SIAM J. Matrix Anal. Appl.
Vol. 41, No. 1, pp. 83–105

PROJECTIONS, DEFLATION, AND MULTIGRID FOR NONSYMMETRIC MATRICES*

LUIS GARCÍA RAMOS†, RENÉ KEHL‡, AND REINHARD NABBEN†

Abstract. Deflation is a well-known technique to accelerate Krylov subspace methods for solving linear systems of equations. In contrast to preconditioning, in deflation methods singular systems have to be solved. The original system is multiplied by a projection which leads to a singular linear system which can be more favorable for a Krylov subspace method. Deflation methods are also closely related to multigrid or multilevel and domain decomposition methods. Here, we continue the analysis of deflation methods with arbitrary deflation subspaces applied to nonsymmetric matrices. We give some characterizations of the spectrum of the deflated matrix, and we prove an equivalence theorem for two types of deflation methods. New results on projections, established here, allow simple proofs for the spectral relation between deflation and multigrid methods.

Key words. projections, deflation, augmentation, multigrid, Krylov subspace methods, preconditioning

AMS subject classifications. 65F10, 65F50, 65N22, 65N55

DOI. 10.1137/18M1180268

1. Introduction.

The solution of linear systems of equations of the form

$$Ax = b,$$

where $A \in \mathbb{C}^{n\times n}$ is nonsingular and $x, b \in \mathbb{C}^n$, is a major component in numerical simulations arising in scientific and engineering applications. When the matrix A is sparse, iteration methods based on Krylov subspaces are among the methods of choice. However, in many cases additional techniques are needed to speed up the convergence of the Krylov subspace methods. The best known technique is preconditioning. Instead of solving the original system, one can solve the equivalent system

$$M^{-1}Ax = M^{-1}b,$$

where M^{-1} is a nonsingular matrix that approximates A^{-1} in some sense. In this setting the matrix M^{-1} is known as the preconditioner. Solving the preconditioned system now might be faster than solving the original system.

Alternatively, deflation or augmentation techniques can be used to speed up convergence. In its original formulation deflation methods eliminate eigenvalues out of the spectrum of the operator A. If the spectrum is spread out or some eigenvalues are close to zero, this additional technique often leads to faster convergence of the Krylov subspace method. In augmentation techniques the Krylov subspace is enlarged from the beginning of the iteration by another subspace. Ideally, the augmentation subspace contains information about the original problem that is only slowly revealed in the Krylov subspace itself. This approach is used in order to overcome recurring convergence slowdowns after restarting the GMRES method.

*Received by the editors April 12, 2018; accepted for publication (in revised form) by L. Giraud November 12, 2019; published electronically January 9, 2020.
https://doi.org/10.1137/18M1180268

†Technische Universität Berlin, Institut für Mathematik, Straße des 17. Juni 136, D-10623 Berlin, Germany (garcia@math.tu-berlin.de, nabben@math.tu-berlin.de).
‡Weierstraß Institute for Applied Analysis and Stochastics, Mohrenstraße 39, D-10117 Berlin, Germany (rene.kehl@wias-berlin.de).

Deflation for solving linear systems of equations was introduced in [13, 36, 46] in the late eighties. Since then the deflation technique has been successfully used in the numerical solution of several practical problems; see, e.g., [9, 32, 33, 50, 59].

In contrast to preconditioning, in deflation methods singular systems have to be solved. Indeed, projection operators are the major ingredients in these methods. The original system is multiplied by a projection which leads to a singular linear system whose spectrum might be more favorable for the Krylov subspace method. Often the projections are formed with some knowledge of the spectrum of A and some relevant eigenvectors. The so-called deflation subspaces are spanned by eigenvectors or approximate eigenvectors.

Deflation techniques are also closely related to multigrid or multilevel and domain decomposition methods, which are among the most efficient preconditioners. This relation was already observed in [46]. In [44, 45] the authors pointed out several further similarities between deflation and domain decomposition methods. This comparison is extended to multigrid methods in [62, 63] (see also [47, 48]). It turns out that from an abstract point of view all these methods have a lot of common structure and properties. Using this common structure, deflation subspaces can also be built by the columns of the restriction and the prolongation operator used in multigrid; see [63] for an extensive comparison. Combining all advantages of the methods leads to the multilevel Krylov methods introduced in [18, 20]. For some difficult problems, e.g., linear systems that arise in the numerical solution of the Helmholtz equation, multilevel Krylov methods are very efficient solvers [19, 54, 55].

The analysis of deflation type methods started for Hermitian positive definite matrices. In this case, many properties of the deflation methods have been shown [13, 31, 34, 38, 44, 45, 46, 53]. In [11] an inner-outer method (GCRO) for nonsymmetric matrices was introduced in 1996, where in the inner iterations GMRES is applied to a projected (deflated) linear system. Later, in [8, 51] general deflated and augmented Krylov subspace techniques are discussed. Nonsymmetric deflation is also used in combination with restart techniques for GMRES and other Krylov subspace methods [15, 27, 39, 41, 42]. For sequences of (nonsymmetric) linear systems augmentation and deflation became a useful technique. Information from previous linear systems is used to augment the Krylov subspace to compute the actual solution more efficiently. This technique is known as recycling [1, 2, 3, 4, 7, 16, 32, 50, 57, 58, 64].

Subspaces spanned by eigenvectors, (harmonic) Ritz vectors, or approximations of these, are often used as a deflation subspace [10, 16, 21, 27, 39, 41, 42, 56]. For arbitrary deflation subspaces some spectral properties of the deflated systems are given in [17]. A general framework for deflation and augmentation methods is established in [26, 28, 29]. In the recent paper [24] the authors related the spectrum of deflated matrices to the field of values and proved that the GMRES bound based on the field of values is improved when deflation is applied.

Here we continue the analysis of deflation methods for nonsymmetric matrices and their relation to multigrid methods. The usual subspaces used for nonsymmetric matrices are a subspace $\mathcal{Z}$ and another subspace $A\mathcal{Z}$. We consider the case of two different deflation subspaces $\mathcal{Y}$ and $\mathcal{Z}$. This more general approach is needed for comparing deflation and multigrid methods. The subspaces $\mathcal{Y}$ and $\mathcal{Z}$ can be seen as the range of the restriction and prolongation operators. If for nonsymmetric matrices the restriction is the adjoint of the prolongation operator, we have $\mathcal{Y} = \mathcal{Z}$. Furthermore, if the operators are completely different we have $\mathcal{Y} \neq \mathcal{Z}$ and $\mathcal{Y} \neq A\mathcal{Z}$.

In this paper, we give new characterizations of the eigenvalues of the (preconditioned) deflated system matrix which generalize a result given in [25]. As mentioned above, projections not only are the main ingredients in deflation methods but also help analyze these methods. Here we prove new results for projections. We describe the spectrum of the product of an arbitrary matrix with a projection in comparison to the spectrum of the product added by the complementary projection. This surprising coupling between the spectra not only allows elegant proofs for the relation between deflation and multigrid methods but also highlights the role of projections in this area. With this new theorem we generalize results given first in [17, 63]. Moreover, we give short proofs for the characterization of the eigenvalues of the multigrid preconditioner presented in [47, 48], and further new characterizations. These characterizations of the eigenvalues of the deflated system (or the multigrid preconditioned system) depend on the subspaces $\mathcal{Y}$ and $\mathcal{Z}$; thus, they may help choose these spaces (or, in other words, to choose restriction and prolongation operators) to make the resulting system more suitable for iterative methods. In [23] this is done for Hermitian positive definite systems.

In [26], an equivalence theorem is given which relates deflation and augmentation approaches. With a generalization of this theorem that uses different deflation subspaces we are able to prove a mathematical equivalence between two deflation techniques. The first one uses a final correction after the deflated system is solved [21, 37, 44, 45]. The second one uses a specific starting vector before solving the deflated system [13, 34, 46]. To our knowledge, there is no such clear and explicit equivalence formulation in the literature so far. The second deflation variant is also closely related to augmentation methods.

The paper is organized as follows. In the next section we list some definitions and basic properties of projections and prove our new results on projections. In section 3, we give a brief review of deflation techniques for nonsymmetric matrices and the conditions needed for the convergence of the GMRES method applied to the deflated system. We then give a new characterization of the eigenvalues of the deflated system. In section 4, we establish an equivalence theorem between augmentation and deflation methods, which is then used to prove the equivalence between the two deflation variants. Finally, in section 5, our results are applied to multigrid methods.

2. Projections. Most books on linear algebra and functional analysis include projections, but they are restricted to orthogonal ones. Here we present some basic facts on general projections (see [22, 35, 52] for more results).

Recall that a matrix $P \in \mathbb{C}^{n\times n}$ is a *projection* if $P^2 = P$. It follows immediately that a projection P is the identity on the range of P, i.e., on $\mathcal{R}(P)$. Moreover, a projection can be characterized by its range and null space. On one hand, $\mathbb{C}^n$ is the direct sum of the range of P and the null space of P, i.e., $\mathbb{C}^n = \mathcal{R}(P) \oplus \mathcal{N}(P)$. On the other hand, for any two subspaces $\mathcal{V}$ and $\mathcal{W}$ of $\mathbb{C}^n$ with $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}$ there exists a unique projection P with $\mathcal{R}(P) = \mathcal{V}$ and $\mathcal{N}(P) = \mathcal{W}$ (see, e.g., [22, Theorem 7.17]).

We will then use the following notation.

DEFINITION 2.1. *For two subspaces $\mathcal{V}, \mathcal{W} \subseteq \mathbb{C}^n$ with $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}$ the operator $P_{\mathcal{V},\mathcal{W}}$ is defined as the unique projection with $\mathcal{R}(P) = \mathcal{V}$ and $\mathcal{N}(P) = \mathcal{W}$. The projection $P_{\mathcal{V},\mathcal{W}}$ is called the projection onto $\mathcal{V}$ along $\mathcal{W}$.*

Let $\mathcal{V}$ be a subspace of $\mathbb{C}^n$ with dimension r. For an arbitrary matrix $B \in \mathbb{C}^{n\times n}$ we define the set

$$B\mathcal{V} = \{Bv : v \in \mathcal{V}\}.$$

If B is nonsingular, then

$$\left(B^H\mathcal{V}\right)^{\perp} = B^{-1}\mathcal{V}^{\perp}$$

for every subspace $\mathcal{V}$, where B^H denotes the Hermitian transpose of B. The orthogonal projection $P_{\mathcal{V},\mathcal{V}^\perp}$ onto $\mathcal{V}$ along $\mathcal{V}^\perp$ is denoted by $P_{\mathcal{V}}$. Note that for a given projection $P_{\mathcal{V},\mathcal{W}}$, the complementary projection is given by $P_{\mathcal{W},\mathcal{V}} = I - P_{\mathcal{V},\mathcal{W}}$. Here, I denotes the identity matrix independently of the order of the matrix.

