A note on constraint preconditioning

Loghin, Daniel

DOI:
10.1137/16M1072036

A NOTE ON CONSTRAINT PRECONDITIONING

DANIEL LOGHIN *

Abstract. We extend the results derived in Keller, Gould and Wathen [3] for constraint preconditioning. In particular, we consider the case where the leading block of the system matrix as well as that of the preconditioner are non-symmetric and have non-trivial kernels. We also analyse the form of the preconditioner with negated constraints, which ensures that under reasonable assumptions the preconditioned system is diagonalisable, while preserving the non-unit eigenvalues and negating some unit eigenvalues.

Key words. Constraint preconditioning, Krylov subspace methods, saddle-point problems.

AMS subject classifications. 65F08, 65F10, 65F15, 65F50.

1. Background. Constraint preconditioners are commonly employed for improving the performance of Krylov methods applied to indefinite linear systems of equations. Various results have been derived over the last two decades and we refer the reader to the descriptions, reviews and references in Benzi, Golub Liesen [1], and Keller, Gould and Wathen [3]. In this note we focus on the eigenvalue problem described in the latter reference; for ease of cross-referencing we preserve some of the notation employed in [3] and summarize its main results below.

Let

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}, \quad \mathcal{G} = \begin{bmatrix} G & B^T \\ B & O \end{bmatrix},$$

with $A, G \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{m\times n}$ and consider the generalised eigenvalue problem

$$\mathcal{A}\mathbf{v} = \lambda \mathcal{G}\mathbf{v}. \tag{1.1}$$

If $\mathcal{A}, \mathcal{G}$ are non-singular, the matrix B has full rank, so that

$$B^T = QR := [Y \ Z] \begin{bmatrix} R \\ O \end{bmatrix}, \tag{1.2}$$

where $R \in \mathbb{R}^{m\times m}$ is upper triangular and non-singular, while the columns of $Z \in \mathbb{R}^{n\times(n-m)}$ form a basis for $\ker B$. The square matrix

$$\begin{bmatrix} Z^T & O \\ Y^T & O \\ O & I_m \end{bmatrix},$$

together with a permutation can be used to generate matrices $\tilde{\mathcal{A}}, \tilde{\mathcal{G}}$ similar to $\mathcal{A}, \mathcal{G}$ respectively, so that the eigenvalue problem (1.1) becomes [3]

$$\tilde{\mathcal{A}}\tilde{\mathbf{v}} := \begin{bmatrix} R^T & O & O \\ Z^TAY & Z^TAZ & O \\ Y^TAY & Y^TAZ & R \end{bmatrix} \tilde{\mathbf{v}} = \lambda \begin{bmatrix} R^T & O & O \\ Z^TGY & Z^TGZ & O \\ Y^TGY & Y^TGZ & R \end{bmatrix} \tilde{\mathbf{v}} =: \lambda \tilde{\mathcal{G}}\tilde{\mathbf{v}}.$$

Hence, $\mathcal{A}, \mathcal{G}$ are non-singular, if and only if Z^TAZ, Z^TGZ are non-singular. This requirement, together with the full rank assumption on B are necessary and sufficient for $\mathcal{A}, \mathcal{G}$ to

*School of Mathematics, University of Birmingham, B15 2TT, UK, e-mail:d.loghin@bham.ac.uk

be non-singular. While this was shown in [2] for the case of symmetric blocks A, G, it is evident that this characterization of non-singularity holds also for general leading blocks.

The transformed preconditioned system takes the form

$$\tilde{\mathcal{P}} := \tilde{\mathcal{G}}^{-1}\tilde{\mathcal{A}} = \begin{bmatrix} I_m & O & O \\ L & P & O \\ M & N & I_m \end{bmatrix}, \tag{1.3}$$

where

$$P = (Z^T GZ)^{-1}(Z^T AZ) \in \mathbb{R}^{(n-m)\times(n-m)} \tag{1.4}$$

and with the expressions for the blocks L, M, N not relevant for our discussion. In the following we denote by $(\mu_j, \mathbf{w}_j)$ the eigenpairs of P, or equivalently, of the following generalized eigenvalue problem

$$Z^T AZ\mathbf{w}_j = \mu_j Z^T GZ\mathbf{w}_j. \tag{1.5}$$

The block form of $\tilde{\mathcal{P}}$ yields the following result [3, Thm 2.1].

