The Group Inverse of the Nivellateur

R.E. Hartwig∗  P. Patrício†

Abstract

We shall derive necessary and sufficient conditions for the Nivellateur to have a group inverse over an algebraically closed field. We then extend these results to arbitrary fields.

Keywords: Nivellateur, group inverse, matrices over a field

AMS classification: 15A09, 15A21

1 The nivellateur

The matrix equation $AX - XB = C$ can be written in column form as $G\text{vec}(X) = \text{vec}(C)$, where $\text{vec}(Y) = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ when $Y = \begin{bmatrix} y_1 & \ldots & y_n \end{bmatrix}$, and

$$G = I \otimes A - B^T \otimes I$$

is the nivellateur of $A$ and $B$.

∗Mathematics Department, N.C.S.U., Raleigh, NC 27695-8205, U.S.A. e-mail: hartwig@unity.ncsu.edu

†CMAT – Centro de Matemática and Departamento de Matemática e Aplicações, Universidade do Minho, 4710-057 Braga, Portugal. e-mail: pedro@math.uminho.pt

Partially supported by FCT-'Fundaçao para a Ciência e a Tecnologia’, within the project UID-MAT-00013/2013.

Corresponding author.
Our aim is to find necessary and sufficient conditions for the existence of the group inverse of this matrix in terms of $A$ and $B$, and to provide expressions for this group inverse.

We shall use $r(X)$, $\nu(X)$, $R(X)$, $RS(X)$, $N(X)$ to denote rank, nullity, range, row-space, nullspace of $X$, respectively.

Throughout let $A$ be $m \times m$ and $B$ be $n \times n$.

A matrix $A$ has a group inverse if there exists a solution to the equations

\[ AXA = A, \quad XAX = X, \quad AX =XA, \]

in which case the solution is unique and is denoted by $A^\#$. We shall refer to this existence as “$A$ is GP”.

We begin with the easiest case, which is that of a closed field.

\section{The closed field Case}

Consider the matrices $A$ and $B$ over a closed field $F$, with characteristic polynomials

\[ \Delta_A(x) = |xI - A| = \prod_{k=1}^{s(A)} (x - \lambda_k)^{n_k(A)} = \prod_{i=1}^{n(A)} (x - \alpha_i) \]

and

\[ \Delta_B(x) = |xI - B| = \prod_{k=1}^{s(B)} (x - \mu_k)^{n_k(B)} = \prod_{i=1}^{n(B)} (x - \beta_i). \]

Here the $\lambda_k$, $\mu_r$ are distinct and the $\alpha_i$, $\beta_i$ may be repeated. Further let $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$ be the spectrum of distinct eigenvalues of $A$ and let $\tau(A) = (\alpha_1, \ldots, \alpha_m)$ be the list of all of its $m$ eigenvalues – repeated or not. Set $T = \sigma(A) \cap \sigma(B)$.

We denote the algebraic and geometric multiplicities of $\lambda_k(A)$ by $n_k(A)$ and $\nu_k(A) = \dim[N(A - \alpha_k I)]$ respectively.

It is clear that $\sum_{k=1}^{s(A)} n_k(A) = n(A) = m$ and $\sum_{j=1}^{s(B)} n_j(B) = n(B) = n$.

Furthermore, suppose that the minimal polynomial of $A$ is given by

\[ \psi_A(x) = \prod_{k=1}^{s(A)} (x - \lambda_k)^{m_k(A)} \]

with $m_k(A) \leq n_k(A)$. We shall refer to the exponent $m_k(A)$ as the index $\text{ind}(\lambda_k)$ of $\lambda_k$.

It is well known that the group inverse exists if and only if the geometric and algebraic multiplicities of the zero eigenvalue are equal.
We shall compute the algebraic multiplicity \( n_0(G) \) and the geometric multiplicity \( \nu_0(G) \) of the zero eigenvalue of \( G \).

