A cancellation property of the Moore-Penrose inverse of triple products

Tobias Damm
Fachbereich Mathematik
TU Kaiserslautern
D-67663 Kaiserslautern
Germany

Harald K. Wimmer
Mathematisches Institut
Universität Würzburg
D-97074 Würzburg
Germany

Keywords: Moore-Penrose generalized inverse, compliance matrix, reverse-order property, weighted generalized inverse, Wedderburn-Guttman theorem, matrix equations

Mathematical Subject Classifications (2000): 15A09, 15A24, 15A90.

Abstract

We study the matrix equation $C(BXC)^{\dagger}B = X^{\dagger}$, where $X^{\dagger}$ denotes the Moore-Penrose inverse. We derive conditions for the consistency of the equation and express all its solutions using singular vectors of B and C. Applications to compliance matrices in molecular dynamics, to mixed reverse-order laws of generalized inverses and to weighted Moore-Penrose inverses are given.

1 Introduction

Let B, C, X, be complex matrices of size $s \times n$, $m \times t$, $n \times m$, respectively, and let $X^{\dagger}$ denote the Moore-Penrose generalized inverse of X. It is the purpose of this paper to characterize all triples (B, C, X) which satisfy

$$C(BXC)^{\dagger}B = X^{\dagger}. \tag{1.1}$$

We say that (B, C, X) has the *cancellation property* if (1.1) holds. If B, C, X are nonsingular $n \times n$ matrices then it is obvious that (1.1) holds. In that case we have $C\,(BXC)^{-1}\,B = X^{-1}$.

Our investigation is motivated by recent applications of compliance matrices in molecular dynamics (for instance [6, 4, 5, 15, 2]). According to [15] the compliance matrix can be defined as the inverse of a force-constant matrix. While the force-constant matrix describes the forces between different parts of a molecule (acting in different directions), the compliance matrix shows how the molecule complies with certain external forces acting on it. In particular centrifugal distortion constants or high-temperature mean-square amplitudes depend directly on compliances rather than on force constants. Very often it is advantageous to model a molecule in a redundant coordinate system, which means using more variables than there are degrees of freedom in the molecule. This is, for instance, the case if all bond lengths and interbond angles are taken as coordinates. In a redundant coordinate system the compliance matrix N_r is defined as the Moore-Penrose inverse of the symmetric force-constant matrix F_r in redundant coordinates. Thus, if F_s is the force-constant matrix in a non-redundant coordinate system, then F_r is related to F_s via $J^\dagger F_r (J^\dagger)^T = F_s$, where J is a linearized coordinate transformation. In general F_r is much larger than F_s, and the question arises, whether $N_r = F_r^\dagger$ can be obtained from $N_s = F_s^\dagger$. This is exactly the case, if the triple (J, J^T, F_r) possesses the cancellation property.

The main result of our paper is Theorem 3.4 in Section 3 with necessary and sufficient conditions for the consistency of (1.1). We shall prove that (1.1) holds if and only if

$$\operatorname{Im} B^*BX = \operatorname{Im} X \quad \text{and} \quad \operatorname{Ker} XCC^* = \operatorname{Ker} X. \tag{1.2}$$

In Section 4 we study (1.1) as a matrix equation. If X is a solution of (1.1) then (1.2) implies that $\operatorname{Im} X$ and $\operatorname{Im} X^*$ are invariant under B^*B and CC^*, respectively. This observation will be used to construct all solutions of (1.1). In Section 5 we consider topics which involve products of the form $C(BXC)^\dagger B$. In particular, we reexamine the issue of compliance matrices and we apply Theorem 3.4 to mixed-type reverse-order laws and to weighted generalized inverses.

2 Notation, basic facts, auxiliary results

Let us first summarize the main issues related to the definition of the Moore-Penrose inverse. Consider a matrix $A \in \mathbb{C}^{n \times m}$ and the corresponding linear mapping $A : \mathbb{C}^m \to \mathbb{C}^n$. Let $\operatorname{Ker} A$ and $\operatorname{Im} A$ denote the kernel and the image of A, respectively. The restriction $A|_{(\operatorname{Im} A^*)} : \operatorname{Im} A^* \to \operatorname{Im} A$ is invertible. Then $A^\dagger$ is defined by $A^\dagger x = \left(A|_{\operatorname{Im} A^*}\right)^{-1} x$ if $x \in \operatorname{Im} A$, and $A^\dagger x = 0$ if $x \in (\operatorname{Im} A)^\perp = \operatorname{Ker} A^*$. This functional definition (see [3, p. 8]) can be illustrated in a diagram:

$$\begin{array}{ccccc}
\mathbb{C}^m & = & \left(\operatorname{Im} A^* = \operatorname{Im} A^\dagger\right) & \oplus & \left((\operatorname{Im} A^*)^\perp = \operatorname{Ker} A\right) \\
 & & A\Big\downarrow \quad \Big\uparrow A^\dagger & & A = 0\Big\downarrow \quad \Big\uparrow A^\dagger = 0 \\
\mathbb{C}^n & = & \operatorname{Im} A & \oplus & \left((\operatorname{Im} A)^\perp = \operatorname{Ker} A^* = \operatorname{Ker} A^\dagger\right).
\end{array}$$

