Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

FOCAL POWER*

ROBERT E. HARTWIG† AND JOHN MAROULAS‡

Abstract. The properties of a 2×2 block matrix M over a field for which $(M^k)_{11} = (M)_{11}^k$ are examined. This "fire-wall" property will be characterized by the vanishing of the sequence of moments $F_i = CD^iB$.

Key words. Matrix powers, Polynomials, (1,1) blocks.

AMS subject classifications. 15A24, 15A09, 15A57.

1. Background material. Let $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ be a square matrix over a field $\mathbb{F}$, with square diagonal blocks. Suppose further that $M^k = \begin{bmatrix} A_k & C_k \\ B_k & D_k \end{bmatrix}$. Then the recurrence relation for the block can be obtained from

$$\begin{bmatrix} A_{k+1} & C_{k+1} \\ B_{k+1} & D_{k+1} \end{bmatrix} = M^{k+1} = MM^k = \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} A_k & C_k \\ B_k & D_k \end{bmatrix} = M^k M,$$

which gives

$$\begin{array}{llllll} & \text{(i)} & A_{k+1} & = AA_k + CB_k & = A_kA + C_kB & (A_0 = I) \\ (1.1) & \text{(ii)} & B_{k+1} & = BA_k + DB_k & = B_kA + D_kB & (B_0 = 0) \\ & \text{(iii)} & C_{k+1} & = AC_k + CD_k & = A_kC + C_kD & (C_0 = 0) \\ & \text{(iv)} & D_{k+1} & = BC_k + DD_k & = B_kC + D_kD & (D_0 = I). \end{array}$$

We call M "focused on the (1,1) position" or (1,1)-focused, if $(M^k)_{11} = (M)_{11}^k$ for all $k = 1, 2, \ldots$ Examples of this are block triangular matrices. In this note we shall be characterizing this property. Our first observation is that if we think of a matrix as a relation, then matrix multiplication corresponds to the composition of relations. This is best addressed by using the digraph associated with M that has **weights** m_{ij} attached to the arc (edge) (v_i, v_j) from node v_i to node v_j. We may loosely speaking, think of m_{ij} as the "flow of information" from v_i to v_j. The entry $(M^k)_{ij}$ is represented by the sum of all k-step path products from v_i to v_j and as such represents the total flow of information from v_i to node v_j. When we partition our matrix M as $\begin{bmatrix} A & C \\ B & D \end{bmatrix}$ then we partition the nodes accordingly into two classes $V_1 = \{S_1, S_2, \ldots, S_m\}$ and $V_2 = \{T_1, T_2, \ldots, T_n\}$ (often called condensed nodes). Moreover the weight of the arcs $(S_i, S_j), (S_i, T_k), (T_k, T_q), (T_q, S_j)$ are respectively given by a_{ij}, c_{ik}, d_{kq} and b_{qj}.

* Received by the editors 18 February 2004. Accepted for publication 11 January 2005. Handling Editor: Ravindra B. Bapat.

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205 USA (hartwig@math.ncsu.edu).

‡Department of Mathematics, National Technical University of Athens, Zografou Campus, Athens, Greece (maroulas@math.ntua.gr).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

We can now think of block matrix multiplication as representing the total flow of information between the condensed nodes V_i corresponding to the diagonal block entries.

In particular the block entry $(M^k)_{11}$ represents the flow of information from V_1 into V_1 after a k-fold application of the linear map M. That is, the flow of information from S_i into S_j, for any (i, j).

The statement that M is (1,1)-focussed precisely means that **no** information flows from V_1 into V_1 via V_2, for any repeated application of the relation M.

The basic fact that we shall show is that for two block rows, the (1,1) focal property occurs precisely when **all the moments** $F_i = CD^iB$ **vanish!** That is when

$$(CD^rB)_{ij} = \sum_{k=1}^{m}\sum_{q=1}^{n} c_{ik}(D^r)_{kq}b_{qj} = 0 \quad \text{for all } i = 1, \ldots, m,\ j = 1, \ldots, n.$$

Alternatively we may think of the (1,1)-focal property as corresponding to a "fire-wall" around the node V_1. The analogy that comes to mind are the two sides of the brain, for which in certain cases, no information flows from the right side into the left side (all the neurons leading into the left half have been cut).

When we have more than two block rows, we may likewise define $M = [A_{ij}]$ to be (k, k) focussed or to be focussed on any principal sub block matrix of M. The diagram given below illustrates this idea

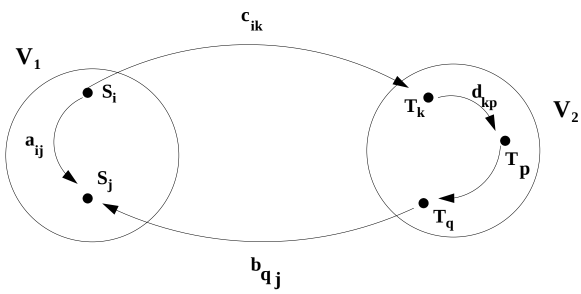

FIG. 1.1. *Basic Flow for a 2 x 2 block matrix*

We shall also need the moments $E_i = BA^iC,\ i = 1, 2, \ldots$, which represent the flow of information into the second node V_2. When A is nonsingular we may also define $E_{-i} = BA^{-i}C,\ \ i = 1, 2, \ldots$ A key role in all of this is played by the cornered matrices

$$\Gamma_k(A, C, D) = A^{k-1}C + A^{k-2}CD + \ldots + CD^{k-1} = \sum_{i=0}^{k-1} A^iCD^{k-i-1}.$$

Throughout this note, the minimal and characteristic polynomials of M will be denoted by $\psi_M(\lambda)$ and $\Delta_A(\lambda)$ respectively, the Drazin inverse of M is denoted by M^D (it is always a polynomial in M), and the index of a square matrix M is denoted by in(M). When $in(M) = 0$ or 1, the Drazin inverse reduces to the group inverse, which will be denoted by $M^{\#}$. The range, nullspace and nullity of a matrix M will

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

be denoted by $R(M)$, $N(M)$ and $\nu(M)$ respectively. If $M = [A_{ij}]$ is an $n \times n$ block matrix we denote the leading $k \times k$ block matrix, $[A_{ij}]$, $i, j = 1, \ldots, k$, by M_k.