Next we consider two subspaces $\mathcal{V}, \mathcal{W}$ such that $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}^\perp$. With the help of a basis of $\mathcal{V}$ and $\mathcal{W}$, the projection $P_{\mathcal{V},\mathcal{W}^\perp}$ can be described easily.

THEOREM 2.2. *Let $\mathcal{V}, \mathcal{W}$ be subspaces of $\mathbb{C}^n$ of dimension r spanned by the colums of $V, W \in \mathbb{C}^{n\times r}$, respectively. The following statements are equivalent:*

(a) $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}^\perp$.

(b) $W^H V$ *is nonsingular and the projection* $P_{\mathcal{V},\mathcal{W}^\perp}$ *has the representation*

$$P_{\mathcal{V},\mathcal{W}^\perp} = V\left(W^H V\right)^{-1} W^H.$$

(c) $\mathcal{V} \cap \mathcal{W}^\perp = \{0\}$.

(d) *The orthogonal complements form a direct sum, i.e.,* $\mathbb{C}^n = \mathcal{V}^\perp \oplus \mathcal{W}$.

The statements in the previous theorem are well known and can be found in, e.g., [22, 52].

Next we will state some new theorems on projections that will be used several times in the following sections. We consider a nonsingular matrix C multiplied by a projection P, i.e.,

$$S = CP.$$

Clearly, if $P \neq I$, then 0 is in the spectrum of S. We then add the complementary projection $I - P$ to the product, i.e., we build

$$T = CP + I - P. \tag{2.1}$$

Now the eigenvalue zero is translated to one, i.e., the eigenvectors of CP in the null space of the projection P are eigenvectors of T corresponding to the eigenvalue one. However, we can prove that the remaining eigenvalues of T are eigenvalues of CP. Further, in Theorem 2.4 we will formulate a necessary and sufficient condition to ensure that T is nonsingular, that is, no extra zeros appear in the spectrum of T.

With help of these theorems the spectra of deflated systems and deflation-based preconditioned systems can be established easily. Moreover, the relation between spectra of deflated systems and spectra of multigrid preconditioned systems as well as characterizations of these spectra can be proven in a short way. In the following the spectrum of $B \in \mathbb{C}^{n\times n}$ is denoted by $\sigma(B)$.

THEOREM 2.3. *Let $C \in \mathbb{C}^{n\times n}$ be nonsingular and let $\mathcal{V}, \mathcal{W}$ be a pair of subspaces of $\mathbb{C}^n$ of dimension r such that $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}^\perp$. Let $P = P_{\mathcal{W}^\perp,\mathcal{V}}$.*

(a) *If*

$$\sigma(CP) = \{0\} \cup \{\mu_{r+1}, \dots, \mu_n\},$$

then

$$\sigma(CP + (I - P)) = \{1\} \cup \{\mu_{r+1}, \dots, \mu_n\}.$$

(b) *Conversely, if*

$$\sigma(CP + (I - P)) = \{1\} \cup \{\mu_{r+1}, \dots, \mu_n\},$$

then

$$\sigma(CP) = \{0\} \cup \{\mu_{r+1}, \dots, \mu_n\}$$

for $\mu_j \in \mathbb{C}$, $r + 1 \leq j \leq n$.

Proof. We begin with part (a). Consider $y \neq 0$ such that

$$CPy = \lambda y. \tag{2.2}$$

Now let $y \in \mathcal{V}$, then $Py = 0$ and $CPy = 0$. This shows that 0 is an eigenvalue of CP and it has algebraic multiplicity at least $r = \dim(\mathcal{V})$.

Next, let $y \notin \mathcal{V}$ so $Py \neq 0$. Multiplying (2.2) with P and adding $0 = (I - P)Py$ gives

$$(PCP + (I - P))Py = \lambda Py,$$

therefore $\lambda \in \sigma(PCP + (I - P))$. Since $PCP + (I - P) = P(C - I)P + I$ it follows that

$$\begin{aligned}\sigma(PCP + (I - P)) &= \sigma(P(C - I)P + I)\\ &= \{1 + z : z \in \sigma(P(C - I)P)\}\\ &= \{1 + z : z \in \sigma((C - I)P)\}\\ &= \sigma(CP + (I - P)),\end{aligned}$$

where we have used that $\sigma(FG) = \sigma(GF)$ for all $F, G \in \mathbb{C}^{n\times n}$ and the property $P^2 = P$. Thus, $\lambda \in \sigma(CP + (I - P))$.

We show now part (b). Consider $y \neq 0$ such that

$$(CP + (I - P))y = \lambda y. \tag{2.3}$$

Suppose that $y \in \mathcal{V}$. Then $Py = 0$ and we have $(CP+(I-P))y = y$ and $CPy = 0$. Now, let $y \notin \mathcal{V}$. Multiplying (2.3) with P and using $P(I - P) = (I - P)P = 0$, we obtain

$$P(CP + (I - P))y = P(CP + (I - P))Py = \lambda Py.$$

Since $y \notin \mathcal{V}$, we have $Py \neq 0$. This gives

$$\lambda \in \sigma(P(CP + (I - P))) = \sigma(PCP) = \sigma(CP^2) = \sigma(CP),$$

which concludes the proof. ∎

Note that some of the μ_i in the above spectra can be zero. The next theorem gives a necessary and sufficient condition to avoid additional zeros.

THEOREM 2.4. *Let $C \in \mathbb{C}^{n\times n}$ be nonsingular and let $\mathcal{V}, \mathcal{W}$ be a pair of subspaces of $\mathbb{C}^n$ of dimension r such that $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}^\perp$. Let $P = P_{\mathcal{W}^\perp,\mathcal{V}}$. Then $CP+(I-P)$ is nonsingular if and only if $C\mathcal{W}^\perp \cap \mathcal{V} = \{0\}$ or, equivalently, if and only if $\mathcal{W}^\perp \cap C^{-1}\mathcal{V} = \{0\}$.*

Proof. Suppose first that $CP+(I-P)$ is nonsingular and let $y \in C\mathcal{W}^\perp \cap \mathcal{V}$. Then $y = Cw$ for some $w \in \mathcal{W}^\perp$ and

$$(CP + (I - P))(w - y) = \big(CP_{\mathcal{W}^\perp,\mathcal{V}}(w - y) + P_{\mathcal{V},\mathcal{W}^\perp}(w - y)\big) = Cw - y = 0.$$

Therefore $w - y = 0$; this gives $y \in \mathcal{W}^\perp \cap \mathcal{V}$ so $y = 0$ and we have $C\mathcal{W}^\perp \cap \mathcal{V} = \{0\}$.

For the other direction, let $y \in \mathbb{C}^n$ such that

$$(CP + (I - P))y = 0,$$

and let $z = CPy = -(I - P)y$. Then $z \in C\mathcal{W}^\perp \cap \mathcal{V}$, so $z = 0$ and $Py = (I - P)y = 0$, which implies $y = 0$; therefore $CP + (I - P)$ is nonsingular. ∎

If $C\mathcal{W}^\perp \cap \mathcal{V} = \{0\}$ then the inverse of $CP_{\mathcal{W}^\perp,\mathcal{V}} + P_{\mathcal{V},\mathcal{W}^\perp}$ can be stated explicitly.

THEOREM 2.5. *Let $C \in \mathbb{C}^{n\times n}$ be nonsingular and let $\mathcal{V}, \mathcal{W}$ be a pair of subspaces of $\mathbb{C}^n$ of dimension r such that $\mathbb{C}^n = \mathcal{V} \oplus \mathcal{W}^\perp$. Let $C\mathcal{W}^\perp \cap \mathcal{V} = \{0\}$; then $CP_{\mathcal{W}^\perp,\mathcal{V}} + P_{\mathcal{V},\mathcal{W}^\perp}$ is nonsingular and*

$$\left(CP_{\mathcal{W}^\perp,\mathcal{V}} + P_{\mathcal{V},\mathcal{W}^\perp}\right)^{-1} = P_{\mathcal{V},C\mathcal{W}^\perp} + C^{-1}P_{C\mathcal{W}^\perp,\mathcal{V}}.$$

Proof. We have

$$\begin{aligned}
&\left(P_{\mathcal{V},C\mathcal{W}^\perp} + C^{-1}P_{C\mathcal{W}^\perp,\mathcal{V}}\right)\left(CP_{\mathcal{W}^\perp,\mathcal{V}} + P_{\mathcal{V},\mathcal{W}^\perp}\right)\\
&= C^{-1}P_{C\mathcal{W}^\perp,\mathcal{V}}CP_{\mathcal{W}^\perp,\mathcal{V}} + P_{\mathcal{V},C\mathcal{W}^\perp}P_{\mathcal{V},\mathcal{W}^\perp}\\
&= P_{\mathcal{W}^\perp,\mathcal{V}} + P_{\mathcal{V},\mathcal{W}^\perp}\\
&= I,
\end{aligned}$$

which completes the proof. ∎

3. Nonsymmetric deflation.

In this section we study deflation operators based on a pair of subspaces $\mathcal{Y}$ and $\mathcal{Z}$, satisfying some conditions, leading to a projection of the form

$$P_{\mathcal{Y}^\perp,A\mathcal{Z}}.$$

Here A is in general nonsingular and nonsymmetric. The nonsymmetry of A allows that we consider a pair of subspaces. Often in deflation methods the subspace $\mathcal{Y}$ is chosen as $\mathcal{Y} = A\mathcal{Z}$. We will study the more general case and also point out further results that hold only for this choice of $\mathcal{Y}$.