It follows that

$$P_A = AA^\dagger : \mathbb{C}^n \to \mathbb{C}^n \quad \text{and} \quad P_{A^*} = A^\dagger A : \mathbb{C}^m \to \mathbb{C}^m$$

are the orthogonal projections on $\operatorname{Im} A$ and $\operatorname{Im} A^*$, respectively. These properties characterize $A^\dagger$ uniquely, so that $W = A^\dagger$ is the unique solution of the two

Moore equations ([3, p. 9], [1, p. 370])

$$AW = P_A \quad \text{and} \quad WA = P_{A^*}\ . \tag{2.1}$$

It is also clear that $W = A^\dagger$ satisfies the four *Penrose equations* ([3, p. 9], [1, p. 40])

$$\begin{aligned} &AWA = A \quad (1) \qquad && WAW = A \quad (2) \\ &(AW)^* = AW \quad (3) \qquad && (WA)^* = WA\ , \quad (4) \end{aligned} \tag{2.2}$$

and in fact these equations determine $A^\dagger$ uniquely. The sets of conditions (2.1) and (2.2) are equivalent such that $A^\dagger$ is rightly named *Moore-Penrose inverse* of A. We shall exploit the equivalence of the three definitions where it is convenient.

Sometimes, we will only consider a subset of the Penrose conditions (2.2). In accordance with [1, p. 40] let $A\{i, j, \ldots, p\}$ denote the set of matrices $W = A^{(i,j,\ldots,p)} \in \mathbb{C}^{n\times m}$ which satisfy equations $(i), (j), \ldots, (p)$ from (2.2). Thus $\{A^\dagger\} = A^{(1,2,3,4)}$.

The following lemma collects some auxiliary results on kernel and image inclusions, matrix products and Moore-Penrose inverses.

Lemma 2.1. *Let $X \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{k\times m}$ and $C \in \mathbb{C}^{n\times p}$. Then we have the following.*

(i) $\operatorname{Im} X \subseteq \operatorname{Im} B^* \iff X = B^\dagger BX \iff X^\dagger = X^\dagger B^\dagger B$.

(ii) $\operatorname{Ker} C^* \subseteq \operatorname{Ker} X \iff X = XCC^\dagger \iff X^\dagger = CC^\dagger X^\dagger$.

(iii) *If $X = B^\dagger BX = XCC^\dagger$ then*

$$X\,[C\,(BXC)^\dagger\,B] = B^\dagger\,P_{BX}\,B \quad \textit{and} \quad [C\,(BXC)^\dagger\,B]\,X = C\,P_{(XC)^*}\,C^\dagger .$$

Proof. *(i) The assertion follows from $B^\dagger B = P_{B^*}$, and $(X^\dagger)^* = (X^*)^\dagger$ together with $\operatorname{Im} X^\dagger = \operatorname{Im} X^*$.*
(ii) It suffices to note that $\operatorname{Ker} C^* \subseteq \operatorname{Ker} X$ is equivalent to $\operatorname{Im} X^* \subseteq \operatorname{Im} C$.
(iii) Note that $X = XCC^\dagger$ implies $\operatorname{Im} BXC = \operatorname{Im} BX$. Hence we have $P_{BXC} = P_{BX}$ and

$$X[C\,(BXC)^\dagger\,B] = B^\dagger[BXC\,(BXC)^\dagger]B = B^\dagger P_{BXC} B = B^\dagger P_{BX} B.$$

□

3 Main results

In this section we characterize those triplets (B, C, X) which possess property (1.1). In particular, we aim at criteria which do not involve pseudoinverses and describe the cancellation property in terms of image and kernel inclusions. Our first criterion, presented in the following lemma, is rather technical and serves as an intermediate step in the derivation of the main Theorem 3.4.

Lemma 3.1. *The following statements are equivalent.*

(i) *We have*

$$C\,(BXC)^{\dagger}\,B = X^{\dagger}. \tag{3.1}$$

(ii) *The matrices* $K = B^{\dagger}\,P_{BX}\,B$ *and* $L = C\,P_{(XC)^*}\,C^{\dagger}$ *are Hermitian, and*

$$X = B^{\dagger}B\,X \quad \text{and} \quad X = X\,CC^{\dagger}. \tag{3.2}$$

Proof. *Put* $W = C\,(BXC)^{\dagger}\,B$.
'(i) ⇒ (ii)': From (3.1) *it follows that* $X^{\dagger} = X^{\dagger}B^{\dagger}B = CC^{\dagger}X^{\dagger}$. *By Lemma 2.1(i) the preceding identity is equivalent to* (3.2). *Then Lemma 2.1(iii) implies* $XW = K$ *and* $WX = L$. *By* $W = X^{\dagger}$ *we have* $K = K^*$ *and* $L = L^*$.
'(ii) ⇒ (i)': Using (3.2) *we obtain* $W \in X\{1\}$ *from*

$$XWX = [B^{\dagger}B\,X\,CC^{\dagger}]\,C(BXC)^{\dagger}B\,[B^{\dagger}B\,X\,CC^{\dagger}] = \\ B^{\dagger}\,[(BXC)(BXC)^{\dagger}(BXC)]\,C^{\dagger} = B^{\dagger}\,BXC\,C^{\dagger} = X.$$