From Stephanos’ theorem (see [6, Theorem 1, page 411]) we know that the eigenvalues of \( G \) have the form \( \lambda_{ij}(G) = \lambda_i(A) - \lambda_j(B) \) with \( i = 1, \ldots, m \) and \( j = 1, \ldots, n \), counted according to multiplicity. This immediately tell us that

\[
\nu_0(G) = \sum_{\gamma \in \sigma(A) \cap \sigma(B)} n_\gamma(A) n_\gamma(B).
\] (2)

To get more information about \( G \), we first reduce \( B \) to its Jordan form, via

\[
Q^{-1}BQ = J_B = \text{diag}(J_{q_1}^1(\beta_1), \ldots, J_{q_u}^u(\beta_u)),
\]

where \( J_k(a) = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ \vdots & \vdots & \ddots & \ddots \\ & & & 1 \\ & & & a \end{bmatrix} \) and \( Q \) is a suitable invertible matrix, made up of Jordan Chains of generalized e-vectors. The \( \beta_j \) may be repeated and \( u \) is the number of Jordan blocks. The associated elementary divisors of \( B \) are given by

\[
\mathcal{E}_B = \{(x - \beta_j)^{q_j}; j = 1, \ldots, u\}.
\]

Likewise the elementary divisors of \( A \) are given by \( \mathcal{E}_A = \{(x - \alpha_i)^{p_i}; i = 1, \ldots, t\} \).

Transforming \( G \) we have

\[
(Q^T \otimes I)G[(Q^T)^{-1} \otimes I] = I \otimes A - J_B^T \otimes I = \text{diag}(G_1, \ldots, G_s),
\]

where

\[
G_i = I \otimes A - J_{q_i}^T(\beta_i) = \begin{bmatrix} A - \beta_i I & 0 \\ -I & A - \beta_i I \\ \vdots & \ddots & \ddots \\ 0 & \ldots & -I & A - \beta_i I \end{bmatrix}_{\text{of block size } q_i \times q_i}
\] (3)

which will also give (2).

We now observe that if \( Au = 0 \) and \( B^Tv = 0 \) then \( G(v \otimes u) = 0 \). This means that

\[
N(B^T) \otimes N(A) \subseteq N(G),
\] (4)
and hence on taking dimensions
\[ \nu(A) \cdot \nu(B) \leq \nu(G). \]

Consequently we have (product rule)
\[ \nu(G) = \nu(A) \cdot \nu(B) \iff N(G) = N(B^T) \otimes N(A). \]

Let us now refine the block form of (3) to obtain:

(i) an expression for \( \nu(G) \) in terms of \( A \) and \( B \),

(ii) conditions for \( G \) to have a group inverse, and

(iii) give a formula for \( G^\# \).

We shall then use the expression for \( \nu(G) \) to show when precisely the product rule holds and when \( \nu(G) = n_0(G) \), i.e. when \( G^\# \) exists.

We begin with

**Lemma 2.1.** Let \( R \) be a ring with unity \( 1 \), and suppose that

\[
J_n(-a) = \begin{bmatrix} a & 0 \\ -1 & a \\ \vdots & \vdots \\ 0 & -1 & a \end{bmatrix} \quad \text{and} \quad K_n(a) = \begin{bmatrix} 1 \\ a \\ a^2 \\ \vdots \\ a^{n-1} & a \\ 0 \end{bmatrix}
\]

are over \( R \) with \( n \geq 2 \). Then

(i) \( K_n(a)^T J_n(-a) = \begin{bmatrix} 0 & a^n \\ I & b \end{bmatrix} \), where \( b^T = [a^{n-1}, a^2, a] \).

(ii) \( J_n(-a)^\# \) exists iff \( a^{-1} \) exists. In which case \( J_n(-a)^\# = J_n(-a)^{-1} = \begin{bmatrix} a^{-1} \\ a^{-2} & a^{-1} & 0 \\ \vdots \\ a^{-n} & \cdots & a^{-1} \end{bmatrix} \).

**Proof.** (i) Clear.

(ii) Equating (2,1) entries in \( J_n(-a)^2 X = J_n(-a) \) and \( (n,n-1) \) entries in \( Y J_n(-a)^2 = J_n(-a) \) we see that \( a \) has both left and right inverses.
From (3) we know that $G^\#$ exists iff each of the blocks $G_i$ has a group inverse. Now when $\beta_i$ is not an eigenvalue of $A$ then $G_i$ is invertible and there is no contribution to $\nu(G)$. So we only need to consider a common eigenvalue $\gamma = \alpha_i = \beta_j$.