Many of the theorems in this section generalize results for deflation applied to symmetric matrices. However, without the assumption of symmetry several new questions need to be answered:

- What conditions on $\mathcal{Z}$ and $\mathcal{Y}$ guarantee that the projection is well-defined?
- When does the GMRES method converge for the deflated (singular) system, $P_{\mathcal{Y}^\perp,A\mathcal{Z}}Ax = P_{\mathcal{Y}^\perp,A\mathcal{Z}}b$?
- Does the deflation procedure speed up the Krylov subspace method?

The first question is answered by Theorem 2.2 using $\mathcal{W} = \mathcal{Y}$ and $\mathcal{V} = A\mathcal{Z}$. Let Y and Z be matrices whose columns form a basis of $\mathcal{Y}$ and $\mathcal{Z}$, respectively; then the condition $\mathbb{C}^n = A\mathcal{Z} \oplus \mathcal{Y}^\perp$ (equivalent to $\mathbb{C}^n = A^{-1}\mathcal{Y}^\perp \oplus \mathcal{Z}$) leads to a nonsingular matrix $Y^HAZ \in \mathbb{C}^{r\times r}$ such that the projection is well-defined.

Nonsymmetric deflation uses the projections

$$P_D := I - AZ\left(Y^HAZ\right)^{-1}Y^H = P_{\mathcal{Y}^\perp,A\mathcal{Z}} \tag{3.1}$$

and

$$Q_D := I - Z\left(Y^HAZ\right)^{-1}Y^HA = P_{A^{-1}\mathcal{Y}^\perp,\mathcal{Z}}. \tag{3.2}$$

We also define for convenience

$$E = Y^HAZ \quad \text{and} \quad Q = Z\left(Y^HAZ\right)^{-1}Y^H. \tag{3.3}$$

If we choose $\mathcal{Y} = A\mathcal{Z}$, then a matrix Z whose columns form a basis of $\mathcal{Z}$ leads to a nonsingular matrix $Y^HAZ = Z^HA^HAZ$ if A is nonsingular.

We immediately obtain the following properties of P_D and Q_D.

PROPOSITION 3.1. *For any $Z, Y \in \mathbb{C}^{n\times r}$ with ranks r such that P_D and Q_D as given in* (3.1) *and* (3.2) *are well-defined the following equalities hold:*

1. $P_D AZ = 0$,
2. $Y^H AQ_D = 0$,
3. $Q_D Z = 0$,
4. $Y^H P_D = 0$,
5. $P_D A = AQ_D$.

If x is the solution of $Ax = b$, then $(I - Q_D)x = Z(Y^H AZ)^{-1} Y^H b$.

Similar to the symmetric case [38], one way to describe deflation for solving linear systems of equations is to first solve

$$P_D Ax = P_D b, \tag{3.4}$$

which gives a solution x^*. Then we have

$$\begin{aligned} & P_D Ax^* = P_D b \\ \Leftrightarrow\ & AQ_D x^* = P_D b \\ \Leftrightarrow\ & AQ_D x^* = (I - AQ)b \\ \Leftrightarrow\ & AQ_D x^* + AQb = b \\ \Leftrightarrow\ & A(Q_D x^* + Qb) = b. \end{aligned}$$

Thus, $Q_D x^* + Qb$ is the solution of the original system $Ax = b$.

We then have the following deflation procedure:

1. Solve: $P_D A\tilde{x} = P_D b$.
2. Update: $x = Q_D \tilde{x} + Qb$.

Of course, the deflated system (3.4) can be solved with the help of a preconditioner M^{-1} resulting in

$$M^{-1} P_D Ax = M^{-1} P_D b. \tag{3.5}$$

Deflation can also be motivated by a splitting of x into

$$x = (I - Q_D)x + Q_D x. \tag{3.6}$$

But $Q_D = P_{A^{-1}\mathcal{Y}^\perp,\mathcal{Z}}$ and thus we have

$$(I - Q_D)x \in \mathcal{Z} \quad \text{and} \quad Q_D x \in A^{-1}\mathcal{Y}^\perp.$$

Since $(I - Q_D)x = Qb$ the first part of (3.6) can be obtained easily. To obtain the second part we solve $P_D A\tilde{x} = P_D b$. One can then show that $Q_D \tilde{x} = Q_D x$ if A is nonsingular. This is stated in the following lemma.

LEMMA 3.2. *Let A be nonsingular. Let P_D and Q_D as given in* (3.1) *and* (3.2) *be well-defined. Let $\tilde{x}$ be a solution of $P_D A\tilde{x} = P_D b$ and let x be the solution of $Ax = b$. Then*

$$Q_D \tilde{x} = Q_D x.$$

Proof. We have

$$Q_D \tilde{x} = A^{-1} AQ_D \tilde{x} = A^{-1} P_D A\tilde{x} = A^{-1} P_D b = A^{-1} P_D Ax = A^{-1} AQ_D x = Q_D x. \quad \square$$

Another version of the deflation procedure uses a specific starting vector for solving the deflated system with a Krylov subspace method [13, 46]. We will show in the next section that both approaches are equivalent.

Next we analyze the spectrum of the deflated operator if an A-invariant subspace is used as a deflation subspace $A\mathcal{Z} = \mathcal{Z} = \mathcal{Y}$. We then have

$$P_D = I - AZ\left(Y^H AZ\right)^{-1} Y^H = I - Z\left(Z^H Z\right)^{-1} Z^H = P_{\mathcal{Z}^\perp}.$$

The following theorem was presented in [26, Theorem 3.3].

Theorem 3.3. *Let $A \in \mathbb{C}^{n\times n}$ have the Jordan decomposition*

$$\begin{aligned} A &= SJS^{-1} \\ &:= \left[\begin{array}{cc} S_1 & S_2 \end{array}\right] \left[\begin{array}{cc} J_1 & 0 \\ 0 & J_2 \end{array}\right] \left[\begin{array}{c} \widehat{S}_1^H \\ \widehat{S}_2^H \end{array}\right], \end{aligned}$$

where $S_1, \widehat{S}_1 \in \mathbb{C}^{n\times r}$, $S_2, \widehat{S}_2 \in \mathbb{C}^{n\times n-r}$, $J_1 \in \mathbb{C}^{r\times r}$, $J_2 \in \mathbb{C}^{n-r\times n-r}$ for $r > 0$ and $[\,\widehat{S}_1 \quad \widehat{S}_2\,]^H = [\,S_1 \quad S_2\,]^{-1}$. If $\mathcal{Z} = \mathcal{Y} = \mathcal{R}(S_1)$ and J_1 is nonsingular, then

(a) $P_D = P_{\mathcal{Z}^\perp}$,

(b) *the deflated operator $P_D A$ has the Jordan decomposition*

$$P_D A = \left[\begin{array}{cc} S_1 & P_D S_2 \end{array}\right] \left[\begin{array}{cc} 0 & 0 \\ 0 & J_2 \end{array}\right] \left[\begin{array}{cc} S_1 & P_D S_2 \end{array}\right]^{-1}$$

$$\textit{with} \quad \left[\begin{array}{cc} S_1 & P_D S_2 \end{array}\right]^{-1} = \left[\begin{array}{cc} S_1(S_1^H S_1)^{-1} & \widehat{S}_2 \end{array}\right]^H,$$

(c) $\sigma(P_D A) = \{0\} \cup \sigma(J_2)$.

Thus the above theorem justifies the name deflation. If one chooses an A-invariant subspace as the deflation subspace, the corresponding eigenvalues are shifted to zero, while the other eigenvalues remain the same. If the deflation subspaces are arbitrary subspaces of dimension r the spectrum of the deflated operator $P_D A$ contains at least r zeros.

Hence, the system that has to be solved by a Krylov subspace method is singular, and the convergence of the Krylov subspace method is not necessarily guaranteed. Properties of GMRES applied to singular systems have been studied by many authors, e.g., [6, 11, 14, 26, 30]. Summarizing, we have the following theorem.

Theorem 3.4. *Let $Bx = b$ be a consistent linear system with $B \in \mathbb{C}^{n\times n}$ (possibly singular) and $b \in \mathbb{C}^n$. Then the following two conditions are equivalent:*

(a) *For every starting vector $x_0 \in \mathbb{C}^n$, the GMRES method applied to the linear system $Bx = b$ converges to a solution.*

(b) $\mathcal{R}(B) \cap \mathcal{N}(B) = \{0\}$.

Now consider the preconditioned deflated linear system

$$M^{-1} P_D A x = M^{-1} P_D b, \tag{3.7}$$

where we assume that $A\mathcal{Z} \oplus \mathcal{Y}^\perp$ such that P_D is well-defined. We have the following result.

Theorem 3.5. *Let $Ax = b$ with a nonsingular $A \in \mathbb{C}^{n\times n}$ and $b \in \mathbb{C}^n$. For subspaces $\mathcal{Z}, \mathcal{Y} \subseteq \mathbb{C}^n$ of dimension r with $A\mathcal{Z} \oplus \mathcal{Y}^\perp = \mathbb{C}^n$ the following statements are equivalent:*

(a) *For every starting vector x_0, the GMRES method applied to the linear system (3.7) is convergent.*

(b) $\mathcal{Z} \cap M^{-1}\mathcal{Y}^\perp = \{0\}$.

(c) $\mathbb{C}^n = \mathcal{Z} \oplus M^{-1}\mathcal{Y}^\perp$.

Proof. The equivalence of condition (b) in Theorem 3.4 with condition (b) of Theorem 3.5 becomes apparent by calculating the range and null space of the operator $M^{-1}P_DA$:

$$\mathcal{R}(M^{-1}P_DA) = \mathcal{R}(M^{-1}P_{\mathcal{Y}^\perp,A\mathcal{Z}}A) = M^{-1}\mathcal{Y}^\perp,$$
$$\mathcal{N}(M^{-1}P_DA) = \mathcal{N}(M^{-1}P_{\mathcal{Y}^\perp,A\mathcal{Z}}A) = \mathcal{Z}.$$

The equivalence of (b) with (c) follows with Theorem 2.2. ∎

Thus, to guarantee convergence of GMRES applied to the deflated system (3.7) we need some additional condition on $\mathcal{Z}$ and $\mathcal{Y}$, namely

$$\mathcal{Z} \oplus M^{-1}\mathcal{Y}^\perp = \mathbb{C}^n. \tag{3.8}$$

Note, that (3.8) is equivalent with

$$M\mathcal{Z} \oplus \mathcal{Y}^\perp = \mathbb{C}^n.$$

If we choose $\mathcal{Y} = A\mathcal{Z}$ and $M = I$ this is exactly the condition for P_D to be well-defined.

