at our example, we have for its eigenvalue $\Lambda_1 = 3$ that

$$x_1 \; y_1 = \begin{bmatrix} 1 \\ -2.9 \\ 3.9 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{3.9} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{1}{3.9} & 0 \\ 0 & 0 & -\frac{2.9}{3.9} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Comparing this with our earlier explicit formula (1) for A^k, we note that the powers of Λ_1 affect only the entries in A^k corresponding to non–zero entries in the corresponding eigenvector dyad $x_1 \; y_1$, i.e., in the third column above its diagonal position with precisely the coefficients of the dyad. Likewise for $\Lambda_2 = 2$ we have

$$x_2 \; y_2 = \begin{bmatrix} \frac{1}{2.9} \\ 0 \\ 0 \\ 0 \end{bmatrix} \left[\; 2.9 \quad 1 \quad 0 \quad 0 \; \right] = \begin{bmatrix} 1 & \frac{1}{2.9} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and the powers of Λ_2 in the expansion (2) of A^k occur only in positions (1,1) and (1,2), again with their proper coefficients. The same holds for $\Lambda_3 = 1$,

$$x_3 \; y_3 = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{1.9} \\ 0 & 0 & 0 & \frac{2.9}{1.9} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and likewise for $\Lambda_4 = -0.9$.

As the powers of larger magnitude eigenvalues dominate those of lesser magnitudes in A^k as $k \to \infty$, we see the need to proceed from the maximum modulus eigenvalue of our example matrix A on down when describing the asymptotic behaviour (disregarding the dyadic coefficients for the moment) of

$$A^k \approx \begin{bmatrix} 2^k & 2^k & 3^k & -1/1.9 \\ & (-0.9)^k & -3^k & 2.9/1.9 \\ & & 3^k & 0 \\ & & & 1 \end{bmatrix}$$

for example.

Before formalizing this process in sections 2.1, 2.2 and 3 below, we note that our example matrix A was in some ways nongeneric: If B is any upper triangular matrix with distinct eigenvalues and the eigenvalue λ_m appears in B's m^{th} diagonal position, then the right and left eigenvectors x_m and y_m for B will have the general form $x_m = \begin{bmatrix} * \\ \vdots \\ * \\ \boxed{*} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow m \\ \\ \\ \\ \end{matrix}$ and $y_m = [0, \ldots, 0, \underset{\underset{m}{\uparrow}}{\boxed{*}}, *, \ldots, *]$, respectively, where the symbol $\boxed{*}$ denotes a strictly nonzero entry and the symbol $*$ denotes possibly nonzero entries. Thus the m^{th} eigenvector dyad for a generic upper triangular matrix B has the form

$$x_m\, y_m = \left[\begin{array}{ccc|cccc} 0 & \ldots & 0 & * & \cdots & \cdots & * \\ & \ddots & \vdots & \vdots & & & \vdots \\ & & 0 & * & & & \vdots \\ & & & \boxed{*} & * & \ldots & * \\ \hline & & & & 0 & \ldots & 0 \\ & & & & & \ddots & \vdots \\ & & & & & & 0 \end{array} \right] \begin{matrix} \\ \\ \\ \leftarrow m \\ \\ \\ \\ \end{matrix} ,$$

where the framed upper right block has dimensions $m \times n{-}m{+}1$. In our example, both eigenvector dyads for $\Lambda_1 = 3$ and $\Lambda_2 = 2$ of A are rather special, allowing structural zeros in their natural range of convergence dominance. These zero entries in $x_1\, y_1$ and $x_2\, y_2$ allow us to "see through" to the effect of the less dominant eigenvector dyad for $\Lambda_3 = 1$ in positions (1,4) and (2,4) of A^k. Precisely one of these special positions apparently is eventually "wiped out" in the numerical Matlab computations, hinting that the subtlety of the zero patterns of the dominant dyads $x_1\, y_1$ and $x_2\, y_2$ make our given matrix A possibly ill–conditioned for these computations.

Following the above heuristics, we shall now introduce c–ads and associated *pane matrices*, and study the uniqueness of the dyadic expansion (2) in full generality first. Thereafter we shall treat elementwise matrix power convergence for the four cases of eigenvalues of A having modulus larger, equal to or less than 1, or being equal to zero. Subsequently we shall deal with the question of elementwise sign convergence of matrix powers.

We can generalize the concept of a dyad as follows.

DEFINITION 2.1. If $x_1, ..., x_c$ are column vectors in $\mathbb{R}^n$ and $y_1, ..., y_c$ are row vectors in $\mathbb{R}^n$ for $c \geq 1$ and $a \in \mathbb{R}^{cc}$ is square, we call the matrix

$$X_c \cdot a \cdot Y_c = \begin{bmatrix} \vdots & & \vdots \\ x_1 & \dots & x_c \\ \vdots & & \vdots \end{bmatrix} \cdot a \cdot \begin{bmatrix} \dots & y_1 & \dots \\ & \vdots & \\ \dots & y_c & \dots \end{bmatrix} \in \mathbb{R}^{nn}$$

the **c–ad** generated by the vectors x_i and y_i and the $c \times c$ 'kernel matrix' a.

Our concept of c–ad generalizes that of a dyad, since any 1–ad in our notation is a standard dyad. In our context the 'kernel matrices' a of eigen c–ads will generally consist of either the identity matrix I_c or of eigenvalue matrices, Jordan blocks or Jordan chains of A.

DEFINITION 2.2. For any matrix $B = (b_{im}) \in \mathbb{R}^{nn}$ we define the **pane matrix** $P = (p_{im}) \in \{0, 1, -1\}^{nn}$ by setting $p_{im} = \begin{cases} 0 & \text{if } b_{im} = 0 \\ 1 & \text{if } b_{im} > 0 \\ -1 & \text{if } b_{im} < 0 \end{cases}$.

The term "pane matrix" was chosen to remind us of a latticed $n \times n$ window with n^2 individual window panes, one for each of its entries. For a pane matrix, "0" signifies a clear window pane, while a "± 1" denotes an opaque, possibly signed one. In our applications below, we shall introduce the notion of a "regular" zero in a Jordan c–ad generated matrix B for actually generating more meaningful "restricted" pane matrices P, see Section 2.2.