The identity $WXW = W$ *is obvious. Hence* $W \in X\{1,2\}$. *We know that* (3.2) *implies both* $K = XW$ *and* $L = WX$. *Since* K *and* L *are Hermitian we have* $W \in X\{3,4\}$. *Therefore* $W = X^{\dagger}$, *which is* (3.1). □

We remark that $W \in X\{1\}$ can be deduced from a more general result. Note that (3.2) implies $\operatorname{rank} X = \operatorname{rank} BX = \operatorname{rank} XC$. According to [9] the preceding rank condition holds if and only if $C\,(BXC)^{(1)}B \in X\{1\}$ for each $(BXC)^{(1)} \in (BXC)\{1\}$.

The following example shows that, in general, condition (3.2) on its own is not sufficiently strong to imply (3.1).

Example 1. Take $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $C = B^*$ and $X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then (3.2) holds and $X^{\dagger} = X$. For $BXC = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ we find $(BXC)^{\dagger} = \frac{1}{4}BXC$ and thus $W = C(BXC)^{\dagger}B = \frac{1}{4}\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} \neq X^{\dagger}$. In fact we have $W \in X\{1\}$, but W does not satisfy any of the conditions (2) – (4) in (2.2).

Corollary 3.2. *We have* $X^{\dagger} = C\,(BXC)^{\dagger}\,B$ *if and only if both*

$$X^{\dagger} = (BX)^{\dagger}B \quad \text{and} \quad X^{\dagger} = C(XC)^{\dagger} \tag{3.3}$$

hold.

Proof. *Consider the special cases of Lemma 3.1 with* $C = I$ *or* $B = I$. *We have* $X^\dagger = (BX)^\dagger B$ *if and only if* $B^\dagger P_{BX} B$ *is Hermitian and* $X = B^\dagger BX$. *Similarly,* $X^\dagger = C(XC)^\dagger$ *is valid if and only if* $CP_{(XC)^*}C^\dagger$ *is Hermitian and* $X^\dagger = C(XC)^\dagger$. □

Clearly, $X^\dagger = C(XC)^\dagger$ holds if and only if the adjoint $Y = X^*$ satisfies $Y^\dagger = (C^*Y)^\dagger C^*$. Hence we can focus on the equation $X^\dagger = (BX)^\dagger B$.

Theorem 3.3. *The following statements are equivalent.*

(i) $(BX)^\dagger B = X^\dagger$.

(ii) $\operatorname{Im} B^*BX = \operatorname{Im} X$.

(iii) $\operatorname{Im} B^*BX \subseteq \operatorname{Im} X \subseteq \operatorname{Im} B^*$.

(iv) $(BX)^\dagger = X^\dagger B^\dagger$ *and* $X = B^\dagger BX$.

(v) $\operatorname{Im} X = \operatorname{Im} (B^*B)^\dagger X$.

Proof. *We prove '(i)⇒(ii)⇒ (iv)⇒ (i)' and '(ii)⇒(iii)⇒ (v)⇒ (ii)'.*
'(i) ⇒ (ii)': From $\operatorname{Ker} X^\dagger = [\operatorname{Im} (X^\dagger)^*]^\perp = (\operatorname{Im} X)^\perp$ *and*

$$\operatorname{Ker}(BX)^\dagger B = [\operatorname{Im} B^*\big((BX)^\dagger\big)^*]^\perp = [B^* \operatorname{Im}\big((BX)^\dagger\big)^*]^\perp = [\operatorname{Im} B^*BX]^\perp$$

we obtain $\operatorname{Im} B^*BX = \operatorname{Im} X$.
'(ii) ⇒ (iv)': Clearly $\operatorname{Im} B^*BX = \operatorname{Im} X$ *implies* $\operatorname{Im} B^* \supseteq \operatorname{Im} X$, *and thus* $X = B^\dagger BX$. *We derive the reverse-order identity* $(BX)^\dagger = X^\dagger B^\dagger$ *from* $\operatorname{Im} B^*BX = \operatorname{Im} X$ *as follows. We first prove* $\operatorname{Im}(BX)^\dagger = \operatorname{Im}(X^\dagger B^\dagger)$ *and* $\operatorname{Ker}(BX)^\dagger = \operatorname{Ker}(X^\dagger B^\dagger)$ *and then show that* $(BX)^\dagger z = X^\dagger B^\dagger z$ *for all* $z \in \operatorname{Im} BX$. *We have*

$$\operatorname{Im} (BX)^\dagger = \operatorname{Im} (BX)^* = \operatorname{Im} X^*(B^*BX) = \operatorname{Im} X^*X = \operatorname{Im} X^*$$

and

$$\operatorname{Im} X^\dagger B^\dagger = \operatorname{Im} X^\dagger B^* = \operatorname{Im} X^\dagger B^*B(X^\dagger)^* = \operatorname{Im} X^\dagger(B^*BX) = \operatorname{Im} X^\dagger X = \operatorname{Im} P_{X^*} = \operatorname{Im} X^*.$$