The next example shows the influence of condition (3.8) on the operator $M^{-1}P_DA$. Additional zeros may occur in the spectrum of $M^{-1}P_DA$ if (3.8) is not satisfied.

Example 1. Let

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad M = I, \quad Y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and let $\mathcal{Y} = \mathcal{R}(Y)$ and $\mathcal{Z} = \mathcal{R}(Z)$. Hence, $\mathcal{Y}^\perp = \mathcal{Z}$ and we obtain

$$P_D = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}, \quad P_DA = \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix} \quad \text{with} \quad \sigma(P_DA) = \{0\}.$$

Since $\mathcal{Y}$ and $\mathcal{Z}$ are just one-dimensional subspaces the theory given so far guarantees the eigenvalue zero only with multiplicity one in the spectrum of $M^{-1}P_DA$.

We will comment on (3.8) later on. Indeed, we will show that for r-dimensional subspaces $\mathcal{Y}$ and $\mathcal{Z}$, (3.8) is a necessary and sufficient condition to have exactly r zeros in the spectrum of $M^{-1}P_DA$.

Summarizing we have our next result.

THEOREM 3.6. *Let $A, M \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Z}$ and $\mathcal{Y}$ be two subspaces of $\mathbb{C}^n$ of dimension r with matrices Z and Y whose columns form bases of $\mathcal{Z}$ and $\mathcal{Y}$. If*

$$A\mathcal{Z} \oplus \mathcal{Y}^\perp = M\mathcal{Z} \oplus \mathcal{Y}^\perp = \mathbb{C}^n,$$

then

- *the GMRES method applied to the linear system*

$$M^{-1}P_DAx = M^{-1}P_Db$$

  is convergent for all starting vectors,
- *letting $\tilde{x}_i$ be the approximation of the GMRES method applied to $M^{-1}P_DAx = M^{-1}P_Db$, then*

$$x_i = Q_D\tilde{x}_i + Qb$$

  satisfies

– *if* $r_i := b - Ax_i$ *and* $\tilde{r}_i := P_D(b - A\tilde{x}_i)$*, then* $r_i = \tilde{r}_i$*,*
– $P_D\tilde{r}_i = \tilde{r}_i$*,*
– $P_D r_i = r_i$*.*

Proof. The first items follow immediately from the above theorems. Moreover we have

$$\begin{aligned} r_i &= b - Ax_i \\ &= b - (AQ_D\tilde{x}_i + AQb) \\ &= b - P_D A\tilde{x}_i - A(I - Q_D)x \\ &= AQ_D x - P_D A\tilde{x}_i \\ &= P_D(Ax - A\tilde{x}_i) \\ &= P_D(b - A\tilde{x}_i) \\ &= \tilde{r}_i. \end{aligned}$$

The last items can be obtained from the above equations since P_D is a projection. □

In the next theorems we characterize the spectrum of the deflated system. Therefore we will use orthonormal bases of $\mathcal{Y}$ and $\mathcal{Z}$ and their orthogonal complements. Since the projections are independent of the choice of the bases, the use of orthonormal bases is no additional requirement.

For the two subsets $\mathcal{Y}$ and $\mathcal{Z}$ of $\mathbb{C}^n$ we will use matrices $Y, \tilde{Y}, Z$, and $\tilde{Z}$ such that

- the columns of Y form an orthonormal basis of $\mathcal{Y}$,
- the columns of $\tilde{Y}$ form an orthonormal basis of $\mathcal{Y}^\perp$,
- the columns of Z form an orthonormal basis of $\mathcal{Z}$,
- the columns of $\tilde{Z}$ form an orthonormal basis of $\mathcal{Z}^\perp$.

Note that the matrices $[\tilde{Y}, Y]$ and $[\tilde{Z}, Z]$ are then unitary. We have the following properties of the nonsymmetric deflation operator.

THEOREM 3.7. *Let* $A \in \mathbb{C}^{n\times n}$ *be nonsingular. Let* $\mathcal{Y}$ *and* $\mathcal{Z}$ *be subspaces of* $\mathbb{C}^n$ *of dimension* r *such that* $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$*. Then*

- *the matrix* $[\tilde{Y}, Y]^H P_D A[\tilde{Z}, Z]$ *has the block* 2×2 *form*

$$[\tilde{Y}, Y]^H P_D A[\tilde{Z}, Z] = \begin{bmatrix} \tilde{Y}^H P_D A\tilde{Z} & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.9}$$

- $\tilde{Y}^H P_D A\tilde{Z}$ *is nonsingular with*

$$\tilde{Y}^H P_D A\tilde{Z} = (\tilde{Z}^H A^{-1}\tilde{Y})^{-1}. \tag{3.10}$$

Proof. We easily obtain the first item, since $P_D = P_{\mathcal{Y}^\perp, A\mathcal{Z}}$. Now consider $(P_D A)$ $(P_{\mathcal{Z}^\perp}A^{-1})$. We have

$$(P_D A)(P_{\mathcal{Z}^\perp}A^{-1}) = P_D A(I - P_{\mathcal{Z}})A^{-1} = P_D A A^{-1} = P_D. \tag{3.11}$$

Thus, we obtain

$$\begin{aligned} &[\tilde{Y}, Y]^H P_D A P_{\mathcal{Z}^\perp} A^{-1}[\tilde{Y}, Y] \\ &= [\tilde{Y}, Y]^H P_D A[\tilde{Z}, Z][\tilde{Z}, Z]^H P_{\mathcal{Z}^\perp} A^{-1}[\tilde{Y}, Y] \\ &= \begin{bmatrix} \tilde{Y}^H P_D A\tilde{Z} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{Z}^H P_{\mathcal{Z}^\perp} A^{-1}\tilde{Y} & * \\ * & * \end{bmatrix} \\ &= [\tilde{Y}, Y]^H P_D[\tilde{Y}, Y] \\ &= \begin{bmatrix} I & * \\ * & * \end{bmatrix}. \end{aligned}$$

Here we have used (3.11) and that $P_D\tilde{Y} = \tilde{Y}$. Thus we obtain

$$\left(\tilde{Y}^H P_D A \tilde{Z}\right)\left(\tilde{Z}^H P_{\mathcal{Z}^\perp} A^{-1}\tilde{Y}\right) = I.$$

Hence

$$\tilde{Y}^H P_D A \tilde{Z} = \left(\tilde{Z}^H P_{\mathcal{Z}^\perp} A^{-1}\tilde{Y}\right)^{-1} = \left(\tilde{Z}^H A^{-1}\tilde{Y}\right)^{-1},$$

which completes the proof. ∎

Next we transfer one basis into the other, or in other words, we transform $[\tilde{Z}, Z]^H$ into $[\tilde{Y}, Y]^H$, i.e., we need to find a nonsingular matrix T such that

$$[\tilde{Z}, Z]^H = T[\tilde{Y}, Y]^H.$$

Since $[\tilde{Y}, Y]^H$ is a unitary matrix, we obtain

$$T = [\tilde{Z}, Z]^H[\tilde{Y}, Y] = \begin{bmatrix} \tilde{Z}^H\tilde{Y} & \tilde{Z}^H Y \\ Z^H\tilde{Y} & Z^H Y \end{bmatrix}. \tag{3.12}$$

As a consequence of the previous theorem, we obtain a characterization of the spectrum of $P_D A$.

THEOREM 3.8. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Y}$ and $\mathcal{Z}$ be subspaces of $\mathbb{C}^n$ of dimension r such that $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. Let $\tilde{Y}$ and $\tilde{Z}$ be matrices whose columns form bases of $\mathcal{Y}^\perp$ and $\mathcal{Z}^\perp$, respectively. Then*

$$\sigma(P_D A) = \{0\} \cup \sigma\left(\tilde{Z}^H\tilde{Y}(\tilde{Z}^H A^{-1}\tilde{Y})^{-1}\right).$$

Proof. We obtain with Theorem 3.7 and (3.12)

$$\begin{aligned}[\tilde{Z}, Z]^H P_D A[\tilde{Z}, Z] &= T[\tilde{Y}, Y]^H P_D A[\tilde{Z}, Z] \\ &= \begin{bmatrix} \tilde{Z}^H\tilde{Y} & \tilde{Z}^H Y \\ Z^H\tilde{Y} & Z^H Y \end{bmatrix}\begin{bmatrix} (\tilde{Z}^H A^{-1}\tilde{Y})^{-1} & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \tilde{Z}^H\tilde{Y}(\tilde{Z}^H A^{-1}\tilde{Y})^{-1} & 0 \\ Z^H\tilde{Y}(\tilde{Z}^H A^{-1}\tilde{Y})^{-1} & 0 \end{bmatrix},\end{aligned}$$

which gives the desired result. ∎

If the matrix $\tilde{Z}^H\tilde{Y}$ is nonsingular, the product $\tilde{Z}^H\tilde{Y}(\tilde{Z}^H A^{-1}\tilde{Y})^{-1}$ is also nonsingular. Thus, the spectrum of $P_D A$ contains exactly r zeros, i.e., the algebraic multiplicity of the eigenvalue zero is r. The nonsingularity of $\tilde{Z}^H\tilde{Y}$ is equivalent to $\mathcal{Z}^\perp \oplus \mathcal{Y} = \mathcal{Z} \oplus \mathcal{Y}^\perp = \mathbb{C}^n$ (see Theorem 2.2). This is exactly the condition we have discussed above. We have then the following corollary.

COROLLARY 3.9. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Y}$ and $\mathcal{Z}$ be subspaces of $\mathbb{C}^n$ of dimension r such that $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. Then the spectrum of $P_D A$ has the eigenvalue zero with exact multiplicity r if and only if $\mathcal{Y}^\perp \oplus \mathcal{Z} = \mathbb{C}^n$, i.e., if and only if $\tilde{Z}^H\tilde{Y}$ is nonsingular.*

Moreover, if in Theorem 3.8 we set $\mathcal{Z} = \mathcal{Y}$ and choose the same orthonormal basis, then we have $\tilde{Z}^H\tilde{Y} = I$ and we get the same result as for symmetric matrices.