THEOREM 1.1. *The preconditioned matrix $\tilde{\mathcal{P}}$ has the following eigenvalues:*

(a) $\lambda = 1$ *with algebraic multiplicity* $2m$*;*

(b) $\lambda = \mu_j, (j = 1, \ldots, n-m)$ *where μ_j are the eigenvalues of* (1.5).

REMARK 1.1. *While the above result was stated in* [3] *under the assumption of symmetry for A and G, it is clear that this is not needed for the result to hold.*

In general, the eigenpairs $(\mu_j, \mathbf{w}_j)$ are complex, possibly defective, unless certain restrictions are placed on A, G. For example, in [3] it is assumed that A, G are symmetric and that $Z^T GZ$ is positive definite, so that μ_j are real and P is diagonalizable. Under the same assumptions, [3] includes an eigenvector analysis and a discussion of the minimal polynomial for $\tilde{\mathcal{P}}$, indicating that the preconditioned matrix $\mathcal{G}^{-1}\mathcal{A}$ is defective, under reasonable assumptions.

In this note, we extend the results in [3] as follows:

- we remove the assumption of symmetry on the leading blocks A, G;
- we allow the leading blocks A, G to be singular and P in (1.4) to be defective;
- we analyze the alternative preconditioner $\mathcal{G}_-$ given by

$$\mathcal{G}_- := \begin{bmatrix} G & B^T \\ -B & O \end{bmatrix}.$$

  We show that replacing the pencil $[\mathcal{A}, \mathcal{G}]$ with $[\mathcal{A}, \mathcal{G}_-]$ achieves the following:
  (i) it preserves the eigenvalues $\mu_j (j = 1, \ldots, n-m)$;
  (ii) it preserves m eigenvalues at 1, while shifting m unit eigenvalues to -1;
  (iii) it allows for a full set of eigenvectors of $\mathcal{G}_-^{-1}\mathcal{A}$, under reasonable assumptions;
  (iv) it preserves the bound on the degree of the minimal polynomial given in [3].
  Thus, the choice $\mathcal{G}_-$ maintains a favourable eigenvalue distribution, while at the same time allowing for analysis that relates the convergence of preconditioned Krylov methods directly to the eigenvalues of the preconditioned system [4]; this is not possible in the defective case.

Throughout the paper we make the following assumptions:

A1 $\mathcal{A}$ and $\mathcal{G}$ are invertible.

A2 $A, G \in \mathbb{R}^{n\times n}$ with $\dim\ker(A \pm G) = k^{\pm}$ and with $\dim\ker(A \pm G) \cap \ker B = s^{\pm}$.

A3 The eigenspace of P corresponding to $\mu_j \neq \pm 1$ has dimension r, with $1 \leq r \leq n-m$.

Assumption A1 is standard: in many applications $\mathcal{A}$ is non-singular, while the action of the inverse of the preconditioner $\mathcal{G}$ is required in iterative methods of Krylov type. We note here that assumption A1 implies that B has full rank and that

$$\ker A \cap \ker B = \ker G \cap \ker B = \{\mathbf{0}\}\,.$$

Assumption A2 is more general than in [3], with both blocks A and G allowed to be nonsymmetric and indefinite and to possess a non-trivial kernel; for the case where this is known a priori, the preconditioner block G may be designed to have the same kernel as A. Moreover, the spaces

$$S_\pm := \ker(A \pm G) \cap \ker B, \tag{1.6}$$

will arise naturally in our analysis. Finally, A3 allows for P to be defective, although for many important classes of problems the matrix P is diagonalizable.

2. Eigenvector analysis. In the following, we will use the notation

$$\mathcal{G}_\pm = \begin{bmatrix} G & B^T \\ \pm B & O \end{bmatrix}, \quad \mathcal{P}_\pm = \mathcal{G}_\pm^{-1}\mathcal{A}.$$

We will also denote by $\mathbf{e}_j$ the jth column of I_m.

We start with the following result concerning the spaces $S_\pm$.

