Proceedings of the Royal Society of Edinburgh, **132A**, 1307–1331, *2002*

Nonlinear eigenvalue–eigenvector problems for STP matrices

Uri Elias and **Allan Pinkus**
Department of Mathematics, Technion, IIT, Haifa 32000, Israel

(MS received 12 July 2001; accepted 6 February 2002)

Let A_i, $i = 1, \dots, m$, be a set of $N_i \times N_{i-1}$ strictly totally positive (STP) matrices, with $N_0 = N_m = N$. For a vector $\boldsymbol{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$ and arbitrary $p > 0$, set

$$\boldsymbol{x}^{p*} = (|x_1|^p \operatorname{sgn} x_1, \dots, |x_N|^p \operatorname{sgn} x_N).$$

We consider the eigenvalue–eigenvector problem

$$A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*} = \lambda \boldsymbol{x}^{r*},$$

where $p_1 \cdots p_{m-1} = r$. We prove an analogue of the classical Gantmacher–Krein theorem for the eigenvalue–eigenvector structure of STP matrices in the case where $p_i \geqslant 1$ for each i, plus various extensions thereof.

1. Introduction

A matrix A is said to be strictly totally positive (STP) if all its minors are strictly positive. STP matrices were independently introduced by Schoenberg in 1930 (see [13, 14]) and by Krein and Gantmacher in the 1930s.

The main results concerning eigenvalues and eigenvectors of STP matrices were proved by Gantmacher and Krein in their 1937 paper [6]. (An announcement appeared in 1935 in [5]. Chapter 2 of their book [7, 8] is a somewhat expanded version of their paper [6].) Among the results proved in that paper is that an $N \times N$ STP matrix has N positive simple eigenvalues, and the eigenvector associated with the ith eigenvalue, in descending order of magnitude, has $i - 1$ sign changes.

To explain this more precisely, let us define for each $\boldsymbol{x} \in \mathbb{R}^N$ two sign-change indices. These are $S^-(\boldsymbol{x})$, which is simply the number of ordered sign changes in the vector $\boldsymbol{x}$, where zero entries are discarded, and $S^+(\boldsymbol{x})$, which is the maximum number of ordered sign changes in the vector $\boldsymbol{x}$, where zero entries are given arbitrary values. Thus, for example,

$$S^-(1, 0, 2, -3, 0, 1) = 2 \quad \text{and} \quad S^+(1, 0, 2, -3, 0, 1) = 4.$$

Note also that $S^-(\mathbf{0}) = 0$, while for convenience we will set $S^+(\mathbf{0}) = N$.

We now can formally state the Gantmacher–Krein theorem.

Theorem GK. *Let A be an $N \times N$ STP matrix. Then A has N positive simple eigenvalues*

$$\lambda_1 > \cdots > \lambda_N > 0.$$

Let $\boldsymbol{x}_i$ be an eigenvector associated with the eigenvalue λ_i. Then, if $1 \leqslant p \leqslant q \leqslant N$ and $c_p, \dots, c_q$ are not all zero, we have

$$p - 1 \leqslant S^-\Bigl(\sum_{i=p}^{q} c_i \boldsymbol{x}_i\Bigr) \leqslant S^+\Bigl(\sum_{i=p}^{q} c_i \boldsymbol{x}_i\Bigr) \leqslant q - 1.$$

In particular, for each $i \in \{1, \dots, N\}$,

$$S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i) = i - 1, \quad i = 1, \dots, N.$$

Equalities of the form $S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i)$ will be frequently encountered, both in our results and their proofs. Such an equality implies that every zero component of $\boldsymbol{x}$, if they exist, is flanked by non-zero components of opposite sign. In particular, neither the first nor the last component of $\boldsymbol{x}$ may vanish. This equality may be viewed, in a sense, as the vector analogue of a function with only simple interior zeros.

In this paper, we will consider various generalizations of this result. To explain our results, we introduce some notation. For a vector $\boldsymbol{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$, and arbitrary $p > 0$, we set

$$\boldsymbol{x}^{p*} = (|x_1|^p \operatorname{sgn} x_1, \dots, |x_N|^p \operatorname{sgn} x_N).$$

Let A be an $M \times N$ STP matrix. In [11,12], Pinkus considered the nonlinear 'self-adjoint' eigenvalue–eigenvector problem

$$A^{\mathrm{T}}(A\boldsymbol{x})^{p*} = \lambda \boldsymbol{x}^{p*}. \tag{1.1}$$

For $p = 1$, this reduces to the self-adjoint equation

$$A^{\mathrm{T}}A\boldsymbol{x} = \lambda \boldsymbol{x}.$$

This problem (for general p) arose in connection with some n-width problems, and also has analogues as certain integral and differential equations. If $(\lambda, \boldsymbol{x})$ satisfies (1.1), with $\boldsymbol{x} \neq \boldsymbol{0}$ and $\lambda \in \mathbb{R}$, $\lambda \neq 0$, then we will say that $(\lambda, \boldsymbol{x})$ is an *eigenvalue–eigenvector pair* for (1.1). It is readily checked that if $(\lambda, \boldsymbol{x})$ satisfies (1.1), then so does $(\lambda, \alpha\boldsymbol{x})$ for every $\alpha \in \mathbb{R}$. As such, we may consider eigenvalue–eigenvector pairs, up to multiplication of the eigenvector by constants. The following result, analogous to the Gantmacher–Krein theorem, was proved in [11].

Theorem P. *Assume A is an $M \times N$ STP matrix. Let $R = \min\{M, N\}$. Then, for each $p > 0$, there exist exactly R eigenvalue–eigenvector pairs $(\lambda_i, \boldsymbol{x}_i)$ satisfying (1.1). Furthermore,*

(a) $\lambda_1 > \cdots > \lambda_R > 0$,

(b) $S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i) = i - 1$ *for* $i = 1, \dots, R$.

Equation (1.1) was generalized by Buslaev [3] to

$$A^{\mathrm{T}}(A\boldsymbol{x})^{p*} = \lambda \boldsymbol{x}^{r*} \tag{1.2}$$

for arbitrary $p, r > 0$. Note that if $(\lambda, \boldsymbol{x})$ satisfies (1.2), then so does $(\lambda|\alpha|^{p-r}, \alpha\boldsymbol{x})$. Because of this nonlinearity of the eigenvalue, when we talk about eigenvalue–eigenvector pairs, we will generally assume that $\|\boldsymbol{x}\|_2 = 1$ (the Euclidean norm equals 1). Buslaev proved the following.

Theorem B. *Assume A is an $M \times N$ STP matrix. Let $R = \min\{M, N\}$. Then, for given $p, r > 0$ and each $i \in \{1, \ldots, R\}$, there exists at least one eigenvalue–eigenvector pair $(\lambda_i, \boldsymbol{x}_i)$ satisfying (1.2), with $\lambda_i > 0$ and $S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i) = i - 1$.*

For $p \neq r$, there may be many more than R eigenvalue–eigenvector pairs. As such, there may, and sometimes will, be more than one eigenvector with any fixed number of sign changes. Some examples are presented in § 5.

Two different proofs of the existence of eigenvalue–eigenvector pairs are given in [11] and [12]. One depends on the Lyusternik–Schnirelman theorem on the category of an n-dimensional projective space. The other follows from an application of the implicit function theorem. The proof given by Buslaev is different and uses both an ingenious fixed-point argument and Borsuk's antipodensatz theorem.

In this paper, we will consider certain generalizations of theorem P. In particular, we will consider the eigenvalue–eigenvector problem

$$A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*} = \lambda\boldsymbol{x}^{r*}, \tag{1.3}$$

where the $p_1, \ldots, p_{m-1}, r$ are arbitrary positive values and A_i is an $N_i \times N_{i-1}$ STP matrix, $i = 1, \ldots, m$, with $N_0 = N_m = N$. Unfortunately, the lack of a certain form of 'self-adjointness', which is present in (1.1) and (1.2), renders the category argument in [11], and the fixed-point and Borsuk arguments in [3], seemingly inapplicable. We will use a variation on the argument from [12] that applies the implicit function theorem. As we will shall see, the existence, and not the uniqueness or characterization, is the major problem when considering (1.3).

Our main result is as follows.

Theorem 1.1. *Assume that the A_i are as above, with $p_1, \ldots, p_{m-1} \geqslant 1$ and $r = p_1 \cdots p_{m-1}$. Let $R = \min\{N_i : i = 1, \ldots, m\}$. Then there exist exactly R eigenvalue–eigenvector pairs $(\lambda_i, \boldsymbol{x}_i)$ satisfying (1.3). Furthermore,*

(a) $\lambda_1 > \cdots > \lambda_R > 0$,

(b) $S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i) = i - 1$ *for* $i = 1, \ldots, R$.

We conjecture that this same result holds for any positive $p_1, \ldots, p_{m-1}$ and $r = p_1 \cdots p_{m-1}$. While our proof uses the fact that the $p_i \geqslant 1$, we will show how certain additional cases follow from this result. For $r \neq p_1 \cdots p_{m-1}$, there may be many more eigenvalue–eigenvector pairs. We will also touch upon this briefly.

Why do we consider only the particular nonlinear operation $\boldsymbol{x}^{p*}$? If we demand that if $\boldsymbol{x}$ is an eigenvector, then $\alpha\boldsymbol{x}$ is also an eigenvector for all $\alpha \in \mathbb{R}$ (but with perhaps different eigenvalues), then we are led to a consideration of the functional equation $f(ab) = f(a)f(b)$ for all $a, b \in \mathbb{R}$. As is known (see, for example, [1, p. 41]), the only continuous solutions to this equation are

$$f(x) = |x|^p$$

and

$$f(x) = |x|^p \operatorname{sgn} x = x^{p*},$$

for some $0 < p < \infty$. This is our reason for the choice of this operator.

Before ending this introductory section, we recall two important properties of STP matrices. These are called *variation diminishing*. This was Schoenberg's initial contribution to the theory. The versions we shall use are the following.

THEOREM VD. *Let A be an $M \times N$ STP matrix. Then, for each vector $\boldsymbol{x} \in \mathbb{R}^N$, $\boldsymbol{x} \neq \mathbf{0}$,*

$$S^+(A\boldsymbol{x}) \leqslant S^-(\boldsymbol{x}).$$

Furthermore if $S^+(A\boldsymbol{x}) = S^-(\boldsymbol{x})$, then the sign of the first (and last) component of $A\boldsymbol{x}$ (if zero, then the sign given in determining $S^+(A\boldsymbol{x})$) agrees with the sign of the first (and last) non-zero component of $\boldsymbol{x}$.

We say that an $M \times N$ matrix is *STP of rank K* if it is of rank K and all of its $r \times r$ minors are strictly positive for all $r = 1, \dots, K$. An associated dual property to theorem VD is contained in this next result.