The application of the deflation operator P_D requires solving a linear system for $Y^H AZ$. In practice, and especially when the deflation subspace has large dimension, this system has to be solved with great care. If it is solved inexactly, this may destroy the zero structure and can introduce very small eigenvalues in the spectrum of the deflated matrix $P_D A$. A possible remedy is to shift some eigenvalues away from zero. With this goal in mind in [18, 63] an additional term is added to the deflation operator resulting in the preconditioned adapted deflation operator P_{ADEF} given by

$$\begin{aligned} P_{ADEF} &= M^{-1}\left(I - AZ\left(Y^H AZ\right)^{-1} Y^H\right) + Z\left(Y^H AZ\right)^{-1} Y^H \\ &= M^{-1} P_D + Q \\ &= M^{-1} P_{\mathcal{Y}^\perp, A\mathcal{Z}} + A^{-1} P_{A\mathcal{Z}, \mathcal{Y}^\perp} \end{aligned} \tag{3.13}$$

for a nonsingular matrix M^{-1}. The operator P_{ADEF} is well-defined if $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. We then easily obtain

$$P_{ADEF} A = M^{-1} A (I - QA) + QA,$$

or in the projection notation

$$\begin{aligned} P_{ADEF} A &= M^{-1} P_{\mathcal{Y}^\perp, A\mathcal{Z}} A + A^{-1} P_{A\mathcal{Z}, \mathcal{Y}^\perp} A \\ &= M^{-1} A (I - P_{\mathcal{Z}, (A^H \mathcal{Y})^\perp}) + P_{\mathcal{Z}, (A^H \mathcal{Y})^\perp} \\ &= M^{-1} A P_{(A^H \mathcal{Y})^\perp, \mathcal{Z}} + P_{\mathcal{Z}, (A^H \mathcal{Y})^\perp} \\ &= M^{-1} A P_{A^{-1}\mathcal{Y}^\perp, \mathcal{Z}} + P_{\mathcal{Z}, A^{-1}\mathcal{Y}^\perp}. \end{aligned} \tag{3.14}$$

In section 4 we describe the relation between P_{ADEF} and preconditioners resulting from multigrid methods. Equation (3.14) shows that the preconditioned system $P_{ADEF} A$ is of the form given in (2.1), since the matrix $M^{-1}A$ is multiplied by a projection and the complementary projection is added to the product. Thus, Theorem 2.4 can be used to obtain necessary and sufficient conditions for the nonsingularity of P_{ADEF}.

THEOREM 3.10. *Let $A, M \in \mathbb{C}^{n \times n}$ be nonsingular. Let $\mathcal{Y}$ and $\mathcal{Z}$ be subspaces of $\mathbb{C}^n$ of dimension r such that $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. Let Y and Z be matrices whose columns form bases of $\mathcal{Y}$ and $\mathcal{Z}$, respectively. Then*

$$P_{ADEF} = M^{-1} P_{\mathcal{Y}^\perp, A\mathcal{Z}} + A^{-1} P_{A\mathcal{Z}, \mathcal{Y}^\perp}$$

is nonsingular if and only if $\mathcal{Y}^\perp \cap M\mathcal{Z} = \{0\}$, i.e., if and only if $\mathcal{Y}^\perp \oplus M\mathcal{Z} = \mathbb{C}^n$, i.e., if and only if $Y^H M Z$ is nonsingular.

Proof. We apply Theorem 2.4 with

$$C = M^{-1}A, \quad \mathcal{W}^\perp = A^{-1}\mathcal{Y}^\perp, \quad \mathcal{V} = \mathcal{Z}.$$

Therefore, $C\mathcal{W}^\perp = M^{-1}\mathcal{Y}^\perp$. Hence, $P_{ADEF} A$ and so P_{ADEF} is nonsingular if and only if $\mathcal{Y}^\perp \cap M\mathcal{Z} = \{0\}$, i.e., if and only if $Y^H M Z$ is nonsingular. ∎

The above Theorem 3.10 is a special case of a fact listed in [48]. However, the relation to projections is new. Moreover, the proof given here seems to be shorter than the one given in [48]. The statement of Theorem 3.10 corrects Theorem 2.9 of [17]. The condition $\mathcal{Y}^\perp \oplus M\mathcal{Z} = \mathbb{C}^n$ is missing there.

If $Y^H M Z$ is nonsingular, then the explicit form of P_{ADEF}^{-1} can also be given using results of the previous section.

THEOREM 3.11. *Let $A, M \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Y}$ and $\mathcal{Z}$ be subspaces of $\mathbb{C}^n$ of dimension r such that $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. Let Y and Z be matrices whose columns form bases of $\mathcal{Y}$ and $\mathcal{Z}$, respectively. If Y^HMZ is nonsingular, then P_{ADEF} is nonsingular and*

$$P_{ADEF}^{-1} = AP_{\mathcal{Z},M^{-1}\mathcal{Y}^\perp} + M(I - P_{\mathcal{Z},M^{-1}\mathcal{Y}^\perp}),$$

which is in matrix notation

$$P_{ADEF}^{-1} = AQ_MM + M(I - Q_MM), \tag{3.15}$$

where

$$Q_M = Z(Y^HMZ)^{-1}Y^H.$$

Proof. The statement follows immediately from Theorem 2.5 with $C = M^{-1}A$, $\mathcal{W}^\perp = A^{-1}\mathcal{Y}^\perp$, and $\mathcal{V} = \mathcal{Z}$. ◻

Note that we have $QMQ_M = Q$ and $QAQ_M = Q_M$. Thus, (3.15) can be proved by a simple calculation

$$(M^{-1}(I - AQ) + Q)(M(I - Q_MM) + AQ_MM) = I.$$

As a special case we obtain for $M = I$

$$P_A := I - AZ\left(Y^HAZ\right)^{-1}Y + Z(Y^HAZ)^{-1}Y = P_D + Q = P_{\mathcal{Y}^\perp,A\mathcal{Z}} + A^{-1}P_{A\mathcal{Z},\mathcal{Y}^\perp}.$$

Now, if $\mathcal{Z} \oplus \mathcal{Y}^\perp = \mathbb{C}^n$ the inverse of P_A is given by

$$P_A^{-1} = P_{\mathcal{Y}^\perp,\mathcal{Z}} + AP_{\mathcal{Z},\mathcal{Y}^\perp}.$$

Theorem 3.11 is also listed in different form in [48]. But again the relation to projections is new and the proof is much shorter.

Next, we characterize the spectrum of $M^{-1}P_DA$ and $P_{ADEF}A$.

THEOREM 3.12. *Let $A, M \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Z}$ and $\mathcal{Y}$ be two subspaces of $\mathbb{C}^n$ with matrices Z and Y whose columns form bases of $\mathcal{Z}$ and $\mathcal{Y}$, respectively. Let $\tilde{Y}$ and $\tilde{Z}$ be matrices whose columns form bases of $\mathcal{Y}^\perp$ and $\mathcal{Z}^\perp$, respectively. Assume that*

$$A\mathcal{Z} \oplus \mathcal{Y}^\perp = M\mathcal{Z} \oplus \mathcal{Y}^\perp = \mathbb{C}^n.$$

Let P_D and P_{ADEF} be defined as in (3.1) and (3.13). Then we have the following:

1. *$P_{ADEF}A$ has the eigenvalue one with multiplicity r; the other eigenvalues are nonzero and are denoted by $\mu_{r+1}, \ldots, \mu_n$, i.e.,*

$$\sigma(P_{ADEF}A) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}.$$

2. *$M^{-1}P_DA$ has the eigenvalue zero with multiplicity r; the other eigenvalues are $\mu_{r+1}, \ldots, \mu_n$, i.e.,*

$$\sigma(M^{-1}P_DA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}.$$

3. *The matrix $A^{-1}M(I - Q_MM) + Q_MM$ has the eigenvalue one with multiplicity r; the other eigenvalues are $\frac{1}{\mu_{r+1}}, \ldots, \frac{1}{\mu_n}$, i.e.,*

$$\sigma(A^{-1}M(I - Q_MM) + Q_M) = \left\{1, \ldots, 1, \frac{1}{\mu_{r+1}}, \ldots, \frac{1}{\mu_n}.\right\}.$$

4. *The matrix* $A^{-1}M(I - Q_M M)$ *has the eigenvalue zero with multiplicity* r *the other eigenvalues are* $\frac{1}{\mu_{r+1}}, \dots, \frac{1}{\mu_n}$, *i.e.,*
$$\sigma(A^{-1}M(I - Q_M M)) = \left\{0, \dots, 0, \frac{1}{\mu_{r+1}}, \dots, \frac{1}{\mu_n}\right\}.$$
5. *The matrix* $\tilde{Z}^H A^{-1}\tilde{Y}$ *is nonsingular and*
$$\sigma\left(\left(\tilde{Z}^H A^{-1}\tilde{Y}\right)^{-1} \tilde{Z}^H M^{-1}\tilde{Y}\right) = \{\mu_{r+1}, \dots, \mu_n\}.$$

Proof. Statements 1 and 2 follow immediately from Theorem 2.3 and (3.1) and (3.13). For 3 consider the eigenvalues μ_i of the nonsingular matrix $P_{ADEF}A$. Then $(\mu_i)^{-1}$ is an eigenvalue of $A^{-1}P_{ADEF}^{-1}$. Using the explicit formula for P_{ADEF}^{-1} given in Theorem 3.11 leads to statement 3.

For 4, consider
$$\begin{aligned} A^{-1}P_{ADEF}^{-1} &= A^{-1}(M - (M - A)Q_M) = A^{-1}\left(MP_{(M^H\mathcal{Y})^\perp,\mathcal{Z}} + AP_{\mathcal{Z},(M^H\mathcal{Y})^\perp}\right) \\ &= A^{-1}MP_{(M^H\mathcal{Y})^\perp,\mathcal{Z}} + P_{\mathcal{Z},(M^H\mathcal{Y})^\perp}. \end{aligned}$$
Statement 4 follows then immediately from Theorem 2.3 and statement 3.