LEMMA 2.1. *Let P be defined in* (1.4) *and let the spaces $S_\pm$ be defined in* (1.6) *with* $\dim S_\pm = s^\pm \geq 0$. *Then P has eigenvalues 1 and -1 with geometric multiplicities s^- and s^+, respectively.*

Proof. Let $\{\mathbf{x}_j, j = 1, \ldots, s^\pm\}$ be a basis set for $S_\pm$. Since $\mathbf{x}_j \in \ker B$, we have $\mathbf{x}_j = Z\mathbf{w}_j$, with the set $\{\mathbf{w}_j, j = 1, \ldots, s^\pm\}$ linearly independent, by the full-rank assumption on B. Since $\mathbf{x}_j \in \ker(A \pm G)$, we have

$$A\mathbf{x}_j = \mp G\mathbf{x}_j \Longleftrightarrow \mathbf{w}_j^T Z^T A Z \mathbf{w}_j = \mp \mathbf{w}_j^T Z^T G Z \mathbf{w}_j,$$

so that $(\mp 1, \mathbf{w}_j)$ is an eigenpair of P. □

REMARK 2.1. *By the above result, the dimension of the eigenspace of P corresponding to $\mu_j \neq \pm 1$ satisfies $r \leq n - m - s^- - s^+$, with equality holding only if P is non-defective. If P has no eigenvalues at ∓ 1, then $s^\pm = 0$ and $S_\pm = \{\mathbf{0}\}$.*

REMARK 2.2. *The spaces $S_\pm$ have trivial intersection: if $\mathbf{x} \in \ker(A + G) \cap \ker(A - G)$ then $\mathbf{x} \in \ker A \cap \ker G$, so that $\mathbf{x} \notin \ker B \setminus \{\mathbf{0}\}$, by assumption A1.*

THEOREM 2.2. *Let* A1–A3 *hold. Then the matrix $\mathcal{P}_+$ has a set of $m + k^- + r$ eigenvectors corresponding to the following eigenpairs:*

(i)

$$\left\{ \left(1, \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_j \end{bmatrix}\right), \quad j = 1, \ldots, m \right\};$$

(ii)

$$\left\{ \left(1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{0} \end{bmatrix}\right), \quad j = 1, \ldots, k^- \right\},$$

where $\{\mathbf{x}_j, j = 1, \ldots, k^-\}$ is a basis for $\ker(A - G)$*;*

(iii)

$$\left\{ \left(\mu_j, \begin{bmatrix} Z\mathbf{w}_j \\ \mathbf{y}_j \end{bmatrix} \right), \quad j = 1, \ldots, r \right\},$$

with

$$\mathbf{y}_j = \frac{1}{\mu_j - 1}(BB^T)^{-1}B(A - \mu_j G)Z\mathbf{w}_j,$$

where $(\mu_j, \mathbf{w}_j)$ *is an eigenpair of* P *with* $\mu_j \neq 1$.

Proof. We seek eigenvectors corresponding to $\lambda = 1$ and $\lambda = \mu_j \neq 1$.

(a) Let $\lambda = 1$. The eigenvalue problem (1.1) becomes

$$A\mathbf{x} + B^T\mathbf{y} = G\mathbf{x} + B^T\mathbf{y}, \tag{2.1a}$$

$$B\mathbf{x} = B\mathbf{x}, \tag{2.1b}$$

which implies that $\mathbf{x} \in \ker(A - G)$, with no restrictions on $\mathbf{y}$. We distinguish the following cases.

i. $\mathbf{x} = \mathbf{0}$. Then (2.1a) is satisfied for any $\mathbf{y} \in \mathbb{C}^m$ and we can choose the following set of eigenpairs:

$$\left\{ \left(1, \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_j \end{bmatrix} \right), \quad j = 1 \ldots, m \right\}.$$

ii. $\mathbf{x} \in \ker(A - G) \setminus \{\mathbf{0}\}$. Then, by A2, another set of eigenpairs can be taken to be

$$\left\{ \left(1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{0} \end{bmatrix} \right), \quad j = 1 \ldots, k^- \right\},$$

where the set $\{\mathbf{x}_j, j = 1, \ldots, k^-\}$ is a basis for $\ker(A - G)$.