Theorem 1.1. *Let* $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ *be a square block matrix with* A *and* D *square and* $M^k = \begin{bmatrix} A_k & C_k \\ B_k & D_k \end{bmatrix}$. *Then*

(a) *the following are equivalent:*

1. $A_k = A^k$ *for all* $k = 1, 2, \ldots$
2. $A_{k+1} = AA_k$ *for all* $k = 0, 1, \ldots$
3. $CB_k = 0$ *for all* $k = 0, 1, \ldots$
4. $F_k = CD^kB = 0$ *for all* $k = 0, 1, \ldots$
5. $C_kB = 0$ *for all* $k = 0, 1, \ldots$
6. $C\,\mathrm{adj}\,(\lambda I - D)B = 0$.

In which case,

(b)

1. $B_k = \Gamma_k(D, B, A)$
2. $C_k = \Gamma_k(A, C, D)$ *and*
3.

$$D_k = D^k + \sum_{i=2}^{k}\sum_{r=0}^{i-2} D^r(BA^{k-i}C)D^{i-2-r}$$

$$= D^k + [I, D, \ldots, D^{k-2}] \begin{bmatrix} E_{k-2} & E_{k-3} & & E_0 \\ E_{k-3} & & E_0 & 0 \\ & \iddots & \iddots & \\ E_0 & 0 & & 0 \end{bmatrix} \begin{bmatrix} I \\ D \\ \vdots \\ D^{k-2} \end{bmatrix}. \tag{1.2}$$

Proof.

(1)⇔(2). Clear since $A_1 = A$ is the initial condition.

(2)⇔ (3). Clear from (1.1)-(i).

(3)⇔ (4). To do this let us first solve for CB_k in term of the F_k and A_k.

From this expression it is clear that when all the moments E_i vanish, then no information will flows into V_2, i.e., $D_k = D^k$ for all $k = 1, 2, \ldots$

The following lemma is needed before we proceed with the proof of the theorem.

Lemma 1.2. *If* $B_{k+1} = BA_k + DB_k$, *then*

$$CB_k = F_0A_{k-1} + F_1A_{k-2} + \ldots + F_{k-2}A_1 + F_{k-1} \tag{1.3}$$

and

$$A_{k+1} = AA_k + F_0A_{k-1} + F_1A_{k-2} + \ldots + F_{k-2}A_1 + F_{k-1}, \tag{1.4}$$

which is the recurrence relation for the (1,1) blocks A_k.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

Proof of lemma. It is easily seen by induction that

$$CB_k = F_0A_{k-1} + F_1A_{k-2} + \ldots + F_{s-1}A_{k-s} + CD^sB_{k-s} \text{ for } s = 0, 1, \ldots, k-1.$$

Indeed, for s = 0 this is clear. So assuming this is valid for s, we have $CB_k = F_0A_{k-1}+F_1A_{k-2}+\ldots+F_{s-1}A_{k-s}+CD^s(BA_{k-s-1}+DB_{k-s-1}) = F_0A_{k-1}+F_1A_{k-2}+\ldots+F_sA_{k-s-1}+CD^{s+1}B_{k-s-1}$, which establishes the validity for $s+1, \ldots$ Setting $s = k-1$ yields (1.3). The rest is clear. □

In block matrix from this gives

$$\begin{bmatrix} CB_1 \\ CB_2 \\ \vdots \\ CB_k \end{bmatrix} = \begin{bmatrix} I & & & 0 \\ A_1 & I & & \\ \vdots & & \ddots & \\ A_{k-2} & \cdots & A_1 & I \end{bmatrix} \begin{bmatrix} F_0 \\ F_1 \\ \vdots \\ F_{k-1} \end{bmatrix}. \tag{1.5}$$

From this it follows at once that $CB_k = 0$ for all $k = 0, 1, 2\ldots$ iff $F_i = 0$ for all $i = 0, 1, \ldots$ By $B \longleftrightarrow C$ symmetry it also follows that $C_mB = 0$ iff $F_i = 0$ for all $i = 0, 1, \ldots$

(4)⇔ (6). Recall that if D is $m \times m$ has characteristic polynomial, $\Delta_D(\lambda) = d_0 + d_1\lambda + \ldots + \lambda^m$, then $\operatorname{adj}(\lambda I - D) = D_0 + D_1\lambda + \ldots + D_{m-1}\lambda^{m-1}$ with $D_{m-1} = I$. Consequently

$$C \operatorname{adj}(\lambda I - D)B = CD_0B + (CD_1B)\lambda + \ldots + (CB)\lambda^{m-1}.$$

The matrices $D_i = f_i(D)$ are the *adjoint* polynomials of D, which are given by

$$[f_0(x), f_1(x), .., f_{m-1}(x)] = [1, x, .., x^{m-1}](H \otimes I), \tag{1.6}$$

where $H = \begin{bmatrix} d_1 & d_2 & \cdots & d_m \\ d_2 & d_3 & & \\ & \iddots & & \\ d_m & & \cdots & 0 \end{bmatrix}$. This means that

$$[CD_0B, CD_1B, \ldots, CD_{m-1}B] = C[D_0, D_1, \ldots, D_{m-1}](I_m \otimes B)$$

$$= C[I, D, \ldots, D^{m-1}](H \otimes I)(I \otimes B) = C[I, D, \ldots, D^{m-1}](I \otimes B)(H \otimes I)$$

$$= [F_0, F_1, \ldots, F_{m-1}](H \otimes I).$$

From this we see that $C \operatorname{adj}(\lambda I - D)B = 0$ iff $CD_iB = 0$ for $i = 0.1, \ldots, m$ iff $F_i = 0$ for all i.

Let us next turn to expressions for the blocks B_k and C_k, for which we again use induction.

It is clear that $B_1 = B = \Gamma_1(D, B, A)$. So let us assume that $B_r = \Gamma_r(D, B, A)$, for $r \leq k$. Then $B_{k+1} = BA_k + DB_k = BA^k + D(D^{k-1}B + D^{k-2}BA + ... + BA^{k-1}) = \Gamma_{k+1}(D, B, A)$, as desired.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

Likewise $C_1 = C = \Gamma_1(A, C, D)$ and $C_{k+1} = A_k C + C_k D = A^k C + (A^{k-1}C + A^{k-2}CD + \ldots + CD^{k-1})D = \Gamma_{k+1}(A, C, D)$.