For statement 5, consider $M^{-1}P_D A$. We have
$$\sigma(M^{-1}P_D A) = \sigma(M^{-1}AQ_D) = \sigma(Q_D M^{-1}A).$$
Next consider the nonzero eigenvalues of $Q_D M^{-1}A$. With the Jordan decomposition of $Q_D M^{-1}A$ and the nonsingularity of P_{ADEF}, there is a nonsingular matrix $J \in \mathbb{C}^{n-r\times n-r}$ and a matrix $S \in \mathbb{C}^{n\times n-r}$ of rank $n - r$ such that
$$Q_D M^{-1}AS = SJ,$$
and the spectrum of J is the same as the spectrum of $Q_D M^{-1}A$ without the zeros. Next, let $W = AS$. Then W is of rank $n - r$ and
$$Q_D M^{-1}W = A^{-1}WJ. \tag{3.16}$$
Multiplying the above with $Y^H A$ gives
$$Y^H WJ = Y^H AQ_D M^{-1}W = 0$$
since $Y^H AQ_D = 0$. But J is nonsingular, thus $Y^H W = 0$. Hence, the columns of W form a basis of $\mathcal{Y}^\perp$. Transforming this basis into $\tilde{Y}$ leads to a nonsingular matrix R such that $W = \tilde{Y}R$. Using this in (3.16) and multiplying it with $\tilde{Z}^H$ leads to
$$\tilde{Z}^H Q_D M^{-1}\tilde{Y}R = \tilde{Z}^H A^{-1}\tilde{Y}RJ.$$
But $\tilde{Z}^H Q_D = \tilde{Z}^H$. So we obtain
$$\tilde{Z}^H M^{-1}\tilde{Y}R = \tilde{Z}^H A^{-1}\tilde{Y}RJ.$$
With (3.10), $\tilde{Z}^H A^{-1}\tilde{Y}$ is nonsingular. Hence
$$(\tilde{Z}^H A^{-1}\tilde{Y})^{-1}\tilde{Z}^H M^{-1}\tilde{Y}R = RJ,$$
and
$$\sigma\left((\tilde{Z}^H A^{-1}\tilde{Y})^{-1}\tilde{Z}^H M^{-1}\tilde{Y}\right) = \sigma\left(R^{-1}(\tilde{Z}^H A^{-1}\tilde{Y})^{-1}\tilde{Z}^H M^{-1}\tilde{Y}R\right) = \sigma(J),$$
which gives the desired result. ∎

Summarizing we have the following corollary.

COROLLARY 3.13. *Let $A, M \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Y}$ and $\mathcal{Z}$ be subspaces of $\mathbb{C}^n$ of dimension r such that $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. Then the spectrum of $M^{-1}P_DA$ has the eigenvalue zero with exact multiplicity r if and only if $\mathcal{Y}^\perp \oplus M\mathcal{Z} = \mathbb{C}^n$, i.e., if and only if Y^HMZ is nonsingular. Moreover, if $\tilde{Y}$ and $\tilde{Z}$ are matrices whose columns form bases of $\mathcal{Y}^\perp$ and $\mathcal{Z}^\perp$, respectively, then $\tilde{Z}^HA^{-1}\tilde{Y}$ is nonsingular and*

$$\sigma(M^{-1}P_DA) = \{0\} \cup \sigma\left((\tilde{Z}^HA^{-1}\tilde{Y})^{-1}\tilde{Z}^HM^{-1}\tilde{Y}\right),$$
$$\sigma(P_{ADEF}A) = \{1\} \cup \sigma\left((\tilde{Z}^HA^{-1}\tilde{Y})^{-1}\tilde{Z}^HM^{-1}\tilde{Y}\right).$$

4. Nonsymmetric deflation and augmentation. Augmentation and deflation of Krylov subspace methods have been proposed in various forms in a large number of publications. As mentioned in [28], many of the methods differ not only algorithmically and numerically but also mathematically. Related names of such methods are spectral deflation, projected methods, augmented methods, recycling techniques, spectral preconditioner, and singular preconditioner. In the previous section we introduced the deflation technique as consisting of two steps. First, a singular system is solved, and next, a solution of this system is updated or corrected to get the unique solution of the original system in a final correction step. This approach follows the approach given in one of first publications on deflation for symmetric positive definite systems [36]. However, in the earlier papers [13] and [46], deflation is introduced in another way. There a certain starting vector is used and then the CG method is applied to a projected system using an A-orthogonal projection. A similar algorithm was later discovered in [16] and in [53]. This approach can be seen more like an augmentation of the Krylov subspace from the first iteration rather than as a deflation approach. In [53] the first link between augmented and deflated Krylov subspace methods seems to be made by showing that the deflated CG algorithm is a generalization of the augmented CG algorithm given in [16] to arbitrary augmentation spaces.

The augmentation approach, i.e., using specific starting vectors, was used in [39] and also in [8] in order to overcome recurring convergence slowdowns after restarting the GMRES method (see also [40, 41, 42]). In [11] an inner-outer GMRES/GCR that uses augmentation is introduced. Later, a general analysis of Krylov subspace methods with augmented basis is given in [51]. This analysis is extended in the survey article [14]. But the augmentation or deflation technique is also combined with other Krylov subspace methods, not only with the CG and GMRES method. In [64] a recycled MINRES or RMINRES method based on augmentation of the Krylov basis is introduced. In [3] a recycling BiCG method is established, followed by a recycling BiCGSTAB method [2]. Both methods use specific starting vectors for augmentation.

Of course the choice of the deflation subspace, or the subspace which augments the Krylov basis, is important for the convergence of the Krylov subspace methods. Originally eigenvectors were used to build the deflation space or augment the Krylov space, but already in [38] vectors based on a coarse grid interpolation are used. In the restarted GMRES setup in [39] and [8] a few harmonic Ritz vectors which correspond to the harmonic Ritz values of smallest magnitude augment the Krylov basis. Recently, POD-based deflation vectors have been used for the deflated CG method [7, 12]. Moreover, the effect of the distance of arbitrary deflation vectors on eigenvectors is analyzed in, e.g., [14, 51, 56]. For sequences of (nonsymmetric) linear systems, augmentation and deflation became a useful technique. Information from previous linear systems is used to augment the Krylov subspace to compute the actual solution more efficiently. This technique is known as recycling [2, 3, 4, 7, 16, 32, 50, 57, 58, 64] and uses again specific start vectors. Recently, the deflation and augmentation technique has been used successfully for block Krylov methods [42, 43, 59, 60, 61, 65].

Here we concentrate on the deflated or augmented GMRES method and, since we can use similar arguments, on the deflated CG method. As in the previous sections, we use two arbitrary deflation spaces, which allow the use of oblique projections. We do not comment on a specific choice of the deflation spaces here. We complement the analysis which is started in [26] and is continued in [28, 29]. Our main result in this section states that both deflation and augmentation variants, i.e., on one hand first solve a projected system and then do a final correction or on the other hand use a specific starting vector first and then solve a projected system, are mathematically equivalent.

Before we prove our main results of this section we first give a short review of projection methods. The idea of general projection methods for solving linear systems of equations is to approximate the solution in a sequence of subspaces with increasing dimensions. In the ith iteration step we use two subspaces $\mathcal{S}_i$ and $\mathcal{C}_i$ of dimension i. The subspace $\mathcal{S}_i$ acts as a so-called search space. For a given starting vector, the approximation x_i is chosen out of $x_0 + \mathcal{S}_i$, i.e.,

$$x_i \in x_0 + \mathcal{S}_i. \tag{4.1}$$

The space $\mathcal{C}_i$ acts as a *constraint space*. The approximation x_i is chosen such that

$$b - Ax_i =: r_i \perp \mathcal{C}_i. \tag{4.2}$$

The method described by (4.1) and (4.2) is called a *Galerkin method* if $\mathcal{S}_i = \mathcal{C}_i$ and a *Petrov–Galerkin method* otherwise. A (Petrov–) Galerkin method is called *well-defined* if there exists a unique approximate solution x_i for each i. If $S_i, C_i \in \mathbb{C}^{n\times i}$ are matrices whose columns form bases of $\mathcal{S}_i$ and $\mathcal{C}_i$, respectively, then the Petrov–Galerkin method is well-defined if and only if $C_i^H AS_i$ is nonsingular for all i. If the $\mathcal{S}_i = \mathcal{C}_i$ is chosen as a sequence of nested Krylov subspaces and A is symmetric positive definite, the projection method describes the CG method. If $\mathcal{C}_i = A\mathcal{S}_i$ is chosen as a sequence of nested Krylov subspaces, the projection method is the GMRES method. Note that in this case ($\mathcal{C}_i = A\mathcal{S}_i$) for a singular matrix A, the condition $\mathcal{N}(A) \cap \mathcal{S}_i = \{0\}$ leads to a nonsingular matrix $C_i^H AS_i$. For more details on projection methods see [35, 52].

The deflation version which uses a specific start vector corresponds to an augmentation method. The main focus is to enlarge the Krylov subspace from the beginning of the iteration by another subspace, i.e., by $\mathcal{Z}$. However, the Krylov space which is used is not $K_i(A, r_0)$ but $K_i(P_D A, P_D r_0)$, where $K_i(B, r_0)$ is the ith Krylov subspace, i.e., $K_i(B, r_0) = \operatorname{span}\{B^{i-1}r_0, \ldots, Br_0, r_0\}$ for $B \in \mathbb{C}^{n\times n}$.

The augmented version can be described by the following Petrov–Galerkin formulation for $i > 0$:

$$\begin{aligned} x_i &\in x_0 + K_i(P_D A, P_D r_0) + \mathcal{Z}, \\ r_i &= b - Ax_i \perp P_D A K_i(P_D A, P_D r_0) + \mathcal{Y}, \end{aligned} \tag{4.3}$$

where the starting vector x_0 has to be found to ensure that $r_0 \perp \mathcal{Y}$.

The solution of (4.3) can be described by the following theorem.