2.1. Uniqueness of c–adic Jordan Chain Expansions. Our purpose here is to show that for complete Jordan chain kernels $C(\lambda)$ used in the c–adic expansion of A, each c–ad $X \cdot C(\lambda) \cdot Y$ is unique for the given matrix A. This will help us study partial and sign convergence of matrix powers A^k in Section 3. It is clear that each Jordan chain $C(\lambda)$ is uniquely defined for every eigenvalue λ of A, once one has settled on one consistent form for Jordan blocks. Why the converse should be true does not seem obvious, and a simpler algebraic–geometric reasoning for this has eluded us so far.

We assume at first that λ is a real eigenvalue of a given matrix $A \in \mathbb{R}^{nn}$ with equal algebraic and geometric multiplicity. Such an eigenvalue shall be called *non–defective*. Let m be the dimension of the associated eigenspace, with $x_1, ..., x_m$ a basis for the right eigenspace and $y_1, ..., y_m$ the corresponding one for the left eigenspace for λ of A. With

$$X = \begin{bmatrix} \vdots & & \vdots \\ x_1 & \dots & x_m \\ \vdots & & \vdots \end{bmatrix}, \quad Y = \begin{bmatrix} \dots & y_1 & \dots \\ & \vdots & \\ \dots & y_m & \dots \end{bmatrix} \quad \text{and } YX = I_m,$$

we have $AX = X\Lambda$ and $YA = \Lambda Y$ for $\Lambda = \text{diag}(\lambda) \in \mathbb{R}^{mm}$, and thus $YAX = \Lambda = \lambda I$.

By enlarging the two corresponding partial eigenvector basis $\{x_j\}$ and $\{y_j\}$ for A to full left and right Jordan Normal Form basis of A, it follows that the powers A^k of A can be expressed as a sum of certain c–ads, much like formula (2) above. In our circumstances, this sum will contain one term of the form $X\ \Lambda^k\ Y = \lambda^k XY$. We will now show that the matrix product XY, used to generate A^k in (2), is unique for A, $\lambda \in \mathbb{R}$ and arbitrary m.

THEOREM 2.3. *If X and Y are any two corresponding column and row eigenvector matrices for a non–defective real eigenvalue λ of multiplicity $1 \leq m \leq n$ of $A \in \mathbb{R}^{nn}$ with $YX = I_m$, then their m–ad product matrix $XY \in \mathbb{R}^{nn}$ is uniquely determined.*

Proof. If $X_0 \in \mathbb{R}^{mm}$ and $Y_0 \in \mathbb{R}^{mm}$ are two other matrices comprised of corresponding column and row eigenvectors for $\lambda \in \mathbb{R}$ and A, then $\text{span}\{X_0\} = \text{span}\{X\}$ and $\text{span}\{Y_0\} = \text{span}\{Y\}$. Hence $X_0 = X \cdot \mathcal{A}$ for some nonsingular matrix $\mathcal{A} \in \mathbb{R}^{mm}$, and likewise $Y_0 = \mathcal{B} \cdot Y$ for some $\mathcal{B} \in \mathbb{R}^{mm}$ nonsingular.
By assumption $I_m = Y_0X_0 = \mathcal{B}YX\mathcal{A} = \mathcal{B}\mathcal{A}$, since $YX = I_m$. Thus since both $\mathcal{A}$ and $\mathcal{B}$ are square $m \times m$ matrices, $\mathcal{A}\mathcal{B} = I_m$ as well, making $X_0Y_0 = X\mathcal{A}\mathcal{B}Y = XY$ uniquely determined. ☐

If $\lambda = a + bi \notin \mathbb{R}$ is a complex and non–defective eigenvalue of $A \in \mathbb{R}^{nn}$ of multiplicity ℓ with $2 \leq 2\ell \leq n$, then there are ℓ linearly independent complex eigenvectors $z_j = p_j + iq_j$ with $p_j,\ q_j \in \mathbb{R}^n$, so that $Az_j = \lambda z_j$. This implies that the following real matrix equations hold for A and all $j = 1, ..., \ell$,

$$A \begin{bmatrix} \vdots & \vdots \\ p_j & q_j \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots \\ p_j & q_j \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

With the symbol Λ now interpreted as the 2×2 real matrix $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ for $\lambda = a + bi \notin \mathbb{R}$, we thus obtain by analogy from above for the corresponding real and imaginary parts 'eigenvector matrices' X and $Y(= X^{-1})$ that

$$YAX = \begin{bmatrix} \Lambda & & \\ & \ddots & \\ & & \Lambda \end{bmatrix}_{2\ell \times 2\ell}.$$

The ideas of Theorem 2.3 can be carried over to the complex non–defective eigenvalue case: The resulting 2ℓ–ad $X\text{diag}(\Lambda, ..., \Lambda)Y \in \mathbb{R}^{nn}$ in the analogous sum representation of A^k in (2) is unique as well for non–defective eigenvalues $\lambda \in \mathbb{C}$ of A of arbitrary multiplicity.

THEOREM 2.4. *If X and Y are any two corresponding real column and row eigenvector matrices for a non–defective nonreal eigenvalue $\lambda = a+bi$ of multiplicity ℓ with $1 \leq 2\ell \leq n$ of $A \in \mathbb{R}^{nn}$ with $YX = I_{2\ell}$, then their 2ℓ–ads $X\,diag(\Lambda, ..., \Lambda)Y \in \mathbb{R}^{nn}$ and XY are uniquely determined, where $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.*

