THE SPECTRAL THEORY OF NON-NEGATIVE MATRICES

M. R. F. Smyth

Department of Health and Social Services (Northern Ireland)

ABSTRACT

A proof of the celebrated Perron-Frobenius results for non-negative matrices is given via classical spectral theory. This approach is based on the properties of a hierarchy of subsets of the non-negative matrices which are of considerable interest in their own right and which are detailed here.

§1 Introduction

The main aim of this document is to present elementary proofs of two classical results, namely the Perron-Frobenius (PF) Theorem and the Compact Group Theorem, using basic spectral theory. For background the reader is only required to have a working knowledge of spectral theory as can be found in any comprehensive treatise on linear operators such as [1].

A brief word on notation. Although a matrix is an object of interest in its own right, we shall often prefer to regard it as a linear operator on a finite-dimensional vector space. If T is such an operator we shall formally denote its matrix relative to a fixed basis by [T] and the actual element in row i, column j by $[T]_{ij}$. Where there is no danger of confusion the square brackets may be omitted. The transpose of T will be denoted by T^T. We will be chiefly interested in non-negative matrices and we shall use $\aleph_m$ to denote the set of all such m-square matrices. Whenever the size of the matrix is immaterial we shall simply write T ∈ ℵ or, more usually, $T \geq 0$. We shall write $\text{diag}(T) > 0$ if every entry on the diagonal of T is positive. The *spectrum* or set of eigenvalues of T is denoted by $\sigma(T)$ and the largest modulus amongst its eigenvalues, the *spectral radius* of T, is denoted by $r(T)$. The *peripheral spectrum* $\pi(T)$ is the set of eigenvalues of T with modulus $r(T)$, and the number of elements in $\pi(T)$ is called the *index of imprimitivity* of T.

If T is any matrix we shall say that T is *zero-free* whenever T does not have either a zero row or a zero column. For our purposes this will be a key property. Another important attribute of any matrix $T \geq 0$ is its *potency* namely the smallest positive integer n such that $\text{diag}(T^n) > 0$. If no such number exists then T is said to be *impotent*. Note that a potent matrix must always be zero-free and have a positive spectral radius. Since a nilpotent matrix must be impotent (easily seen by considering its trace), impotence may be regarded as a rather weak form of nilpotence.

Two standard concepts for dealing with non-negative matrices are *connectedness* and *permutations*. The latter are often used to transform a matrix to make it more tractable. They effectively re-order the basis of the underlying space and by this means alter the positions of rows and columns in the matrix. A valid row/column interchange is to swop row i with row j and then to swop column i and column j. As these swops commute their order is immaterial. It will be seen that whilst such interchanges move rows and columns around, the contents of rows and columns (and also the diagonal) may be re-ordered but never changed. Permutation matrices have many trivial properties. There is a single non-zero entry (one) in each row and each column. They are all invertible and their inverses are the same as their transposes.

A non-negative matrix T is said to be *decomposable* (some authors use *reducible*) if there exists a permutation matrix E such that ETE^T has a 2 x 2 block form where the diagonal blocks are both square and the top right hand corner block is zero. Matrices that do not enjoy this property are called *connected* (also *indecomposable* and *irreducible*). See [2] page 122 for details. There are several equivalent definitions of connectedness and one of particular importance is that $T \geq 0$ is connected if for every i, j there exists n such that $[T^n]_{ij} > 0$. This is presented as exercise 5.2.6 on page 122 of [2]. It is not hard to see that this condition implies connectedness, but the converse is a little tricky. It may be proved using a permutation similar to that of Lemma 7 with suitable alterations; in other words by concentrating on elements i, j such that $[T^n]_{ij} = 0$ for all n instead of the zeros of P. It is also easy to see that if $T \geq 0$ is connected then T is potent and $r(T) > 0$. For to every i there is an n(i) such that $[T^{n(i)}]_{ii} > 0$, and the product of the n(i) is then always a suitable candidate for the potency.

Two fundamental ideas which permeate the following proofs are *hierarchy* and *inheritance*. The latter is demonstrated in the crucial Lemma 4 wherein we see the child operator inheriting connectedness and dominating its parent to a quite astonishing extent. It also surfaces in inductive proofs such as those of Lemma 7 and Proposition 4 where the properties of several small matrices are used to derive similar properties for a larger one. The former is based on the observation that to descend from a positive matrix to a non-negative matrix there are three intermediate steps along the way which are neatly summed up in the following simple hierarchy statement.

positivity $\Rightarrow$ connectedness $\Rightarrow$ potency $\Rightarrow$ zero-freedom

All of these properties will be central to our work. Note that they are all qualitative as opposed to quantitative. In other words they only depend on the entries in the matrix insofar as those entries are either zero or non-zero; to them the magnitude of the various entries in the matrix is of no consequence whatsoever. Note also that each of them is invariant under any permutation of the basis of the underlying space.