(b) Let $\lambda = \mu_j \neq 1$. The eigenvalue problem (1.1) becomes

$$(A - \mu_j G)\mathbf{x} = (\mu_j - 1)B^T\mathbf{y}; \tag{2.2a}$$

$$(\mu_j - 1)B\mathbf{x} = 0. \tag{2.2b}$$

Hence, any component $\mathbf{x}$ of an eigenvector is required to satisfy $\mathbf{x} \in \ker B \setminus \{\mathbf{0}\}$, so that $\mathbf{x} = Z\mathbf{w}$ for some $\mathbf{w} \in \mathbb{C}^{n-m}$. Multiplying (2.2a) by $\mathbf{x}^*$ we obtain the following expression for μ_j

$$\mathbf{w}^* Z^T A Z\mathbf{w} = \mu_j \mathbf{w}^* Z^T G Z\mathbf{w}.$$

Thus, we can identify $\mathbf{w}$ as an eigenvector $\mathbf{w}_j$ of P for $\mu_j \neq 1$ with $\mathbf{x} = \mathbf{x}_j := Z\mathbf{w}_j$. Using (2.2a), for each $\mathbf{x}_j$ there are corresponding vectors $\mathbf{y}_j$ given by

$$\mathbf{y}_j = \frac{1}{\mu_j - 1}(BB^T)^{-1}B(A - \mu_j G)Z\mathbf{w}_j. \tag{2.3}$$

Hence, using A3, $\mathcal{P}_+$ has the following set of eigenpairs

$$\left\{ \left(\mu_j, \begin{bmatrix} Z\mathbf{w}_j \\ \mathbf{y}_j \end{bmatrix} \right), \quad j = 1 \ldots, r \right\}. \tag{2.4}$$

□

We note here that we can take s^- vectors $\mathbf{x}_j$ in part (a)ii. of the proof to be a basis of S_-. By Lemma 2.1, these $\mathbf{x}_j$ are in a one-to-one correspondence with the s^- eigenpairs $(1, \mathbf{w}_j)$ of P, so that s^- of the k^- unit eigenvalues of $\mathcal{P}_+$ correspond to unit eigenvalues of P.

REMARK 2.3. *By Remark 2.1, if P is non-defective, $r = n - m - s^-$ and the dimension of the eigenspace of $\mathcal{P}_+$ is $n + k^- - s^-$. This indicates that, under reasonable assumptions, the matrix $\mathcal{P}_+$ is defective, as the next example shows.*

EXAMPLE 2.1. *If $\ker(A - G) = \{\mathbf{0}\}$, and P is non-defective, the eigenspace of $\mathcal{P}_+$ has dimension n. This may arise in many situations of interest, for example, when A, G are symmetric and positive definite matrices with nonsingular $A - G$.*

The matrix $\mathcal{P}_+$ is diagonalisable only if $k^- = m + s^-$; an example is included below. However, we note here that this restriction excludes some important classes of applications.

EXAMPLE 2.2. *The preconditioned matrix $\mathcal{P}_+$ is non-defective if $s^- = 0$ and $k^- = m$. The latter implies that $\ker(A - G)^\perp = \ker B$. For example, one could have non-singular $\mathcal{A}, \mathcal{G}_+$ of the form*

$$\mathcal{A} = \left[\begin{array}{c|c} & I \\ \hat{A} & \\ \hline I & \end{array}\right], \quad \mathcal{G}_+ = \left[\begin{array}{c|c} & I \\ \hat{G} & \\ \hline I & \end{array}\right],$$

with the pencil $[\hat{A}, \hat{G}]$ non-defective and with no unit eigenvalues.

We now turn to the analysis of $\mathcal{P}_-$. The eigenvalue problem $\mathcal{A}\mathbf{v} = \lambda\mathcal{G}_-\mathbf{v}$ can be transformed as before into the following problem

$$\tilde{\mathcal{A}}\tilde{\mathbf{v}} := \begin{bmatrix} R^T & O & O \\ Z^TAY & Z^TAZ & O \\ Y^TAY & Y^TAZ & R \end{bmatrix} \tilde{\mathbf{v}} = \lambda \begin{bmatrix} -R^T & O & O \\ Z^TGY & Z^TGZ & O \\ Y^TGY & Y^TGZ & R \end{bmatrix} \tilde{\mathbf{v}} =: \lambda\tilde{\mathcal{G}}_-\tilde{\mathbf{v}}.$$

We immediately derive the following result concerning the eigenvalues of $\mathcal{P}_-$.