Lastly, let us show by induction that (1.2) holds. Let X_k denote the RHS of equation (1.2). Since $D_2 = BC + D^2$ we see that (1.2) holds for $k = 2$. Now $D_{k+1} = BC_k + DD_k = B(A^{k-1}C + \ldots + CD^{k-1}) + D^{k+1} + \sum_{i=2}^{k}\sum_{r=0}^{i-2} D^{r+1}(BA^{k-i}C)D^{i-2-r}$. Setting $j = i + 1$, reduces the last sum to

$$\sum_{j=3}^{k+1}\sum_{r=0}^{j-1} D^{r+1}(BA^{k-j+1}C)D^{j-r-3}.$$

Next we set s = r + 1, giving

$$\sum_{j=3}^{k+1}\sum_{s=1}^{j} D^{s}(BA^{k-j+1}C)D^{j-s-2}.$$

The term with $j = 2$ is empty because $j - 2 - s$ will be negative. Likewise the terms with $s = j - 1$ or $s = j$ will be absent. We may thus write the sum as

$$\sum_{j=2}^{k+1}\sum_{s=1}^{j-2} D^{s}(BA^{(k+1)-j}C)D^{j-s-2}.$$

The term with $s = 0$ is precisely $BA^{k-1}C + \ldots + BCD^{k-1}$, and so we end up with

$$D^{k+1} + \sum_{j=2}^{k+1}\sum_{s=0}^{j-2} D^{s}(BA^{(k+1)-j}C)D^{j-s-2},$$

which precisely equals X_{k+1}. This completes the proof of the theorem. $\square$

We note that we may also write D_k in terms of chains as

$$(1.7)\quad D_k = D^k + [B, DB, \ldots, D^{k-2}B]\begin{bmatrix} A^{k-2} & A^{k-3} & & I \\ A^{k-3} & \cdots & I & 0 \\ \vdots & \iddots & \iddots & \\ I & 0 & & 0 \end{bmatrix}\begin{bmatrix} C \\ CD \\ \vdots \\ CD^{k-2} \end{bmatrix}.$$

It is of interest to observe that the cornered matrices $\Gamma_k(A, C, D)$ and $\Gamma_k(D, B, A)$ appear in

$$\begin{bmatrix} A & C \\ 0 & D \end{bmatrix}^k = \begin{bmatrix} A^k & \Gamma_k(A, C, D) \\ 0 & D^k \end{bmatrix} \text{ and } \begin{bmatrix} A & 0 \\ B & D \end{bmatrix}^k = \begin{bmatrix} A^k & 0 \\ \Gamma_k(D, B, A) & D^k \end{bmatrix}.$$

Summing these we see that for any polynomial $f(x)$:

$$f(\begin{bmatrix} A & C \\ 0 & D \end{bmatrix}) = \begin{bmatrix} f(A) & \Gamma_f(A, C, D) \\ 0 & f(D) \end{bmatrix} \text{ and } f(\begin{bmatrix} A & 0 \\ B & D \end{bmatrix}) = \begin{bmatrix} f(A) & 0 \\ \Gamma_f(D, B, A) & f(D) \end{bmatrix},$$

where $\Gamma_f(A, C, D) = \sum_{i=1}^{k} f_i\Gamma_i(A, C, D)$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

2. Consequences. Let us now derive several of the properties of focused matrices, as related to (i) convergence, (ii) graphs, (iii) invariant subspaces, and (iv) chains.

Throughout this section we assume that $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ is (1,1)-focused.

Our first observation is

Corollary 2.1. *If M is (1,1) focussed and M^k converges (to zero) as $k \to \infty$ then A^k also converges (to zero).*

Next we have

Corollary 2.2. *If M if (1,1)-focused, then for any polynomial $f(x) = f_0 + f_1 x + \ldots + f_s x^s$, with adjoint polynomials $f_i(x)$,*

$$f(M) = f(\begin{bmatrix} A & C \\ B & D \end{bmatrix}) = \begin{bmatrix} f(A) & \Gamma_f(A, C, D) \\ \Gamma_f(D, B, A) & D_f \end{bmatrix}, \tag{2.1}$$

where $D_f =$

$$f(D) + [B, DB, \ldots, D^{s-2}B] \begin{bmatrix} f_1(A) & f_2(A) & & f_{s-1}(A) \\ f_2(A) & \cdots & f_{s-1}(A) & 0 \\ \vdots & \iddots & \iddots & \\ f_{s-1}(A) & 0 & & 0 \end{bmatrix} \begin{bmatrix} C \\ CD \\ \vdots \\ CD^{s-2} \end{bmatrix}.$$

In particular if $q(M) = 0$, then $q(A) = 0$. From this it follows that

Corollary 2.3. *If M is (1,1)-focused, then $\psi_A | \psi_M$.*

This tells us that the eigenvalue of A (if any) must be among the e-values of M. In addition

Corollary 2.4. *If M is (1,1)-focused, then for any scalar β: $in(A-\beta I) \leq in(M-\beta I)$.*

We further know that for any polynomial $f(x)$,

$$f(M) = g(M) \quad \Rightarrow \quad f(A) = g(A). \tag{2.2}$$

This simple observation has numerous consequences.

Corollary 2.5. *For complex matrices, if $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ is (1,1)-focused and normal, then A is also normal.*

Proof. M is normal exactly when $M^* = f(M)$ for some polynomial $f(x)$. In this case $\begin{bmatrix} A^* & B^* \\ C^* & D^* \end{bmatrix} = M^* = f(M) = \begin{bmatrix} f(A) & ? \\ ? & ? \end{bmatrix}$, and thus $A^* = f(A)$ and A is normal. $\square$

Corollary 2.6. *Suppose that M is (1,1)-focused.*
(i) If M is invertible, then A is also invertible.
(ii) If $M^{-1} = f(M)$, then $A^{-1} = f(A)$.

Proof. If M is invertible then M^{-1} is a polynomial in M , say $f(M)$. As such $Mf(M) = I$, and hence $Af(A) = I$, with $f(A) = A^{-1}$. $\square$

Let us next extend this to Drazin and group inverses.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

Corollary 2.7. *If* $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ *is (1,1) focussed, then*

$$M^D = \begin{bmatrix} A^D & ? \\ ? & ? \end{bmatrix}. \tag{2.3}$$

If $M^{\#}$ *exists then*

$$M^{\#} = \begin{bmatrix} A^{\#} & ? \\ ? & ? \end{bmatrix}. \tag{2.4}$$

Proof. To show (2.3), let X = g(M) = M^D. Then, using the notation of generalized inverses, we have

(1k) $M^{k+1}X$ =M$\Leftrightarrow M^{k+1}g(M) = M^k \Rightarrow A^{k+1}g(A) = A^k$.
(2) $XMX = X \Leftrightarrow g(M)Mg(M) = g(M) \Rightarrow g(A)Ag(A) = g(A)$.
(5) $MX = XM \Leftrightarrow Mg(M) = g(M)M \Rightarrow Ag(A) = g(A)A$.