So let $\gamma \in T = \sigma(A) \cap \sigma(B)$ and assume that the associated elementary divisors are

$$\mathcal{E}_A = \{(x - \gamma)^{p_1(\gamma)}, \ldots, (x - \gamma)^{p_k(\gamma)}\}$$

and

$$\mathcal{E}_B = \{(x - \gamma)^{q_1(\gamma)}, \ldots, (x - \gamma)^{q_t(\gamma)}\},$$

respectively, where $p_1(\gamma) \geq p_2(\gamma) \geq \cdots \geq p_k(\gamma) \geq 1$ and $q_1(\gamma) \geq q_2(\gamma) \geq \cdots \geq q_t(\gamma) \geq 1$.

There are two cases that can happen.

(i) If $q_i > 1$ then by Lemma 2.1 we know that $G_i^\#$ exists iff $(A - \gamma I)^{-1}$ exists, that is, iff $\gamma \notin \sigma(A)$. So this case cannot occur.

(ii) If $q_i = 1$, i.e. when we have a linear elementary divisor $x - \gamma$ in $\mathcal{E}_B$, then $G_i^\#$ exists iff $(A - \gamma I)^\#$ exists. This happens exactly when $\gamma$ is a simple root of $\psi_A(x)$.

Thus,

**Theorem 2.1.** $G^\#$ exists if and only if for every $\gamma \in \sigma(A) \cap \sigma(B)$ with $q_i = 1$ (a $1 \times 1$ Jordan block) we have $\text{ind}_A(\gamma) = 1$.

In other words, for a common eigenvalue all associated elementary divisors for $A$ and $B$ must be linear.

As a by-product we can compute the nullity of $G$ [5]. Indeed, suppose that $A$ is in Jordan form, say $A = A_\gamma \oplus X$, where $A_\gamma = \text{diag}(J_{p_1(\gamma)}, \ldots, J_{p_k(\gamma)})$, and $X$ contains Jordan blocks with non common eigenvalues. Note that $\nu(A_\gamma) = r$. Then $I \otimes A_\gamma - J_{q_j(\gamma)} \otimes I$ takes the form

$$G_{i,j} = \begin{bmatrix} J_{p_1(\gamma)}(0) & \cdots & 0 \\ -I & J_{p_2(\gamma)}(0) & \cdots \\ \vdots & \ddots & \ddots \\ 0 & \cdots & -I & J_{p_r(\gamma)}(0) \end{bmatrix}_{q_j \text{ blocks}}$$

(6)

Now because $\nu[J_n(0)]^k = \min(n, k)$ we see that

$$\nu(G_{ij}) = \sum_{i=1}^r \min\{p_i, q_j\}$$

(7)

Repeating this for all common eigenvalues we arrive at, c.f. [5],
\[ \nu(G) = \sum_{\gamma \in T} \sum_{j=1}^{r} \sum_{i=1}^{r} \min\{p_i, q_j\}. \] (8)

Let us now use this result to derive a couple of special cases.

If \( T = \emptyset \), there are no common eigenvalues and \( \nu(G) = 0 \). In particular \( 0 \notin T \) and either \( A \) or \( B \) is invertible. Hence \( \nu(A) \cdot \nu(B) = 0 \) and the product rule holds.

If there are common eigenvalues, but 0 is not one of them, then \( \nu(A) \cdot \nu(B) = 0 < \nu(G) \).

Lastly, if 0 is a common eigenvalue, then separating off the common zero eigenvalue we get

\[ \nu(G) = \sum_{0 \neq \alpha \in T} \sum_{i=1}^{k(\alpha)} \sum_{j=1}^{t(\alpha)} \min\{p_i(\alpha), q_j(\alpha)\} \geq \nu(A) \cdot \nu(B). \]

This we rewrite as

\[ \nu(G) - \nu(A) \nu(B) = \sum_{i=1}^{k(\alpha)} \sum_{j=1}^{t(\alpha)} \min\{p_i(\alpha), q_j(\alpha)\} - 1 \] \[ \sum_{0 \neq \alpha \in T} \sum_{i=1}^{k(\alpha)} \sum_{j=1}^{t(\alpha)} \min\{p_i(\alpha), q_j(\alpha)\} \geq 0. \] (9)

Since all terms are non-negative, we see that \( \nu(G) = \nu(A) \cdot \nu(B) \) if and only if there are no common eigenvalues besides zero and for the zero eigenvalue

\[ \sum_{i=1}^{k(\alpha)} \sum_{j=1}^{t(\alpha)} [\min\{p_i(0), q_j(0)\} - 1]. \]

That is, \( \min(p_i, q_j) = 1 \) for all \( i = 1, \ldots, k \), \( j = 1, \ldots, t \). Hence if some \( p_i(0) > 1 \) then all \( q_j(0) > 1 \) or if some \( q_j(0) = 1 \) then all \( p_i(0) = 1 \). That is, either all elementary divisors of \( A \) associated with zero are linear or all those of \( B \) are. Thus the product rule holds if and only if either \( \psi_B(x) = xf(x) \) or \( \psi_B(x) = xg(x) \), where \( (x,f) = 1 = (x,g) \). In other words, the product rule holds if and only if \( A \) and \( B \) have at most the zero eigenvalue in common and either \( A^\# \) or \( B^\# \) or both, exist.