Last but not least the author would like to record his thanks to T. T. West for his help, advice, and occasional inspiration during the preparation of this document.

§2 Perron-Frobenius

Our first target is to prove the PF-Theorem. Some preliminary results are needed and we deal with them before giving the proof proper. Matrices with positive diagonals are discussed in more detail in [4]. Indeed Lemma 2 below is an easy corollary of the eccentric disk theorem (see [4] proposition 6).

- **Lemma 1**

If $0 \neq T \geq 0$ and T commutes with a connected non-negative matrix then T is zero-free.

Proof : Suppose $S \geq 0$ is connected and ST = TS. Choose i, j such that $[T]_{ij} > 0$. Then given k by connectivity there is an n such that $[S^n]_{ki} > 0$ so $[S^nT]_{kj} > 0$. By commutativity $[TS^n]_{kj} > 0$ hence row k of T is non-zero. Thus T has no zero rows and a similar argument shows that it has no zero columns either ☺

- **Lemma 2**

If $T \geq 0$ and diag(T) > 0 then $\pi(T) = \{r(T)\}$.

Proof : Clearly r(T) > 0 so we may assume without loss of generality that r(T) = 1. Let $C = \{e^{i\theta} ; \pi/2 \leq \theta \leq 3\pi/2\}$. Fix n and let ε be the smallest entry on the diagonal of T^n. Since reducing the diagonal elements clearly cannot increase the spectral radius we have $r(T^n - \varepsilon 1) \leq r(T^n) = 1$ and $\sigma(T^n - \varepsilon 1)$ is just $\sigma(T^n)$ shifted left by $\varepsilon > 0$ so T^n cannot have eigenvalues on C. As n is arbitrary $\pi(T) = \{1\}$ by the spectral mapping theorem ☺

- **Lemma 3** (the peripheral projection)

Suppose $T \geq 0$ is connected with potency c and spectral radius r. The spectral projection (called *the peripheral projection of T*) given by $P = P(\pi(T);T)$ satisfies the following :-

(i) TP = PT and $P = P(r^c;T^c)$;

(ii) $P \geq 0$;

(iii) $T^cP = r^cP$;

(iv) diag(P) > 0.

Proof : The results of (i) are standard properties of spectral projections. Since $r > 0$ we may assume, using appropriate scaling, that $r = 1$. Then $\pi(T^c) = \{1\}$ by Lemma 2, and by the usual spectral decomposition $T^c = U \oplus V$ where $U = T^c |PX$ and $V = T^c |(1 - P)X$. If ν is chosen such that $r(V) < \nu < 1$ then ultimately the norm of V^k is bounded by ν^k and so $V^k \to 0$ as $k \to \infty$. Now suppose that 1 is a pole of T^c of order $q + 1$ where $q \geq 1$. Then $U = 1 + N$ where $N^q \neq 0$, $N^{q+1} = 0$ so $U^k = 1 + kN + ... + {}^kC_qN^q$ which is a polynomial in k of degree q, the coefficient of k^q being $N^q/q!$ which is a non-zero operator, W say. Moreover it follows immediately that $UW = W$ and $W^2 = 0$. Then $k^{-q} U^k \to W$ and writing $E = W \oplus 0$ shows that $k^{-q} T^{kc} = k^{-q} U^k \oplus k^{-q} V^k \to E$. Note that $E \geq 0$ since $T \geq 0$ and also $TE = ET$. Now $E^2 = W^2 \oplus 0 = 0$. It follows that $ET^kE = 0$ for all k. But if $[E]_{ij} \neq 0$ then by connectedness there would exist k such that $[T^k]_{ji} \neq 0$ which would mean that $ET^kE \neq 0$ which is false. Therefore 1 is a simple pole of T^c and $U = 1$. Hence $T^cP = P$ which proves (iii). Note that $T^{kc} = 1 \oplus V^k \to 1 \oplus 0 = P$ so P is real and $P \geq 0$ proving (ii). Finally to prove (iv) we show $[P]_{ii} > 0$. Since $P \geq 0$ is non-zero and commutes with the connected $T \geq 0$ by Lemma 1 there exists j such that $[P]_{ij} > 0$. Again by connectedness choose k so that $[T^k]_{ji} > 0$. This implies $[PT^k]_{ii} > 0$ so raising to the c^{th} power $[PT^{kc}]_{ii} > 0$, but $PT^{kc} = P$ hence $[P]_{ii} > 0$ ☺