The latter show that $g(A)$ is unique solution to the three equations

$$A^{r+1}Y = A^r, \quad AY = YA, \quad YAY = Y \text{ (for some r)},$$

and thus must equal the Drazin inverse of A.

Equation (2.4) is a special case of (2.3), when $k = 0$ or 1.

Corollary 2.8. *Over* $\mathbb{C}$*, if* M *is (1,1)-focused and EP (i.e.* $R(M) = R(M^*)$*), then* A *is also EP.*

Proof. If M is EP, then $M^\dagger = M^{\#} = g(M)$ is the group inverse of M. This means that $MM^{\#} = Mg(M) = \begin{bmatrix} Ag(A) & ? \\ ? & ? \end{bmatrix}$ and $M^{\#}M = g(M)M = \begin{bmatrix} g(A)A & ? \\ ? & ? \end{bmatrix}$. Since the latter two matrices are Hermitian it follows that $Ag(A) = g(A)A$ is also Hermitian, ensuring that $g(A) = A^{\#} = A^\dagger$. □

In the closed field case, the spectral projection associated with the zero eigenvalue is a polynomial in M, and satisfies:

Corollary 2.9.

$$Z_M = I - MM^D = I - Mg(M) = \begin{bmatrix} I - Ag(A) & ? \\ ? & ? \end{bmatrix} =$$
$$\begin{bmatrix} I - AA^D & ? \\ ? & ? \end{bmatrix} = \begin{bmatrix} Z_A & ? \\ ? & ? \end{bmatrix}.$$

Corollary 2.10. *The spectral components associated with the zero eigenvalue are given by*

$$Z_M^j = M^j Z_M = M^j(I - MM^D) = M^j(I - Mg(M)) =$$

$$\begin{bmatrix} A^j(I - Ag(A)) & ? \\ ? & ? \end{bmatrix} =$$

$$\begin{bmatrix} A^j(I - AA^D) & ? \\ ? & ? \end{bmatrix} = \begin{bmatrix} Z_A^j & ? \\ ? & ? \end{bmatrix}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

Corollary 2.11. *If M is (1,1)-focused, then $\Delta_M = \Delta_A.\Delta_D$*

Proof. We shall use the following polynomial form identity, in which D is $k \times k$ and M is $n \times n$.

$$\begin{bmatrix} \Delta_D I & -C\,\mathrm{adj}\,(\lambda I - D) \\ 0 & I \end{bmatrix} \begin{bmatrix} \lambda I - A & -C \\ -B & \lambda I - D \end{bmatrix}$$
$$= \begin{bmatrix} (\lambda I - A)\Delta_D - C\,\mathrm{adj}\,(\lambda I - D)B & 0 \\ -B & \lambda I - D \end{bmatrix}. \tag{2.5}$$

Taking determinants gives

$$\Delta_D^{n-k}\Delta_M = \Delta_D \det[\Delta_D(\lambda I - A) - C\,\mathrm{adj}\,(\lambda I - D)B] \tag{2.6}$$

Now because of (1.6), $CD_iB = 0$ for $i = 0, 1, \ldots, k-1$ iff $CD^iB =$ for $i = 0, 1, \ldots, k-1$. Hence we see that

if $F_i = 0$ for all i, then $C\,\mathrm{adj}\,(\lambda I - D)B = 0$, and (2.6) simplifies to

$$\Delta_D^{n-k}\Delta_M = \Delta_D^{n-k+1}\Delta_A$$

in which we may cancel the Δ_D^{n-k} to give the desired result. □

Corollary 2.12. *Let M be A-focussed. Then, as $k \to \infty$, $M^k \to 0$ iff $A^k \to 0$ and $D^k \to 0$.*

We remark that when M is A-focussed and M^k converges, then D^n need not converge, as seen from the matrix $M = \left[\begin{array}{c|cc} 1/2 & 0 & 1 \\ \hline -1/2 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]$, with D $= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

3. Transitivity.

When we have more than two block rows, the term "(1,1)-focussed" becomes ambiguous, since it depends on the particular partitioning one has in mind. To avoid this we shall explicitly refer to the principal sub-block matrix that is used. For example if $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} = \left[\begin{array}{cc|c} A_1 & A_3 & C_1 \\ A_2 & A_4 & C_2 \\ \hline B_1 & B_2 & D \end{array}\right]$, we shall say that M is A-focussed, or M is A_1-focussed, but shall not use "(1,1)- focussed" for either of these. In this vein it is best to think of this as a relation and define

Definition 3.1. *$A \preccurlyeq M$ iff A is a principal submatrix of M and M is A-focussed.*

It is clear that this relation is reflexive and anti-symmetric. We shall now show that it is transitive as well, making it a partial order. It suffices to consider the case of three block rows. The general case then follows by induction and by permutation similarity.

Theorem 3.2. *Let $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} = \left[\begin{array}{cc|c} A_1 & A_3 & C_1 \\ A_2 & A_4 & C_2 \\ \hline B_1 & B_2 & D \end{array}\right]$. Then $A_1 \preccurlyeq A$ and $A \preccurlyeq M$ imply $A_1 \preccurlyeq M$.*

Proof. From theorem (1.1) we know that

$$(a) \quad \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} D^k[B_1, B_2] = 0 \quad for \quad all \quad k = 0, 1, \ldots \tag{3.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

and

$$(b)\quad A_3A_4^{\ell}A_2 = 0, \quad for \;\; all \;\; \ell = 0, 1, 2, \ldots \tag{3.2}$$

Our aim is to show that

$$[A_3, C_1]\begin{bmatrix} A_4 & C_2 \\ B_2 & D \end{bmatrix}^r \begin{bmatrix} A_2 \\ B_1 \end{bmatrix} = 0 \text{ for all } r = 0, 1, 2, \ldots$$

We begin by setting $\begin{bmatrix} A_4 & C_2 \\ B_2 & D \end{bmatrix}^k = \begin{bmatrix} \alpha_k & \gamma_k \\ \beta_k & \delta_k \end{bmatrix}$ for $k = 1, 2, \ldots$ In terms of this we must establish that

$$A_3\alpha_kA_2 + A_3\gamma_kB_1 + C_1\beta_kA_2 + C_1\delta_kB_1 = 0, \quad k = 0, 1, \ldots \tag{3.3}$$

We shall use the recurrence of (1.1) and apply the conditions (3.1) and (3.2) to show by induction that each term in this sum vanishes.