THEOREM VD2. *Let A be an $M \times N$ STP matrix of rank K. If $A\boldsymbol{y} = \mathbf{0}$, then either $\boldsymbol{y} = \mathbf{0}$ or $S^-(\boldsymbol{y}) \geqslant K$.*

The two above results, and variants thereof, may be found in [2, 9, 13].

This paper is organized as follows. In § 2 we provide a new proof of the Gantmacher–Krein theorem. This is the 'linear' version of our problem. The standard method of proving the Gantmacher–Krein theorem is by the use of Perron's theorem regarding the largest eigenvalue of a positive matrix, and its eigenvector, and Kronecker's theorem on the eigenvalues and eigenvectors of the associated compound matrices. In § 2 we will present an alternative method of proof for this theorem, which depends on neither of these results. It instead depends on the variation diminishing property of STP matrices. In § 3, we consider generalizations of the method of proof given in § 2, which permits us to prove the uniqueness and other properties of the solutions of the nonlinear problem (1.3) (always under the assumption that $r = p_1 \cdots p_{m-1}$). In § 4, we consider the problem of existence of solutions to (1.3). It is here that we apply the Gantmacher–Krein theorem (which is applicable in the case $p_1 = \cdots = p_{m-1} = r = 1$) and the implicit function theorem, which permits us to continue the solutions that exist in the former case to $p_1, \dots, p_{m-1} > 1$ (and $r = p_1 \cdots p_{m-1}$). This is used in the proof of theorem 1.1. We may totally forego this restriction on the p_i when dealing with the largest eigenvalue, or the smallest eigenvalue if all the A_i are invertible. We also rotate, and in certain cases, invert, the p_i, and thus generalize theorem 1.1 to cover additional cases. In § 5 we briefly discuss the case where $r \neq p_1 \cdots p_{m-1}$, proving, for example, a Perron-type theorem if $r \geqslant p_1 \cdots p_{m-1}$.

Somewhat analogous results for ordinary differential equations may be found in [4].

2. An alternative proof of the Gantmacher–Krein theorem

In this section, we provide a new and alternative proof of the Gantmacher–Krein theorem. Let A be an $N \times N$ STP matrix. We start by proving that A has N simple and positive eigenvalues. This part of the proof may be found in [10]. We use the notation

$$A_\lambda \begin{pmatrix} i_1, \ldots, i_k \\ j_1, \ldots, j_k \end{pmatrix}$$

to denote the $k \times k$ minor of the matrix $A - \lambda I$, obtained by choosing the rows $i_1, \ldots, i_k$ and columns $j_1, \ldots, j_k$. From Sylvester's determinant identity (see, for example, [9, p. 3]), we have

$$\begin{aligned} &A_\lambda \begin{pmatrix} 1, \ldots, N \\ 1, \ldots, N \end{pmatrix} A_\lambda \begin{pmatrix} 2, \ldots, N-1 \\ 2, \ldots, N-1 \end{pmatrix} \\ &\quad = A_\lambda \begin{pmatrix} 1, \ldots, N-1 \\ 1, \ldots, N-1 \end{pmatrix} A_\lambda \begin{pmatrix} 2, \ldots, N \\ 2, \ldots, N \end{pmatrix} - A_\lambda \begin{pmatrix} 1, \ldots, N-1 \\ 2, \ldots, N \end{pmatrix} A_\lambda \begin{pmatrix} 2, \ldots, N \\ 1, \ldots, N-1 \end{pmatrix}. \end{aligned} \tag{2.1}$$

It is easily verified that

$$A_\lambda \begin{pmatrix} 1, \ldots, N-1 \\ 2, \ldots, N \end{pmatrix}, \; A_\lambda \begin{pmatrix} 2, \ldots, N \\ 1, \ldots, N-1 \end{pmatrix} > 0 \tag{2.2}$$

for all $\lambda > 0$. Expand each determinant as a polynomial in λ and note that all the coefficients of this polynomial are strictly positive, since all minors of A are strictly positive.

We now use an induction argument to prove our result. We also simultaneously prove that the eigenvalues of the two principal minors obtained by deleting either the first row and column, or the last row and column, strictly interlace the eigenvalues of the original matrix. The case $N = 2$ is easily checked.

For notational ease, set

$$p(\lambda) = A_\lambda \begin{pmatrix} 1, \ldots, N \\ 1, \ldots, N \end{pmatrix},$$

$$q_1(\lambda) = A_\lambda \begin{pmatrix} 2, \ldots, N \\ 2, \ldots, N, \end{pmatrix},$$

$$q_2(\lambda) = A_\lambda \begin{pmatrix} 1, \ldots, N-1 \\ 1, \ldots, N-1 \end{pmatrix}$$

and

$$r(\lambda) = A_\lambda \begin{pmatrix} 2, \ldots, N-1 \\ 2, \ldots, N-1 \end{pmatrix}.$$

We assume, by the induction hypothesis, that all the $N-2$ zeros of r are positive and simple, and interlace the $N-1$ positive simple zeros of q_1 and q_2. Let

$$\mu_1 > \cdots > \mu_{N-1} > 0$$

denote the zeros of either q_1 or q_2. From (2.1) and (2.2), it follows that

$$p(\mu_i)r(\mu_i) < 0, \quad i = 1, \dots, N-1.$$

Furthermore, as $r(0) > 0$ and by the induction hypothesis,

$$r(\mu_i)(-1)^{i+N-1} > 0, \quad i = 1, \dots, N-1,$$

implying

$$p(\mu_i)(-1)^{i+N} > 0, \quad i = 1, \dots, N-1.$$

As $p(0) > 0$ and $p(\mu_{N-1}) < 0$, the polynomial p has an additional zero in $(0, \mu_{N-1})$. As p has leading coefficient $(-1)^N$ and $p(\mu_1)(-1)^{N+1} > 0$, p also has a zero in (μ_1, ∞). Thus p has N positive simple zeros, which are interlaced by the $N-1$ zeros of both q_1 and q_2. This proves the induction step, and thus A has N positive simple eigenvalues.

We now assume that $\lambda > \mu > 0$ are eigenvalues, with associated linearly independent eigenvectors $\boldsymbol{x}$ and $\boldsymbol{y}$. Then, from theorem VD, we have, for all $(\alpha, \beta) \neq (0, 0)$,

$$S^+(\lambda\alpha\boldsymbol{x} + \mu\beta\boldsymbol{y}) = S^+(A(\alpha\boldsymbol{x} + \beta\boldsymbol{y})) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}). \tag{2.3}$$

Note that this implies

$$S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) \quad \text{and} \quad S^+(\boldsymbol{y}) = S^-(\boldsymbol{y}).$$

We now assume that both $\alpha \neq 0$ and $\beta \neq 0$. We may rewrite (2.3) as

$$S^+\left(\alpha\boldsymbol{x} + \left(\frac{\mu}{\lambda}\right)\beta\boldsymbol{y}\right) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}).$$

Iterating this process, we have

$$S^+\left(\alpha\boldsymbol{x} + \left(\frac{\mu}{\lambda}\right)^n\beta\boldsymbol{y}\right) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$$

for every positive integer n. As $n \to \infty$, and since $\mu/\lambda < 1$ and $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x})$, it is easily verified that

$$\lim_{n\to\infty} S^+\left(\alpha\boldsymbol{x} + \left(\frac{\mu}{\lambda}\right)^n\beta\boldsymbol{y}\right) = S^-(\boldsymbol{x}).$$

The above follows from the general fact (which we will use again) that if

$$\lim_{n\to\infty} \boldsymbol{z}^n = \boldsymbol{z}$$

and $\boldsymbol{z} \in \mathbb{R}^N$ satisfies $S^+(\boldsymbol{z}) = S^-(\boldsymbol{z})$, then

$$S^+(\boldsymbol{z}^n) = S^-(\boldsymbol{z}^n) = S^+(\boldsymbol{z}) = S^-(\boldsymbol{z}) \tag{2.4}$$

for n sufficiently large. Note that if $\varepsilon > 0$ satisfies

$$\varepsilon < \min\{|z_k| : z_k \neq 0\},$$

then, for every n for which

$$\max\{|z_k^n - z_k| : k = 1, \dots, N\} < \varepsilon,$$

we necessarily have

$$S^+(\boldsymbol{z}^n) = S^-(\boldsymbol{z}^n) = S^+(\boldsymbol{z}) = S^-(\boldsymbol{z}).$$

Similarly, from (2.3),

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-\left(\left(\frac{\mu}{\lambda}\right)\alpha\boldsymbol{x} + \beta\boldsymbol{y}\right)$$

and, iterating this process, we obtain

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-\left(\left(\frac{\mu}{\lambda}\right)^n\alpha\boldsymbol{x} + \beta\boldsymbol{y}\right)$$

for every positive integer n. As $n \to \infty$, and since $\mu/\lambda < 1$ and $S^-(\boldsymbol{y}) = S^+(\boldsymbol{y})$, it follows that

$$\lim_{n\to\infty} S^-\left(\left(\frac{\mu}{\lambda}\right)^n\alpha\boldsymbol{x} + \beta\boldsymbol{y}\right) = S^-(\boldsymbol{y}).$$

Thus, for every $(\alpha, \beta) \neq (0, 0)$,

$$S^-(\boldsymbol{x}) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-(\boldsymbol{y}).$$

If $S^-(\boldsymbol{x}) = S^-(\boldsymbol{y})$, then equality holds throughout our sequence of inequalities, and, in particular,

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$$

for all $(\alpha, \beta) \neq (0, 0)$. This is impossible, since there are choices of α, β for which the first or last coefficient of the vector $\alpha\boldsymbol{x} + \beta\boldsymbol{y}$ vanishes, in which case the above equality does not hold. Therefore, $S^-(\boldsymbol{x}) < S^-(\boldsymbol{y})$.