THEOREM 4.1. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Let $\mathcal{Y}$ and $\mathcal{Z}$ be two subspaces of $\mathbb{C}^n$ of dimension r such that $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$. Let Y and Z be matrices whose columns* *form bases of* $\mathcal{Y}$ *and* $\mathcal{Z}$*, respectively. Let* P_D *and* Q_D *be defined as in* (3.1) *and* (3.2)*. For every* $x_0 \in \mathbb{C}^n$ *and* $r_0 = b - Ax_0$*, the two pairs of conditions*

$$\begin{aligned} &\hat{x}_i \in x_0 + K_i(P_D A, P_D r_0) + \mathcal{Z} \\ &\hat{r}_i = b - A\hat{x}_i \perp P_D A K_i(P_D A, P_D r_0) + \mathcal{Y} \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} &\tilde{x}_i \in x_0 + K_i(P_D A, P_D r_0), \\ &\tilde{r}_i = P_D(b - A\tilde{x}_i) \perp P_D A K_i(P_D A, P_D r_0) \end{aligned} \tag{4.5}$$

are equivalent for $i \geq 1$ *in the sense that*

$$\hat{x}_i = Q_D \tilde{x}_i + Qb, \quad \text{and} \quad \hat{r}_i = \tilde{r}_i, \tag{4.6}$$

if $\tilde{x}_i$ *is uniquely determined.*

Proof. The proof is almost the same as the proof given in [26] for the case $\mathcal{Y} = A\mathcal{Z}$. It is shown that every solution of (4.4) gives a solution of (4.5). But due to the uniqueness assumption the reverse holds also, leading to a unique $\hat{x}_i$. ∎

With Theorem 4.1 we have the following corollary.

COROLLARY 4.2. *Let* $A \in \mathbb{C}^{n\times n}$ *be nonsingular. Let* $\mathcal{Y}$ *and* $\mathcal{Z}$ *be two subspaces of* $\mathbb{C}^n$ *of dimension* r *such that* $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathbb{C}^n$*. Let* Y *and* Z *be matrices whose columns form bases of* $\mathcal{Y}$ *and* $\mathcal{Z}$*, respectively. Let* P_D *and* Q_D *be defined as* (3.1) *and* (3.2)*. For an arbitrary* $x_0 \in \mathbb{C}^n$ *let* $r_0 = b - Ax_0$*,* $x_0^* = Q_D x_0 + Qb$*, and* $r_0^* = b - Ax_0^*$*. Then the four pairs of conditions*

$$\begin{aligned} &x_i^* \in x_0^* + K_i(P_D A, P_D r_0^*) + \mathcal{Z} \\ &r_i^* = b - Ax_i^* \perp P_D A K_i(P_D A, P_D r_0^*) + \mathcal{Y}, \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} &\bar{x}_i \in x_0^* + K_i(P_D A, P_D r_0^*), \\ &\bar{r}_i = P_D(b - A\bar{x}_i) \perp P_D A K_i(P_D A, P_D r_0^*), \end{aligned} \tag{4.8}$$

and

$$\begin{aligned} &\hat{x}_i \in x_0 + K_i(P_D A, P_D r_0) + \mathcal{Z} \\ &\hat{r}_i = b - A\hat{x}_i \perp P_D A K_i(P_D A, P_D r_0) + \mathcal{Y}, \end{aligned} \tag{4.9}$$

and

$$\begin{aligned} &\tilde{x}_i \in x_0 + K_i(P_D A, P_D r_0), \\ &\tilde{r}_i = P_D(b - A\tilde{x}_i) \perp P_D A K_i(P_D A, P_D r_0) \end{aligned} \tag{4.10}$$

are equivalent for $i \geq 1$ *in the sense that*

$$x_i^* = \hat{x}_i = Q_D \tilde{x}_i + Qb = Q_D \bar{x}_i + Qb \quad \text{and} \quad r_i^* = \hat{r}_i = \tilde{r}_i = \bar{r}_i \tag{4.11}$$

if $\tilde{x}_i$ *and* $\bar{x}_i$ *are uniquely determined.*

Proof. With Theorem 4.1 we have that conditions (4.7) and (4.8) as well as conditions (4.9) and (4.10) are equivalent. But Theorem 4.1 gives equivalent statements for every starting vector x_0. We have

$$\begin{aligned} x_0 - x_0^* &= x_0 - Q_D x_0 - Qb = x_0 - (I - QA)x_0 - Qb = Q(Ax_0 - b) \\ &= Z\left(Y^H AZ\right)^{-1} Y^H (Ax_0 - b) \in \mathcal{R}(Z). \end{aligned}$$

Moreover,

$$r_0^* = P_D\left(b - AP_D^T x_0 + AQb\right) = P_D(b - P_D Ax_0) + P_D AQb = P_D(b - Ax_0) = \hat{r}_0.$$

Hence, all the search and constraint spaces are the same. Thus

$$\begin{aligned} \hat{x}_i &\in x_0^* + P_D AK_i(P_D A, P_D r_0^*) + \mathcal{Z}, \\ x_i^* &\in x_0 + P_D AK_i(P_D A, P_D r_0) + \mathcal{Z}, \end{aligned}$$

which completes the proof. ∎

Note that conditions (4.7) and conditions (4.9) are Petrov–Galerkin formulations. But in this form it is not clear that GMRES applied to the linear system $P_D Ax = P_D b$ realizes this Petrov–Galerkin formulation. However, conditions (4.8) and conditions (4.10) are Petrov–Galerkin formulations that can be realized by the GMRES method applied to $P_D Ax = P_D b$. Since $P_D A$ is singular, we have to analyze whether a projection method applied to (4.8) is well-defined or not. Therefore it suffices to show that $\mathcal{S}_i \cap \mathcal{N}(P_D A) = \{0\}$, where $\mathcal{S}_i$ is the search space of the projection method that realizes the Petrov–Galerkin formulation (here the GMRES method). For GMRES we have $\mathcal{S}_i = K_i(P_D A, P_D r_0)$. Since $P_D = P_{\mathcal{Y}^\perp, A\mathcal{Z}}$ we have $S_i \subset \mathcal{Y}^\perp$ and $S_i \cap \mathcal{N}(P_D A) = S_i \cap \mathcal{Z}$. So if we assume that $\mathcal{Y}^\perp \cap \mathcal{Z} = \{0\}$, i.e., $\mathcal{Y}^\perp \oplus \mathcal{Z} = \mathbb{C}^n$, a projection method that realizes (4.8) is well-defined and this projection method is the GMRES method. Hence, it is not surprising that $\mathcal{Y}^\perp \cap \mathcal{Z} = \{0\}$ is just the condition which is needed for the convergence of GMRES applied to $P_D Ax = P_D b$.

Moreover, when the starting vectors are chosen, condition $\mathcal{Y}^\perp \cap \mathcal{Z} = \{0\}$ leads to uniquely defined approximations $\bar{x}_i$ and $\tilde{x}_i$, and thus by the equivalence, we obtain also unique approximations x_i^* and $\hat{x}_i$. This discussion leads to our next result.

THEOREM 4.3. *Let* $A \in \mathbb{C}^{n\times n}$ *be nonsingular. Let* $\mathcal{Y}$ *and* $\mathcal{Z}$ *be two subspaces of* $\mathbb{C}^n$ *of dimension* r *such that* $\mathcal{Y}^\perp \oplus A\mathcal{Z} = \mathcal{Y}^\perp \oplus \mathcal{Z} = \mathbb{C}^n$. *Let* Y *and* Z *be matrices whose columns form the basis of* $\mathcal{Y}$ *and* $\mathcal{Z}$, *respectively. Let* P_D *and* Q_D *be defined as in* (3.1) *and* (3.2). *For an arbitrary starting vector* $x_0 \in \mathbb{C}^n$ *we have that*

- *the approximations* x_i^* *given by the GMRES method applied to the linear system* $P_D Ax = P_D b$ *with starting vector* $x_0^* = Q_D x_0 + Qb$ *converge to the solution of* $Ax = b$,
- *the approximations* $\tilde{x}_i$ *given by the GMRES method applied to the linear system* $P_D Ax = P_D b$ *with starting vector* x_0 *converge to a solution of* $P_D Ax = P_D b$,
- *it holds that* $x_i^* = Q_D \tilde{x}_i + Qb$.

Proof. As seen above, the condition $\mathcal{Y}^\perp \oplus \mathcal{Z} = \mathbb{C}^n$ leads to approximations that are uniquely determined by one of the sets of conditions (4.7) to (4.10) of Corollary 4.2. But since $\tilde{r}_i = \bar{r}_i = r_i^*$ and with Theorem 3.6 $P_D \tilde{r}_i = \tilde{r}_i$ we have $P_D \bar{r}_i = \bar{r}_i$. Since P_D is a projection onto $\mathcal{Y}^\perp$, the residual $\bar{r}_i$ is orthogonal to $\mathcal{R}(Y)$, thus $\bar{x}_i = x_i^*$ and no final correction is needed for $\bar{x}_i$. ∎

The statement of the above theorem can also be proved for Hermitian positive definite matrices A and the CG method. We then use $\mathcal{Y} = \mathcal{Z}$ and the residuals must be orthogonal to $K_i(P_D A, P_D r_0)$. Moreover, $Q_D = P_D^H$. In this case the proofs of Theorem 4.1, Corollary 4.2, and Theorem 4.3 are similar to the given ones. Thus we have the following theorem.

THEOREM 4.4. *Let* $A \in \mathbb{C}^{n\times n}$ *be Hermitian positive definite. Let* $\mathcal{Z}$ *be a subspace of* $\mathbb{C}^n$ *of dimension* r. *Let* Z *be a matrix whose columns form a basis of* $\mathcal{Z}$. *Let* P_D *be defined as in* (3.1) *with* $Y = Z$. *For an arbitrary starting vector* $x_0 \in \mathbb{C}^n$ *we have that*

- *the approximations* x_i^* *given by the CG method applied to the linear system* $P_D A x = P_D b$ *with starting vector* $x_0^* = P_D^H x_0 + Qb$ *converge to the solution of* $Ax = b$,
- *the approximations* $\tilde{x}_i$ *given by the CG method applied to the linear system* $P_D A x = P_D b$ *with starting vector* x_0 *converge to a solution of* $P_D A x = P_D b$,
- *it holds that* $x_i^* = P_D^H \tilde{x}_i + Qb$.