(a) $C_i\beta_k = C_i(B_2\alpha_{k-1} + D\beta_{k-1}) = C_iD\beta_{k-1} = \ldots = C_iD^r\beta_{k-r} = \ldots = C_iD^{k-1}\beta_1 = C_iD^{k-1}B_2 = 0$.

(b) $C_i\delta_kB_j = C_i(B_2\delta_{k-1} + D\delta_{k-1})B_j = C_iD\delta_{k-1}B_j = \ldots = C_iD^r\delta_{k-r}B_j = \ldots = C_iD^{k-1}\delta_1B_j = C_iD^{k-1}DB_j = 0$ for $i, j = 1, 2$.

(c) $\gamma_kB_j = (\alpha_{k-1}C_2 + \gamma_{k-1}D)B_j = \gamma_{k-1}DB_j = \ldots = \gamma_{k-r}D^rB_j = .. = \gamma_1D^{k-1}B_j = C_2D^{k-1}B_j = 0$.

(d) $A_3\alpha_kA_2 = A_3(A_4\alpha_{k-1} + C_2\beta_{k-1})A_2 = A_3A_4\alpha_{k-1}A_2$
$= A_3A_4(A_4\alpha_{k-2} + C_2\beta_{k-1})A_2$, in which the latter term vanishes by part (a). Hence, by (3.2), we have

$$\begin{aligned} A_3\alpha_kA_2 &= A_3A_4^2\alpha_{k-2}A_2 = \ldots = A_3A_4^r\alpha_{k-r}A_2 \\ &= \ldots = A_3A_4^{k-1}\alpha_1A_2 = A_3A_4^{k-1}A_4A_2 = 0. \;\Box \end{aligned}$$

This result is not surprising in term of information flow. If no information can flow from M into A and no information can flow from A into A_1 then **no** information can flow from M into A_1. If M_k is the leading principal submatrix in M, then

Corollary 3.3. *$M_1 \preccurlyeq M$ iff $M_1 \preccurlyeq M_{k_1} \preccurlyeq M_{k_2} \preccurlyeq \ldots \preccurlyeq M$ for some increasing sequence $(1, k_1, k_2, \ldots, n)$.*

Corollary 3.4. *If $A_{kk} \preccurlyeq M_k$ and $M_k \preccurlyeq M$ then $A_{kk} \preccurlyeq M$.*

Using permutations, this may be extended to any nested sequence of principal block matrices containing A_{kk}.

4. Special cases.

Let us now turn to some special cases.

Proposition 4.1. $\begin{bmatrix} A & B^* \\ B & D \end{bmatrix}$ *is A-focussed iff $B = 0$.*

Proof. The condition $CB = B^*B = 0$, forces $B = 0$. $\Box$

Proposition 4.2. *If $\begin{bmatrix} A & \boldsymbol{c} \\ \boldsymbol{b}^T & d \end{bmatrix}$, where $\boldsymbol{b}$ and $\boldsymbol{c}$ are columns, then M is (1,1)-focused iff either $\boldsymbol{c} = \boldsymbol{0}$, or $\boldsymbol{b} = \boldsymbol{0}$. In which case M is block triangular*

Proof. Suppose that M is (1,1)-focused. Then with $k = 2$ we see that $A_2 = A^2 + \mathbf{cb}^T$ and thus $\mathbf{cb}^T = 0$. This means that either $\mathbf{c} = \mathbf{0}$, or $b = \mathbf{0}$. In which case M is block triangular. No other conditions are necessary. The converse is clear. $\Box$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

As an application of this consider the companion matrix

$$\mathrm{L(f)} = \begin{bmatrix} 0 & & & & -f_0 \\ 1 & 0 & & & -f_1 \\ 0 & 1 & & & \\ & & \ddots & \ddots & \vdots \\ & 0 & & & \\ 0 & & & 1 & -f_{n-1} \end{bmatrix} = \begin{bmatrix} N & \mathbf{c} \\ \mathbf{b}^T & d \end{bmatrix},$$

associated with the monic polynomial $f(t) = f_0 + f_1 t + \ldots + t^n$, with $\mathbf{b} = \mathbf{e}_{n-1}$ and $d = -f_{n-1}$. If L is N-focussed, then $\mathbf{c} = \mathbf{0}$, and L reduces to $L = \begin{bmatrix} N & 0 \\ \mathbf{e}_{n-1}^T & d \end{bmatrix}$. It follows at once by induction that

(a) for $k \leq n-1$, $L^k = \left[\begin{array}{ccc|ccc} 0 & & & & & 0 \\ \vdots & & & & & \\ 0 & & & & & \\ \hline 1 & & & & & \\ & \ddots & & & & \\ 0 & & 1 & d & \cdots & d^k \end{array}\right]$

and

(b) for $r \geq 0$, $L^{n+r} = \mathbf{e}_n[d^{r+1}, \ldots, d^{n+r}]$.

Given the vector $\mathbf{a}^T = [a_0, a_1, \ldots, a_{n-1}]$, the coefficients $b_k = \mathbf{a}^T L^k \mathbf{e}_1$, $k = 0, 1, 2, \ldots$ can be computed as $b_k = a_k$ for $k = 0, \ldots, n-1$ and $b_{n+r} = a_{n-1} d^{r+1}$ for $r = 0, 1, \ldots$ That is, the tail forms a geometric progression. This construction finds use in the search for minimal realizations of order n [6].