We have proved that A has N simple positive eigenvalues, and thus N associated eigenvectors. For each of these N eigenvectors, we have

$$S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) \in \{0, 1, \dots, N-1\},$$

and if $\lambda > \mu > 0$ are eigenvalues with associated eigenvectors $\boldsymbol{x}$ and $\boldsymbol{y}$, then $S^-(\boldsymbol{x}) < S^-(\boldsymbol{y})$. This proves that if

$$\lambda_1 > \cdots > \lambda_N > 0$$

are the eigenvalues of A, and $\boldsymbol{x}_i$ is an eigenvector associated with the eigenvalue λ_i, then

$$S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i) = i - 1.$$

Now assume $1 \leqslant p \leqslant q \leqslant N$ and we are given real constants $c_p, \dots, c_q$, not all zero. Then, following the above analysis, repeated application of

$$S^+\left(\sum_{i=p}^{q} c_i\lambda_i\boldsymbol{x}_i\right) = S^+\left(\sum_{i=p}^{q} c_iA\boldsymbol{x}_i\right) \leqslant S^-\left(\sum_{i=p}^{q} c_i\boldsymbol{x}_i\right)$$

implies that

$$S^-\left(\sum_{i=p}^{q} c_i\left(\frac{\lambda_i}{\lambda_p}\right)^n\boldsymbol{x}_i\right) \leqslant S^-\left(\sum_{i=p}^{q} c_i\boldsymbol{x}_i\right) \leqslant S^+\left(\sum_{i=p}^{q} c_i\boldsymbol{x}_i\right) \leqslant S^+\left(\sum_{i=p}^{q} c_i\left(\frac{\lambda_q}{\lambda_i}\right)^n\boldsymbol{x}_i\right).$$

We may assume that $c_p, c_q \neq 0$ and then apply

$$\lim_{n\to\infty} S^-\left(\sum_{i=p}^{q} c_i \left(\frac{\lambda_i}{\lambda_p}\right)^n \boldsymbol{x}_i\right) = S^-(\boldsymbol{x}_p) = p-1,$$

while

$$\lim_{n\to\infty} S^+\left(\sum_{i=p}^{q} c_i \left(\frac{\lambda_q}{\lambda_i}\right)^n \boldsymbol{x}_i\right) = S^-(\boldsymbol{x}_q) = q-1.$$

This implies

$$p-1 \leqslant S^-\left(\sum_{i=p}^{q} c_i\boldsymbol{x}_i\right) \leqslant S^+\left(\sum_{i=p}^{q} c_i\boldsymbol{x}_i\right) \leqslant q-1,$$

which completes the proof of this theorem. □

Remark 2.1. There is an additional result that is sometimes stated in connection with the Gantmacher–Krein theorem. It has to do with the 'interlacing of the zeros' of $\boldsymbol{x}_{i+1}$ and $\boldsymbol{x}_i$, which is important in the continuous (integral) analogue of this theorem. One form of this is the following. Let $x_i(t)$ be the continuous function defined on $[1, N]$, which is linear on each $[j, j+1]$, $j = 1, \dots, N-1$, and such that $x_i(j)$ is the jth component of $\boldsymbol{x}_i$. Since

$$S^-(\boldsymbol{x}_i) = S^+(\boldsymbol{x}_i) = i-1,$$

the function $x_i(t)$ has exactly $i-1$ zeros, which are each strict sign changes. Since

$$i-1 \leqslant S^-(\alpha\boldsymbol{x}_i + \beta\boldsymbol{x}_{i+1}) \leqslant S^+(\alpha\boldsymbol{x}_i + \beta\boldsymbol{x}_{i+1}) \leqslant i,$$

for every choice of $(\alpha, \beta) \neq (0, 0)$ it can now be shown that the $i-1$ zeros of $x_i(t)$ strictly interlace the i zeros of $x_{i+1}(t)$.

A similar argument proves the following variant, which we will need in § 4.

Theorem GK2. *Let A be an $N\times N$ STP matrix of rank K. Then A has K positive simple eigenvalues*

$$\lambda_1 > \cdots > \lambda_K > 0$$

and, for the essentially unique eigenvector $\boldsymbol{x}_i$ associated with the eigenvalue λ_i,

$$S^+(\boldsymbol{x}_i) = S^-(\boldsymbol{x}_i) = i-1, \quad i = 1, \dots, K.$$

3. Uniqueness and other properties

In this section, we will discuss uniqueness and certain associated properties for the problem

$$A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*} = \lambda\boldsymbol{x}^{r*}, \tag{3.1}$$

where the p_i are positive values and $p_1\cdots p_{m-1} = r$. We show that the latter half of the proof of theorem GK may be applied in this situation.

For ease of notation, we set

$$T(\boldsymbol{x}) = A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*},$$

where each A_i is an $N_i \times N_{i-1}$ STP matrix, $i = 1, \dots, m$, with $N_0 = N_m = N$, and $R = \min\{N_i : i = 1, \dots, m\}$. Note that

$$T(\alpha \boldsymbol{x}) = \alpha^{p_1 \cdots p_{m-1} *} T(\boldsymbol{x}) = \alpha^{r*} T(\boldsymbol{x}).$$

Thus $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair for (3.1) if

$$T(\boldsymbol{x}) = \lambda \boldsymbol{x}^{r*},$$

and if $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair, then, for every $\alpha \in \mathbb{R} \setminus \{0\}$, so is $(\lambda, \alpha \boldsymbol{x})$.

PROPOSITION 3.1. *Under the above assumptions, let $(\lambda, \boldsymbol{x})$ and $(\mu, \boldsymbol{y})$ be eigenvalue–eigenvector pairs for (3.1). Assume $\boldsymbol{x}$, $\boldsymbol{y}$ are linearly independent. Then, for all $(\alpha, \beta) \neq (0, 0)$,*

$$S^+(\lambda \alpha^{r*} \boldsymbol{x}^{r*} + \mu \alpha^{r*} \boldsymbol{y}^{r*}) \leqslant S^-(\alpha \boldsymbol{x} + \beta \boldsymbol{y}).$$

Proof. In this proof, we will use the easily checked

$$\operatorname{sgn}(a^{p*} + b^{p*}) = \operatorname{sgn}(a + b) \tag{3.2}$$

for any $a, b \in \mathbb{R}$ and $p > 0$. We also recall that we are assuming that $\lambda, \mu \neq 0$.

We start by noting that since $(\lambda, \boldsymbol{x})$ and $(\mu, \boldsymbol{y})$ are eigenvalue–eigenvector pairs for (3.1),

$$\lambda \alpha^{r*} \boldsymbol{x}^{r*} + \mu \alpha^{r*} \boldsymbol{y}^{r*} = T(\alpha \boldsymbol{x}) + T(\beta \boldsymbol{y}).$$

Thus

$$\begin{aligned} S^+(\lambda \alpha^{r*} \boldsymbol{x}^{r*} + \mu \alpha^{r*} \boldsymbol{y}^{r*}) &= S^+(T(\alpha \boldsymbol{x}) + T(\beta \boldsymbol{y})) \\ &= S^+(A_m(\cdots (A_2(A_1(\alpha \boldsymbol{x}))^{p_1*})^{p_2*} \cdots)^{p_{m-1}*} \\ &\qquad + A_m(\cdots (A_2(A_1(\beta \boldsymbol{y}))^{p_1*})^{p_2*} \cdots)^{p_{m-1}*}). \end{aligned}$$

From theorem VD applied to the matrix A_m, the above quantity is less than or equal to

$$\begin{aligned} S^-((A_{m-1} \cdots (A_2(A_1(\alpha \boldsymbol{x}))^{p_1*})^{p_2*} \cdots)^{p_{m-1}*} \\ + (A_{m-1} \cdots (A_2(A_1(\beta \boldsymbol{y}))^{p_1*})^{p_2*} \cdots)^{p_{m-1}*}) \end{aligned}$$

and, by (3.2), is equal to

$$\begin{aligned} S^-(A_{m-1}(\cdots (A_2(A_1(\alpha \boldsymbol{x}))^{p_1*})^{p_2*} \cdots)^{p^*_{m-2}} \\ + A_{m-1}(\cdots (A_2(A_1(\beta \boldsymbol{y}))^{p_1*})^{p_2*} \cdots)^{p^*_{m-2}}). \end{aligned}$$

As $S^-(\boldsymbol{w}) \leqslant S^+(\boldsymbol{w})$ for every vector $\boldsymbol{w}$, we can now continue this process, each time peeling away an A_k and a power, until we reach

$$S^-(\alpha \boldsymbol{x} + \beta \boldsymbol{y}).$$

□

Let us now consider some consequences of this simple inequality.

Corollary 3.2. *Assume $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair for (3.1). Then $\lambda > 0$, and*

$$S^-(\boldsymbol{x}) = S^+(\boldsymbol{x}) \leqslant R - 1,$$

where $R = \min\{N_i : i = 1, \dots, m\}$.

Proof. Put $\alpha = 1$ and $\beta = 0$ in the proof of proposition 3.1. Then, starting with $S^+(\lambda \boldsymbol{x}^{r*})(= S^+(\boldsymbol{x}))$, we get a whole series of increasing inequalities, ending with $S^-(\boldsymbol{x})$. As $S^-(\boldsymbol{x}) \leqslant S^+(\boldsymbol{x})$, it therefore follows that this series of inequalities are equalities. That is, $S^-(\boldsymbol{x}) = S^+(\lambda \boldsymbol{x}^{r*}) = S^+(\boldsymbol{x})$. Furthermore, from the case of equality as stated in the second part of theorem VD, we obtain $\lambda > 0$.

From these series of equalities we also have, for each $i \in \{1, \dots, m\}$,

$$S^+(\boldsymbol{x}) = S^+(A_i(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p^*_{i-1}}), \tag{3.3}$$

where $p_0 = 1$. As $A_i(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p^*_{i-1}}$ is a non-zero vector in $\mathbb{R}^{N_i}$, we must have $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x}) \leqslant N_i - 1$. This is true for each i, and therefore $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x}) \leqslant R - 1$. □

We have so far only considered eigenvalue–eigenvector pairs with positive (non-zero) eigenvalues. What, if anything, can we say about those vectors satisfying $T(\boldsymbol{x}) = \boldsymbol{0}$?

Corollary 3.3. *If $T(\boldsymbol{x}) = \boldsymbol{0}$, then $S^-(\boldsymbol{x}) \geqslant R$.*

Proof. Starting with

$$T(\boldsymbol{x}) = A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*} = \boldsymbol{0},$$

we have, from theorem VD2, that either

$$A_{m-1}(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p^*_{m-2}} = \boldsymbol{0}$$

or this vector is not identically zero and

$$S^-(A_{m-1}(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p^*_{m-2}}) \geqslant N_m.$$

In the latter case, we can repeatedly apply theorem VD and (3.2), as in the proof of proposition 3.1, to obtain

$$S^-(\boldsymbol{x}) \geqslant S^-(A_{m-1}(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p^*_{m-2}}) \geqslant N_m \geqslant R,$$

since $R = \min\{N_i : i = 1, \dots, m\}$. If

$$A_{m-1}(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p^*_{m-2}} = \boldsymbol{0},$$

then we continue as above (stripping away the A_{m-1}). It follows, in all cases, that $S^-(\boldsymbol{x}) \geqslant R$. □

The main consequence of proposition 3.1 is the following result.