Proof. The matrices

$$X = \begin{bmatrix} \vdots & \vdots & & \vdots & \vdots \\ p_1 & q_1 & \dots & p_\ell & q_\ell \\ \vdots & \vdots & & \vdots & \vdots \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} \dots & r_1 & \dots \\ \dots & s_1 & \dots \\ & \vdots & \\ \dots & r_\ell & \dots \\ \dots & s_\ell & \dots \end{bmatrix}$$

made up of the real and imaginary parts $p_i, q_i \in \mathbb{R}^n$ and $t_i, s_i \in \mathbb{R}^n$ of the complex right or left eigenvectors of A for λ are unique only up to right (and left) multiplication by nonsingular $2\ell \times 2\ell$ matrices $\mathcal{A}$ (and $\mathcal{B}$ for Y), which are composed of "complex type" 2×2 blocks of the form $\begin{bmatrix} x & y \\ -y & x \end{bmatrix} \in \mathbb{R}^{22}$ throughout. Note that such matrices naturally commute with $\mathrm{diag}(\Lambda, ..., \Lambda)$ for any real $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$; see, e.g., [9, Theorem 6]. With X_0 and Y_0 denoting any thus modified Jordan basis for Λ and since $YX = I$ implies $Y_0 X_0 = \mathcal{B}\mathcal{A} = I_{2\ell}$, we have that

$$\begin{aligned} X_0 \ \mathrm{diag}(\Lambda, ..., \Lambda)\, Y_0 &= X \ \mathcal{A} \ \ \mathrm{diag}(\Lambda, ..., \Lambda)\, \mathcal{B}\, Y \\ &= X \ \ \mathrm{diag}(\Lambda, ..., \Lambda)\, \mathcal{A}\mathcal{B}\, Y \\ &= X \ \ \mathrm{diag}(\Lambda, ..., \Lambda)\, Y, \end{aligned}$$

and clearly $X_0 Y_0 = X\mathcal{A}\mathcal{B}Y = XY$. ◻

We can summarize these two results as follows. For each *simple eigenvalue* $\Lambda \in \mathbb{R}$ or $\mathbb{C}$ of $A \in \mathbb{R}^{nn}$ of arbitrary multiplicity c there is a unique real c–ad (or 2c–ad in the complex case) generated by two sets of corresponding right and left, possibly generalized, real eigenvectors, that generate one term of the matrix power A^k in the expansion (2) as

$$\begin{bmatrix} \vdots & & \vdots \\ x_1 & \dots & x_d \\ \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \Lambda^k & & \\ & \ddots & \\ & & \Lambda^k \end{bmatrix} \begin{bmatrix} \dots & y_1 & \dots \\ & \vdots & \\ \dots & y_d & \dots \end{bmatrix},$$

where $\Lambda \in \mathbb{R}$ or $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in \mathbb{R}^{2,2}$, respectively, and $d = c$ if $\Lambda \in \mathbb{R}$ or $d = 2c$ if $\Lambda \notin \mathbb{R}$.

For *defective eigenvalues* of A, i.e., for those with differing algebraic and geometric multiplicities, the same qualitative results hold, as can easily be seen by rephrasing Theorems 2.3 and 2.4 for Jordan blocks instead of (possibly block) diagonal eigenvalue matrices. For example we have the following theorem.

THEOREM 2.5. *If X and Y are any two corresponding column and row principal vector matrices for a defective real eigenvalue λ of $A \in \mathbb{R}^{nn}$ of algebraic multiplicity $1 \le m \le n$ and geometric multiplicity 1 with $YX = I_m$, then their m–ads $XJ_m(\lambda)Y \in \mathbb{R}^{nn}$ and XY are uniquely determined, where $J_m(\lambda)$ denotes the m–dimensional Jordan block for λ.*

Proof. To find X with $YAX = J_m(\lambda)$ and $YX = I_m$, one usually starts with a nonzero highest order principal vector v_m with $(A-\lambda I)^{m-1}v_m \neq 0$ and $(A-\lambda I)^m v_m = 0$. From v_m one constructs $v_{m-1} = (A - \lambda I)v_m$, etc., and obtains the principal vector matrix X as $\begin{bmatrix} \vdots & & \vdots \\ v_1 & \dots & v_m \\ \vdots & & \vdots \end{bmatrix}$ for an upper triangular Jordan block $J_m(\lambda)$, and likewise for Y. As the geometric multiplicity of λ is assumed to be one here, v_m is unique up to scaling. Thus any other right principal vector matrix X_0 for λ has the form $X \cdot cI_m$. Likewise, any other left principal vector matrix Y_0 has the form $dI_m \cdot Y$, and $Y_0X_0 = I = YX$ makes $d = c^{-1}$. Thus $X_0J_mY_0 = XJ_mY$ and $X_0Y_0 = X\mathcal{AB}Y = XY$ are indeed unique. □

A similar argument carries through for a complex defective geometric multiplicity version of Theorem 2.4.

THEOREM 2.6. *If X and Y are any two corresponding real column and row principal vector matrices for a defective nonreal eigenvalue $\lambda = a + bi$ of geometric multiplicity one and algebraic multiplicity ℓ with $1 \leq 2\ell \leq n$ of $A \in \mathbb{R}^{nn}$ with $YX = I_{2\ell}$, then their 2ℓ–ads $XJ_{2\ell}(\Lambda)Y \in \mathbb{R}^{nn}$ and XY are uniquely determined, where $J_{2\ell}(\Lambda)$ denotes the real Jordan block with ℓ 2×2 blocks $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ on its main block diagonal and $\ell-1$ I_2 blocks on its upper block co–diagonal.*

Finally we need to consider defective eigenvalues with multiple Jordan blocks of dimensions greater than one for one eigenvalue λ of A.

We shall indicate how uniqueness of the c–ads made up of the principal left and right vectors for λ and its associated full Jordan chain $C(\lambda)$ can be proven. For brevity, we shall not state explicit analogues of our previous theorems 2.5 or 2.6, but rather just indicate a proof for the case of two Jordan blocks. This will be followed by Theorem 2.7, which is a summary statement on the uniqueness of a full Jordan chain c–ad expansion of A such as (2).

Assume that $\lambda \in \mathbb{R}$ is a real eigenvalue of A that is defective and whose Jordan chain contains precisely two Jordan blocks $J_m(\lambda)$ and $J_k(\lambda)$, $m \geq k$. There are two chains of principal column vectors $v_1, ..., v_m$ and $u_1, ..., u_k$ and two chains of principal row vectors $w_1, ..., w_m$ and $z_1, ..., z_k$ that obtain $J_m(\lambda)$ and $J_k(\lambda)$, respectively, from A, where v_1 and u_1, and w_1 and z_1 are the (only) eigenvectors in the lot. (The others are principal vectors of level j.)