- **Lemma 4** (the child operator)

Let $T \geq 0$ be connected with potency c, spectral radius r, peripheral projection P, and index of imprimitivity h. Then *the child operator of T* given by $R = TP$ satisfies :-

(i) $\pi(R) = \pi(T) = \sigma(R)\backslash\{0\}$ and $P = P(\pi(R);R)$;

(ii) $R \geq 0$ is connected ;

(iii) If (for any n) an entry in $[R^n]$ is zero the corresponding entry in $[T^n]$ is also zero ;

(iv) h is equal to the potency of R ;

(v) $R^h = r^hP$ and $R^{h+1} = r^hR$;

(vi) c is divisible by h ;

(vii) If n is divisible by h then $\mathrm{diag}(R^n) > 0$; otherwise $\mathrm{diag}(R^n) = 0$;

(viii) If $[R^n]_{ij} > 0$ then $[R^m]_{ji} > 0 \Leftrightarrow (m+n)$ is divisible by h .

Proof : (i) is standard. By Lemma 3(ii) it follows that R is real ≥ 0. If $[R^n]_{ij} = 0$ then $[T^nP]_{ij} = 0$ so $\Sigma\, [T^n]_{ik}[P]_{kj} = 0$ and as all terms are non-negative and $[P]_{jj} > 0$ this means that $[T^n]_{ij} = 0$. Hence R is connected proving (ii) and (iii). (iv) and (v) now follow by applying Lemma 3 to R and using (i). If $[R^n]_{jj} > 0$ by connectedness there exist u, v with $[R^u]_{ij} > 0$ and $[R^v]_{ji} > 0$. Then $[R^{u+v}]_{ii} \geq [R^u]_{ij}[R^v]_{ji} > 0$ so $[R^u]_{ij}[R^n]_{jj}[R^v]_{ji}[R^{(u+v)(h-1)}]_{ii} > 0$ so $[R^{h(u+v)+n}]_{ii} > 0$ so $[R^n]_{ii} > 0$ by (v). This implies $\mathrm{diag}(R^n) > 0$ hence $n \geq h$ and $\mathrm{diag}(R^n) = 0$ for $n < h$. This proves (vi) and (vii). Finally if $[R^n]_{ij} > 0$ and (m+n) is not a multiple of h then $[R^m]_{ji} = 0$ by (vii). But as R is connected we know by (v) that $[R^m]_{ji}$ is periodically positive and hence $[R^m]_{ji} > 0$ whenever (m+n) is a multiple of h. This proves (viii) ☺

- **Lemma 5**

A positive projection P always has rank one and may be written as u⊗f where u is the first column of P and f is the top row of P divided by $[P]_{11}$. This gives f(u) = 1.

Proof : Clearly Pu = u and u > 0. Let v be any other column and λ be the largest real number such that $w = v - \lambda u \geq 0$. Then Pw = w and at least one component, say number k, of w is zero so $\sum [P]_{kj}w_j = 0$ which means that w = 0 and hence every column of P is a multiple of the first one. Thus P has rank one and we can write it in the given form ☺

- **Lemma 6** (the Perron projection)

Suppose that T ≥ 0 is connected with spectral radius r, peripheral projection P, index of imprimitivity h, and child R. Then for any $\lambda \in \pi(T)$ the spectral projection P(λ;T) is given by the formula $P(\lambda;T) = h^{-1}(P + \lambda^{-1}R + \lambda^{-2}R^2 + ... + \lambda^{-(h-1)}R^{h-1})$. In particular P(r;T) which will be called the *Perron projection of T* is positive and has rank one.

Proof : Let $Q = h^{-1}(P + \lambda^{-1}R + \lambda^{-2}R^2 + ... + \lambda^{-(h-1)}R^{h-1})$. As $R^h = r^hP$ by Lemma 4(v) it follows that Q is a projection. Obviously QR = λQ = RQ so Q reduces R. Note that (R-λ)|QX is zero, and also $(R-\lambda)(1-Q)x = 0 \Rightarrow Rx = \lambda x \Rightarrow Qx = x \Rightarrow (1-Q)x = 0$ so (R-λ)|(1-Q)X is one-to-one and hence invertible. Therefore Q = P(λ;R) = P(λ;T) as required. Now $P(r;T) \geq h^{-1}r^{-1}R \geq 0$ (note that if h = 1 then $P = r^{-1}R$) so P(r;T) is a connected projection hence it must be positive. By Lemma 5 its rank must be one ☺

Having scaled the lesser peaks we are now finally ready to tackle the big one. The observant reader will notice that statement (IV) below is much stronger than the version that is usually quoted, namely that $\sigma(T)$ is invariant under rotation by an angle $2\pi/h$. It implies that spectral characteristics such as pole order and algebraic and geometric multiplicities are faithfully preserved under rotation. Used in conjunction with (I) it also shows that if $\lambda \in \pi(T)$ then $P(\lambda;T)$ has rank one and consequently that all points of the peripheral spectrum are simple roots of the characteristic equation of T.