Theorem 3.4. *In the eigenvalue–eigenvector problem (3.1), to each eigenvalue there corresponds an essentially unique eigenvector. Furthermore, if $(\lambda, \boldsymbol{x})$ and $(\mu, \boldsymbol{y})$ are eigenvalue–eigenvector pairs for (3.1) and $\lambda > \mu$, then $S^-(\boldsymbol{x}) < S^-(\boldsymbol{y}) \leqslant R - 1$.*

Proof. From proposition 3.1, we have

$$S^+(\lambda\alpha^{r*}\boldsymbol{x}^{r*} + \mu\alpha^{r*}\boldsymbol{y}^{r*}) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}),$$

which we can rewrite, by (3.2) (and since $\lambda, \mu > 0$), as

$$S^+(\lambda^{1/r}\alpha\boldsymbol{x} + \mu^{1/r}\alpha\boldsymbol{y}) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}). \tag{3.4}$$

Assume $\boldsymbol{x}$, $\boldsymbol{y}$ are linearly independent eigenvectors associated with the same eigenvalue $\lambda > 0$. Then, from (3.4),

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$$

for all $(\alpha, \beta) \neq (0, 0)$. But this is impossible, as this would imply, for example, that the first and last coefficients of the vector $\alpha\boldsymbol{x} + \beta\boldsymbol{y}$ never vanish. This contradiction implies that the eigenvector associated with the eigenvalue λ is unique (up to multiplication by a constant).

Assume now that $(\lambda, \boldsymbol{x})$ and $(\mu, \boldsymbol{y})$ are eigenvalue–eigenvector pairs for (3.1), with $\lambda > \mu$. We apply the same argument as may be found in our proof of theorem GK. For $\alpha \neq 0$, $\beta \neq 0$, we have, by (3.4),

$$S^+\left(\alpha\boldsymbol{x} + \left(\frac{\mu}{\lambda}\right)^{1/r}\alpha\boldsymbol{y}\right) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}).$$

Iterating this process, we have

$$S^+\left(\alpha\boldsymbol{x} + \left(\frac{\mu}{\lambda}\right)^{n/r}\alpha\boldsymbol{y}\right) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$$

for every positive integer n. As $n \to \infty$ and since $\mu/\lambda < 1$ and $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x})$, it follows by (2.4) that

$$\lim_{n\to\infty} S^+\left(\alpha\boldsymbol{x} + \left(\frac{\mu}{\lambda}\right)^{n/r}\alpha\boldsymbol{y}\right) = S^-(\boldsymbol{x}).$$

Similarly,

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-\left(\left(\frac{\mu}{\lambda}\right)^{1/r}\alpha\boldsymbol{x} + \beta\boldsymbol{y}\right)$$

and, iterating this process, we obtain

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-\left(\left(\frac{\mu}{\lambda}\right)^{n/r}\alpha\boldsymbol{x} + \beta\boldsymbol{y}\right)$$

for every positive integer n. As $n \to \infty$ and since $\mu/\lambda < 1$ and $S^-(\boldsymbol{y}) = S^+(\boldsymbol{y})$, it follows that

$$\lim_{n\to\infty} S^-\left(\left(\frac{\mu}{\lambda}\right)^{n/r}\alpha\boldsymbol{x} + \beta\boldsymbol{y}\right) = S^-(\boldsymbol{y}).$$

Thus, for all $(\alpha, \beta) \neq (0, 0)$,

$$S^-(\boldsymbol{x}) \leqslant S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \leqslant S^-(\boldsymbol{y}).$$

If $S^-(\boldsymbol{x}) = S^-(\boldsymbol{y})$, then equality holds throughout and, in particular,

$$S^+(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = S^-(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$$

for all $(\alpha, \beta) \neq (0, 0)$, which is again impossible. Therefore, $S^-(\boldsymbol{x}) < S^-(\boldsymbol{y})$. From corollary 3.2, we have $S^-(\boldsymbol{y}) \leqslant R - 1$. □

4. Existence

In this section, we prove the existence of R distinct eigenvalue–eigenvector pairs for the problem

$$T(\boldsymbol{x}) = A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*} = \lambda\boldsymbol{x}^{r*}, \tag{4.1}$$

where the $p_i \geqslant 1$ and $p_1 \cdots p_{m-1} = r$. This result, together with theorem 3.4, implies the existence of exactly R such pairs and proves theorem 1.1.

In the proof of the existence, we will use a method of proof based on the implicit function theorem. This proof is rather technical. We shall assume, in what follows, that each $p_i > 1$. We may do this because, if $p_i = 1$, then we simply reduce the number of terms in the above equation. For each $t \in [0, 1]$, we will set

$$p_i(t) = p_i t + (1 - t)$$

for $i = 1, \ldots, m - 1$, and

$$r(t) = p_1(t) \cdots p_{m-1}(t).$$

Note that

$$(p_1(1), \ldots, p_{m-1}(1), r(1)) = (p_1, \ldots, p_{m-1}, r),$$

while

$$(p_1(0), \ldots, p_{m-1}(0), r(0)) = (1, \ldots, 1, 1).$$

By theorem GK2, we know the solutions of (4.1) for $(p_1(0), \ldots, p_{m-1}(0), r(0))$. We continue these solutions as functions of t until we reach $(p_1(1), \ldots, p_{m-1}(1), r(1))$.

Define, for $k = 1, \ldots, N$,

$$G_k(\boldsymbol{x}, \lambda, t) = [A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1(t)*})^{p_2(t)*}\cdots)^{p_{m-1}(t)*} - \lambda\boldsymbol{x}^{r(t)*}]_k, \tag{4.2}$$

and

$$G_{N+1}(\boldsymbol{x}, \lambda, t) = \sum_{k=1}^{N} |x_k|^2 - 1. \tag{4.3}$$

Solving (4.1) is equivalent to solving $G_k(\boldsymbol{x}, \lambda, t) = 0$ for $k = 1, \ldots, N$. The condition $G_{N+1}(\boldsymbol{x}, \lambda, t) = 0$ is a normalization condition.

Now,

$$\boldsymbol{G}(\boldsymbol{x}, \lambda, t) = (G_1(\boldsymbol{x}, \lambda, t), \ldots, G_{N+1}(\boldsymbol{x}, \lambda, t)) \in C(\mathbb{R}^{N+2}, \mathbb{R}^{N+1}).$$

As we will use the implicit function theorem for $\boldsymbol{G}(\boldsymbol{x},\lambda,t)=\mathbf{0}$ with respect to the $N+1$ variables $x_1,\dots,x_N,\lambda$ as functions of t, we must show that

$$\boldsymbol{G}(\boldsymbol{x},\lambda,t)=(G_1(\boldsymbol{x},\lambda,t),\dots,G_{N+1}(\boldsymbol{x},\lambda,t))\in C^1(\mathbb{R}^{N+2},\mathbb{R}^{N+1}).$$

We will do this explicitly by calculating most of the partial derivatives, since we need them later.

We introduce the following notation. For $k=1,\dots,m$, we let $D_k(\boldsymbol{x})$ be the $N_k\times N_k$ diagonal matrix whose diagonal entries are the components of the vector

$$A_k(A_{k-1}\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{k-1}*},$$

which is in $\mathbb{R}^{N_k}$ (and we set $p_0=1$ so that $D_1(\boldsymbol{x})$ is the $N_1\times N_1$ diagonal matrix whose entries are the components of $A_1\boldsymbol{x}$). Thus the diagonal entries of $D_m(\boldsymbol{x})$ are the components of the vector $T(\boldsymbol{x})$. When we write

$$|D_k(\boldsymbol{x})|^{p_k-1},$$

it is to be understood that the operator $|\cdot|^{p_k-1}$ is being applied to each element of this diagonal matrix. We will also drop the t from the $p_i(t)$, except when necessary.

The above notation is convenient since, for example, for $p_1>1$,

$$\frac{\partial}{\partial x_\ell}((A_1\boldsymbol{x})^{p_1*})_k=p_1|(A_1\boldsymbol{x})_k|^{p_1-1}(A_1)_{k,\ell}=p_1(|D_1(\boldsymbol{x})|^{p_1-1}A_1)_{k,\ell}.$$

Using this notation, and the calculations as above, we can write, for $k,\ell=1,\dots,N$ (and all $p_i>1$),

$$\frac{\partial G_k}{\partial x_\ell}(\boldsymbol{x},\lambda,t)=r(A_m|D_{m-1}(\boldsymbol{x})|^{p_{m-1}-1}A_{m-1}|D_{m-2}(\boldsymbol{x})|^{p_{m-2}-1}\cdots A_2|D_1(\boldsymbol{x})|^{p_1-1}A_1)_{k,\ell}-\lambda r|x_\ell|^{r-1}\delta_{\ell,k}. \tag{4.4}$$

For $k=N+1$ and $\ell=1,\dots,N$, we have

$$\frac{\partial G_{N+1}}{\partial x_\ell}(\boldsymbol{x},\lambda,t)=2x_\ell. \tag{4.5}$$

As the variables are $x_1,\dots,x_N$ and λ, it will sometimes be convenient to use the notation $\lambda=x_{N+1}$. For $k=1,\dots,N$ and $\ell=N+1$, we see that

$$\frac{\partial G_k}{\partial \lambda}(\boldsymbol{x},\lambda,t)=\frac{\partial G_k}{\partial x_{N+1}}(\boldsymbol{x},\lambda,t)=-x_k^{r*}, \tag{4.6}$$

while for $k,\ell=N+1$,

$$\frac{\partial G_{N+1}}{\partial x_{N+1}}(\boldsymbol{x},\lambda,t)=0. \tag{4.7}$$

It is also true that

$$\frac{\partial G_k}{\partial t}(\boldsymbol{x},\lambda,t)$$

is a continuous function. As we do not need its explicit value, we will not precisely calculate it.

Note that, for all $t \in (0,1]$, we have $p_i(t) > 1$, and thus $\boldsymbol{G}(\boldsymbol{x}, \lambda, t)$ is a C^1 function in all its variables. At the $p_i(0) = 1$, there is a problem in that $\boldsymbol{G}(\boldsymbol{x}, \lambda, 0)$ is not necessarily C^1. To be more precise, we have a problem with the continuity of the elements of $|D_k(\boldsymbol{x})|^{p_k(t)-1}$ and $|x_k|^{r(t)-1}$ if some of their components tend to zero as $t \to 0$. However, it is readily verified that $\boldsymbol{G}(\boldsymbol{x}, \lambda, 0)$ is locally C^1 at those values of $\boldsymbol{x}$ for which the components $\boldsymbol{x}$ and the diagonal components of $D_k(\boldsymbol{x})$, all $k = 1, \ldots, m$, are not zero. We will consider the case where the $p_i(t) > 1$, i.e. $t \in (0,1]$, although, as we will note later, the analysis also applies to the case $p_i = 1$ (and also to the cases $p_i < 1$) at those $\boldsymbol{x}$ for which it and the diagonal entries of the $D_k(\boldsymbol{x})$ have no zero components.