Next we consider

\[ n_0(G) - \nu(G) = \sum_{\alpha \in T} \sum_{i=1}^{k(\alpha)} \sum_{j=1}^{t(\alpha)} [p_i q_j - \min(p_i, q_j)] \geq 0. \]

It thus follows that \( n_0(G) = \nu(G) \), i.e. \( G^\# \) exists, if and only if for each common eigenvalue \( \gamma \), \( p_i q_j = \min(p_i, q_j) \geq 1 \), for all \( i = 1, \ldots, k \), \( j = 1, \ldots, t \). Next we note that if \( r, s \geq 1 \), then

\[ rs = \min\{r, s\} \text{ if and only if } r = s = 1 \] (10)
and conclude that $G^\#$ exists if and only if for each common eigenvalue $\alpha$, the elementary divisors are linear. In other words, if and only if $\gamma \in T \Rightarrow \psi_A(x) = (x - \gamma)f(x)$ and $\psi_B(x) = (x - \gamma)g(x)$, where $\gamma$ is not a root of $f(x)$ or $g(x)$.

Remarks

(i) If $G^\#$ exists then $\gamma \in T$ implies $(A - \gamma I)^\#$ and $(B - \gamma I)^\#$ both exist, yet $A^\#$ and/or $B^\#$ may not exist. For example, if $A$ is invertible and $\psi_B = x^2f(x)$ where $gcd(\Delta_A, f) = 1$, then the condition for $G^\#$ to exist are satisfied, yet $B^\#$ does not exist.

On the other hand, if $A^\#$ and $B^\#$ both exist, then $G^\#$ need not exist since they could have common e-values other than zero.

(ii) We know that if $G^\#$ exists then it is a polynomial in $G$, the coefficients of which can be derived from $\Delta(G)$, which in turn can be found from the eigenvalues of $A$ and $B$. Since this becomes intractable, we shall proceed differently. First an alternative proof of the above which is based on the property of Jordan blocks.

(iii) Since $G^T$ is similar to $(A^T \otimes I - I \otimes B)$ and $\psi_A = \psi_A^T$, we may interchange the roles of $A$ and $B$ to deduce the desired symmetry of Theorem 2.1.

To compute $G^\#$ suppose that $\beta_i \notin \sigma(A)$, for $i = 1, \ldots, t$, and $\beta_i \in \sigma(A)$, for $i = t + 1, \ldots, v$. Next let $Q = [Q_1, \cdots, Q_v]$ and $Y = (Q^T)^{-1} = [Y_1, \cdots, Y_v]$ so that $BQ_i = Q_iJ_{q_i}(\beta_i)$ and $B_i^T = Y_iJ_{q_i}^T(\beta_i)$. Then

$$G^\# = (Y \otimes I) \begin{bmatrix} G_1^{-1} & \cdots & 0 \\ 0 & G_t^{-1} \\ \vdots & \ddots & \ddots \\ 0 & \cdots & G_{t+1}^\# \end{bmatrix} (Q^T \otimes I)$$

$$= \sum_{i=1}^t Y_iG_i^{-1}Q_i^T + \sum_{i=t+1}^v Y_iG_i^\#Q_i^T.$$

Now $G_i^{-1}$ is given as in (2.1) in which $(A - \beta_i I)^{-r}$ can be calculated from the spectral theorem [3]. Indeed,

$$(A - \beta_i I)^{-r} = \sum_{k=1}^s \sum_{j=0}^{m_k-1} (x - \beta_i)^{-r} [y_j Z_k^j] = \sum_{k=1}^s \sum_{j=0}^{m_k-1} (-1)^j (r + j - 1)! (\lambda_k - \beta_i)^{-r-j} Z_k^j. \quad (11)$$
Furthermore \((A - \beta_i I)^\# = g(A)\) where \(g(x) = \begin{cases} 0 & x = \beta_i \\ 1/(x - \beta_i) & x \neq \beta_i \end{cases}\) and so

\[
(A - \beta_i I)^\# = \sum_{k=1}^s \sum_{j=0}^{m_k-1} g^{(j)}(\lambda_k) Z_k^j = \sum_{\lambda_k \neq \beta_j}^{m_k-1} \sum_{j=0} \frac{(-1)^j}{(\lambda_k - \beta_j)^{j+1}} Z_k^j. \tag{12}
\]

Substituting these in the above yields \(G^\#\).