As our next example we consider the linear system

$$\overset{\bullet}{\mathbf{x}}(t) = D\mathbf{x}(t) + B\mathbf{u}(t) \text{ and } \mathbf{y}(\mathrm{t}) = \mathrm{C}\mathbf{x}(\mathrm{t}).$$

Differentiating this repeatedly, gives [5]

$$\begin{bmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(n-1)}(t) \end{bmatrix} = \begin{bmatrix} C \\ CD \\ \vdots \\ CD^{n-1} \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 & & & 0 \\ F_0 & 0 & & \\ \vdots & & \ddots & \\ F_{n-2} & \cdots & F_0 & 0 \end{bmatrix} \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix}, \tag{4.1}$$

where the moments $F_k = CD^k B$ are now called the **Markov parameters** of the system. These are uniquely determined by the transfer matrix $H = C(\lambda I - D)^{-1}B = \sum_{k=0} F_k \lambda^{-k}$. When M is (1,1)-focused, we know that all the Markov parameters vanish, and consequently $\mathcal{O}B = \begin{bmatrix} C \\ CD \\ \vdots \\ CD^{n-1} \end{bmatrix} B = 0$. If in addition, the pair (C, D) is *observable*, then $\mathcal{O}$ has a left inverse and B must vanish. In this case M will be upper

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

triangular.

Lastly, we next turn our attention to the **dilation** of a (1,1)-focused matrix.

Proposition 4.3. *If* $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ *is (1,1) focused, then so is the* $(N+1)\times(N+1)$ *block dilation*

$$\hat{M}_{N-1} = \left[\begin{array}{c|cccc} A & C & 0 & \cdots & 0 \\ \hline 0 & 0 & I & & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & & & I \\ B & D & & \cdots & 0 \end{array}\right] = \begin{bmatrix} A & \mathcal{C} \\ \mathcal{B} & \Omega \end{bmatrix}. \tag{4.2}$$

Proof. *It suffices to note that* $\Omega^k = \begin{bmatrix} 0 & & I & & \\ & & & \ddots & \\ D & & & & I \\ & \ddots & & & \\ & & D & & 0 \end{bmatrix}$*, for* $k = 0, 1, \ldots, N-1$ *so that* $\Omega^N = I \otimes D$ *and* $\Omega^{N+1} = (I_N \otimes D)\Omega$*. We next compute* $[C, 0, .., 0]\Omega^k \begin{bmatrix} 0 \\ \vdots \\ B \end{bmatrix} =$ $C(\Omega^k)_{1N}B$*. Since the* $(1, N)$ *block in* Ω^k *is either zero or a power of* D*, we see that* $\mathcal{C}\Omega^k\mathcal{B} = 0$*, for all* $k = 0, 1, \ldots,$ *which on account of theorem (1.1) suffices.* □

5. Schur complements. When A is invertible, we may diagonalize the matrix $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ as follows:

$$\begin{bmatrix} I & 0 \\ -BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} I & -A^{-1}C \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D - BA^{-1}C \end{bmatrix}.$$

Because of this form, the matrix $Z = D - BA^{-1}C$ is called the Schur Complement (SC) of A in M, and is denoted by M/A. Likewise, when D is nonsingular, we have the factorization

$$\begin{bmatrix} I & -CD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}B & I \end{bmatrix} = \begin{bmatrix} A - CD^{-1}B & 0 \\ 0 & D \end{bmatrix},$$

giving the Schur complement $\zeta = A - CD^{-1}B = M/D$.

When neither A nor D are invertible, we have the following fundamental form

$$PMQ = \begin{bmatrix} I & 0 \\ -BX & I \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} I & -XC \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & (I - AX)C \\ B(I - XA) & Z \end{bmatrix} = N$$

and

$$P'MQ' = \begin{bmatrix} I & -CY \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -YB & I \end{bmatrix} = \begin{bmatrix} \zeta & C(I - YD) \\ (I - DY)B & D \end{bmatrix} = N',$$

where

$$Z = D - B(2X - XAX)C \quad \text{and} \quad \zeta = A - C(2Y - YDY)B$$

are the “generalized“ Schur complements. In order to mimic the nonsingular case, it stands to reason that we want X and Y, to be some kind of generalized inverse.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

Indeed, if $X = \hat{A}$ is a 2-inverse of A (i.e., $XAX = X$), then $Z = D - B(2X - XAX)C$ reduces to $Z = D - B\hat{A}C$, which we shall denote by M/A. In particular this holds when $X = A^+$ is a reflexive (1-2) inverse of A (i.e. $AXA = A, XAX = X$). On the other hand if $X = A^-$ is an inner inverse of A (i.e. $AXA = A$), then Z reduces to $Z = D - B(2A^- - A^-AA^-)C$. Similarly if $Y = \hat{D}$ is a 2-inverse of D, then we see that $\zeta = A - C(2Y - YDY) = A - C\hat{D}B = M/D$. Needless to say, M/D and M/A depend on the choice of 2-inverse in general.

The matrices N and N' have properties that are similar to those of M. Indeed, using 2-inverses we may state

$$N'/D = M/D \quad and \quad M/A = N/A. \tag{5.1}$$

We next turn to the case where A is partitioned further, say

$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} = \left[\begin{array}{cc|c} A_1 & A_3 & C_1 \\ A_2 & A_4 & C_2 \\ \hline B_1 & B_2 & D \end{array}\right] = \left[\begin{array}{c|cc} A_1 & A_3 & C_1 \\ \hline A_2 & A_4 & C_2 \\ B_1 & B_2 & D \end{array}\right] = \left[\begin{array}{c|c} A_1 & ? \\ \hline ? & Y \end{array}\right]. \tag{5.2}$$

We shall first present a generalization of the Haynsworth Quotient formula [1, Eq. 3.22].

Theorem 5.1. *For each choice of 2-inverse $\hat{A}_1$ and $(A_4 - A_2\hat{A}_1A_3)^\wedge$, there exists a 2-inverse $\hat{A}$, depending only on these choices, such that*

$$(M/A_1)/(A/A_1) = M/A, \tag{5.3}$$

where $M/A = D - B\hat{A}C$.