Proposition 4.1. *Let* $\tilde{\boldsymbol{x}} \in \mathbb{R}^N$, $\tilde{\lambda} > 0$ *and* $\tilde{t} \in (0,1]$ *be such that*

$$\boldsymbol{G}(\tilde{\boldsymbol{x}}, \tilde{\lambda}, \tilde{t}) = \boldsymbol{0}. \tag{4.8}$$

Then

$$\det\left(\frac{\partial G_k}{\partial x_\ell}\right)_{k,\ell=1}^{N+1}\Bigg|_{(\tilde{\boldsymbol{x}}, \tilde{\lambda}, \tilde{t})} \neq 0. \tag{4.9}$$

Proof. Set

$$b_{k\ell} = \frac{1}{r}\frac{\partial G_k}{\partial x_\ell}\Bigg|_{(\tilde{\boldsymbol{x}}, \tilde{\lambda}, \tilde{t})} = (A_m|D_{m-1}(\tilde{\boldsymbol{x}})|^{p_{m-1}-1}A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1} \cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1)_{k,\ell}$$

and $B = (b_{k\ell})$ for $k, \ell = 1, \ldots, N$. Note that B is the product of STP matrices and diagonal matrices with non-negative diagonal entries.

Assume (4.9) does not hold. Thus there exists a $\boldsymbol{z} = (z_1, \ldots, z_{N+1}) \neq \boldsymbol{0}$ in $\mathbb{R}^{N+1}$ for which

$$\left(\frac{\partial G_k}{\partial x_\ell}\right)_{k,\ell=1}^{N+1}\Bigg|_{(\tilde{\boldsymbol{x}}, \tilde{\lambda}, \tilde{t})} \boldsymbol{z} = \boldsymbol{0}.$$

We will contradict this fact.

Assuming the above Jacobian is singular, then the above system of equations may be rewritten, using (4.4)–(4.7), as

$$\sum_{\ell=1}^N b_{k\ell} z_\ell = \tilde{\lambda}|\tilde{x}_k|^{r-1} z_k - \frac{\tilde{x}_k^{r*}}{r} z_{N+1}, \tag{4.10}$$

which we will rewrite as

$$\sum_{\ell=1}^N b_{k\ell} z_\ell = \tilde{\lambda}|\tilde{x}_k|^{r-1}\left(z_k - \frac{z_{N+1}\tilde{x}_k}{\tilde{\lambda} r}\right), \tag{4.11}$$

for $k = 1, \ldots, N$, and

$$\sum_{\ell=1}^N \tilde{x}_\ell z_\ell = 0. \tag{4.12}$$

We also have that $(\tilde{\boldsymbol{x}}, \tilde{\lambda}, \tilde{t})$ is a solution of (4.8). This we may rewrite as

$$\sum_{\ell=1}^{N} b_{k\ell}\tilde{x}_\ell = \tilde{\lambda}|\tilde{x}_k|^{r-1}\tilde{x}_k, \tag{4.13}$$

for $k = 1, \ldots, N$, and

$$\sum_{\ell=1}^{N} \tilde{x}_\ell\tilde{x}_\ell = \sum_{\ell=1}^{N} \tilde{x}_\ell^2 = 1. \tag{4.14}$$

Equations (4.13) are obtained via

$$\begin{aligned}\sum_{\ell=1}^{N} b_{k\ell}\tilde{x}_\ell &= \sum_{\ell=1}^{N}(A_m|D_{m-1}(\tilde{\boldsymbol{x}})|^{p_{m-1}-1}A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1} \\ &\qquad \cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1)_{k,\ell}\tilde{x}_\ell \\ &= (A_m|D_{m-1}(\tilde{\boldsymbol{x}})|^{p_{m-1}-1}A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1}\cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1\tilde{\boldsymbol{x}})_k.\end{aligned}$$

Now we apply

$$|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1\tilde{\boldsymbol{x}} = (A_1\tilde{\boldsymbol{x}})^{p_1^*},$$

and

$$|D_2(\tilde{\boldsymbol{x}})|^{p_2-1}A_2(A_1\tilde{\boldsymbol{x}})^{p_1^*} = (A_2(A_1\tilde{\boldsymbol{x}})^{p_1^*})^{p_2^*},$$

etc.

Set

$$\tilde{\boldsymbol{z}} = (z_1, \ldots, z_N)$$

(leaving off z_{N+1}) and recall that

$$\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_N).$$

If $\tilde{\boldsymbol{z}} = (z_1, \ldots, z_N) = \mathbf{0}$, then (4.10) reduces to

$$0 = -\frac{\tilde{x}_k^{r*}}{r} z_{N+1}$$

for $k = 1, \ldots, N$, which, in turn, implies that $z_{N+1} = 0$. This would contradict the singularity of the Jacobian and prove (4.9). We will prove that $\tilde{\boldsymbol{z}} = \mathbf{0}$.

Suppose $\tilde{\boldsymbol{z}} \neq \mathbf{0}$. From (4.12) and (4.14), it follows that $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{z}}$ are linearly independent. We divide the remaining proof of this proposition into the two cases: $z_{N+1} = 0$ and $z_{N+1} \neq 0$.

Case 1 ($z_{N+1} = 0$). In this case, equation (4.11) reduces to

$$\sum_{\ell=1}^{N} b_{k\ell} z_\ell = \tilde{\lambda}|\tilde{x}_k|^{r-1} z_k \tag{4.15}$$

for $k = 1, \ldots, N$. We recall that we also have

$$\sum_{\ell=1}^{N} b_{k\ell}\tilde{x}_\ell = \tilde{\lambda}|\tilde{x}_k|^{r-1}\tilde{x}_k \tag{4.13}$$

for $k = 1,\dots,N$. Next we shall follow an analysis similar to that in the proof of proposition 3.1 to obtain, for all $(\alpha,\beta) \neq (0,0)$,

$$S^-(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}}) = S^+(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}}). \tag{4.16}$$

As we noted in the previous section, equation (4.16) cannot hold for all $(\alpha,\beta) \neq (0,0)$, and thus we are led to a contradiction.

To obtain (4.16), we argue as follows. Firstly,

$$S^+(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}}) \leqslant S^+(\{\tilde{\lambda}|\tilde{x}_k|^{r-1}(\alpha\tilde{x}_k + \beta z_k)\}_{k=1}^N),$$

since we have multiplied each component of the vector on the left by a positive value, or made it zero. From (4.13) and (4.15), the above is equal to

$$S^+(B(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}})) = S^+(A_m|D_{m-1}(\tilde{\boldsymbol{x}})|^{p_{m-1}-1}A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1} \cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}})).$$

From theorem VD (applied to the matrix A_m), we obtain an upper bound that is less than or equal to

$$S^-(|D_{m-1}(\tilde{\boldsymbol{x}})|^{p_{m-1}-1}A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1} \cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}})).$$

As $|D_{m-1}(\tilde{\boldsymbol{x}})|^{p_{m-1}-1}$ is a diagonal matrix with non-negative entries, it follows that the above is less than or equal to

$$S^-(A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1} \cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}})),$$

which is less than or equal to

$$S^+(A_{m-1}|D_{m-2}(\tilde{\boldsymbol{x}})|^{p_{m-2}-1} \cdots A_2|D_1(\tilde{\boldsymbol{x}})|^{p_1-1}A_1(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}})).$$

We continue, peeling away the A_k followed by the $|D_{k-1}(\tilde{\boldsymbol{x}})|^{p_{k-1}-1}$, until we finally reach

$$S^-(\alpha\tilde{\boldsymbol{x}} + \beta\tilde{\boldsymbol{z}}).$$

The above argument is valid if none of the vectors along the way are identically zero. This is because theorem VD only holds for non-zero vectors. Let us therefore suppose that, somewhere along the way, one of these vectors is identically zero. Then, necessarily,

$$\tilde{\lambda}|\tilde{x}_k|^{r-1}(\alpha\tilde{x}_k + \beta z_k) = 0$$

for $k = 1,\dots,N$, which we rewrite (since $\tilde{\lambda} \neq 0$ and $r > 1$) as

$$\tilde{x}_k(\alpha\tilde{x}_k + \beta z_k) = 0 \tag{4.17}$$

for $k = 1,\dots,N$. Thus, if $\tilde{x}_k \neq 0$, then

$$z_k = -\frac{\alpha}{\beta}\tilde{x}_k$$

(and $\beta \neq 0$). From (4.12) (and (4.14)),

$$0 = \sum_{\ell=1}^N \tilde{x}_\ell z_\ell = \sum_{\ell=1}^N \tilde{x}_\ell\biggl(-\frac{\alpha}{\beta}\tilde{x}_\ell\biggr) = -\frac{\alpha}{\beta}\sum_{\ell=1}^N \tilde{x}_\ell^2 = -\frac{\alpha}{\beta},$$

and thus $\alpha = 0$. Therefore, equation (4.17) reduces to

$$\tilde{x}_k z_k = 0 \tag{4.18}$$

for all $k = 1, \ldots, N$.

Each A_k is STP and each $|D_k(\tilde{\boldsymbol{x}})|^{p_k - 1}$ is a diagonal matrix with positive or zero entries. As such, it follows that B is an $N \times N$ matrix of rank K, which is STP of order K, i.e. all its $r \times r$ minors are strictly positive for all $r \in \{1, \ldots, K\}$. K is, in fact, the minimum of the N_i and the number of non-zero diagonal entries in each of the $D_k(\tilde{\boldsymbol{x}})$. As $B\tilde{\boldsymbol{z}} = \mathbf{0}$, this implies, from theorem VD2, that $S^-(\tilde{\boldsymbol{z}}) \geqslant K$.

We set $\alpha = 1$ and $\beta = 0$ in the above series of inequalities (in which case, we always have equality and the vectors never identically vanish), and obtain

$$\begin{aligned} S^-(\tilde{\boldsymbol{x}}) &= S^+(\tilde{\boldsymbol{x}}) \\ &= S^-(A_k(\cdots(A_2(A_1(\boldsymbol{x}))^{p_1 *})^{p_2 *}\cdots)^{p_{k-1}^*}) \\ &= S^+(A_k(\cdots(A_2(A_1(\boldsymbol{x}))^{p_1 *})^{p_2 *}\cdots)^{p_{k-1}^*}) \end{aligned} \tag{4.19}$$

for $k = 1, \ldots, m$. Equivalently, equation (4.19) may be obtained from (3.3) in the proof of corollary 3.2, applied to the eigenvalue–eigenvector pair $(\tilde{\lambda}, \tilde{\boldsymbol{x}})$. As the $D_k(\tilde{\boldsymbol{x}})$ are $N_k \times N_k$ diagonal matrices whose diagonal entries are the components of the vector in (4.19), it follows that

$$S^-(\tilde{\boldsymbol{x}}) = S^+(\tilde{\boldsymbol{x}}) \leqslant K - 1.$$

Thus, from (4.18) and the above,

$$K + 1 \leqslant S^-(\tilde{\boldsymbol{z}}) + 1 \leqslant \#\{i : z_i \neq 0\} \leqslant \#\{i : \tilde{x}_i = 0\} \leqslant S^+(\tilde{\boldsymbol{x}}) \leqslant K - 1.$$

This is a contradiction. Thus (4.18) cannot hold. This completes the proof of (4.16), which, in turn, is a contradiction. Thus we have proved the result in the case $z_{N+1} = 0$.