Let us now turn to the case of an arbitrary field.

### 3 The Arbitrary Field Case

We shall now give conditions for \(G^\#\) to exist in term of the invariant factors \(\{a_1(x), \ldots, a_r(x)\}\) of \(A\), and \(\{b_1(x), \ldots, b_s(x)\}\) of \(B\), and compute \(G^\#\) in terms of polynomial matrices associated with \(A\) and/or \(B\).

We begin by reducing \(A\) and \(B\) to their respective rational canonical forms and as such reduce the problem to one where we have two companion matrices \([3, p. 163]\), i.e.,

\[
P^{-1}AP = A_c = \text{diag}[L(a_1(x)), \ldots, L(a_r(x))]\]

and

\[
Q^{-1}BQ = B_c = \text{diag}[L(b_1(x)), \ldots, L(b_s(x))].
\]

The nivellateur becomes

\[
(Q^T \otimes P^{-1})G(Q^{-T} \otimes P) = I_n \otimes A_c - B_c^T \otimes I_m
\]

We permute the diagonal blocks using the “universal flip” matrix – see [3] – to get

\[
G \approx \oplus_{i=1}^r \oplus_{j=1}^s G_{ij},
\]

where \(G_{ij} = I_{n_i} \otimes L[a_i(x)] - L^T[b_j(x)] \otimes I_{m_j}\).

We now replace \(G\) by \(G_{ij}\) and consider the “two-companion” case where \(G = I_n \otimes L[a(x)] - L^T[b(x)] \otimes I_m\), with \(b(x) = b_0 + b_1 x + \cdots + b_n x^n\).

Following [3] we reduce \(xI - L^T[b(x)]\) to its Smith Normal Form via

\[
R(x)[xI - L^T(b)]K(x) = \begin{bmatrix} b(x) & 0 \\ 0 & I_{n-1} \end{bmatrix}, \tag{13}
\]

where \(R(x) = \begin{bmatrix} \beta^T(x) & 1 \\ -I & 0 \end{bmatrix}\), \(K(x)\) is as in lemma (2.1) and \([\beta^T(x), 1] = [b_0(x), \ldots, b_{n-2}(x), 1]\).

In this the \(b_i(x)\) are the adjoint polynomials defined by \([\beta^T(x), 1] = [b_1, \ldots, b_n] K(x)\). We recall in passing that \(\text{adj}(xI - B) = \sum_{i=0}^{n-1} b_i(B) x^i\). Solving this gives

\[
[xI - L^T(b)] = R(x)^{-1} \begin{bmatrix} b(x) & 0 \\ 0 & I_{n-1} \end{bmatrix} K(x)^{-1}, \tag{14}
\]
and subsequently replacing $x$ by $A = L[a(x)]$ throughout, these polynomial identities we arrive at

$$G = R(A)^{-1} \begin{bmatrix} b(A) & 0 \\ 0 & I_{n-1} \end{bmatrix} K(A)^{-1} = PDQ. \quad (15)$$

Since $P$ and $Q$ are invertible we may use [10, Corollary 2], which says that $(PDQ)^\#$ exists if and only if $U = DQPDD^+ + I - DD^+$ is invertible. Since

$$(1 - ab)^{-1} = 1 + a(1 - ba)^{-1}b,$$

this is equivalent to $U' = DQP + I - DD^-$ being invertible, i.e. to $W = D + (I - DD^-)R(A)K(A)$ being invertible.