Unfortunately most existing proofs of the PF-Theorem are neither straightforward nor intuitive. Indeed [2] declines to give one because it would be "too long and too involved". A new and attractive proof using semigroups has however recently been published in [3]. By comparison our proof is less general but more direct.

- **The Perron-Frobenius Theorem** (see [2] page 124)

Let $T \geq 0$ be connected with spectral radius r. Then :-

(I) r is a real positive eigenvalue (the Perron root of T) which is a simple root of the characteristic equation of T.

(II) There exists a positive eigenvector corresponding to r.

(III) If T has h eigenvalues of modulus r these are the h distinct roots of $z^h - r^h = 0$.

(IV) If $\omega = e^{2\pi i/h}$ then T is similar to ωT.

(V) If $h > 1$ there exists a permutation matrix E such that ETE^T has a representation consisting of h square blocks on the main diagonal and all blocks except those directly above the main diagonal and the one in the bottom left corner are zero.

Proof : By Lemma 6 the Perron projection is rank one which proves (I), moreover its first column is a positive eigenvector corresponding to the Perron root. This proves (II).

We next prove (V). In order to do so we look at the underlying complex vector space X with basis $\{e_1 , e_2 , ... , e_{\dim(X)}\}$ where e_i has a 1 in its i^{th} coordinate and 0 elsewhere. R imposes a strict regime on this basis to the extent that X is expressible as the direct sum of h subspaces $V_1 , V_2 , ... , V_h$ by the formula $V_n = \text{span}\{e_j \,;\, [R^n]_{1j} > 0\}$. Since $R^{h+1} = r^h R$

every base vector lies in at least one of these subspaces, moreover no base vector e_i can belong to more than one subspace, for if we had $[R^m]_{1i} > 0$, $[R^n]_{1i} > 0$ with $1 \le m \le h$ and $1 \le n < h$ then by Lemma 4(viii) we have $[R^{h-n}]_{i1} > 0$ so $[R^{h+m-n}]_{11} > 0$ so $h + m - n = h$ by Lemma 4(vii) which implies $m = n$. Thus $X = V_1 \oplus V_2 \oplus .... \oplus V_h$ and we now show that $R(V_{n+1}) \subseteq V_n$. [Here we are adopt a local convention that $V_0 = V_h$]. Suppose $e_j \in V_{n+1}$ and consider $Re_j = \Sigma_i[R]_{ij}e_i$. If $[R]_{ij} > 0$ then by Lemma 4(viii) we have $[R^{h-1}]_{ji} > 0$ so $[R^{n+1}]_{1j}[R^{h-1}]_{ji} > 0$ so $[R^{h+n}]_{1i} > 0$ thus $[R^n]_{1i} > 0$ so $e_i \in V_n$. Therefore $Re_j \in V_n$ and it follows that $R(V_{n+1}) \subseteq V_n$. Now by Lemma 4(iii) we see that T shares exactly the same property, namely $T(V_{n+1}) \subseteq V_n$ which forces the block decomposition of T given in (V).

Now let $\omega = e^{2\pi i/h}$. As X is a direct sum we can define a linear operator D on X by the formula $D(\Sigma v_n) = \Sigma\omega^{n-1}v_n$ where $v_n \in V_n$ for $1 \le n \le h$. Clearly D has a diagonal matrix and $D^h = 1$. If $v_n \in V_n$ then $Dv_n = \omega^{n-1}v_n \in V_n$ hence $TDv_n = \omega^{n-1}Tv_n \in V_{n-1}$ therefore $D^{-1}TDv_n = \omega^{-(n-2)}\omega^{n-1}Tv_n = \omega Tv_n$. Hence $D^{-1}TD = \omega T$ which proves (IV) and also shows that the roots of $z^h - r^h = 0$ lie in $\pi(T)$. However as $R^{h+1} = r^hR$ and $\pi(T) = \pi(R)$ there are at most h points in $\pi(T)$. This establishes (III) and completes the proof ☺☺

It is interesting to compare the peripheral projection and the Perron projection. The former is real, non-negative, and potent, the latter is real, positive, and rank one. Of course the spectral projections of general eigenvalues aren't normally real, but by taking non-real eigenvalues in conjugate pairs it is easy to see, using the standard complex integral formula, that the spectral projections of conjugate pairs of eigenvalues are real, and so too are the spectral projections corresponding to real eigenvalues. However none of these projections can be non-negative since they must all be orthogonal to the Perron projection.