Case 2 ($z_{N+1} \neq 0$). We recall that, in this case, we have (4.11) and (4.13), i.e.

$$\sum_{\ell=1}^{N} b_{k\ell} z_\ell = \tilde{\lambda} |\tilde{x}_k|^{r-1} \left(z_k - \frac{z_{N+1} \tilde{x}_k}{\tilde{\lambda} r} \right) \tag{4.11}$$

and

$$\sum_{\ell=1}^{N} b_{k\ell} \tilde{x}_\ell = \tilde{\lambda} |\tilde{x}_k|^{r-1} \tilde{x}_k \tag{4.13}$$

for $k = 1, \ldots, N$. We also still have (4.12) and (4.14), which, of course, imply that the $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{z}}$ are linearly independent.

We follow, in general, the analysis of the previous case to obtain, for all $(\alpha, \beta) \neq (0, 0)$,

$$S^+\left(\left(\alpha - \frac{\beta z_{N+1}}{\tilde{\lambda} r} \right) \tilde{\boldsymbol{x}} + \beta \tilde{\boldsymbol{z}} \right) \leqslant S^-(\alpha \tilde{\boldsymbol{x}} + \beta \tilde{\boldsymbol{z}}). \tag{4.20}$$

Let us assume for the moment that (4.20) is valid. Why does this lead to a contradiction?

Assuming (4.20) holds for all $(\alpha,\beta)\neq(0,0)$, we have the inequalities

$$\begin{aligned}S^-\biggl(\biggl(\alpha-\frac{\beta(k+1)z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)&\leqslant S^+\biggl(\biggl(\alpha-\frac{\beta(k+1)z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)\\&\leqslant S^-\biggl(\biggl(\alpha-\frac{\beta k z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)\\&\leqslant S^+\biggl(\biggl(\alpha-\frac{\beta k z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr),\end{aligned}$$

which are valid for all $k\in\mathbb{Z}$. As $S^+(\boldsymbol{y}),S^-(\boldsymbol{y})\in\{0,1,\dots,N-1\}$ for every non-zero vector $\boldsymbol{y}\in\mathbb{R}^N$, it follows that

$$S^-\biggl(\biggl(\alpha-\frac{\beta k z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)=S^+\biggl(\biggl(\alpha-\frac{\beta k z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)=m_1 \tag{4.21}$$

for all $k>K_1$, and

$$S^-\biggl(\biggl(\alpha-\frac{\beta k z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)=S^+\biggl(\biggl(\alpha-\frac{\beta k z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)=m_2 \tag{4.22}$$

for all $k<K_2$. Assume $\beta\neq0$ (if $\beta=0$, there is nothing to prove), divide the vectors in (4.21) and (4.22) by k, and let $k\to\infty$ and $k\to-\infty$, respectively. Then, recalling that $(\tilde\lambda,\tilde{\boldsymbol{x}})$ is an eigenvalue–eigenvector pair, and, from corollary 3.2,

$$S^-(\tilde{\boldsymbol{x}})=S^+(\tilde{\boldsymbol{x}}),$$

we obtain

$$S^-(\tilde{\boldsymbol{x}})=S^+(\tilde{\boldsymbol{x}})=m_1=m_2.$$

(See (2.4) in the proof of theorem GK.)

Thus $m_1=m_2$ and, as a consequence, we have equality throughout the above series of inequalities for all $k\in\mathbb{Z}$. In particular, we obtain

$$S^-(\alpha\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}})=S^+(\alpha\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}})=m_1$$

for all $(\alpha,\beta)\neq(0,0)$. But this is impossible for linearly independent $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{z}}$, and a contradiction ensues. Thus it remains to prove that (4.20) holds for all $(\alpha,\beta)\neq(0,0)$.

To obtain (4.20), we argue as follows. From (4.11) and (4.13),

$$\begin{aligned}S^+\biggl(\biggl(\alpha-\frac{\beta z_{N+1}}{\tilde\lambda r}\biggr)\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}\biggr)&\leqslant S^+\biggl(\biggl\{\tilde\lambda|\tilde x_k|^{r-1}\biggl(\alpha\tilde x_k+\beta\biggl(z_k-\frac{z_{N+1}}{\tilde\lambda r}\tilde x_k\biggr)\biggr)\biggr\}_{k=1}^N\biggr)\\&=S^+(B(\alpha\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}})).\end{aligned}$$

We now apply the series of inequalities as in case 1 until we finally reach

$$S^-(\alpha\tilde{\boldsymbol{x}}+\beta\tilde{\boldsymbol{z}}).$$

As in case 1, the above is valid if none of the vectors along the way are identically zero. Again, this is because theorem VD only holds for non-zero vectors. The initial vectors

$$\tilde{\boldsymbol{x}} \quad \text{and} \quad \tilde{\boldsymbol{z}} - \frac{z_{N+1}}{\tilde{\lambda} r}\tilde{\boldsymbol{x}}$$

are linearly independent, since $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{z}}$ are linearly independent. Thus, if any of the intermediate vectors are zero, then, necessarily,

$$\tilde{\lambda}|\tilde{x}_k|^{r-1}\left[\alpha\tilde{x}_k + \beta\left(z_k - \frac{z_{N+1}}{\tilde{\lambda} r}\tilde{x}_k\right)\right] = 0$$

for $k = 1, \ldots, N$, which we rewrite (since $\tilde{\lambda} \neq 0$) as

$$\tilde{x}_k\left[\left(\alpha - \frac{\beta z_{N+1}}{\tilde{\lambda} r}\right)\tilde{x}_k + \beta z_k\right] = 0 \tag{4.23}$$

for $k = 1, \ldots, N$.

From (4.23) and (4.14), it follows that if $\beta = 0$, then $\alpha = 0$. Assuming $\beta \neq 0$, we have

$$z_\ell = -\frac{1}{\beta}\left(\alpha - \frac{\beta z_{N+1}}{\tilde{\lambda} r}\right)\tilde{x}_\ell$$

if $\tilde{x}_\ell \neq 0$. But then, from (4.12) and (4.14),

$$0 = \sum_{\ell=1}^{N}\tilde{x}_\ell z_\ell = -\frac{1}{\beta}\left(\alpha - \frac{\beta z_{N+1}}{\tilde{\lambda} r}\right)\sum_{\ell=1}^{N}\tilde{x}_\ell^2 = -\frac{1}{\beta}\left(\alpha - \frac{\beta z_{N+1}}{\tilde{\lambda} r}\right).$$

Thus we must have

$$\alpha - \frac{\beta z_{N+1}}{\tilde{\lambda} r} = 0.$$

Substituting in (4.23) we see that this implies

$$\tilde{x}_k z_k = 0 \tag{4.24}$$

for all $k = 1, \ldots, N$.

Applying (4.24) to (4.11) we have

$$\sum_{\ell=1}^{N} b_{k\ell} z_\ell = -\tilde{\lambda}|\tilde{x}_k|^{r-1}\frac{z_{N+1}\tilde{x}_k}{\tilde{\lambda} r}, \tag{4.25}$$

for $k = 1, \ldots, N$. From (4.25) it follows that

$$S^-(\tilde{\boldsymbol{x}}) = S^+(\tilde{\boldsymbol{x}}) = S^+(B\tilde{\boldsymbol{z}}) \leqslant S^-(\tilde{\boldsymbol{z}}).$$

On the other hand, from (4.24),

$$S^-(\tilde{\boldsymbol{z}}) + 1 \leqslant \#\{i : z_i \neq 0\} \leqslant \#\{i : \tilde{x}_i = 0\} \leqslant S^+(\tilde{\boldsymbol{x}}),$$

and thus

$$S^-(\tilde{\boldsymbol{z}}) < S^+(\tilde{\boldsymbol{x}}).$$

We have arrived at a contradiction. This proves (4.20) for all $(\alpha, \beta) \neq (0, 0)$ and completes the proof of the proposition. ☐

As we noted before the statement of proposition 4.1, $G(\boldsymbol{x}, \lambda, t)$ is not necessarily a C^1 function at $t = 0$. This is a consequence of the fact that $|y|^{p-1}$ (which appears in various guises in (4.4)) does not necessarily have a limit as both $y \to 0$ and $p \to 1$. Of course, if $y \to y^* \neq 0$, then this limit does exist (and equals 1). The proof of proposition 4.1 in the case where $\tilde{\boldsymbol{x}}$ and $A_k \cdots A_1 \tilde{\boldsymbol{x}}$, $k = 1, \dots, m$, have no zero components is much simpler since $\tilde{x}_k z_k = 0$ for $k = 1, \dots, N$ implies $\tilde{\boldsymbol{z}} = \boldsymbol{0}$. We will not repeat this proof here. In fact, under the assumption of the non-vanishing of the components of the above vectors we can also apply the above analysis for any $p_i > 0$, for all i, with $r = p_1 \cdots p_{m-1}$. We state this in the special case of $(p_1(0), \dots, p_{m-1}(0)) = (1, \dots, 1)$ as we will first use it to prove theorem 1.1.

PROPOSITION 4.2. *Let $\tilde{\boldsymbol{x}} \in \mathbb{R}^N$ and $\tilde{\lambda} > 0$ be such that*

$$\boldsymbol{G}(\tilde{\boldsymbol{x}}, \tilde{\lambda}, 0) = \boldsymbol{0}. \tag{4.26}$$

Further assume that each of the vectors $\tilde{\boldsymbol{x}}$ and $A_k \cdots A_1 \tilde{\boldsymbol{x}}$, $k = 1, \dots, m$, has no zero component. Then

$$\det\left(\frac{\partial G_k}{\partial x_\ell}\right)_{k,\ell=1}^{N+1}\Bigg|_{(\tilde{\boldsymbol{x}}, \tilde{\lambda}, 0)} \neq 0. \tag{4.27}$$

We can now prove theorem 1.1.

Proof of theorem 1.1. Let $n \in \{0, \dots, R-1\}$. From theorem GK2 the linear problem has a solution. That is, there exists an $\boldsymbol{x} \in \mathbb{R}^N$ and $\lambda > 0$ for which $\boldsymbol{G}(\boldsymbol{x}, \lambda, 0) = \boldsymbol{0}$ and $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x}) = n$.

Assume $\boldsymbol{x}$ and $A_k \cdots A_1 \boldsymbol{x}$, $k = 1, \dots, m$, have no zero components. Then from proposition 4.2 and the implicit function theorem there exists a $t^* > 0$ such that for all $t \in [0, t^*)$ there are continuously differentiable functions of t, namely $\boldsymbol{x}(t)$ and $\lambda(t)$ for which

$$\boldsymbol{G}(\boldsymbol{x}(t), \lambda(t), t) = \boldsymbol{0}.$$

Furthermore, since $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x}) = n$ and $\lambda > 0$, we have $S^-(\boldsymbol{x}(t)) = S^+(\boldsymbol{x}(t)) = n$ for t near 0.