**Theorem 3.1.** $W$ is invertible if and only if $G^\#$ exists.

To compute $R(x)K(x)$ we define $T(x) = \begin{bmatrix} b^T(x) & 1 \\ -K_{n-1}^{-1} & 0 \end{bmatrix}$, where $b^T = [b_1, \ldots, b_n]$. Then

$$T(x)K_n(x) = R(x) = \begin{bmatrix} b^T(x) & 1 \\ -I_{n-1} & 0 \end{bmatrix}$$

and

$$R(x)K(x) = T(x)K(x)^2 = \begin{bmatrix} b^T & 1 \\ -K_{n-1}^{-1} & 0 \end{bmatrix} \begin{bmatrix} K_{n-1}^2(x) & 0 \\ ? & 1 \end{bmatrix} = \begin{bmatrix} \gamma^T(x) & 1 \\ -K_{n-1}(x) & 0 \end{bmatrix}, \quad (16)$$

in which $\gamma^T(x) = [b'(x), \rho^T(x)]$ and $\rho^T = [b'_0(x), \ldots, b'_{n-3}(x)]$. These contain the formal derivatives of the adjoint polynomials.

We next form

$$(I - DD^-)R(A)K(A) = \begin{bmatrix} I - b(A)b(A)^- & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} [b'(A), \rho^T(A)] & 1 \\ ? & ? \end{bmatrix}$$

$$= \begin{bmatrix} [I - b(A)b(A)^-]b'(A) & C \\ 0 & 0 \end{bmatrix},$$

where $C = [I - b(A)b(A)^-][\rho^T(A), I]$. Adding in $D = \begin{bmatrix} b(A) & 0 \\ 0 & I_{n-1} \end{bmatrix}$ we arrive at

$$W = \begin{bmatrix} b(A) + [I - b(A)b(A)^-]b'(A) & C \\ 0 & I \end{bmatrix}. \quad (17)$$

This will be invertible exactly when $b(A) + [I - b(A)b(A)^-]b'(A)$ is invertible. Note that $b(A)$ and $b'(A)$ commute, but that $b(A)^-$ need not be a polynomial in $A$.

We now need
Lemma 3.1. Suppose $R$ is a von Neumann finite regular ring and $ah = ha$. If $a + (1 - aa^-)h$ is a unit then $a\#$ must exist.

Proof. Let $u = a + (1 - aa^-)h$. Then $ua = a^2 + (1 - aa^-)ha = a^2 + (1 - aa^-)ah = a^2$ and thus $a = u^{-1}a^2$. Since $R$ is finite we may conclude that $a\#$ exists. $\square$

Suppose now that $W$ is invertible. Then $b(A)$ is GP and we can replace $b(A)^{-}$ by $b(A)^\# = g(A)$ in $W$, implying that

Theorem 3.2. $W$ is a unit if and only if $b(A)$ is GP and $f(A) = b(A) + [I - b(A)b(A)^\#]b'(A)$ is a unit.

We shall now reduce these conditions to suitable polynomial results.

First we recall the trivial gcd result

Lemma 3.2. $(u,d) = 1$ if and only if $(dm + u,d) = 1$.

and the group inverse result

Lemma 3.3. Suppose $M$ has minimal polynomial $\psi_M(x)$, and let $f(x)$ be a polynomial with $d(x) = \gcd(f(x), \psi_M(x))$. The following are equivalent:

(i) $f(M)^\#$ exists (ii) $d(M)^\#$ exists (iii) $(d, \psi/d) = 1$ (iv) $(f, \psi/d) = 1$.

The proof is left as an exercise.

The latter says that if $f = p^r\tilde{f}$ and $\psi = p^s\tilde{\psi}$ for some prime factor $p$, with $(\tilde{f}, p) = 1 = (p, \tilde{\psi})$, then $r \geq s$. In other words, common factors of $f$ and $\psi$ occur with minimal degree in $\psi_M$.

Since we may interchange $L(a)$ and $L(b)$ we must actually have that $r = s$. In other words the common prime factors of any invariant factor $a(x)$ of $A$ and any invariant factor $b(x)$ of $B$ must have the same multiplicity.

Now recall that $\psi_A = a(x)$ and set $(a, b) = d$. Then $b = db$ and $a = d\tilde{a}$ for some $\tilde{b}, \tilde{a}$, with $(\tilde{a}, \tilde{b}) = 1$. Moreover $b(A)$ has a group inverse if and only if $(d, \tilde{a}) = 1$ or if $(b, \tilde{a}) = 1$.