THEOREM 2.3. *The spectrum of $\mathcal{P}_-$ consists of*

(a) $\lambda = 1$ *with algebraic multiplicity m;*

(b) $\lambda = -1$ *with algebraic multiplicity m;*

(c) $\lambda = \mu_j$ $(j = 1, \dots, n - m)$, *where μ_j satisfy* (1.5).

While the spectrum of $\mathcal{P}_-$ is not as clustered as that of $\mathcal{P}_+$, due to m unit eigenvalues being relocated at -1, there is an advantage in using $\mathcal{G}_-$ to precondition $\mathcal{A}$: the preconditioned matrix is diagonalisable, under reasonable assumptions.

THEOREM 2.4. *Let* A1–A3 *hold. Then the matrix $\mathcal{P}_-$ has a set of $2m + s^- + s^+ + r$ eigenvectors corresponding to the following eigenpairs:*

(i)

$$\left\{ \left(1, \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_j \end{bmatrix}\right), \quad j = 1, \dots, m \right\};$$

(ii)

$$\left\{ \left(1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{0} \end{bmatrix}\right), \quad j = 1, \dots, s^- \right\},$$

with $\{\mathbf{x}_j,\ j = 1, \dots, s^-\}$ a basis for S_-;

(iii)

$$\left\{ \left(-1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{0} \end{bmatrix} \right) \quad j = 1 \ldots, k^+ \right\},$$

where $\{\mathbf{x}_j,\ j = 1, \ldots, k^+\}$ *a basis for* $\ker(A+G)$*;*

(iv)

$$\left\{ \left(-1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{y}_j \end{bmatrix} \right), \quad j = 1, \ldots, m - k^+ + s^+ \right\},$$

where $\{\mathbf{x}_j,\ j = 1, \ldots, m - k^+ + s^+\}$*, is a basis for* $S_+^\perp$.

(v)

$$\left\{ \left(\mu_j, \begin{bmatrix} Z\mathbf{w}_j \\ \mathbf{y}_j \end{bmatrix} \right), \quad j = 1, \ldots, r \right\},$$

with

$$\mathbf{y}_j = \frac{1}{\mu_j - 1}(BB^T)^{-1}B(A - \mu_j G)Z\mathbf{w}_j$$

where $(\mu_j, \mathbf{w}_j)$ *is an eigenpair of* P *with* $\mu_j \neq \pm 1$.

REMARK 2.4. *By Remark 2.2, the vectors* $\mathbf{x}_j \in S_-$ *arising in the set of eigenpairs in* (ii) *satisfy* $\mathbf{x}_j \notin \ker(A+G)$ *and are thus linearly independent of the eigenvectors in part* (iii)*. Thus, all the eigenvectors listed in the theorem are linearly independent.*

Proof.

(a) Let $\lambda = 1$. The eigenvalue problem (1.1) becomes

$$(A - G)\mathbf{x} = \mathbf{0}; \tag{2.5a}$$
$$2B\mathbf{x} = \mathbf{0}, \tag{2.5b}$$

so that $\mathbf{x} \in S_-$ (cf. (1.6)). We distinguish the following cases.

i. $\mathbf{x} = \mathbf{0}$. Then both (2.5a), (2.5b) are satisfied, with $\mathbf{y} \in \mathbb{C}^m$ arbitrary. Hence, we can choose the following set of m eigenpairs:

$$\left\{ \left(1, \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_j \end{bmatrix} \right), \quad j = 1 \ldots, m \right\}.$$

ii. $\mathbf{x} \in S_- \setminus \{\mathbf{0}\}$. In this case, we can choose the eigenpairs

$$\left\{ \left(1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{0} \end{bmatrix} \right), \quad j = 1 \ldots, s^- \right\},$$

where the set $\{\mathbf{x}_j, j = 1 \ldots, s^-\}$ is a basis for S_-.