Let $\tilde{t} > 0$ be the largest value for which for all $t \in [0, \tilde{t})$ there exist $\boldsymbol{x}(t)$ and $\lambda(t)$ such that $\boldsymbol{G}(\boldsymbol{x}(t), \lambda(t), t) = \boldsymbol{0}$ and $S^-(\boldsymbol{x}(t)) = S^+(\boldsymbol{x}(t)) = n$. We wish to prove our result for $t = 1$, and we thus want to show that $\tilde{t} > 1$. (In fact $\tilde{t} = \infty$.) Assume $\tilde{t} \leqslant 1$. There then exists a subsequence $t^k \uparrow \tilde{t}$ such that along this subsequence $\boldsymbol{x}(t^k) \to \tilde{\boldsymbol{x}}$ and $\lambda(t^k) \to \tilde{\lambda}$. (The boundedness of $\boldsymbol{x}(t)$ is given by (4.3), and that of $\lambda(t)$ easily proved.) From the limiting process it easily follows that $S^-(\tilde{\boldsymbol{x}}) \leqslant n \leqslant S^+(\tilde{\boldsymbol{x}})$ and $\boldsymbol{G}(\tilde{\boldsymbol{x}}, \tilde{\lambda}, \tilde{t}) = \boldsymbol{0}$. As $(\tilde{\lambda}, \tilde{\boldsymbol{x}})$ is an eigenvalue–eigenfunction pair we have from corollaries 3.2 and 3.3 that $S^-(\tilde{\boldsymbol{x}}) = S^+(\tilde{\boldsymbol{x}}) = n$ and $\tilde{\lambda} > 0$. We can now again apply the implicit function theorem using proposition 4.1, contradicting the maximality of $\tilde{t}$.

It remains for us to consider the case where some of the components of $\boldsymbol{x}$ and $A_k \cdots A_1 \boldsymbol{x}$, $k = 1, \dots, m$, are equal to zero. This is a technical problem which we overcome by perturbation.

We first appeal to lemma 6.6 in [12]. From the method of proof therein there exists, for $\varepsilon > 0$, small, $N_k \times N_{k-1}$ STP matrices A_k^ε which tend to A_k as $\varepsilon \downarrow 0$, and $\boldsymbol{x}^\varepsilon \in \mathbb{R}^N$ which tends to $\boldsymbol{x}$ as $\varepsilon \downarrow 0$, such that

$$A_m^\varepsilon \cdots A_1^\varepsilon \boldsymbol{x}^\varepsilon = \lambda \boldsymbol{x}^\varepsilon,$$

where $S^-(\boldsymbol{x}^\varepsilon) = n$, and $\boldsymbol{x}^\varepsilon$ and $A_k^\varepsilon \cdots A_1^\varepsilon \boldsymbol{x}^\varepsilon$, $k = 1, \ldots, m$, have no zero components. We can now apply the previous analysis to obtain a vector $\tilde{\boldsymbol{x}}^\varepsilon$ which satisfies the result of theorem 1.1 ($S^-(\tilde{\boldsymbol{x}}^\varepsilon) = S^+(\tilde{\boldsymbol{x}}^\varepsilon) = n$) for $p_1, \ldots, p_{m-1}$, but with the matrices A_k replaced by A_k^ε. We now let $\varepsilon \downarrow 0$, and obtain (at least on a subsequence) a limiting vector $\tilde{\boldsymbol{x}}$ and value $\tilde{\lambda}$ which necessarily satisfies

$$\boldsymbol{G}(\tilde{\boldsymbol{x}}, \tilde{\lambda}, 1) = \boldsymbol{0},$$

and $S^-(\tilde{\boldsymbol{x}}) = S^+(\tilde{\boldsymbol{x}}) = n$. This proves theorem 1.1. □

If some (or all) of the $p_i < 1$, then the above method of proof is valid at those $\boldsymbol{x}$ for which none of the components of $\boldsymbol{x}$ nor the diagonal entries of $D_k(\boldsymbol{x})$, $k = 1, \ldots, m$, vanish (see (4.4)). If $S^-(\boldsymbol{x}) = S^+(\boldsymbol{x}) = 0$, and $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair, then $\boldsymbol{x}$ has no zero components, nor do the diagonal entries of $D_k(\boldsymbol{x})$ for any $k = 1, \ldots, m$. Thus we immediately have the following result (which we will generalize in § 5).

Proposition 4.3. *Assume the A_i are as in theorem 1.1 and $r = p_1 \cdots p_{m-1}$. There exists exactly one eigenvalue–eigenvector pair $(\lambda, \boldsymbol{x})$ satisfying (4.1) with $\boldsymbol{x}$ having all positive components. Furthermore, if μ is any other eigenvalue of (4.1), then $\lambda > \mu > 0$.*

If $\boldsymbol{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and $S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) = N - 1$, then $x_i x_{i+1} < 0$ and in particular none of its components vanish. This will also be true for the diagonal entries of $D_k(\boldsymbol{x})$ for every $k = 1, \ldots, m$, if each is in $\mathbb{R}^N$ and

$$\begin{aligned} S^+(A_k(A_{k-1} \cdots (A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*} \cdots)^{p_{k-1}*}) \\ &= S^-(A_k(A_{k-1} \cdots (A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*} \cdots)^{p_{k-1}*}) \\ &= N - 1. \end{aligned}$$

These latter equalities will hold when $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair, $S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) = N - 1$, and each of the A_i is an $N \times N$ STP matrix. (See, for example, (3.3) and the method of proof of corollary 3.2.) As such we have the following.

Proposition 4.4. *Assume the A_i are $N \times N$ STP matrices as in theorem 1.1 and $r = p_1 \cdots p_{m-1}$. Then there exists exactly one eigenvalue–eigenvector pair $(\lambda, \boldsymbol{x})$ satisfying (4.1) with $S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) = N - 1$. Furthermore, if μ is any other eigenvalue of (4.1), then $0 < \lambda < \mu$.*

Let us rewrite (4.1) as

$$(A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*} \cdots)^{p_{m-1}*})^{p_m*} = \lambda \boldsymbol{x}, \tag{4.28}$$

where we have simply set $p_m = 1/r$. (This λ is the previous λ to the power $1/r$.) The condition $p_1 \cdots p_{m-1} = r$ is rewritten as

$$p_1 \cdots p_{m-1} p_m = 1. \tag{4.29}$$

We have proved theorem 1.1 under the assumption that $p_i \geqslant 1$ for $i = 1, \dots, m-1$ (and thus $p_m \leqslant 1$). We conjecture that theorem 1.1 holds for all choices of $p_i > 0$ satisfying (4.29). We can generalize theorem 1.1 to the following cases.

PROPOSITION 4.5. *Assume $p_1 \cdots p_m = 1$, and exactly one of the p_i is strictly less than 1. Then the results of theorem 1.1 hold.*

Proof. We show that starting with (4.28) we can 'rotate' the $\{p_i\}_{i=1}^m$. We do this by setting

$$\boldsymbol{y} = (A_1 \boldsymbol{x})^{p_1 *}.$$

Multiplying (4.28) by A_1 and raising both sides thereof to the power p_1*, we get

$$(A_1(A_m(\cdots(A_2\boldsymbol{y})^{p_2*}\cdots)^{p_{m-1}*})^{p_m*})^{p_1*} = \lambda^{p_1}\boldsymbol{y}. \tag{4.30}$$

If we assume that $p_k < 1$, while all the other $p_i \geqslant 1$, then we can rotate as above until the powers are in the order $p_{k+1}, \dots, p_m, p_1, \dots, p_k$. We now raise both sides of the equation to the power $1/p_k*$ and apply theorem 1.1.

Furthermore, it is easily checked that if $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair for (4.28) with $S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) = r$, then $(\lambda^{p_1}, \boldsymbol{y})$ is an eigenvalue–eigenvector pair for (4.30) with $S^+(\boldsymbol{y}) = S^-(\boldsymbol{y}) = r$. Since we may continue to rotate, eventually arriving back at (4.28), it is also true that if $(\lambda^{p_1}, \boldsymbol{y})$ is an eigenvalue–eigenvector pair for (4.30) with $S^+(\boldsymbol{y}) = S^-(\boldsymbol{y}) = r$, then $(\lambda, \boldsymbol{x})$ is an eigenvalue–eigenvector pair for (4.28). We apply this type of argument to translate our N eigenvalue–eigenvector pairs obtained from our application of theorem 1.1 back to (4.28). □

If we assume that each A_i is an $N \times N$ STP matrix (and thus invertible), then we also have the following.

PROPOSITION 4.6. *Assume $p_1 \cdots p_m = 1$, and exactly one of the p_i is strictly greater than 1. If each A_i is an $N \times N$ STP matrix, then the results of theorem 1.1 hold.*

Proof. As the A_i are invertible, we shall invert (4.28) and then apply proposition 4.5.

Inverting each of the A_i, and then raising it to the appropriate power, it is readily verified that (4.28) is equivalent to

$$B_1(\cdots(B_m(\boldsymbol{x})^{q_m*})^{q_{m-1}*}\cdots)^{q_1*} = \frac{1}{\lambda}\boldsymbol{x}, \tag{4.31}$$

where $B_i = A_i^{-1}$ and $q_i = 1/p_i$, $i = 1, \dots, m$. If A is any $N \times N$ STP matrix and $B = A^{-1}$, then JBJ is an $N \times N$ STP matrix, where J is the diagonal matrix with ith diagonal entry equal to $(-1)^i$. Set $C_i = JB_iJ$, $i = 1, \dots, m$, and $\boldsymbol{y} = J\boldsymbol{x}$. Then (4.31) (and thus (4.28)) is equivalent to

$$C_1(\cdots(C_m(\boldsymbol{y})^{q_m*})^{q_{m-1}*}\cdots)^{q_1*} = \frac{1}{\lambda}\boldsymbol{y}. \tag{4.32}$$

Furthermore, $S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) = i - 1$ if and only if $S^+(\boldsymbol{y}) = S^-(\boldsymbol{y}) = N - i$. As exactly one of the p_i is in value greater than 1, then exactly one of the q_i is in value less than 1, and we can apply proposition 4.5 to obtain the desired result. □

5. The non-homogeneous case

In this section we consider the equation

$$T(\boldsymbol{x}) = A_m(\cdots(A_2(A_1\boldsymbol{x})^{p_1*})^{p_2*}\cdots)^{p_{m-1}*} = \lambda\boldsymbol{x}^{r*}, \tag{5.1}$$

where $p_1\cdots p_{m-1} \neq r$. This equation for $m = 2$ and $A_2 = A_1^{\mathrm{T}}$ (see (1.2)) was considered by Buslaev in [3]. Buslaev proved (theorem B) for each $i \in \{1,\dots,R\}$ the existence of an eigenvalue–eigenvector pair with exactly $i-1$ sign changes. The proof used the inherent 'adjointness' of the equation. We conjecture that this result holds in general. Moreover, as was pointed out by Buslaev, even in this case we may have non-uniqueness.