From $\operatorname{Ker}(BX)^\dagger = \operatorname{Ker}(BX)^*$ *follows* $\big(\operatorname{Ker}(BX)^\dagger\big)^\perp = \operatorname{Im} BX$. *On the other hand*

$$\big(\operatorname{Ker} X^\dagger B^\dagger\big)^\perp = \big(\operatorname{Ker} X^* B^\dagger\big)^\perp = \operatorname{Im} (B^\dagger)^* X = \operatorname{Im} (B^\dagger)^* B^*BX = \operatorname{Im} (BB^\dagger)^* BX = P_B \operatorname{Im} BX = \operatorname{Im} BX.$$

If $z \in \operatorname{Im} BX$ *then there is a unique* $u \in \operatorname{Im}(BX)^* = \operatorname{Im} X^*$ *such that* $z = BXu$. *Hence* $(BX)^\dagger z = u$ *and*

$$X^\dagger B^\dagger z = X^\dagger(B^\dagger BX)u = X^\dagger Xu = P_{X^*}u = u\,.$$

'(iv) ⇒ (i)': According to Lemma 2.1(i) the identity $B^\dagger BX = X$ *is equivalent to* $X^\dagger B^\dagger B = X^\dagger$. *Hence we obtain* $X^\dagger B^\dagger B = (BX)^\dagger B = X^\dagger$.
'(ii) ⇒ (iii)': This is obvious.
'(iii) ⇒ (v)': Set $\beta = B^*B_{|\mathrm{Im}\, B^*}$. *Then* $\beta : \mathrm{Im}\, B^* \to \mathrm{Im}\, B^*$ *is invertible and* $\beta^{-1} = (B^*B)^\dagger{}_{|\mathrm{Im}\, B^*}$. *Since* $\mathrm{Im}\, X \subseteq \mathrm{Im}\, B^*$ *the inclusion* $\mathrm{Im}\, B^*BX \subseteq \mathrm{Im}\, X$ *can be written as* $\beta(\mathrm{Im}\, X) \subseteq \mathrm{Im}\, X$. *Hence* $\beta(\mathrm{Im}\, X) = \mathrm{Im}\, X$, *and we obtain* $\mathrm{Im}\, X = \beta^{-1}(\mathrm{Im}\, X) = \mathrm{Im}\,(B^*B)^\dagger X$.
'(v) ⇒ (ii)': We know that $\mathrm{Im}\, B^*BX = \mathrm{Im}\, X$ *implies* $\mathrm{Im}(B^*B)^\dagger X = \mathrm{Im}\, X$. *The converse implication follows from* $(B^*B)^\dagger = B^\dagger (B^\dagger)^*$. □

According to [1, p. 160, Example 22] or [3, p. 23, Theorem 1.4] the *reverse-order property* $(BX)^\dagger = X^\dagger B^\dagger$ is equivalent to

$$\mathrm{Im}\, XX^*B^* \subseteq \mathrm{Im}\, B^* \quad \text{and} \quad \mathrm{Im}\, B^*BX \subseteq \mathrm{Im}\, X. \tag{3.4}$$

We did not take advantage of that result in order to make the proof of Theorem 3.3 self-contained. Using (3.4) we could have deduced (iv) from (iii) as follows. Since $\mathrm{Im}\, X \subseteq \mathrm{Im}\, B^*$ implies $\mathrm{Im}\, XX^*B^* \subseteq \mathrm{Im}\, B^*$ both conditions of (3.4) are satisfied. Hence $(BX)^\dagger = X^\dagger B^\dagger$.

Using Corollary 3.2 we can combine Theorem 3.3 with the analogous results for $C(XC)^\dagger = X^\dagger$ to obtain equivalence of the first six statements in the following theorem.

Theorem 3.4. *The following statements are equivalent.*

(i) $C(BXC)^\dagger B = X^\dagger$.

(ii) $(BX)^\dagger B = X^\dagger$ *and* $C(XC)^\dagger = X^\dagger$.

(iii) $\mathrm{Im}\, B^*BX = \mathrm{Im}\, X$ *and* $\mathrm{Ker}\, XCC^* = \mathrm{Ker}\, X$.

(iv) $\mathrm{Im}\, B^*BX \subseteq \mathrm{Im}\, X \subseteq \mathrm{Im}\, B^*$ *and* $\mathrm{Ker}\, C^* \subseteq \mathrm{Ker}\, X \subseteq \mathrm{Ker}\, XCC^*$.

(v) $(BX)^\dagger = X^\dagger B^\dagger$, $(XC)^\dagger = C^\dagger X^\dagger$, $X = B^\dagger BX$, *and* $X = XCC^\dagger$.

(vi) $\mathrm{Im}\,(B^*B)^\dagger X = \mathrm{Im}\, X$ *and* $\mathrm{Ker}\, X(CC^*)^\dagger = \mathrm{Ker}\, X$.