§3 General non-negative matrices

We now drop the connectedness assumption and consider non-negative matrices in general. In this context it is vital for the reader to appreciate how Proposition 1 below and the PF-Theorem interact. The former gives a lower triangular form for any non-negative matrix, moreover the diagonal blocks of this form are connected matrices to which the latter can naturally be applied. Thus these two results complement one another, respectively giving macroscopic and microscopic views of the subject. Alternatively one may like to think of the problem in terms of crossing a river. In this analogy connectedness is a stepping stone that one can reach from one bank using Proposition 1, and which then gives access to the opposite bank by using the PF-Theorem.

We first describe the spectrum of a general non-negative matrix. This is an immediate consequence of the definition of connectedness which shows that any matrix that is not connected (except the trivial 1 x 1 zero matrix) is always decomposable. Note that it depends purely on connectedness and not the PF-Theorem. What the PF-Theorem does do is to provide a detailed description of the diagonal blocks of the matrix.

- **Proposition 1**

If $T \geq 0$ then there exists a permutation matrix E such that ETE^T has the lower triangular form

$$\begin{pmatrix} T_1 & 0 & 0 & \dots & 0 \\ * & T_2 & 0 & \dots & 0 \\ * & * & T_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ * & * & * & \dots & T_n \end{pmatrix}$$

where each $T_k \geq 0$ is a square matrix that is either connected or zero.

Prompted by the PF-Theorem we shall call a set of complex numbers S a *PF-type spectrum* if

(1) S is finite and non-empty and self-conjugate

(2) The outer ring of S consists of all solutions of the equation $z^h - r^h = 0$ for some r and h where $r \geq 0$ is real and h is a positive integer.

(3) S is invariant under rotations by $2\pi/h$.

From Proposition 1 we see that $\sigma(T)$ is the union of the $\sigma(T_k)$ and by the PF-Theorem each $\sigma(T_k)$ has a PF-type spectrum. This shows immediately that each element of $\pi(T)$ is cyclic. However unlike the connected case this does not mean that the outer ring is complete, since if $\omega = e^{\pi i/3}$, then 1, -1, ω^2, ω^4 could be a valid peripheral spectrum.

The 2 x 2 matrix with 1's in its left-hand column and 0's in its right trivially shows that all of the PF properties, (I) - (V) may still hold in the absence of connectedness. However another property $(II)^T$ namely that T^T has a positive eigenvector corresponding to r is also clearly true, and we now show (informally) that if $T \geq 0$ has spectral radius r and properties (I), (II), and $(II)^T$ hold then T must be connected.

To see this decompose T as in Proposition 1 and note that by (I) there exists precisely one diagonal block T_k such that $r(T_k) = r$. Consider the block form of the spectral projection given by $P = P(\pi(T);T)$. By the standard complex integral formula for P it follows that every super- diagonal block is zero, and every diagonal block except P_k is also zero. However the range of P contains the eigenspace of r which by (II) contains a positive vector. This implies P has no zero rows. Similarly by $(II)^T$, P^T contains a positive vector so P has no zero columns, hence P is zero-free. It follows that P collapses to the single block P_k , and hence in similar fashion T collapses to the single block T_k . But T_k is connected by Proposition 1, and hence so is T ☺

We now consider several other standard results on non-negative matrices. The traditional proofs generally need to build eigenvectors in order to establish the existence of eigenvalues. Our spectral theoretic approach almost turns this on its head since it is primarily concerned with eigenvalues, and regards eigenvectors as being of somewhat secondary importance.

- **Proposition 2**

If $T \geq 0$ then $\pi(T)$ is cyclic and the spectral radius $r = r(T)$ is an eigenvalue of T with at least one eigenvector $x \geq 0$.

Proof : We have already seen that $\pi(T)$ is cyclic and r is an eigenvalue of T. Decompose T as in Proposition 1 and let k be the highest suffix such that $r \in \sigma(T_k)$. Then T has the form $\begin{pmatrix} * & 0 & 0 \\ * & T_k & 0 \\ * & B & C \end{pmatrix}$ where $r \notin \sigma(C)$. By the PF-Theorem T_k has an eigenvector $v \geq 0$, $T_k v = rv$. Write $w = (r - C)^{-1}Bv \geq 0$. If $x = \begin{pmatrix} 0 \\ v \\ w \end{pmatrix}$ then $Tx = \begin{pmatrix} 0 \\ T_k v \\ Bv + Cw \end{pmatrix} = \begin{pmatrix} 0 \\ rv \\ rw \end{pmatrix}$

so $Tx = rx$, $x \geq 0$, and x is non-zero since v is. Thus x is an eigenvector of T ☺

- **Proposition 3**

If $T \geq 0$ then $r = r(T)$ is a pole of maximal order among the points of $\pi(T)$.