The existence of $b(A)^\#$ also says that $b(A)^2g(A) = b(A)$ which holds iff $a|b(1 - bg)$ iff $d\tilde{a}|db(1 - gb)$ iff $\tilde{a}|\tilde{b}(1 - gb)$. But $(\tilde{a}, \tilde{b}) = 1$ and thus $\tilde{a}|(1 - gb)$ and conversely. We may as such write $1 - gb = \tilde{a}h$, for some $h(x)$. This ensures that $(\tilde{a}, b) = 1 = (\tilde{a}, g)$ and gives $f = b + \tilde{a}hb'$.

Next recall, by Hensel’s theorem [8, p. 21, Theorem 15.5], that $f(A)$ is invertible if and only if $(f, a) = 1$, i.e. if and only if $(f, d) = 1 = (f, \tilde{a})$. First we observe that $(f, d) = 1$ if
and only if \((b + (1 - bg)b', d) = 1\) if and only if \((d\tilde{b}(1 - gb') + b', d) = 1\). By Lemma (3.2) this happens precisely when \((b', d) = 1\).

Next we note that because \(b = d\tilde{b}\) we have \(b' = d\tilde{b} + d(\tilde{b})'\) and thus again by the lemma, \((b', d) = 1\) if and only if \((d\tilde{b} + d(\tilde{b})', d) = 1\) if and only if \((d\tilde{b}, d) = 1\) if and only if \((d, d') = 1 = (\tilde{b}, d) = 1\).

Since \((\tilde{a}, \tilde{b}) = 1\) it follows that \((a, \tilde{b}) = (\tilde{d}a, \tilde{b}) = 1\) so that \(\tilde{b}(A)\) is invertible.

We now cancel \(\tilde{b}(A)\) in \(d(A)^2\tilde{b}(A)^2g(A) = b(A)^2g(A) = b(A) = d(A)d(\tilde{b}')(A)\). This implies that
\[
d(A)^2\tilde{b}(A)g(A) = d(A),
\]
so that \(d(A)^\#\) exists and
\[
d(A)^\# = g(A)d(\tilde{b}(A))\quad\text{and}\quad b(A)b(\tilde{A})^\# = d(A)d(A)^\#.
\]

The surprising fact is that the condition \((f, \tilde{a}) = 1\) automatically follows if \(b(A)\) is GP.

Indeed, we have
\[
b(a)^\#\quad\Rightarrow\quad(b, \tilde{a}) = 1\quad\Rightarrow\quad(b + \tilde{a}hb', \tilde{a}) = 1\quad\Rightarrow\quad(b + (1 - bg)b', \tilde{a}) = 1\quad\Rightarrow\quad(f, \tilde{a}) = 1.
\]

We recap in

**Theorem 3.3.** If \(G = I_n \otimes L[a(x)] - L^T[b(x)] \otimes I_m\), then \(G^\#\) exists if and only if \((d, \tilde{a}) = 1 = (d, d')\), where \(d = (a, b)\) and \(a = \tilde{d}a\).

Now \((d, d') = 1\) means that \(d\) only has simple prime factors. As a consequence, the common invariant factors have simple prime factors. For the closed field case, this says that all elementary divisors corresponding to common eigenvalues must be linear – as we met in the previous section.

To compute the actual inverse of \(f(A)\) we observe that because \((d, d') = 1\), we can find \(s\) and \(t\) by Euclid’s algorithm, such that \(d(x)s(x) + d'(x)t(x) = 1\). This means that
\[
d'(A)t(A) = 1 - d(A)s(A).
\]

Substituting for \(b'\) we may rewrite \(f(A) = b(A) + [I - b(A)g(A)]b'(A)\) as \(f(A) = b(A) + [I - d(A)d(\tilde{A})^\#]d'(A)d(\tilde{b}')(A)\), which we may invert to give
\[
f(A)^{-1} = b(A)^\# + [I - d(A)d(\tilde{A})^\#]d(\tilde{b}(A))^{-1}t(A).
\]

Indeed, this follows because
\[
[I - d(A)d(\tilde{A})^\#]d'(A)d(\tilde{b}(A))t(A) = [I - d(A)d(\tilde{A})^\#]d'(A)t(A)
\]
\[
= [I - d(A)d(\tilde{A})^\#][I - d(A)s(A)]
\]
\[
= I - d(A)d(\tilde{A})^\#.
\]
Remark We could have used the fact that $(b', d) = 1$ which gives $b'u = 1 - dv$ for some $v(x)$ and write $f(A)^{-1} = b(A)^# + [I - b(A)b(A)^#]u(A)$. The computation of $u$, however, is more difficult than that of $t(x)$.