As before, we need to consider the right and left principal vector matrices

$$X = \begin{bmatrix} \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ v_1 & \dots & v_m & u_1 & \dots & u_k \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} \dots & w_m & \dots \\ & \vdots & \\ \dots & w_1 & \dots \\ \dots & z_k & \dots \\ & \vdots & \\ \dots & z_1 & \dots \end{bmatrix}$$

for A and λ which satisfy $YX = I_{m+k}$. Note specifically the reverse ordering of the principal vectors in the left principal vector matrix Y here. This is necessary when using the upper triangular form Jordan blocks and when the index ℓ of w_ℓ and z_ℓ, etc., is to indicate vector level consistently.

Recall that the v_i and u_i are determined from $\mathrm{Ker}(A-\lambda I)^j$ for decreasing exponents j. If $m \geq k$, then any other two such Jordan reducing matrices $X_0 \in \mathbb{R}^{n,m+k}$ and $Y_0 \in \mathbb{R}^{m+k,n}$ for A and λ must have the form

$$X_0 = X \cdot H = \begin{bmatrix} \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ v_1 & \dots & v_m & u_1 & \dots & u_k \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} \cdot \begin{array}{c} \left[\begin{array}{ccc|ccc} c & & & e & & \\ & \ddots & & & \ddots & \\ & & \ddots & & & e \\ & & c & & & \\ \hline & & & d & & \\ & O & & & \ddots & \\ & & & & & d \end{array}\right] \end{array} \begin{array}{l} \\ \\ \leftarrow k \\ \leftarrow m \\ \\ \\ \\ \end{array}$$

$$\text{and} \qquad Y_0 = G \cdot Y = \left[\begin{array}{ccc|ccc} p & & & & & \\ & \ddots & & & O & \\ & & \ddots & & & \\ & & p & & & \\ \hline & r & & q & & \\ & & \ddots & & \ddots & \\ & & r & & & q \end{array}\right] \cdot \begin{bmatrix} \dots & w_m & \dots \\ & \vdots & \\ \dots & w_1 & \dots \\ \dots & z_k & \dots \\ & \vdots & \\ \dots & z_1 & \dots \end{bmatrix}$$

for nonzero coefficients c, d, p and q, and yet unspecified scalar entries e and r. To reduce A to partial Jordan form $J_m(\lambda) \oplus J_k(\lambda)$ via X_0 and Y_0, we must have

$$I_{m+k} (= YX) = Y_0 X_0 = GH = \left[\begin{array}{ccc|ccc} pc & & & pe & & \\ & \ddots & & & \ddots & \\ & & \ddots & & & pe \\ & & pc & & & \\ \hline rc & & & & & \\ & \ddots & & & * & \\ & & rc & & & \end{array}\right],$$

implying $e = r = 0$, and thus $p = c^{-1}$ and ultimately $d = q^{-1}$. Thus

$$X_0 \left(J_m(\lambda) \oplus J_k(\lambda) \right) Y_0 = X \left(J_m(\lambda) \oplus J_k(\lambda) \right) Y \quad \text{and} \quad X_0 Y_0 = XY,$$

since clearly $H \left(J_m(\lambda) \oplus J_k(\lambda) \right) G = J_m(\lambda) \oplus J_k(\lambda)$ and $HG = I_{m+k}$.

The ideas that we have detailed in the three proofs above completely suffice to establish the following result which itself will not be proved simply for fear of triple and quadruple indices.

ELA

THEOREM 2.7. *If $\Lambda \in \mathbb{R}$ or $\mathbb{C}$ is an eigenvalue of algebraic multiplicity ℓ for A and*

$$C(\Lambda) = J_{m_1}(\Lambda) \oplus \ldots \oplus J_{m_s}(\Lambda)$$

is the full Jordan chain associated with Λ, $\sum_{i=1}^{s} m_i = \ell$, where the Jordan blocks $J_{..}(\Lambda)$ are taken over the reals, as 2–block generalized Jordan blocks if necessary, then the ℓ–ads

$$X \cdot C(\Lambda) \cdot Y \quad \text{and} \quad X \cdot Y$$

are unique, no matter how the Jordan form producing eigen– or principal vectors making up X and Y have been chosen as long as $YX = I_\ell$.

On an historical note, eigen–dyads and their generalizations have apparently not entered much into the standard literature such as Gantmacher [2] or the monographs of Horn and Johnson [6]. However, they appear in the more numerically inclined literature such as in [4, p. 545] and in [1, p. 111, 170 ff, 273 ff and 307 ff], but not with as detailed an analysis as we have provided here. Of course, a different kind of dyadic matrix expansion is standard with the singular value decomposition of A.

2.2. Large Powers of Jordan Chain c–ads. The aim of this section is to study large powers

$$(XC_m(\lambda)Y)^k = X(C_m(\lambda))^k Y$$

of Jordan chain m–ads $XC_m(\Lambda)Y$ associated with a real matrix A_{nn}, where 'large' means $k \gg n$, in order to enable us to understand partial and sign convergence of matrix powers theoretically.

Case 1. If $\lambda = 0$ is an eigenvalue of A, then its associated Jordan chain $C(0)$ is nilpotent of order at most n. Hence $C^k(0) = 0$ for all relevant $k \gg n$ and its c–ad eventually does not contribute to A^k in the c–adic expansion analogous to (2). Thus we are left to consider Jordan chain c–ad powers for $\lambda \neq 0$ only.

Next we deal with non–defective nonzero eigenvalues $\Lambda \in \mathbb{R}$ or $\mathbb{C}$ of A, i.e., those eigenvalues of A for which the associated Jordan chain has the form $C_m(\Lambda) = \text{diag}(\Lambda, ..., \Lambda) \in \mathbb{R}^{mm}$ and $\Lambda \neq 0 \in \mathbb{R}$, or $C_{2\ell}(\Lambda) = \text{diag}(\Lambda, ..., \Lambda) \in \mathbb{R}^{2\ell,2\ell}$ and $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in \mathbb{R}^{22}$ if $\Lambda = a + bi \notin \mathbb{R}$.