The standard proof of this is very elegant, namely if $r = 1$ and $\lambda \in \pi(T)$ then for every real number $\delta > 1$ and positive integer n we have $|(\lambda\delta - \lambda)^n(\lambda\delta - T)^{-1}| \leq (\delta - 1)^n(\delta - T)^{-1}$. So if we now let $\delta \downarrow 1$ then it is easily seen that the order of the pole at 1 is at least as great as the order of the pole at λ .

A less esoteric reason why this is so can be inferred from Proposition 1. For $\lambda \in \pi(T)$ to be a non-simple pole it is clearly necessary that λ must appear in at least two diagonal blocks, say T_i and T_j with $i > j$, and the block U at the intersection of row block i and column block j must satisfy $P(\lambda;T_i)UP(\lambda;T_j) \neq 0$. But then $U \geq 0$, $U \neq 0$ and since $P(r;T_i)$ and $P(r;T_j)$ are both positive blocks we have $P(r;T_i)UP(r;T_j) \neq 0$ which shows that r can't be a simple pole of T. Example 4 gives a matrix in which the orders of the poles on $\pi(T)$ are not all the same.

We end this section by giving three applications of an important induction technique for establishing potency. Note the striking similarity between Lemma 1 and Proposition 4. The next result will be a key element in the proof of the Compact Group Theorem in the next section. It has been independently derived by Radjavi (see [3] lemma 4).

- **Lemma 7**

If P is a non-negative zero-free projection then there exists a permutation matrix E such that EPE^T is a direct sum of rank(P) positive projections. There exist $u_k \geq 0$, $f_k \geq 0$ such that $EPE^T = \sum u_k \otimes f_k$ and $f_i(u_j) = \delta_{ij}$ for all i and j.

Proof : Let P be m-square and assume the result holds for all smaller matrices. It is trivially true for m = 1. When P > 0 the result follows immediately from Lemma 5, so assume $P \not> 0$. Choose any row of P that contains at least one zero. Suppose this row has k zeros in total; note that $1 \leq k \leq m-1$. Now apply a permutation so that the k zeros move to the right hand side of the row, and let Q be the resultant matrix. Of course the row itself may be moved up or down by this operation but its overall contents won't change. Then taking each of its k zeros in turn shows that in all of the right hand k columns the top (m-k) entries must be zero. Thus $Q = \begin{pmatrix} U & 0 \\ W & V \end{pmatrix}$ where U is (m-k)-square and V is k-square. Clearly U and V must both be projections and in addition W = WU + VW so multiplying on the left by V gives VWU = 0. As Q is clearly zero-free, U has no zero rows and V has no zero columns and so W = 0. Therefore $Q = U \oplus V$ so U and V must be zero-free and since they are both smaller than P then by the inductive hypothesis each of them is a direct sum of positive projections. The proof is completed by repeated applications of Lemma 5 ☺

Recall the hierarchy statement for non-negative matrices from the opening section :-

positivity $\Rightarrow$ connectedness $\Rightarrow$ potency $\Rightarrow$ zero-freedom

For projections positivity and connectedness are clearly equivalent, and Lemma 7 shows that potency and zero-freedom are also equivalent. Thus we can almost (but not quite) reverse the implications. The problem is of course that potency generally does not imply connectedness; identity matrices are trivial exponents of this. The 3 x 3 L-matrix given in example 2 is the simplest case of a zero-free matrix that is not potent.

- **Proposition 4**

Every invertible non-negative matrix is potent. Every non-zero non-negative matrix that commutes with a positive matrix is potent.

Proof : Both results are clearly true for all 1 x 1 matrices. Each proof is by induction on the size of the matrix. Suppose then that T is a non-negative square matrix and that the results hold for all matrices smaller than T. If T is connected then we already know that T is potent so assume that T is not connected and write $ETE^T = \begin{pmatrix} U & 0 \\ W & V \end{pmatrix}$.

If T is invertible then so are both U and V and by the inductive hypothesis they are both potent, say $\text{diag}(U^m) > 0$ and $\text{diag}(V^n) > 0$. But then $\text{diag}(ET^{mn}E^T) > 0$ so T is potent.