Proof. First we compute

$$M/A_1 = \begin{bmatrix} A_4 & C_2 \\ B_2 & D \end{bmatrix} - \begin{bmatrix} A_2 \\ B_1 \end{bmatrix} \hat{A}_1[A_3, C_1] = \begin{bmatrix} A_4 - A_2\hat{A}_1A_3 & C_2 - A_2\hat{A}_1C_1 \\ B_2 - B_1\hat{A}_1A_3 & D - B_1\hat{A}_1C_1 \end{bmatrix}.$$

Next observe that Z $= A/A_1 = A_4 - A_2\hat{A}_1A_3$ and hence that

$$(M/A_1)/(A/A_1) = (D - B_1\hat{A}_1C_1) - (B_2 - B_1\hat{A}_1A_3)\hat{Z}(C_2 - A_2\hat{A}_1C_1),$$

which reduces to

$$\begin{aligned}(M/A_1)/(A/A_1) =& D - B_1(\hat{A}_1 + \hat{A}_1A_3\hat{Z}A_2\hat{A}_1)C_1 \\ & -B_2\hat{Z}C_2 + B_1\hat{A}_1A_3\hat{Z}C_2 + B_2\hat{Z}A_2\hat{A}_1C_1.\end{aligned} \tag{5.4}$$

Next consider the matrix

$$X = \begin{bmatrix} \hat{A}_1 + \hat{A}_1A_3\hat{Z}A_2\hat{A}_1 & -\hat{A}_1A_3\hat{Z} \\ -\hat{Z}A_2\hat{A}_1 & \hat{Z}. \end{bmatrix} \tag{5.5}$$

It is easily checked that $XAX = X$. Using this 2-inverse in $M/A = D - [B_1, B_1]\hat{A}$ $\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$ we again arrive at (5.4), completing the proof. $\square$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

It should be remarked that when we select reflexive inverses A_1^+ and $(A_4 - A_2A_1^+A_3)^+$, then the matrix X in (5.5) also becomes a 1-2 inverse.

Let us next turn to the question of focal power and the quotient formulae. We shall use $M//A$ to denote the special Schur complement $M//N = D - BA^dC$, where A^d denotes the Drazin inverse of A. It will be shown that the quotient formula holds for this type of Schur complements, when we add focal power.

Theorem 5.2. *Let M be as in (5.2). If $D \preccurlyeq Y$ and $Y \preccurlyeq M$ then*

$$M//A = (M//A_1)//(A//A_1) \quad \text{and} \quad M//Y = (M//D)//(Y//D) \tag{5.6}$$

Proof. If $Y \preccurlyeq M$ then $\begin{bmatrix} A_2 \\ B_1 \end{bmatrix} A_1^k[A_3, C_1] = 0$ for all $k = 1, 2 \ldots$, which implies that $\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} A_1^d[A_3, C_1] = 0$ and thus $A_2A_1^dA_3 = 0$ as well. As such we obtain

$$M//A_1 = Y - \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} A_1^d[A_3, C_1] = Y \text{ and } A//A_1 = A_4 - A_2A_1^dA_3 = A_4.$$

Next we note that if D $\preccurlyeq$ Y then $B_2A_4^kC_2 = 0$, for all $k = 0, 1, \ldots$ and thus $B_2A_4^dC_2 = 0$. This means that $Y//A = D - B_2A_4^dC_2 = D$. We may conclude that

$$(M//A_1)//(A//A_1) = Y//A_4 = D.$$

Lastly, by transitivity, $D \preccurlyeq M$, which tells us that $BA^kC = 0$ for all $k = 0, 1, \ldots$ and hence $BA^dC =$ as well. Substituting this into $M//A = D - BA^dC$, we arrive at $M//A = D$, completing the proof of the first identity. The remaining identity follows by symmetry. $\square$

Let us now return to the case where A is invertible, and $Z = D - BA^{-1}C = D - E$. We shall examine a different completion.

Proposition 5.3. *Let $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$ and suppose A is invertible. Further let $Z = D - BA^{-1}C$ be a Schur complement of M and set $N" = \begin{bmatrix} A & C \\ B & Z \end{bmatrix}$. Then the following are equivalent.*
(i) M is (1,1)-focused
(ii) $CD^kZ^rB = 0$ for all $k, r = 0, 1, \ldots$
(iii) $CZ^rB = 0$ for all $r = 0, 1, \ldots$
(iv) N" is (1,1)-focused.

In which case,
(a) $Z^rB = D^rB = D_rB$
(b) $CZ^r = CD^r = CD_r$
(c) $D^r = Z^r + \Gamma_r(Z, E, Z)$
(d) $B_r = \Gamma_r(Z, B, A)$ and $C_r = \Gamma_r(A, C, Z)$
(e)

$$D_k = Z^k + [B, ZB, \ldots, Z^{k-1}B] \begin{bmatrix} A^{k-2} & A^{k-3} & & I & A^{-1} \\ A^{k-3} & \cdots & I & A^{-1} & 0 \\ \vdots & \iddots & \iddots & & \\ I & \iddots & & & \\ A^{-1} & 0 & & \cdots & 0 \end{bmatrix} \begin{bmatrix} C \\ CZ \\ \vdots \\ CZ^{k-1} \end{bmatrix} \tag{5.7}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

$$(5.8) \quad = Z^k + [I, Z, \ldots, Z^{k-1}] \begin{bmatrix} E_{k-2} & E_{k-3} & & E_0 & E_{-1} \\ E_{k-3} & \cdots & & E_{-1} & 0 \\ \vdots & \iddots & \iddots & & \\ E_0 & \iddots & & & \\ E_{-1} & 0 & & \cdots & 0 \end{bmatrix} \begin{bmatrix} I \\ Z \\ \vdots \\ Z^{k-1} \end{bmatrix}.$$

Proof.
(i) ⇒(ii). It holds for $r = 0$. So assume it also holds for $r = r$ and all k. Then $CD^kZ^{r+1}B = CD^kZ.Z^rB = CD^k(D - BA^{-1}C)Z^rB = CD^{k+1}.Z^rB = 0$.
(ii) ⇒ (iii) Clear.
(iii) ⇒ (i). When $r = 0$ we see that $F_0 = CB = 0$. We now claim that $D^kZB = Z^{k+1}B$. This is clear for $k = 0$. So assuming this holds for $k = k$, we have $D^{k+1}ZB = D(D^kZB) = DZ^kB = (Z + BA^{-1}C)Z^kB = Z^{k+1}B$ since $CZ^kB = 0$. We then arrive at $CD^{k+1}B = CD^k.DB = CD^k(Z + BA^{-1}C)B = CD^kZB = CZ^{k+1}B$.
(i) ⇔ (iv). This is clear from theorem (1.1)-(4).
(a) Both equalities are clearly true for $r = 1$, so suppose they hold for $r = r$. Then $D^{r+1}B = DD^rB = (Z + E)Z^rB = Z^{r+1}B + EZ^rB = Z^{r+1}B$, since $EZ^rB = 0$ by part (iii). Also $D_{r+1}B = (BC_r + DD_r)B = B(C_rB) + D(D_rB) = 0 + D(Z^rB) = D.D^rB = D^{r+1}B = Z^{r+1}B$, where we used (a) twice.
(b) This follows by symmetry.
(c) First note that $E^2 = 0$ and $EZ^rE = 0$ and thus $\Gamma_r(Z, E, Z)E = 0 = E\Gamma_r(Z, E, Z)$. Next we observe that the results clearly holds for $r = 1$. Assuming its validity for $r = r$, we arrive at $D^{r+1} = D^rD = [Z^r + \Gamma_{r-1}(Z, E, Z)](Z + E) = Z^{r+1} + (Z^{r-1}EZ + {} + EZ^r) + Z^rE + 0 = Z^{r+1} + \Gamma_r(Z, E, Z)$.
(d) From part (a), $C_k = \sum_{i=0}^{k-1} A^{k-i-1}CD^i = \sum_{i=0}^{k-1} A^{k-i-1}CZ^i = \Gamma_k(A, C, Z)$ and $B_k = \sum_{i=0}^{k-1} D^{k-i-1}BA^i = \sum_{i=0}^{k-1} Z^{k-i-1}BA^i = \Gamma_k(Z, B, A)$.
(e) Note that (c) can be written as