Case 2. If $\Lambda = \lambda \in \mathbb{R}$, then $C_m(\Lambda) = \lambda I_m$ and $XC_m^k(\Lambda)Y = \lambda^k XY$. This Jordan chain m–ad will have a zero entry for all k precisely when there is a row in the right eigenvector matrix X that is orthogonal to a column of the left eigenvector matrix Y for λ and A. We will call such a zero in XY "regular". Note further that the sign pattern of $XC_m^k(\Lambda)Y$ will be equal to that of XY if $\lambda > 0$, and it will reverse itself for alternate powers k if $\lambda < 0$.
If $|\lambda| < 1$ then this m–ad will converge to zero, possibly alternating in sign in the nonzero positions of XY if $\lambda < 0$. If $\lambda = 1$, the m–ad powers are stationary for all k,

and if $|\lambda| > 1$, then the m–ad powers will diverge, possibly alternatingly (if $\lambda \leq -1$) in all nonzero positions of XY.

Case 3. If $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ for a non–defective eigenvalue $\lambda = a + bi \notin \mathbb{R}$ of A, then $m = 2\ell$ and

$$XC_m^k(\Lambda)Y = \begin{bmatrix} \vdots & \vdots & & \vdots & \vdots \\ p_1 & q_1 & \dots & p_\ell & q_\ell \\ \vdots & \vdots & & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \Lambda^k & & \\ & \ddots & \\ & & \Lambda^k \end{bmatrix} \begin{bmatrix} \dots & r_1 & \dots \\ \dots & s_1 & \dots \\ & \vdots & \\ \dots & r_\ell & \dots \\ \dots & s_\ell & \dots \end{bmatrix}$$

for the real and imaginary parts of the left and right generalized real eigenvector pairs r_i, s_i and q_j, p_j for λ of A, see Section 2.1. Note that each power $\Lambda^k = \begin{bmatrix} A_k & B_k \\ -B_k & A_k \end{bmatrix}$ is the 2×2 real matrix representation of $\lambda^k = A_k + B_k i \in \mathbb{C}$. If $B_1 := b \neq 0$ as assumed, then the set $\left\{\lambda^k\right\}_{k=1}^{\infty} \subset \mathbb{C}$ lies on a 'spiral' that tightens to 0 if $|\lambda| < 1$, or which loosens to ∞ if $|\lambda| > 1$, and it is contained on the unit circle in $\mathbb{C}$ otherwise. If $\lambda \notin \mathbb{R}$ and $\dfrac{\lambda}{|\lambda|}$ is a root of unity, then there are only finitely many, but at least four ratios for $\dfrac{A_k}{B_k}$ and all powers λ^k lie on the finitely many rays generated by these ratios through the origin of $\mathbb{C}$. Otherwise $\left|\left\{\dfrac{A_k}{B_k}\right\}\right| = \infty$. Associating the 2ℓ–ad as follows, $X\left(C_m^k(\Lambda)Y\right)$, we note that a partial right product

$$\begin{bmatrix} A_k & B_k \\ -B_k & A_k \end{bmatrix} \begin{bmatrix} \dots & r_h & \dots \\ \dots & s_h & \dots \end{bmatrix}$$

will have at least two linearly independent 2–vectors (the worst case happens if $\lambda = \alpha \cdot i$ with $\alpha \neq 0 \in \mathbb{R}$) in each column as $k \to \infty$, even if $\dfrac{\lambda}{|\lambda|}$ is a root of unity, which cannot all be perpendicular to every row in $\mathbb{R}^2$ of any $\begin{bmatrix} \vdots & \vdots \\ p_j & q_j \\ \vdots & \vdots \end{bmatrix}$, unless such a column or row is zero from the start. Due to the 'spiraling' in or out of the powers Λ^k, a zero can thus occur in position (h, j) of $XC_m^k(\Lambda)Y$ only if the leading part of a pair of rows of X and the matching trailing part of a pair of columns of Y is zero, i.e., if $p_{h,1} = q_{h,1} = \dots = p_{h,u} = q_{h,u} = 0 = r_{u+1,j} = s_{u+1,j} = \dots = r_{m,j} = s_{m,j} = 0$ for some index u. We again call such a zero occurrence in XY "regular", thus discounting zeros in XY that are obtained by mere orthogonality of rows and columns. With no such regular zero row and column pattern matching present in XY, all entries of the m–ad powers $XC_m^k(\Lambda)Y$ will converge to zero with a diverging sign pattern if $|\lambda| < 1$, $\lambda \notin \mathbb{R}$, and they will simply diverge otherwise.

Finally, we treat defective eigenvalues. If $J_m(\Lambda)$ is a real Jordan block for a real or complex eigenvalue $|\lambda| \geq 1$ of A_{nn}, then for the associated left and right, possibly generalized principal vector matrices $Y = (y_{ij})$ and $X = (x_{ij})$ with $YX = I_m$, the m–ad $XJ_m^k(\Lambda)Y$ will diverge to infinity as $k \to \infty$ in all positions, except in those positions (i,j) for which $x_{i,1} = \ldots = x_{i,p} = y_{p+1,j} = \ldots = y_{m,j} = 0$ for some index p, resulting in a "regular" zero. This follows readily from the known Toeplitz matrix expansion of

$$J_m^k(\Lambda) = \begin{bmatrix} \Lambda^k & k\Lambda^{k-1} & \frac{k(k-1)\Lambda^{k-2}}{2!} & \ldots & \ldots & \frac{k(k-1)\ldots(k-m+1)\Lambda^{k-m+1}}{(m-1)!} \\ & \ddots & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \ddots & \vdots \\ & O & & \ddots & \ddots & \frac{k(k-1)\Lambda^{k-1}}{2!} \\ & & & & \ddots & k\Lambda^{k-1} \\ & & & & & \Lambda^k \end{bmatrix}$$

for $k > m$ and $\Lambda = \lambda \in \mathbb{R}$, see e.g. [4, Theorem 11.1.1], which easily generalizes to $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in \mathbb{R}^{22}$ as well.

Case 4. If $|\lambda| \geq 1$, then the upper upper–triangular entries of $J_m^k(\Lambda)$ diverge in nondependent ways, so that the second m–ad factor $\left(J_m^k(\Lambda)Y\right)$ of $X\left(J_m^k(\Lambda)Y\right)$ will diverge in every position, unless a row in X has leading zeros that match the trailing zeros of a column of Y as described above and called "regular" there, in which case this zero entry will be preserved in all c–ad powers. As the off–diagonal entries in $J_m^k(\Lambda)$ become dominant in magnitude from the right upper corner $(1,n)$ element on in, there will be eventual sign convergence for $X\left(J_m^k(\Lambda)Y\right)$ from some exponent k_0 on in the positive real case $\Lambda \in \mathbb{R}$, and sign divergence for $\Lambda \notin \mathbb{R}$.