Consider, for example, the simple equation

$$\begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}^{r*}. \tag{5.2}$$

For each solution we have $x, y \neq 0$. Therefore, we may assume, without loss of generality, that $y = 1$. Thus (5.2) is equivalent to the equation

$$\frac{3x+1}{2x+4} = x^{r*}.$$

For $r > 1$, there is exactly one positive root x_1, $0 < x_1 < 1$, and at least one negative root. There is also a value $\tilde{r} \approx 2.5$ such that if $1 \leqslant r < \tilde{r}$, then there is precisely one negative root, while for $r > \tilde{r}$ there are precisely three negative roots (all of which lie in $(-2, -\frac{1}{3})$). For $0 < r < 1$ there is precisely one negative root and at least one positive root. There is also a value $\hat{r} \approx 0.112$ such that, if $1 \geqslant r > \hat{r}$, there is precisely one positive root, while for $0 < r < \hat{r}$ there are precisely three positive roots. We do not understand exactly what can be said in general. However, we can partly explain the above in the next two results.

We will first prove that a Perron theorem argument can be applied in the case $r \geqslant p_1\cdots p_m$. Moreover, this holds without the STP property of the A_i. It is only necessary to assume that each A_i is strictly positive (i.e. all the elements of A_i are strictly positive).

Theorem 5.1. *Assume each matrix A_i is strictly positive and $r \geqslant p_1\cdots p_{m-1}$. Then there exists exactly one eigenvalue–eigenvector pair $(\lambda, \boldsymbol{x})$ for (5.1) with the eigenvector $\boldsymbol{x}$ having all positive components. Furthermore, if $(\mu, \boldsymbol{y})$ is any other eigenvalue–eigenvector pair for (5.1), and $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$, then $\lambda > |\mu|$.*

Proof. Let Λ denote the set of all $\lambda > 0$ for which there exists a non-negative vector $\boldsymbol{x}$ satisfying $\|\boldsymbol{x}\|_2 = 1$ and $T(\boldsymbol{x}) \geqslant \lambda\boldsymbol{x}^r$. It is easily verified that Λ is non-empty and bounded. As such it has a supremum which we will denote by λ_0. Let λ_k be a sequence of values in Λ which increase to λ_0. By definition there exist non-negative vectors $\boldsymbol{x}_k$ satisfying $\|\boldsymbol{x}_k\|_2 = 1$ and $T(\boldsymbol{x}_k) \geqslant \lambda_k\boldsymbol{x}_k^r$. Taking subsequences, if necessary, we may assume that $\lim_{k\to\infty}\boldsymbol{x}_k = \boldsymbol{x}_0$. Thus $\|\boldsymbol{x}_0\|_2 = 1$ and $T(\boldsymbol{x}_0) \geqslant \lambda_0\boldsymbol{x}_0^r$.

We claim that $T(\boldsymbol{x}_0) = \lambda_0\boldsymbol{x}_0^r$. Assume not. Set

$$\boldsymbol{y}_0 = (T(\boldsymbol{x}_0))^{1/r} \geqslant \lambda_0^{1/r}\boldsymbol{x}_0.$$

As $\boldsymbol{y}_0$ is not equal to $\lambda_0^{1/r}\boldsymbol{x}_0$ it follows that

$$T(\boldsymbol{y}_0) > T(\lambda_0^{1/r}\boldsymbol{x}_0) = \lambda_0^{p_1\cdots p_{m-1}/r}T(\boldsymbol{x}_0) = \lambda_0^{p_1\cdots p_{m-1}/r}\boldsymbol{y}_0^r.$$

However, $\|\boldsymbol{y}_0\|_2 \neq 1$ and so this inequality does not yet tell us much.

Set

$$\boldsymbol{z}_0 = \alpha\boldsymbol{y}_0,$$

where $\|\boldsymbol{z}_0\|_2 = 1$, i.e. $\alpha = \|\boldsymbol{y}_0\|_2^{-1}$. Thus

$$T(\boldsymbol{z}_0) = \alpha^{p_1\cdots p_{m-1}}T(\boldsymbol{y}_0) > \alpha^{p_1\cdots p_{m-1}}\lambda_0^{p_1\cdots p_{m-1}/r}\boldsymbol{y}_0^r = \alpha^{p_1\cdots p_{m-1}-r}\lambda_0^{p_1\cdots p_{m-1}/r}\boldsymbol{z}_0^r.$$

Now $\boldsymbol{y}_0 \geqslant \lambda_0^{1/r}\boldsymbol{x}_0 \geqslant \boldsymbol{0}$ and therefore

$$\frac{1}{\alpha} = \|\boldsymbol{y}_0\|_2 > \lambda_0^{1/r}\|\boldsymbol{x}_0\|_2 = \lambda_0^{1/r},$$

i.e. $\alpha\lambda_0^{1/r} < 1$. By assumption $p_1\cdots p_{m-1} - r < 0$. Thus

$$\alpha^{p_1\cdots p_{m-1}-r}\lambda_0^{p_1\cdots p_{m-1}/r} > \lambda_0.$$

We have contradicted our definition of λ_0. Thus $(\lambda_0, \boldsymbol{x}_0)$ is an eigenvalue–eigenvector pair for (5.1) with $\boldsymbol{x}_0$ strictly positive and $\|\boldsymbol{x}_0\|_2 = 1$. This completes the existence proof.

If $T(\boldsymbol{y}) = \mu\boldsymbol{y}^{r*}$ and $\|\boldsymbol{y}\|_2 = 1$, then setting

$$|\boldsymbol{y}| = (|y_1|, \ldots, |y_N|),$$

it easily follows that

$$T(|\boldsymbol{y}|) \geqslant |\mu||\boldsymbol{y}|^r,$$

and of course $\||\boldsymbol{y}|\|_2 = 1$. Thus by definition $\lambda_0 \geqslant |\mu|$. If $\lambda_0 = |\mu|$, then we must have $T(|\boldsymbol{y}|) = \lambda_0|\boldsymbol{y}|^r$ and $\boldsymbol{y}$ must be a vector of one sign.

It therefore remains to prove that there can only be one eigenvalue–eigenvector pair having all positive components.

Assume $(\lambda, \boldsymbol{x})$ and $(\mu, \boldsymbol{y})$ are eigenvalue–eigenvector pairs having all positive components, with $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$, $\boldsymbol{x}$, $\boldsymbol{y}$ linearly independent, and $\lambda \leqslant \mu$. Choose $\alpha > 0$ such that $\alpha\boldsymbol{x} - \boldsymbol{y}$ remains a positive vector, but $\gamma\boldsymbol{x} - \boldsymbol{y}$ is not positive for every $\gamma < \alpha$. Note that as $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$ and $\boldsymbol{x}$, $\boldsymbol{y}$ are linearly independent, it follows that $\alpha > 1$. Since the A_i are strictly positive matrices we have

$$\alpha^{p_1\cdots p_{m-1}}\lambda\boldsymbol{x}^r = T(\alpha\boldsymbol{x}) \geqslant T(\boldsymbol{y}) = \mu\boldsymbol{y}^r.$$

Thus, the vector $\lambda\alpha^{p_1\cdots p_{m-1}}\boldsymbol{x}^r - \mu\boldsymbol{y}^r$ is strictly positive, implying

$$\left(\frac{\lambda}{\mu}\right)^{1/r}\alpha^{p_1\cdots p_{m-1}/r} > \alpha.$$

As $r \geqslant p_1\cdots p_{m-1}$ and $\alpha > 1$, this implies that

$$\lambda > \mu.$$

But this contradicts our assumption. Thus there is at most one eigenvalue–eigenvector pair $(\lambda, \boldsymbol{x})$ with the vector $\boldsymbol{x}$ having all positive components. □

If the A_i are $N \times N$ STP matrices, then we can invert (5.1), as in proposition 4.6, and then apply theorem 5.1 to the smallest eigenvalue. We then obtain the following corollary.

COROLLARY 5.2. *Assume each A_i is an $N \times N$ STP matrix and $p_1 \cdots p_{m-1} \geqslant r$. Then there exists exactly one eigenvalue–eigenvector pair $(\lambda, \boldsymbol{x})$ for (5.1) with $S^+(\boldsymbol{x}) = S^-(\boldsymbol{x}) = N-1$. Furthermore, if $(\mu, \boldsymbol{y})$ is any other eigenvalue–eigenvector pair for (5.1), and $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$, then $\lambda < |\mu|$.*

Acknowledgments

This research was supported by the Fund for the Promotion of Research at the Technion.

References

1 J. Aczél. *Lectures on functional equations and their applications* (Academic, 1966).

2 T. Ando. Totally positive matrices. *Linear Algebra Applic.* **90** (1987), 165–219.

3 A. P. Buslaev. A variational description of the spectra of totally positive matrices, and extremal problems of approximation theory. *Mat. Zametki* **47** (1990), 39–46. (Transl. *Math. Notes* **47** (1990), 26–31.)

4 U. Elias and A. Pinkus. Nonlinear eigenvalue problems for a class of ordinary differential equations. *Proc. R. Soc. Edinb.* A **132** (2002), 1333–1359. (This issue.)

5 F. R. Gantmacher and M. G. Krein. Sur les matrices oscillatoires. *C. R. Hebd. Seanc. Acad. Sci. Paris* **201** (1935), 577–579.

6 F. R. Gantmacher and M. G. Krein. Sur les matrices complètement non négatives et oscillatoires. *Compositio Math.* **4** (1937), 445–476.

7 F. R. Gantmacher and M. G. Krein. *Oscillation matrices and small oscillations of mechanical systems* (Moscow: Gostekhizdat, 1941). (In Russian.)

8 F. R. Gantmacher and M. G. Krein. *Ostsillyatsionye Matritsy i Yadra i Malye Kolebaniya Mekhanicheskikh Sistem* (Moscow: Gosudarstvenoe Izdatel'stvo 1950). (Transl. *Oscillation matrices and kernels and small vibrations of mechanical systems* (USAEC, 1961).)

9 S. Karlin. *Total positivity*, vol. 1 (Stanford University Press, 1968).

10 D. M. Koteljanskii. Some sufficient conditions for reality and simplicity of the spectrum of a matrix. *Am. Math. Soc. Transl. 2* **27** (1963), 35–42.

11 A. Pinkus. Some extremal problems for strictly totally positive matrices. *Linear Algebra Applic.* **64** (1985), 141–156.

12 A. Pinkus. n-widths of Sobolev spaces in L^p. *Constr. Approx.* **1** (1985), 15–62.

13 I. J. Schoenberg. Über variationsvermindernde lineare Transformationen. *Math. Z.* **32** (1930), 321–328.

14 I. J. Schoenberg. *I. J. Schoenberg: selected papers* (ed. C. de Boor) (Birkhäuser, 1988).

(*Issued 13 December 2002*)