(vii) $(BXC)^\dagger = C^\dagger X^\dagger B^\dagger$ *and* $X = B^\dagger BXCC^\dagger$.

Proof. *It remains to include* (vii) *in the graph of equivalences. We will exploit the equivalence*

$$X = B^\dagger BXCC^\dagger \quad \Longleftrightarrow \quad X = B^\dagger BX \quad \text{and} \quad X = XCC^\dagger. \tag{3.5}$$

'(vii) ⇒ (i)': Put $W = C(BXC)^\dagger B$. *Then* $W = CC^\dagger X^\dagger B^\dagger B$. *Using* (3.5), *we find* $XWX = X$, $WXW = W$, $WX = CC^\dagger X^\dagger XCC^\dagger = (WX)^*$, *and* $XW = B^\dagger BXX^\dagger B^\dagger B = (XW)^*$, *which means* $W = X^\dagger$.
'(v) ⇒ (vii)': By (3.5) *we have* $X = B^\dagger BXCC^\dagger$. *Note that* $X = XCC^\dagger$ *implies* $\mathrm{Im}\, BXC = \mathrm{Im}\, BX$. *Hence* $P_{BX} = P_{BXC}$. *Analogously, we have*

$P_{(XC)^*} = P_{(BXC)^*}$. *Put* $A = BXC$ *and* $W = C^\dagger X^\dagger B^\dagger$. *Then the Moore equations are easily verified as follows:*

$$\begin{aligned} AW = &\ B(XCC^\dagger)(X^\dagger B^\dagger) = &\ BX(BX)^\dagger = &\ P_{BX} = &\ P_{BXC} = &\ P_A, \\ WA = &\ (C^\dagger X^\dagger)(B^\dagger BX)C = &\ (XC)^\dagger XC = &\ P_{(XC)^*} = &\ P_{(BXC)^*} = &\ P_{A^*}. \end{aligned}$$

□

Now let us assume that X is Hermitian and $C = B^*$. Theorem 3.3(iii) yields a sufficient condition for the cancellation property, which will be applied to compliance matrices in Section 5.

Corollary 3.5. *If* $X = X^*$ *and* $\operatorname{Im} B^* = \operatorname{Im} X$ *then*

$$X^\dagger = B^*(BXB^*)^\dagger B\ .$$

4 The matrix equation $\mathbf{C(BXC)^\dagger B = X^\dagger}$

In this section we consider the matrix equation

$$C(BXC)^\dagger B = X^\dagger\ , \tag{4.1}$$

where $B \in \mathbb{C}^{s\times n}$ and $C \in \mathbb{C}^{m\times t}$ are given and $X \in \mathbb{C}^{n\times m}$ is unknown. We want to determine all solutions of (4.1). Set $\eta = \min\{s, t, m, n\}$. Then (4.1) implies $\operatorname{rank} X \le \eta$. We know from Theorem 3.4 that X satisfies (4.1) if and only if there exists a B^*B-invariant subspace S of $\mathbb{C}^n$ and a CC^*-invariant subspace T of $\mathbb{C}^m$ such that $\operatorname{Im} X = S$ and $\operatorname{Im} X^* = T$. Therefore the spectral decompositions of B^*B and CC^* should play a rôle in the following. Suppose $\operatorname{rank} B = r$ and $\operatorname{rank} C = q$. Let β_i, $i = 1, \ldots, k$, be the different nonzero eigenvalues of B^*B, and let ν_i be the multiplicity of β_i. Correspondingly, let γ_j, $j = 1, \ldots, \ell$, be the different nonzero eigenvalues of CC^*, and let μ_j be their multiplicities. Then there exists a matrix $U \in \mathbb{C}^{n\times r}$ with

$$U = (U_1, \ldots, U_k),\ \ U_i \in \mathbb{C}^{n\times \nu_i},\ i = 1, \ldots, k,\ \ U^*U = I_r, \tag{4.2}$$

such that

$$B^*B\ = \sum_{i=1}^{k} \beta_i\, U_i U_i^*, \tag{4.3}$$

and a matrix $V \in \mathbb{C}^{q\times m}$ with

$$V = (V_1^*, \ldots, V_\ell^*)^*,\ \ V_j \in \mathbb{C}^{\mu_j\times m},\ j = 1, \ldots, \ell,\ \ VV^* = I_q, \tag{4.4}$$

such that

$$CC^*\ = \sum_{j=1}^{\ell} \gamma_j V_j^* V_j. \tag{4.5}$$

Suppose the columns of a matrix $G \in \mathbb{C}^{n\times p}$ are an orthogonal basis of a subspace $S \subseteq \operatorname{Im} B^*$. Then S is invariant under B^*B if and only if for some permutation matrix P we have

$$GP = (U_1M_1, \dots, U_kM_k) = U\ \operatorname{diag}(M_1, \dots, M_k) \tag{4.6}$$

with

$$M_i \in \mathbb{C}^{\nu_i\times\tau_i},\ \ 0 \le \tau_i \le \nu_i,\ \ M_i^*M_i = I_{\tau_i},\ \ i = 1, \dots, k, \tag{4.7}$$

and we have $\operatorname{Im} X = S$ if and only if

$$X = GL \quad \text{for some} \quad L \in \mathbb{C}^{p\times m} \quad \text{with} \quad \operatorname{rank} L = p. \tag{4.8}$$