$$D^k = Z^k + [I, Z, .., Z^{k-1}] \begin{bmatrix} 0 & 0 & & E_{-1} \\ 0 & \cdots & E_{-1} & 0 \\ & \iddots & & \\ E_{-1} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} I \\ Z \\ \vdots \\ Z^{k-1} \end{bmatrix},$$

which we substitute in

$$D_k = D^k + [I, Z, .., Z^{k-1}] \begin{bmatrix} E_{k-2} & E_{k-3} & & E_0 & 0 \\ E_{k-3} & & E_0 & 0 & 0 \\ & \iddots & \iddots & & \\ E_0 & 0 & & & 0 \\ 0 & & & & \end{bmatrix} \begin{bmatrix} I \\ Z \\ \vdots \\ Z^{k-1} \end{bmatrix}$$

to give the desired expression (5.7) □.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

6. Remarks. (i) The condition $F_k = 0$ means that the controllability space $W = R[B, DB, D^2B, ...]$ is a D-invariant subspace of $N(C)$, of dimension $d \leq \nu(C)$. Let Q be a $m \times d$ basis matrix for W. Since W is D-invariant we know that $DQ = QR$ for some $d \times d$ matrix R, and because $R(B) \subseteq W$, we also know that $B = QT$ for some $d \times n$ matrix T (here A is $n \times n$). This means that

$$\begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} A & 0 \\ B & DQ \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} A & 0 \\ T & R \end{bmatrix}.$$

The matrix $Y = \begin{bmatrix} I_n & 0 \\ 0 & Q \end{bmatrix}$ has full column rank, and thus is left invertible. It is not clear if any further subspace properties can be obtained.

(ii) If A is nonsingular, we may form the Schur complement of N, i.e. $Z_1 = Z - BA^{-1}C = D - 2BA^{-1}C$. It is now clear that if M is (1,1)- focused then so is $\begin{bmatrix} A & C \\ B & Z_1 \end{bmatrix}$. Needless to say we may repeat this for $Z_k = D - kBA^{-1}C$.

(iii) The (1,1) entries $A_k = A^k + \mathcal{E}_k$ in M^k satisfy the recurrence (1.4). This however, does not show the "dominant" term A^k nor the "error" term $\mathcal{E}_k$. The latter can be expressed as [2]

$$(6.1)\ \mathcal{E}_k = [C, AC, .., A^{k-2}C] \begin{bmatrix} Y_{k-2} & Y_{k-3} & \cdots & Y_0 \\ Y_{k-3} & \cdots & Y_0 & 0 \\ \vdots & \iddots & \iddots & \\ Y_0 & 0 & & 0 \end{bmatrix} \begin{bmatrix} B \\ BA \\ \vdots \\ BA^{k-2} \end{bmatrix}, \quad k = 1, 2,$$

where $Y_{k+1} = DY_k + E_0Y_{k-1} + \ldots + E_{k-1}Y_0$, with $Y_0 = I$.

It then follows by induction that $\mathcal{E}_k = 0$ iff $CD^kB = 0$ iff $CY_kB = 0$ for all $k = 0, 1, \ldots$

(iv) M is diagonally focussed if $(M^k)_{ii} = (M_{ii})^k$ for $i = 1, 2$ and all $k = 0, 1, \ldots$ This happens exactly when E_k and F_k vanish for all k. Such matrices act like block diagonal matrices without being of this form.

Let us close by posing some open problems.

7. Open questions. When M is (1,1)-focused, there are several concepts that present themselves, such as minimal polynomials, graphs, convergence and pencils. For example it would be of interest to know when $\psi_D|\psi_M$? or how one could relate focal power to Roth's theorem: $AX - XD = C \Rightarrow \begin{bmatrix} A & C \\ 0 & D \end{bmatrix} \approx \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$. Likewise it would be of interest to see how the (condensed) graph of M is affected by this property, or what invariance the pencil condition $C\,\mathrm{adj}\,(\lambda I - D)B = 0$, corresponds to? On a different note we could only require that $A_k = A^k$ for $k = 0, \ldots, \ell$, with ℓ **fixed**. What would be the the smallest value of ℓ, for which we obtain (1,1) focal power? Could the (1,1)-focal property shed light on the question of when $\psi_A|\psi_M$, for a general block matrix $M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$? Are there any other dilation that preserve the focal property?

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 13, pp. 56-71, February 2005

ELA
www.math.technion.ac.il/iic/ela

REFERENCES

[1] R.E. Hartwig, Block Generalized Inverses. *Arch. Rat. Mech. Anal.*, 61:197–251, 1976.
[2] R.E. Hartwig. Theory and applications of matrix perturbations with respect to Hankel matrices and power formulae. *Int. J. Syst. Sci.*, 10:437–460, 1979.
[3] R.E. Hartwig and P.Semrl. Constrained Convergence. *Rocky Mountain J. of Math.*, 29:177–195, 1999.
[4] R.A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.
[5] T. Kailath. *Linear Systems.* Prentice Hall, Englewood Cliffs, NJ, 1980. p. 353.
[6] J. Maroulas. Contractive partial Realization. *Int. J. Control*, 49:681–689, 1989.