Case 5. If $|\lambda| < 1$, then clearly $J_m^k(\Lambda) \to O_{mm}$, with sign alternation in an eventually fixed alternating pattern if $-1 < \lambda < 0$ is real.

Clearly for complex Jordan chain m–ads, there can again be no sign pattern convergence. For the simpler global convergence alternatives of matrix powers; see, e.g., [5, Sect. 2, first paragraph].

3. Partial and Sign Convergence of Matrix Powers. Having established the uniqueness of the Jordan chain c–adic expansion that generalizes (2) for real matrices A and having studied the behaviour of large powers of such Jordan chain c–ads, we observe from Section 2.2 that in our context it is enough to study the "regular" zero–nonzero pattern of the unique matrices XY of the right and left generalized eigenvectors associated with each eigenvalue Λ of A for each full Jordan chain $C(\Lambda)$ to decide on elementwise convergence in A^k as $k \to \infty$. We shall use the concept of our pane matrices and the well known dominant eigenvalue behaviour to establish our main results on partial and sign matrix power convergence theoretically.

For this we shall order the eigenvalues λ_i of A_{nn} by decreasing magnitudes, where we set $r = n$ for nonsingular matrices A and define

$$\alpha_1 = |\lambda_1| = ... = |\lambda_k| > \alpha_2 = |\lambda_{k+1}| \geq ... \geq |\lambda_r| > \lambda_{r+1} = ... = \lambda_n = 0.$$

To each set of equal magnitude eigenvalues $|\lambda_{..}| = \alpha \neq 0$ of A we shall associate one restricted pane matrix. It shall be the superposition of the restricted pane matrices, where a pane entry of zero derives from a "regular" zero only, for the generalized eigen–dyads XY associated with all Jordan chain c–ads for eigenvalues $\lambda_{..}$ of the same magnitude α of A. We shall then arrange these Jordan chain–derived pane matrices in order of their decreasing eigenvalue magnitudes. This allows us to interpret the $n \times n$ panes from the front, or from the largest magnitude eigenvalues of A on down. If a pane is 'opaque' for the first time, i.e., if a restricted pane matrix entry contains a 1 or –1 for the largest magnitude $\alpha_i < \alpha_1$ eigenvalues only, then the entry for the powers of A^k at that position will behave as warranted by the combination of associated maximum modulus Jordan chain c–ad powers for the first dominant magnitude level α_i from Section 2.2.

If one maximum modulus pane is 'clear' for α_1, i.e., if its entry is a "regular" zero for all eigenvalues of A of maximum modulus α_1, then we can look through this pane to the next in magnitude set of eigenvalue pane matrices for A that generate an opaque pane at that position. The eigenvalues and eigen–dyads with the first magnitude opaque pane then govern the partial convergence of A^k at that entry.

Note that a pane can remain 'clear' for the whole set of restricted pane matrices and that superimposed 'opaque' panes for identical magnitude eigenvalues need special care which we shall not address completely here, except by example as follows. If $\alpha_1 = |\lambda_1| = |\lambda_2|$ and $\lambda_1 \neq \lambda_2, \overline{\lambda_2}$ for A, then there are two unique Jordan chain c–ads $X_1 C_m(\lambda_1) Y_1$ and $X_2 C_\ell(\lambda_2) Y_2$ for α_1 that may 'cloud' the pane matrix for A^k jointly, depending on their mutual interaction. For example the matrix $A = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$ has the same magnitude eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -2$ with easily computed right and left eigenvalue 1–ads $\frac{1}{2}\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ and $\frac{1}{2}\begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix}$, respectively. Their powers will alternately add up to a diagonal matrix with equal diagonal entries or to a counterdiagonal matrix with a one in the left lower corner, which conform for alternating powers with A^k.

We shall now state our first main result without proof, which follows readily from the above, and then give several illustrative examples in the next section.

THEOREM 3.1. (Partial Convergence of Matrix Powers) *For $A \in \mathbb{R}^{nn}$ the (i, j) entry $a_{ij}^{(k)}$ of A^k converges to a nonzero real number a as $k \to \infty$ if and only if the following three conditions hold.*

1. *1 is an eigenvalue of A with at least one one–dimensional Jordan block in its associated Jordan chain $C(1)$,*
2. *the restricted pane matrices for all Jordan chain c–ads XY of A with $|\Lambda| \geq 1$, $\Lambda \neq 1$, $\Lambda \in \mathbb{R}$ or $\mathbb{C}$, and the restricted pane matrix for the partial Jordan chain c–ad comprised of all Jordan blocks of dimensions larger than one in $C(1)$ are "regularly" clear at position (i, j), and*
3. *the (i, j) entry of the partial Jordan chain c–ad XY of A generated by the*

one–dimensional Jordan blocks in $C(1)$ *has the value* $a \neq 0$.
The (i,j) *entry* $a_{ij}^{(k)}$ *of* A^k *converges to zero as* $k \to \infty$ *if and only if all Jordan chain c–ads* XY *of* A *for eigenvalues* $\Lambda \in \mathbb{R}$ *or* $\mathbb{C}$ *with* $|\Lambda| \geq 1$ *are "regularly" clear at position* (i,j).

We have a simple corollary.

COROLLARY 3.2. *If one entry in* A^k *converges to a nonzero value* a *as* $k \to \infty$, *then 1 is an eigenvalue of the given matrix* $A \in \mathbb{R}^{nn}$.

Regarding partial sign convergence of matrix powers, we recall from Section 2.2 that dominant complex opaque panes generated by real Jordan chain c–ads for complex eigenvalues cannot sign converge, while those generated by real eigenvalue Jordan chain c–ads will eventually sign converge if the pane is dominated in its c–ad by a positive eigenvalue and will alternately sign converge otherwise. Here we call a position in A^k *sign convergent*, if its entries eventually all have one sign + or –, and *alternate sign convergent* if its entries will eventually alternate in sign with each subsequent power of A. Similarly *sign convergence to zero* for one entry of A^k means that all entries in that position will eventually become stationary at zero as $k \to \infty$. This last notion clearly differs from our notion of partial convergence to zero in being a much more stringent requirement. Again we recuse ourselves from a detailed proof, but rather state our main result and refer to the analysis of the previous section and to the examples below.