Similarly, if $T \subseteq \operatorname{Im} C$ is a subspace of $\mathbb{C}^m$ with an orthogonal basis given by the columns of a matrix $H \in \mathbb{C}^{m\times p}$, then T is invariant under CC^* if and only if for some permutation matrix Q we have

$$QH = (V_1^*N_1^*, \dots, V_\ell^*N_\ell^*)^* = \operatorname{diag}(N_1, \dots, N_\ell)\, V \tag{4.9}$$

with

$$N_j \in \mathbb{C}^{\mu_j\times\omega_j},\ \ 0 \le \omega_j \le \mu_j,\ \ N_jN_j^* = I_{\omega_j},\ \ j = 1, \dots, \ell. \tag{4.10}$$

Moreover $\operatorname{Im} X^* = T$ if and only if

$$X = KH \quad \text{for some} \quad K \in \mathbb{C}^{n\times p} \quad \text{with} \quad \operatorname{rank} k = p. \tag{4.11}$$

Theorem 4.1. *Let U_i, $i = 1, \dots, k$, and V_j, $j = 1, \dots, \ell$, be given as in* (4.2), (4.3)*, and* (4.4), (4.5)*, respectively. Suppose $p \le \eta$. Then X is a solution of* (4.1) *and* $\operatorname{rank} X = p$*, if and only if*

$$\begin{aligned} X &= U\ \operatorname{diag}(M_1, \dots, M_k)\, Z\ \operatorname{diag}(N_1, \dots, N_\ell)\, V \\ &= (U_1M_1, \dots, U_kM_k)\, Z \begin{pmatrix} N_1V_1 \\ \dots \\ N_\ell V_\ell \end{pmatrix}, \end{aligned} \tag{4.12}$$

where $M_i \in \mathbb{C}^{\nu_i\times\rho_i}$, $i = 1, \dots, k$, and $N_j \in \mathbb{C}^{\omega_j\times\mu_j}$, $j = 1, \dots, \ell$, are as in (4.7) *and* (4.10)*, and $\sum \rho_i = \sum \omega_j = p$, and $Z \in \mathbb{C}^{p\times p}$ is nonsingular.*

Proof. *Suppose X is given as in* (4.12)*. Then*

$$L = Z\ \operatorname{diag}(N_1, \dots, N_\ell)\, V \in \mathbb{C}^{p\times m}$$

has full row rank. Hence we have (4.8)*, and therefore* $\operatorname{Im} B^*BX = \operatorname{Im} X$*. Similarly,*

$$K = U\ \operatorname{diag}(M_1, \dots, M_k)\, Z \in \mathbb{C}^{n\times p}$$

has full column rank. Then (4.11) *yields* $\operatorname{Im} CC^*X^* = \operatorname{Im} X^*$*. Hence X satisfies* (4.1).

On the other hand, if X is a solution of (4.1) *then we have seen that $X = GL = KH$ with G and H given by* (4.6) *and* (4.9)*. Thus*

$$\operatorname{rank} L = \operatorname{rank} K = \operatorname{rank} X = p, \tag{4.13}$$

and

$$X = XX^\dagger X = GL(GL)^\dagger KH = G[LL^\dagger G^\dagger K]H = GZH.$$

It follows from (4.13) *that $Z = LL^\dagger G^\dagger K \in \mathbb{C}^{p\times p}$ has full rank.* □

5 Applications

We first discuss an issue related to compliance matrices. Then we consider reverse-order laws and weighted generalized inverses.

5.1 Compliance matrices

In Section 1 a compliance matrix N_r was introduced. Recall that N_r is the Moore-Penrose inverse of a symmetric matrix F_r, which is related to a Hessian matrix F_s by $F_s = J^\dagger F_r (J^\dagger)^T$. With regard to [2] or [8] it is important to retrieve information on $F_r^\dagger$ from the matrix $F_s^\dagger$. In a chemical set-up Brandhorst [2] makes the assumption that

$$\theta = \operatorname{rank} F_s = \operatorname{rank} F_r, \tag{5.1}$$

where θ represents the maximal degree of freedom in a molecule. Let us show that the additional assumption

$$\operatorname{Im} J \subseteq \operatorname{Im} F_r \tag{5.2}$$

is sufficient to recover $F_r^\dagger$ completely from $F_s^\dagger$. Note that (5.1) implies $\operatorname{rank} J \geq \operatorname{rank} F_r$. Thus (5.2) implies $\operatorname{Im} J = \operatorname{Im} F_r$. Hence the following observation is an immediate consequence of Corollary 3.5.

Proposition 5.1. *Assume* (5.1) *and* (5.2). *Then* $F_r^\dagger = (J^\dagger)^T F_s^\dagger J^\dagger$.