(b) Let $\lambda = \mu_j \neq \pm 1$. In this case, a similar approach to the proof of Thm 2.2 yields the following set of eigenpairs

$$\left\{ \left(\mu_j, \begin{bmatrix} Z\mathbf{w}_j \\ \mathbf{y}_j \end{bmatrix} \right), \quad j = 1 \ldots, r \right\},$$

where $\mathbf{y}_j$ is given by (2.3).

(c) Let $\lambda = -1$. The eigenvalue problem for $\mathcal{P}_-$ becomes

$$A\mathbf{x} + B^T\mathbf{y} = -G\mathbf{x} - B^T\mathbf{y};$$
$$B\mathbf{x} = B\mathbf{x}.$$

so that we require

$$(A+G)\mathbf{x} = -2B^T\mathbf{y}. \tag{2.7}$$

Noting that we cannot have $\mathbf{x} = \mathbf{0}$, we distinguish the following possibilities.

i. $\mathbf{x} \in \ker(A+G) \setminus \{\mathbf{0}\}$. In this case we find $\mathbf{y} = \mathbf{0}$ and we can therefore choose the following set of eigenpairs

$$\left\{ \left(-1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{0} \end{bmatrix} \right), \quad j = 1 \ldots, k^+ \right\},$$

where $\{\mathbf{x}_j, j = 1 \ldots, k^+\}$ form a basis for $\ker(A+G)$.

ii. $\mathbf{x} \in \ker(A+G)^\perp \setminus \{\mathbf{0}\}$. First, we note that we cannot have $\mathbf{x} \in \ker B$, since (2.7) yields

$$\mathbf{x}^*(A+G)\mathbf{x} = -2\mathbf{x}^*B^T\mathbf{y} = \mathbf{0}$$

and hence $\mathbf{x} \in \ker(A+G)$, which is a contradiction. Therefore,

$$\mathbf{x} \in \ker(A+G)^\perp \setminus \ker B = S_+^\perp.$$

Moreover,

$$\dim S_+^\perp = n - k^+ - \dim\ker B + \dim S_+ = n - k^+ - (n-m) + s^+ = m - k^+ + s^+.$$

Hence, we can choose the remaining eigenpairs in the form

$$\left\{ \left(-1, \begin{bmatrix} \mathbf{x}_j \\ \mathbf{y}_j \end{bmatrix} \right), \quad j = 1 \ldots, m - k^+ + s^+ \right\},$$

where $\mathbf{x}_j$ is a basis element of $S_+^\perp$ and

$$\mathbf{y}_j = -\frac{1}{2}(BB^T)^{-1}B(A+G)\mathbf{x}_j.$$

□

REMARK 2.5. *If P is non-defective, $r = n - m - s^- - s^+$ (cf. Remark 2.1), so that the dimension of the eigenspace of $\mathcal{P}_-$ is $n+m$ and hence $\mathcal{P}_-$ is diagonalisable. In particular, this holds independently of $k^\pm$, while allowing non-trivial spaces $S_\pm$, unlike the case for $\mathcal{P}_+$ (cf. Remark 2.3).*

3. Minimal polynomials. The degree of the minimal polynomial of $\mathcal{P}_+$ was shown in [3, Thm 3.5] to be at most $n - m + 2$ under assumption A1 and under the further assumption of symmetry on A, G. We examine this bound for both $\mathcal{P}_+$ and $\mathcal{P}_-$, under assumptions A1–A3.

We note that by Lemma 2.1, there are $s^\pm$ Jordan blocks in the Jordan normal form of P corresponding to eigenvalues ± 1. Let $\beta^\pm$ denote the size of the largest such blocks.

We consider first $\mathcal{P}_+$ for which the eigenvalues are either 1 or the eigenvalues μ_j of P. By assumption A3, the form of the minimal polynomial of P is

$$p(t) = (t-1)^{\beta^-} \prod_{\ell=1}^{r} (t-\mu_{j_\ell})^{\alpha_\ell},$$

where $\{\mu_{j_\ell}, \ell = 1, \dots, r\}$ represent distinct non-unit eigenvalues of P with algebraic multiplicities α_ℓ. If $\lambda = 1$ is not defective, $\beta^- = 1$; if $S_- = \{\mathbf{0}\}$, then $\beta^- = 0$.