Since $d(x)$ only has simple prime factors, the computation of $t(A)$ can be done via the gcd algorithm and the Chinese remainder theorem. Indeed, suppose $d = p_1p_2 \cdots p_k$, where the $p_i$ are distinct prime polynomials. Further set $M_i = \frac{d}{p_i}$ and $g_i = M_i^{-1} \mod p_i$. Next we observe that if $sd + td' = 1$, then $t = (d')^{-1} \mod d$, which is equivalent to $t = (d')^{-1} \mod p_i$ for all $i = 1, \ldots, k$. Because $d' = p_1M_1 + p_2M_2 + \ldots$ we see that $(d')^{-1} \mod p_i = (p_iM_i)^{-1} \mod p_i$. Using the Chinese remainder theorem we may conclude that

$$t = \sum_{i=1}^{k} g_i^2 M_i (p_i')^{-1} \mod p_i.$$  \hfill (20)

4 Computation of $G^#$

We may compute the actual group inverse of $G$ via the formula [10],

$$G^# = PU^{-2}DQ = R(A)^{-1}[I + (I - DK(A)^{-1}R(A)^{-1})(U')^{-1}DD^{-1}]DK(A)^{-1}$$

$$= R(A)^{-1}[I + (RK - D)W^{-1}DD^{-1}]DK(A)^{-1},$$

in which $(U')^{-1} = P^{-1}Q^{-1}W^{-1} = R(A)K(A)W^{-1}$ and $W^{-1} = \begin{bmatrix} f(A)^{-1} & -f(A)^{-1}C \\ 0 & I \end{bmatrix}$.

First we see that

$$W^{-1}DD^{-1} = \begin{bmatrix} f(A)^{-1}b(A)b(A)^# & -f(A)^{-1}C \\ 0 & I \end{bmatrix}.$$
Hence

\[
R(A)K(A)W^{-1}DD^{-1} = \begin{bmatrix}
\begin{bmatrix}
\frac{b'(A)}{b'(A)}
\end{bmatrix} & \rho^T(A) & I \\
-\frac{A}{A^2} & 0 & 0 \\
\vdots & -\frac{A^{n-2}}{A^{n-2}} & 0 \\
-\frac{A}{A^2} & 0 & 0 \\
\end{bmatrix} & \begin{bmatrix}
f(A)^{-1}b(A)b(A)^\# & -f(A)^{-1}C
\end{bmatrix}
\]

Recalling the definition of \( C \) we see that the (1,2) entry becomes

\[
\sigma^T = [I - b'(A)f(A)^{-1}(I - b(A)b(A)^\#)][\rho(A)^T, I].
\]

On the other hand,

\[
DW^{-1}DD^{-1} = \begin{bmatrix}
f(A)^{-1}b(A)b(A)^\# & f(A)^{-1}b(A)b(A)^\# \\
0 & I
\end{bmatrix} = \begin{bmatrix}
f(A)^{-1}b(A) & 0 \\
0 & I
\end{bmatrix},
\]

because \( b(A)C = 0 \).

Whence \( U^{-1} = I + (RK - D)W^{-1}DD^{-1} \) takes the form

\[
U^{-1} = \begin{bmatrix}
\frac{I + f(A)^{-1}b(A)[b'(A)b(A)^\# - I]}{I - f(A)^{-1}C + [0, 0]} & \sigma^T(A)
\end{bmatrix}
\]

This we substitute in

\[
G^\# = R(A)^{-1}[I + (R(A)K(A) - D)W^{-1}DD^{-1}][I + (R(A)K(A) - D)W^{-1}DD^{-1}]DK(A)^{-1},
\]

which is not conducive to simplification.
5 Open Questions and remarks

We end with some pertinent questions and remarks.

1. Squaring the matrix $U^{-1}$ does not look appealing!

2. The expression for $G^\#$ should be “symmetric” in $L(a)$ and $L(b)$, i.e $a(x) - b(x)$ symmetric, and as such there should be some simplification.

3. Can we find a good representation for $(p')^{-1} \mod p$ for a prime polynomial $p(x)$?

4. Can we find the polynomial $g(A) = A^\#$?

5. Can Lemma (3.1) be extended to regular rings?

6. Can we use the invertibility of $ag + 1 - aa^-$ to get a better result?

Acknowledgment

The authors thank an anonymous referee for his/her valuable corrections.

References