5.2 Reverse-order laws

Let $R \in \mathbb{C}^{m \times n}$ and $S \in \mathbb{C}^{n \times p}$. According to [13, p. 3110] the pair (R, S) fulfills the *reverse-order law* $(RS)^\dagger = S^\dagger R^\dagger$ if and only if both *mixed-type reverse-order laws*

$$(RS)^\dagger = S^\dagger (R^\dagger RSS^\dagger)^\dagger R^\dagger \quad \text{and} \quad (RS)^\dagger = S^* (R^* RSS^*)^\dagger R^* \tag{5.3}$$

are satisfied. Both equations in (5.3) are of the form $X^\dagger = C(BXC)^\dagger B$ if we set $X = RS$ and $(B, C) = (R^\dagger, S^\dagger)$ or $(B, C) = (R^*, S^*)$. Hence we can use Theorem 3.4 to obtain the results on mixed-type reverse-order laws.

Theorem 5.2. ([12, Theorem 1]) *The following statements are equivalent.*

(i) $(RS)^\dagger = S^\dagger (R^\dagger \, RS \, S^\dagger)^\dagger R^\dagger$.

(ii) $\operatorname{Im} (R^\dagger)^* R^\dagger RS = \operatorname{Im} RS$ *and* $\operatorname{Im} S^\dagger (S^\dagger)^* (RS)^* = \operatorname{Im}(RS)^*$.

(iii) $(R^\dagger RS)^\dagger R^\dagger = S^\dagger (RSS^\dagger)^\dagger$.

Proof. *'(i)$\Leftrightarrow$ (ii)': We apply part (iii) of Theorem 3.4.*
'(iii)$\Leftrightarrow$(i)': Set $A = RS$ *and* $W = (R^\dagger RS)^\dagger R^\dagger$.
Because of $\operatorname{Im} RSS^\dagger = \operatorname{Im} RS$ *we have* $P_{RSS^\dagger} = P_{RS}$. *If (iii) holds then*

$$AW = RSS^\dagger (RSS^\dagger)^\dagger = P_{RSS^\dagger} = P_{RS}.$$

Similarly,

$$WA = (R^\dagger RS)^\dagger R^\dagger RS = P_{(R^\dagger RS)^*} = P_{(RS)^*}.$$

Thus $W = (RS)^\dagger$. *Setting* $(B, C, X) = (R^\dagger, S^\dagger, RS)$ *we find that (iii) is equivalent to* (3.3)*, and, by Corollary 3.2 also to (i).* □

The rank formula in (5.4) below is due to Mazko [7]. In the case where BXC is an invertible square matrix the result is known as Wedderburn-Guttman theorem (see [10, 11]).

Theorem 5.3. *Let* $X \in \mathbb{C}^{n\times m}$, $B \in \mathbb{C}^{s\times n}$ *and* $C \in \mathbb{C}^{m\times t}$ *so that* $\operatorname{rank} X = r$ *and* $\operatorname{rank} BXC = h$. *If* $(BXC)^- \in (BXC)\{1\}$ *then*

$$\operatorname{rank}[X - XC(BXC)^- BX] = r - h. \tag{5.4}$$

In our context the special case with $r = h$ and $(BXC)^- = (BXC)^\dagger$ is of interest. Note that (5.5) in the following proposition can be regarded as another cancellation property of the triple (B, C, X). As we have seen in Example 1, it is weaker than $X^\dagger = C(BXC)^\dagger B$.

Proposition 5.4. *We have*

$$XC(BXC)^\dagger BX = X \tag{5.5}$$

if and only if

$$\operatorname{rank} X = \operatorname{rank} BXC. \tag{5.6}$$

Proof. *The left hand side of* (5.5) *can be written as*

$$XC(BXC)^\dagger BX = XC(BXX^\dagger XC)^\dagger BX.$$

Thus we are in the setting of Theorem 3.4 with $X = X^\dagger$, $B = BX$ *and* $C = XC$. *Because of*

$$\operatorname{Im} X^*B^*BXX^\dagger = \operatorname{Im} X^*B^*BX = \operatorname{Im} X^*B^*$$

and $\operatorname{Im} X^\dagger = \operatorname{Im} X^*$, *the condition* $\operatorname{Im} B^*BX = \operatorname{Im} X$ *can be expressed as* $\operatorname{Im} X^*B^* = \operatorname{Im} X^*$, *which implies* $\operatorname{rank} BX = \operatorname{rank} X$. *Similarly we have* $\operatorname{Im} CC^*X^* = \operatorname{Im} X^*$ *is equivalent to* $\operatorname{rank} XC = \operatorname{rank} X$. *Since*

$$\operatorname{rank} X = \operatorname{rank} BX = \operatorname{rank} XC$$

is equivalent to (5.6) *the proof is complete.* □

5.3 Weighted Moore-Penrose inverses

Our third application deals with a generalization of the Moore-Penrose inverse. Let $M \in \mathbb{C}^{n\times n}$ and $N \in \mathbb{C}^{m\times m}$ be positive definite Hermitian matrices. If $A \in \mathbb{C}^{n\times m}$ then

$$A^{\dagger}_{M,N} = N^{-\frac{1}{2}}\left(M^{\frac{1}{2}}AN^{-\frac{1}{2}}\right)^{\dagger}M^{\frac{1}{2}}$$

is the *weighted* Moore-Penrose inverse of A with respect to M and N. The following observation, which is contained in [1, p. 121, Example 42], is an immediate consequence of Theorem 3.4.