THEOREM 3.3. (Sufficient Conditions for Sign Convergence of Matrix Powers) *For* $A \in \mathbb{R}^{nn}$ *the* (i,j) *entry* $a_{ij}^{(k)}$ *of* A^k *converges in the* $+,-,0$ *sign sense as* $k \to \infty$ *if the following conditions hold.*

1. *for all complex eigenvalue Jordan chain c–ads* XY *of* A, *the* (i,j) *pane is "regularly" clear, and*

2a. *for sign convergence to zero, all restricted pane matrices associated with Jordan chain c–ads* XY *for nonzero real eigenvalues of* A *are "regularly" clear at position* (i,j), *or*

2b. *for sign convergence to + or –,*
- (i) *the largest magnitude eigenvalue* λ *of* A *is real positive and non–defective and the* (i,j) *position in its Jordan chain c–ad* XY *is positive (for sign convergence to +) or negative (for –), or*
- (ii) *the largest magnitude real eigenvalue* λ *of* A *is positive and defective and the* (i,j) *position in its Jordan chain c–ad powers* $XC(\lambda)^kY$ *becomes sign stationary at + or at –.*

Conditions 1. and 2a. are clearly necessary as well for sign convergence to zero. In fact we have as a corollary that an entry in A^k sign converges to zero, i.e., it equals zero for all powers A^k for $k \geq k_0$, if and only if 1. and 2a. hold and k_0 is the dimension of the maximal Jordan block for the eigenvalue zero of A; see our introductory example.

In case of +,– sign convergence, various troublesome compensations with ± Jordan chain c–ads might happen for dominant real eigenvalues of opposite signs that we shall not sort out here. Similarly we do not investigate alternating +,– sign convergence here, although the results are within reach if, for example, $\alpha_1 = \max|\lambda_i| = -\lambda_1$

is the only eigenvalue of A of this maximal magnitude.

The above results clearly show how sensitive partial or sign convergence is for matrix powers. In essence, entries of A^k either converge to zero or ∞ in magnitude. Any other elementwise convergence of matrix powers is only possible if A has the eigenvalue 1 with proper multiplicities and if all Jordan chain c–ads of A cooperate to show this eigenvalue 1 c–ad off properly. This, undoubtedly, is the cause of so much numerical trouble in the computed examples above and below, and in [5].

4. Numerical Examples for Partial and Sign Convergence of Matrix Powers. While the spectral approach of Jordan chain c–ads has proven valuable to determine partial matrix convergence theoretically in this paper, it has nonetheless been found computationally at best dubious in our introductory triangular example. Computing all eigenvalues and eigenvectors of a diagonalizable real $n \times n$ matrix A costs around 25 n^3 operations with the QR algorithm; see, e.g., [4, p. 380]. Evaluating A^k as $X \operatorname{diag}(\lambda_i^k) Y$ will thus take around 26 n^3 operations since one $n \times n$ matrix multiplication takes n^3 operations if done the old fashioned way. The eigenvalue rounding error and ill–conditioning effects inherent in finding the eigenvalue and eigenvectors of A are well documented as they affect the matrix exponential, e.g., in [8].

For our investigation these problems need not affect us at all, however, if we just multiply A k times with itself to reach A^k, though other rounding and truncation errors would be committed in finite precision arithmetic. Such a method is the one chosen for the analysis in [5]. A much better economy can be obtained by using *binary powering* instead. If the exponent k is expressed as $k = \sum_{i=1}^{m} \beta_i 2^i$ with $\beta_i = 0$ or 1 and $\beta_m = 1$, then the binary powers of A are computed as $A^2 = A * A$, $A^4 = A^2 * A^2$, etc., until $A^{2^m} = A^{2^{m-1}} * A^{2^{m-1}}$. For any $k \leq 16,383 = 2^{14} - 1$, e.g., one would need to compute at most twelve powers of two of A involving 12 n^3 operations, and at most twelve subsequent matrix products to obtain A^k from A and the binary expansion of k. This approach thus weights in at worst at ≤ 24 n^3 operations for all exponents $k < 2^{14}$. The operations count would be much less on average since most $k \ll 2^{14}$ and the average binary expansion of k will have around half its coefficients $\beta_i = 0$; see, e.g., [4, Ch. 11.2.5].

Note finally that each of the above matrix power algorithms has quite different numerical convergence properties in floating point arithmetic; see, e.g., [5].

In our examples below, we have relied on a binary powering scheme to evaluate A^k instead of using the Matlab function '$A \wedge k$' or multiplying one factor at a time.

Example 4.1. The matrix $A = \begin{bmatrix} -5 & 8 & 32 \\ 2 & 1 & -8 \\ -2 & 2 & 11 \end{bmatrix}$ has non–defective eigenvalues 1 and 3, with 3 double. The complete eigen–dyads XY for the two eigenvalues 1 and 3 of A are

$$E(1) = \begin{bmatrix} 4 & -4 & -16 \\ -1 & 1 & 4 \\ 1 & -1 & -4 \end{bmatrix} \quad \text{and} \quad E(3) = \begin{bmatrix} -9 & 12 & 48 \\ 3 & 0 & -12 \\ -3 & 3 & 15 \end{bmatrix},$$

respectively. Here the dominant pane matrix for A has a regularly clear pane at the

(2,2) position and the (2,2) entry of A^k must remain equal to 1 from the eigenvalue 1 dyad $E(1)$ for all powers of A. Computationally, the (2,2) entry of A^k is 1 for all exponents $k \leq 32$ in both the Matlab and the binary power algorithms. For all $k > 32$ it becomes 0 due to the fact that the real number 1 falls below the machine constant in double precision relative to the other entries of A^k. In fact, the powers of $E(1)$ all remain stationary, while for $E(3)^k$ the (2,2) entry jumps to 12 for $k = 33$ in Matlab powering and to 13 in binary powering from its 'numerically more correct' entry of 0. Finally, the computed large powers A^k keep the sign pattern of $E(3)$, except for the troublesome (2,2) spot.

We note that different implementation of floating point arithmetic (IEEE compliant or non–compliant) will generally affect all matrix power computations differently.