THEOREM 3.1. *Let $q^+(t) = (t-1)^2 p(t)$, where $p(t)$ is the minimal polynomial of P. Then $q^+(\mathcal{P}_+) = O$.*

Proof. The permuted form of $\mathcal{P}_+$ is (cf. (1.3), (1.4))

$$\tilde{\mathcal{P}}_+ = \tilde{\mathcal{G}}_+^{-1} \tilde{\mathcal{A}} = \begin{bmatrix} I & O & O \\ L & P & O \\ M & N & I \end{bmatrix}.$$

We first consider the case $\alpha_\ell = 1$; the general case will follow similarly. Under this assumption,

$$q^+(t) = (t-1)^{2+\beta^-} \prod_{\ell=1}^{r} (t-\mu_{j_\ell}).$$

Define

$$p_i(t) = \prod_{\ell=1}^{i} (t-\mu_\ell), \quad q_i^+(t) = (t-1)^{2+\beta^-} p_i(t)$$

so that $q_r^+(t) = q^+(t)$. Let

$$R_0^- = P - I, \quad R_\ell = P - \mu_{j_\ell} I \qquad (\ell = 1, \dots, r).$$

With this notation in place, a straightforward calculation shows that

$$q_i^+(\tilde{\mathcal{P}}_+) = \begin{bmatrix} O & O & O \\ (R_0^-)^{1+\beta^-} p_i(P) L & (R_0^-)^{2+\beta^-} p_i(P) & O \\ N(R_0^-)^{\beta^-} p_i(P) L & N(R_0^-)^{1+\beta^-} p_i(P) & O \end{bmatrix} \tag{3.1}$$

and the result follows since

$$q^+(\tilde{\mathcal{P}}_+) = q_r^+(\tilde{\mathcal{P}}_+) = O,$$

as all the block entries are either O or have the factor $(P-1)^{\beta^-} p_r(P) = p(P) = O$. It remains to show that (3.1) holds. This can be proved by induction. We perform the ith inductive step only, as the base step is straightforward. Noting that $p_i(P)$ and R_0^- commute, we find

$$
\begin{aligned}
q_{i+1}^{+}(\tilde{\mathcal{P}}_+) &= q_i^{+}(\tilde{\mathcal{P}}_+)(\tilde{\mathcal{P}}_+ - \mu_{i+1} I) \\
&= \begin{bmatrix} O & O & O \\ (R_0^-)^{1+\beta^-} p_i(P) L & (R_0^-)^{2+\beta^-} p_i(P) & O \\ N(R_0^-)^{\beta^-} p_i(P) L & N(R_0^-)^{1+\beta^-} p_i(P) & O \end{bmatrix} \begin{bmatrix} (1-\mu_{i+1}) I & O & O \\ L & R_{i+1} & O \\ M & N & (1-\mu_{i+1}) I \end{bmatrix} \\
&= \begin{bmatrix} O & O & O \\ (R_0^-)^{1+\beta^-} p_i(P) L (1-\mu_{i+1}) + (R_0^-)^{2+\beta^-} p_i(P) L & (R_0^-)^{2+\beta^-} p_i(P) R_{i+1} & O \\ N(R_0^-)^{\beta^-} p_i(P) L (1-\mu_{i+1}) + N(R_0^-)^{1+\beta^-} p_i(P) L & N(R_0^-)^{1+\beta^-} p_i(P) R_{i+1} & O \end{bmatrix} \\
&= \begin{bmatrix} O & O & O \\ (R_0^-)^{1+\beta^-} p_i(P)[(1-\mu_{i+1}) I + R_0^-] L & (R_0^-)^{2+\beta^-} p_{i+1}(P) & O \\ N(R_0^-)^{\beta^-} [(1-\mu_{i+1}) I + R_0^-] p_i(P) L & N(R_0^-)^{1+\beta^-} p_{i+1}(P) & O \end{bmatrix} \\
&= \begin{bmatrix} O & O & O \\ (R_0^-)^{1+\beta^-} p_{i+1}(P) L & (R_0^-)^{2+\beta^-} p_{i+1}(P) & O \\ N(R_0^-)^{\beta^-} p_{i+1}(P) L & N(R_0^-)^{1+\beta^-} p_{i+1}(P) & O \end{bmatrix}
\end{aligned}
$$