If, on the other hand, T is non-zero and commutes with a positive matrix S then ESE^T is positive, say $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and commutes with ETE^T. However this means that UA = AU + BW so trace(UA) = trace(AU) + trace(BW) hence trace(BW) = 0 as trace(UA) = trace (AU). Since B > 0 this implies W = 0. Now ETE^T is zero-free by Lemma 1, so U and V are also zero-free. As UA = AU and VD = DV the inductive hypothesis shows U and V must both be potent. Hence, as above, T itself is potent ☺

§4 Groups of non-negative matrices

We now turn our attention to the Compact Group Theorem. The aim here is to show that every compact group of non-negative matrices is finite.

- **Lemma 8**

Let $\Im$ be a compact group in $\aleph_m$. If any element of $\Im$ has a zero row or column then $\Im$ is isomorphic to a compact group in $\aleph_{m-1}$.

Proof : Note that if $S \in \Im$ has a zero row then since $T = SS^{-1}T$ that row will be globally zero for all $T \in \Im$. Suppose that it is row n that is globally zero. For $T \in \Im$ define θ(T) to be the matrix T with column n set to zero. It is easy to see that θ is a homomorphism, moreover if P is the identity of $\Im$ and θ(T) = θ(P) then T and P have the same matrix except for column n so $[P]_{ij} = \Sigma[P]_{ik}[P]_{kj} = \Sigma[T]_{ik}[P]_{kj} = [TP]_{ij}$ giving P = TP = T. Hence θ is a group isomorphism.
Now $\theta(\Im)$ is a group of non-negative matrices in which row n and column n are both zero and compactness has also been preserved as θ is clearly norm reducing, so $\Im$ is isomorphic to a compact group in $\aleph_{m-1}$ ☺

- **Lemma 9**

Let $\Im$ be a bounded group in $\aleph_m$ which contains the identity m x m matrix. Then $\Im$ is isomorphic to a subgroup of the symmetric group of degree m.

Proof : Since $1 \in \Im$ it follows that every $T \in \Im$ is zero-free. Suppose that $T \in \Im$ and $[T]_{ij} > 0$ and $[T]_{ik} > 0$. As T^{-1} is zero-free pick p, q such that $[T^{-1}]_{jp} > 0$ and $[T^{-1}]_{kq} > 0$. Then $[TT^{-1}]_{ip} > 0$ and $[TT^{-1}]_{iq} > 0$ so p = i = q. As T is zero-free choose r with $[T]_{ir} > 0$. Then $[T^{-1}T]_{jr} > 0$ and $[T^{-1}T]_{kr} > 0$ so j = r = k. This shows that no row of T can have more than one positive element. The same is true for each column of T and hence as T is zero-free every row and every column of T contains precisely one positive element.

Now define the mapping $\theta : \Im \to \aleph_n$ by $\begin{cases} [\theta(T)]_{ij} = 0 & \text{if} \quad [T]_{ij} = 0 \\ [\theta(T)]_{ij} = 1 & \text{if} \quad [T]_{ij} > 0 \end{cases}$

The argument above shows $\theta(T)$ is the unique permutation matrix such that there exists $\delta > 0$ with $T \geq \delta\theta(T)$. It follows easily that θ is a group homomorphism. In general θ is not an isomorphism because $\theta^{-1}(1)$ is non-trivial, however when $\mathfrak{I}$ is bounded no element $T \in \mathfrak{I}$ can have an eigenvalue of modulus greater than 1 and the same argument applied to T^{-1} shows that all eigenvalues of all elements of $\mathfrak{I}$ must have modulus 1. From this $\theta^{-1}(1) = \{1\}$ and hence θ is an isomorphism as required. ☺

- **The Compact Group Theorem**

Every compact group of non-negative matrices is finite. More precisely, if P is the identity of the group then the group contains at most rank(P)! elements.

Proof : Suppose first that P is zero-free. Decompose it according to Lemma 7. Pick any $T \in \mathfrak{I}$ and let U be the block located at row-block i, column-block j of T. Since T = PTP we see that $U = (u_i\otimes f_i)U(u_j\otimes f_j)$ which when multiplied out gives $f_i(Uu_j)(u_i\otimes f_j)$ which is a non-negative multiple of $u_i\otimes f_j$ that is zero if and only if U = 0.