EXAMPLE 4.2. The matrix $A = \begin{bmatrix} 1 & 4 & -2 \\ 4 & -3 & 3 \\ 8 & -12 & 9 \end{bmatrix}$ has one non–defective eigenvalue 1 and one defective eigenvalue 3. The complete eigen–dyads XY for the two eigenvalues 1 and 3 of A are

$$E(1) = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ -4 & 8 & -4 \end{bmatrix} \quad \text{and} \quad E(3) = \begin{bmatrix} 0 & 6 & -3 \\ 6 & -7 & 5 \\ 12 & -20 & 13 \end{bmatrix},$$

respectively. The dominant pane matrix for A has a regularly clear pane at the (1,1) position and hence A's (1,1) entry must remain equal to 1 from the eigenvalue 1 dyad $E(1)$ for all powers of A. Computationally, the (1,1) entry of A^k is 1 for all exponents $k \leq 32$ in both the Matlab and the binary power algorithms. For all $k > 32$ it should be computed as 0 due to the fact that the real number 1 falls below the machine constant in double precision relative to the other entries of A^k. In Matlab powering, the (1,1) entry of A^k stays zero for $k = 33$ and 34 only, but then increases from 16 for $k = 35$ to $4.5\ 10^4$ and $5.4\ 10^9$ for $k = 40$ and 50, respectively. In binary powering the (1,1) entry remains 'numerically correct' at 0 for all powers $k \geq 33$. In terms of sign convergence, $E(3)^8$ is the first power of $E(3)$ that exhibits its limiting sign distribution $\begin{bmatrix} 0 & + & - \\ + & + & - \\ + & + & - \end{bmatrix}$ and all powers from A^8 on exhibit that same sign distribution, except of course for the troublesome (1,1) entry of A^k for $k > 34$ via Matlab powering.

EXAMPLE 4.3. Our last example deals with the matrix $A = \begin{bmatrix} -5 & 3 & -3 \\ 6 & 4 & 0 \\ 18 & 3 & 4 \end{bmatrix}$ which has one real eigenvalue 1 and one complex conjugate eigenvalue pair $1 \pm 3i$. The two complete real eigen–dyads XY for the eigenvalues $1 \pm 3i$ and 1 of A are $E(1 \pm 3i) = \begin{bmatrix} -6 & 5 & -4 \\ 8 & 0 & 2 \\ 22 & -5 & 8 \end{bmatrix}$ and $E(1) = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ -4 & 8 & -4 \end{bmatrix}$, respectively. Here the dominant eigenvalue pane matrix for A has a clear pane at the (2,2) position which is not regularly zero. Consequently the (2,2) entry cannot remain zero for the powers of $E(1 \pm 3i)$ and thus of A^k. Computational tests bear this out; the entries of A^k all diverge. Likewise, the sign pattern in A^k does not appear to settle down at all.

This example leads us to the following open questions.

Open Question 4.4. *There are $2^9 = 512$ possible $\pm$ sign patterns for the matrix powers A^k of a 3×3 matrix A. Do all of these actually occur in the sequence $\{A^k\}$ in Example 4.3, and analogously for general non–sign converging matrix powers? Do the sign patterns occurring in $\{A^k\}$ have equal probability? If any of the 512 possible sign patterns are excluded, which ones and why?*

Open Question 4.5. *For a given sign pattern, say that of the original A in Example 4.3, what is the sequence of exponents k_i for which the sign pattern of A^{k_i} and that of A coincide? Likewise for A^2 etc. What number theoretic relations hold for those integer sequences k_i that give A^{k_i} one fixed sign pattern?*

5. Outlook. This paper has dealt with elementwise convergence issues of real matrix powers. Further results can be obtained by specializing to powers of structured matrices, such as to real symmetric, orthogonal, nonnegative, stochastic, positive, etc., matrices. In many ways the ideas underlying this paper can shed new light on old standards such as on stochastic matrices.

For example [1, Satz 2, p. 348] contains the following result and a one page proof for it.

Theorem 5.1. *If A is a nonnegative indecomposable matrix with constant column sum 1 and eigenvalues $\lambda_1 = 1 > |\lambda_i|$ for $i = 2, ..., n$, then the powers A^k converge to the matrix X, comprised in each column of the normalized (column sum equal to 1) right eigenvector of $\lambda_1 = 1$ for A*

This follows instantly from (2) and our matrix power convergence results for the dominant right eigenvector x, properly normalized, of A, since the corresponding left eigenvector for A is (1,...,1).

Or on nonnegative matrices, we have the following reesult.

Theorem 5.2. *If $A^k > 0$ for all $k > k_0$ and A has a single dominant real eigenvalue $|\lambda| = r(A)$, then $\lambda > 0$ with elementwise positive right and left eigenvectors x and y.*

Using (2) on such a matrix A and λ makes $A^k \approx \lambda^k xy$ elementwise positive. As A^k sign converges to $+$ in every position, $\lambda > 0$ by necessity and the corresponding right and left eigenvectors x and y cannot have elements equal to zero, nor elements of opposite signs.

REFERENCES

[1] S. Falk and R. Zurmühl, *Matrizen und ihre Anwendungen 1*, Seventh ed., Springer, Berlin, 1997.

[2] F. R. Gantmacher, *Theory of Matrices*, vols. 1, 2, Chelsea, New York, 1959.

[3] W. Gautschi, The asymptotic behaviour of powers of matrices, *Duke Math. J.*, 20:127-140 (1953).

[4] G. H. Golub and C. F. Van Loan, *Matrix Computations*, second ed., Johns Hopkins U. Press, Baltimore, 1989.

[5] N. J. Higham and P. A. Knight, Matrix powers in finite precision arithmetic, *SIAM J. Matrix Anal. Appl.*, 16:343-358 (1995).

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U. Press, Cambridge, 1985; *Topics in Matrix Analysis*, Cambridge U. Press, Cambridge, 1991.
[7] *The Student Edition of Matlab*, Version 4, User's Guide, Prentice Hall, Englewood Cliffs, 1995.
[8] C. Moler and C. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, *SIAM Review*, 20:801-836 (1978).
[9] D. A. Suprunenko and R. I. Tyshkevich, *Commutative Matrices*, Academic Press, London and New York, 1968.