Proposition 5.5. *We have* $A^{\dagger}_{M,N} = A^{\dagger}$ *if and only if*

$$\operatorname{Im} MA = \operatorname{Im} A \quad \textit{and} \quad \operatorname{Ker} AN = \operatorname{Ker} A.$$

We indicate without proof a condition for the cancellation property in the case of a weighted generalized inverse.

Theorem 5.6. *Let B, C and X be of size $n\times n$, $m\times m$ and $n\times m$, respectively. Then*

$$C\,(BXC)^{\dagger}_{M,N}\,B = X^{\dagger}_{M,N}$$

if and only if

$$\operatorname{Im} M^{-1}B^{*}MBX = \operatorname{Im} X \quad \textit{and} \quad \operatorname{Ker} XCN^{-1}C^{*}N = \operatorname{Ker} X.$$

6 Open problems

Consider the matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{6.1}$$

and let

$$S_1 = A - BD^{\dagger}C \quad \text{and} \quad S_2 = D - CA^{\dagger}B \tag{6.2}$$

be the associated Schur complements. If $A = BXC$ and $D = X^{\dagger}$ then we have $S_1 = BXC - B(X^{\dagger})^{\dagger}C = 0$ and $S_2 = X^{\dagger} - C(BXC)^{\dagger}B$. Thus the cancellation property of (B, C, X) is equivalent to $S_2 = 0$. The following problem arises: Let M be the matrix in (6.1) with Schur complements (6.2). When does $S_1 = 0$ imply $S_2 = 0$?

A more detailed investigation of cancellation properties would require an understanding of relations of the from

$$C(BXC)^{(i,\dots,j)}B = X^{(i,\dots,j)}.$$

We remark that a comprehensive study of triple matrix products and mixed-type reverse-order properties of the form

$$(BXC)^{(i,\dots,j)} = (XC)^{(i,\dots,j)}X(BX)^{(i,\dots,j)}.$$

can be found in [14].

References

[1] A. Ben-Israel and Th. N. E. Greville, Generalized Inverses, 2nd ed., Springer, Berlin, 2003.

[2] K. Brandhorst, Computations of Non-covalent Interactions in Quantum Mechanics, Thesis, Technical University of Braunschweig, Department of Life Sciences, 2006 (in German).

[3] St. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Dover Publications, New York, 1991.

[4] J. Grunenberg and N. Goldberg, How strong is the Gallium≡Gallium triple bond? Theoretical compliance matrices as a probe for intrinsic bond strengths, J. Am. Chem. Soc. 122, 6045–6047 (2000).

[5] J. Grunenberg, R. Streubel, G. von Frantzius, and W. Marten, The strongest bond in the universe? Accurate calculation of compliance matrices for the ions N_2H^+, HCO^+, and HOC^+, J. Chem. Phys. 119, 165–169 (2003).

[6] E. Martínez-Torres, J. J. López-Gonzáles, and M. Fernández-Gómez, Unambiguous formalism of molecular vibrations: Use of redundant coordinates and canonical matrices, J. Chem. Phys. 110, 3302–3308 (1999).

[7] A. G. Mazko, Semiinversion and properties of matrix invariants, (Russian), Ukr. Mat. Zh. 40, 525–528 (1988).

[8] C. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch, Using redundant internal coordinates to optimize equilibrium geometries and transition states, J. Comp. Chem. 17(1), 49–56 (1995).

[9] Y. Takane, Y. Tian, and H. Yanai, On constrained generalized inverses of matrices and their properties, Ann. Inst. Stat. Math., to appear.

[10] Y. Takane and H. Yanai, On the Wedderburn-Guttman theorem, Linear Algebra Appl. 410, 267–278 (2005).

[11] Y. Takane and H. Yanai, Alternative characterizations of the extended Wedderburn-Guttman theorem. Linear Algebra Appl. 422, 701–711 (2007).

[12] Y. Tian, The reverse-order law $(AB)^\dagger = B^\dagger(A^\dagger ABB^\dagger)^\dagger A^\dagger$ and its equivalent equalities, J. Math. Kyoto Univ. 45, 841–850 (2005).

[13] Y. Tian, The equivalence between $(AB)^\dagger = B^\dagger A^\dagger$ and other mixed-type reverse-order laws, Int. J. Math. Educ. Sci. Technol. 37, 331–339 (2006).

[14] Y. Tian and Y. Liu, On a group of mixed-type reverse-order laws for generalized inverses of a triple matrix product with applications, Electron. J. Linear Algebra 16, 73–89 (2007).

[15] J. K. G. Watson, A comment on the use of redundant vibrational constants, J. Mol. Struct. 695/696, 71–75 (2004).