since $(1-\mu_{i+1})I + R_0^- = R_{i+1}$ and $p_{i+1}(P) = p_i(P) R_{i+1}$. ∎

REMARK 3.1. *If P is non-defective,* $\deg p = r + \beta^-$ *so that the degree of the minimum polynomial of $\mathcal{P}_+$ is at most $n - m - s^- + \beta^- + 2$. If additionally S_- is trivial, we have* $\deg q^+ = n - m + 2$.

Let us now consider $\mathcal{P}_-$, with eigenvalues ± 1 or μ_j. As before, the minimal polynomial of P will play a role; by assumption A3, its form is

$$p(t) = (t-1)^{\beta^-}(t+1)^{\beta^+} \prod_{\ell=1}^{r} (t - \mu_{j_\ell})^{\alpha_\ell},$$

where $\{\mu_{j_\ell}, \ell = 1, \ldots, r\}$ represents the set of distinct eigenvalues of P not equal to ± 1 and with corresponding algebraic multiplicities α_ℓ.

THEOREM 3.2. *Let $p(t)$ be the minimal polynomial of P and let $q^-(t) = (t-1)(t+1)p(t)$. Then $q^-(\mathcal{P}_-) = O$.*

Proof. The proof follows similarly; as before, we only present the case $\alpha_\ell = 1$. The permuted form of $\mathcal{P}_-$ is (cf. (1.3), (1.4))

$$\tilde{\mathcal{P}}_- = \tilde{\mathcal{G}}_-^{-1} \tilde{\mathcal{A}} = \begin{bmatrix} -I & O & O \\ L & P & O \\ M & N & I \end{bmatrix}.$$

We define

$$p_i(t) = \prod_{\ell=1}^{i} (t - \mu_\ell), \qquad q_i^-(t) = (t-1)^{1+\beta^-}(t+1)^{1+\beta^+} p_i(t),$$

so that $q_r^-(t) = q^-(t)$. Let

$$R_0^- = P - I, \quad R_0^+ = P + I, \quad R_\ell = P - \mu_{j_\ell} I \qquad (\ell = 1, \ldots, r).$$

Then one can show by induction that

$$q_i^-(\tilde{\mathcal{P}}_-) = \begin{bmatrix} O & O & O \\ (R_0^-)^{\beta^-+1}(R_0^+)^{\beta^+} p_i(P) L & (R_0^-)^{\beta^-+1}(R_0^+)^{\beta^++1} p_i(P) & O \\ N(R_0^-)^{\beta^-}(R_0^+)^{\beta^+} p_i(P) L & N(R_0^-)^{\beta^-}(R_0^+)^{\beta^++1} p_i(P) & O \end{bmatrix}$$

and the result follows since

$$q^{-}(\tilde{\mathcal{P}}_{-}) = q_r^{-}(\tilde{\mathcal{P}}_{-}) = O,$$

as all the block entries are either O or have the factor $(P-I)^{\beta^-}(P+I)^{\beta^+}p_r(P) = p(P) = O$. $\square$

REMARK 3.2. *If P is non-defective, then $\deg p = r + \beta^- + \beta^+$ so that the degree of of the minimum polynomial of $\mathcal{P}_-$ is also at most $n - m + 2$.*

Acknowledgments: I am grateful to the referees for recommending the general approach presented in this paper and for several other useful suggestions.

REFERENCES

[1] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.

[2] W. N. Gansterer, J. Schneid, and C. W. Ueberhuber, *Mathematical properties of equilibrium systems*, Technical Report AURORA TR2003-13, University of Vienna and Vienna University of Technology, 2003.

[3] C. Keller, N. I. M. Gould, and A. J. Wathen, *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1300–1317.

[4] J. Pestana and A. J. Wathen, *On the choice of preconditioner for minimum residual methods for non-Hermitian matrices*, J. Comput. Appl. Math., 249 (2013), pp. 57–68.