Now set $\mathfrak{R}_k$ to be the algebra of k-square matrices with real entries and consider the mapping $\theta : \mathfrak{R}_m \to \mathfrak{R}_n$ given by $[T]_{ij} \to [T]_{ij}(u_i\otimes f_j)$ for all i and j. This is an algebra homomorphism since row-block i, column-block j of the product $\theta(S)\theta(T)$ is $\Sigma_k[S]_{ik}(u_i\otimes f_k)[T]_{kj}(u_k\otimes f_j)$ which is $\Sigma_k[S]_{ik}[T]_{kj}(u_i\otimes f_j)$ as $f_k(u_k) = 1$ for all k, and this is simply $[ST]_{ij}(u_i\otimes f_j)$ which is just row-block i, column-block j of $\theta(ST)$. As θ is clearly one-to-one $\mathfrak{R}_m$ must be isomorphic to its image in $\mathfrak{R}_n$. But as we have seen above this image must contain $\mathfrak{I}$, and θ also preserves both non-negativity and compactness. So $\theta^{-1}(\mathfrak{I})$ is a compact group in $\aleph_m$ containing $\theta^{-1}(P)$ which is the identity m x m matrix. From Lemma 9 it now follows that $\theta^{-1}(\mathfrak{I})$ and $\mathfrak{I}$ must both contain at most rank(P)! elements.

Finally if P is not zero-free then by repeatedly applying Lemma 8 to eliminate any zero rows and columns, the original group reduces to one whose identity P is a zero-free matrix. These applications clearly have no effect upon rank(P). Thus the end product of combining the various isomorphisms is a group of rank(P)! elements at most ☺☺

---

§5 Examples

We conclude with four examples from $\aleph_3$ to reveal some of the pathology of the situation. Whilst all of them are elementary each is included in order to make a specific point. They serve as useful first ports of call when formulating or testing hypotheses.

Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ $C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ $D = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

From the PF-Theorem (V) it follows that if trace(T) > 0 then h = 1. The matrix A shows that the converse is false; even when trace(T) = 0 it is still possible to have h = 1 ☺

Secondly we examine the situation whenever T is not connected using the matrix B which shall be referred to as the T-matrix for obvious reasons. Set T = B and note that T^n is the same as T except for the element in the middle of its top row which is 2n-1. Clearly the eigenvalues of T are {1,1,0} and its peripheral projection P is in essence just T^0 ; that is, the entry in the middle of its top row is -1 and all other entries are the same as in T. So P isn't non-negative and its diagonal isn't positive. As R = PT = T the child is in $\aleph_3$ but a small refinement of the example shows that it too can misbehave. Let $\delta > 0$ be a small number and set $S = (1-\delta) + \delta T$. Then $S \in \aleph_3$ has a positive diagonal and eigenvalues $\{1, 1, 1-\delta\}$. The spectral projections for S are exactly the same as those for T and it can easily be computed that the n^{th} power of its child is the same as T except that it has $2n\delta-1$ in the middle of its top row. Thus although the powers of this child eventually go non-negative it may take a long time for them to do so.

This poses an interesting general question regarding the computability of the peripheral projection and the child matrix. If $T \geq 0$ is connected then there is no problem since if T has potency c then the peripheral projection is the limit of T^{nc} and consequently it is easy to compute. However when T is not connected the peripheral projection is generally not a limit of powers of T and there is no very obvious good way of computing it.

Finally this example gives a nice visual demonstration of how the decomposition in Proposition 1 works in practice. It transforms the T-matrix into the L-matrix; that is, 1's in the left hand column and the bottom row and 0's elsewhere ☺

The third example shows that there may be gaps in the inner rings of the spectrum and also that the arguments of the eigenvalues do not have to be rational multiples of 2π. Let $\delta > 0$ and set $T = \delta + C$. Then $h = 1$ and the eigenvalues of T are $\delta + 1$, $\delta + \omega$, $\delta + \omega^2$ where $\omega = e^{2\pi i/3}$. If, for example, $\delta = 1/2$ then the eigenvalues are 3/2 and $\pm i\sqrt{3}/2$. Thus the spokes connecting the origin to the eigenvalues on the outer ring do not necessarily intersect the inner rings at eigenvalues, and by letting δ increase it is clear that the argument θ of the eigenvalue $\delta + \omega$ can be anything in the range $0 < \theta \leq 2\pi/3$ ☺

Our final example aims to show that even for matrices that appear to be "reasonably well connected" the orders of the poles on $\pi(T)$ may vary quite unexpectedly. Let T = D. The minimum polynomial of T is $(T - 1)^2(T + 1) = 0$. This shows that the eigenvalue 1 is a pole of order 2 whereas -1 is a pole of order 1 ☺

References

[1] N. Dunford & J. T. Schwartz : *Linear operators II*, Interscience Publishers, New York, 1963

[2] M. Marcus & H. Minc : *A survey of matrix theory and matrix inequalities*, Dover, Mineola, N.Y.,1992

[3] H. Radjavi : *The Perron-Frobenius Theorem revisited* (pre-print)

[4] M. R. F. Smyth & T. T. West : *The minimum diagonal element of a positive matrix*, Studia Mathematica 131 (1) (1998)