5. Multigrid. In this section we transform the results of the previous sections to multigrid methods. The results in this section are motivated and partially proved in [47, 48]. See also [49] for a convergence analysis of multigrid for the nonnormal case. Here we change the notation to that which is typically used in the theory of multigrid methods. The $n \times r$ matrix Z is replaced by a matrix P, which is the *prolongation operator*. The $r \times n$ matrix Y^H is replaced by R, the *restriction operator*. Note that we do not use projections explicitly here. We then define the coarse grid matrix

$$A_C := RAP. \tag{5.1}$$

Here we always assume that A and A_C are nonsingular. Equivalent conditions for this assumption have been given in the previous sections. Then let

$$Q := PA_C^{-1}R \quad \text{and} \quad P_D := I - AQ. \tag{5.2}$$

The typical multigrid iteration or error propagation matrix is given by

$$T_M = \left(I - M_2^{-1}A\right)^{\nu_2}(I - QA)\left(I - M_1^{-1}A\right)^{\nu_1}, \tag{5.3}$$

where $M_1^{-1} \in \mathbb{C}^{n\times n}$ and $M_2^{-1} \in \mathbb{C}^{n\times n}$ are *smoothers*, ν_1 and ν_2 are the number of smoothing steps, and Q is the *coarse grid correction* matrix.

If the spectral radius of T_M is less than one, there is a nonsingular matrix B such that

$$T_M = I - BA; \tag{5.4}$$

see, e.g., [5].

The matrix B also acts as the multigrid preconditioner. Therefore eigenvalue estimates of BA are of interest. Note that the eigenvalues of T_M are just one minus the eigenvalues of BA.

In order to use the results of the previous section we assume that there is a nonsingular matrix X such that

$$(I - X^{-1}A) = (I - M_1^{-1}A)^{\nu_1}(I - M_2^{-1}A)^{\nu_2}. \tag{5.5}$$

Note that if the spectral radius of $(I - M_1^{-1}A)^{\nu_1}(I - M_2^{-1}A)^{\nu_2}$ is less than one, then there exists such a matrix X. We then define

$$T_D := (I - X^{-1}A)(I - QA). \tag{5.6}$$

Then we obtain

$$T_D = I - X^{-1}A - QA + X^{-1}AQA \tag{5.7}$$
$$= I - (X^{-1}(I - AQ) + Q)A \tag{5.8}$$
$$= I - (X^{-1}P_D + Q)A. \tag{5.9}$$

Thus, $X^{-1}P_D + Q$ is the adapted deflation preconditioner, where the matrix M is replaced by the matrix X. Since $\sigma(CS) = \sigma(SC)$ for any $C, S \in \mathbb{C}^{n\times n}$ we obtain that the spectrum of T_M is the same as the spectrum of T_D. With

$$\begin{aligned} T_D &= I - (X^{-1}P_D + Q)A \\ &= I - \tilde{B}A, \end{aligned}$$

where

$$\tilde{B} = X^{-1}P_D + Q,$$

we obtain

$$\sigma(BA) = \sigma(\tilde{B}A),$$

where B is defined in (5.4).

As an application of the theory in the previous sections we obtain the following results.

THEOREM 5.1. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Let $P \in \mathbb{C}^{n\times r}$ and $R \in \mathbb{C}^{r\times n}$ such that RAP is nonsingular. Moreover, let $M_1 \in \mathbb{C}^{n\times n}$ and $M_2 \in \mathbb{C}^{n\times n}$ such that the matrix X in* (5.5) *is nonsingular. Then the matrix B in* (5.4) *is nonsingular if and only if RXP is nonsingular.*

Proof. We have $\sigma(BA) = \sigma(\tilde{B}A)$. Thus, for a nonsingular matrix A, the matrix B is nonsingular if and only if $\tilde{B}$ is nonsingular. But $\tilde{B} = X^{-1}P_D + Q$; hence with Theorem 3.10 we get the desired result. ∎

We also have the following theorem.

THEOREM 5.2. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Let $P \in \mathbb{C}^{n\times r}$ and $R \in \mathbb{C}^{r\times n}$ such that RAP is nonsingular. Moreover, let $M_1 \in \mathbb{C}^{n\times n}$ and $M_2 \in \mathbb{C}^{n\times n}$ such that that the matrices X in* (5.5) *and RXP are nonsingular. Let $Q_X = P(RXP)^{-1}R$. Then B in* (5.4) *is nonsingular and we have the following:*

1. *BA has the eigenvalue one with multiplicity r; the other eigenvalues are nonzero and are denoted by $\mu_{r+1}, \dots, \mu_n$, i.e.,*
$$\sigma(BA) = \{1, \dots, 1, \mu_{r+1}, \dots, \mu_n\}.$$
2. *$X^{-1}A(I - QA) = X^{-1}P_DA$ has the eigenvalue zero with multiplicity r, and the other eigenvalues are $\mu_{r+1}, \dots, \mu_n$, i.e.,*
$$\sigma(X^{-1}A(I - QA)) = \{0, \dots, 0, \mu_{r+1}, \dots, \mu_n\}.$$
3. *The matrix $A^{-1}X(I - Q_XX) + Q_XX$ has the eigenvalue one with multiplicity r, and the other eigenvalues are $\frac{1}{\mu_{r+1}}, \dots, \frac{1}{\mu_n}$, i.e.,*
$$\sigma(A^{-1}X(I - Q_XX) + Q_XX) = \left\{1, \dots, 1, \frac{1}{\mu_{r+1}}, \dots, \frac{1}{\mu_n}\right\}.$$

4. *The matrix* $A^{-1}X(I - Q_X X)$ *has the eigenvalue zero with multiplicity* r*, and the other eigenvalues are* $\frac{1}{\mu_{r+1}}, \ldots, \frac{1}{\mu_n}$*, i.e.,*

$$\sigma(A^{-1}X(I - Q_X X)) = \left\{0, \ldots, 0, \frac{1}{\mu_{r+1}}, \ldots, \frac{1}{\mu_n}\right\}.$$

5. *Let* $\tilde{P}, \tilde{R} \in \mathbb{C}^{n \times n-r}$ *such that the columns of* $\tilde{P}$ *and* $\tilde{R}$ *build orthonormal bases of* $(\mathcal{R}(P))^{\perp}$ *and* $(\mathcal{R}(R^H))^{\perp}$*. Then* $\tilde{P}^H A^{-1} \tilde{R}$ *is nonsingular and*

$$\sigma\left(\tilde{P}^H X^{-1} \tilde{R} (\tilde{P}^H A^{-1} \tilde{R})^{-1}\right) = \{\mu_{r+1}, \ldots, \mu_n\}.$$

Proof. All statements follow immediately from Theorem 3.12. ∎

Theorem 5.1 is given in [48] for the case $R = P$. The first four statements of Theorem 5.2 are established in [47]; see also [48] for the case $R = P$. However, statement 5 is new. Moreover, the proofs given here seem to be much shorter and highlight the role of projections in this area (see Theorem 2.3). For Hermitian positive definite matrices A, statement 5 of Theorem 5.1 can be used to obtain optimal prolongation operators [23].

6. Conclusion. Deflation is a well-known technique to accelerate Krylov subspace methods for solving linear systems of equations. Moreover, from an abstract point of view deflation methods are also closely related to multigrid or multilevel and domain decomposition methods. Here we analyzed deflation methods that use two different deflation subspaces applied to nonsymmetric matrices. Projections are the major ingredient of deflation techniques. Here we presented new results on arbitrary projections. These new results allow not only elegant proofs for the relation between deflation and multigrid but also for the study of spectral properties of multigrid methods. Moreover, we established some characterizations of the spectrum of the deflated matrix. We also proved an explicit equivalence theorem for two types of deflation methods.

Acknowledgments. The authors wish to thank the anonymous referees and D. B. Szyld for their comments, which helped to improve the presentation.

REFERENCES

[1] A. M. Abdel-Rehim, R. B. Morgan, D. A. Nicely, and W. Wilcox, *Deflated and restarted symmetric Lanczos methods for eigenvalues and linear equations with multiple right-hand sides*, SIAM J. Sci. Comput., 32 (2010), pp. 129–149.

[2] K. Ahuja, P. Benner, E. de Sturler, and L. Feng, *Recycling BiCGSTAB with an application to parametric model order reduction*, SIAM J. Sci. Comput., 37 (2015), pp. S429–S446.

[3] K. Ahuja, E. de Sturler, S. Gugercin, and E. R. Chang, *Recycling BiCG with an application to model reduction*, SIAM J. Sci. Comput., 34 (2012), pp. A1925–A1949.

[4] A. Amritkar, E. de Sturler, K. Świrydowicz, D. Tafti, and K. Ahuja, *Recycling Krylov subspaces for CFD applications and a new hybrid recycling solver*, J. Comput. Phys., 303 (2015), pp. 222–237.

[5] M. Benzi and D. B. Szyld, *Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods*, Numer. Math., 76 (1997), pp. 309–321.

[6] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.

[7] K. Carlberg, V. Forstall, and R. Tuminaro, *Krylov-subspace recycling via the POD-augmented conjugate-gradient method*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 1304–1336.

[8] A. Chapman and Y. Saad, *Deflated and augmented Krylov subspace techniques*, Numer. Linear Algebra Appl., 4 (1997), pp. 43–66.

[9] M. Clemens, M. Wilke, R. Schuhmann, and T. Weiland, *Subspace projection extrapolation scheme for transient field simulations*, IEEE Trans. Magnetics, 40 (2004), pp. 934–937.
[10] D. Darnell, R. B. Morgan, and W. Wilcox, *Deflated GMRES for systems with multiple shifts and multiple right-hand sides*, Linear Algebra Appl., 429 (2008), pp. 2415–2434.
[11] E. de Sturler, *Nested Krylov methods based on GCR*, J. Comput. Appl. Math., 67 (1996), pp. 15–41.
[12] G. B. Diaz Cortes, C. Vuik, and J. D. Jansen, *On POD-based deflation vectors for DPCG applied to porous media problems*, J. Comput. Appl. Math., 330 (2018), pp. 193–213.
[13] Z. Dostál, *Conjugate gradient method with preconditioning by projector*, Int. J. Comput. Math., 23 (1988), pp. 315–323.
[14] M. Eiermann, O. G. Ernst, and O. Schneider, *Analysis of acceleration strategies for restarted minimal residual methods*, J. Comput. Appl. Math., 123 (2000), pp. 261–292.
[15] J. Erhel, K. Burrage, and B. Pohl, *Restarted GMRES preconditioned by deflation*, J. Comput. Appl. Math., 69 (1996), pp. 303–318.
[16] J. Erhel and F. Guyomarc'h, *An augmented conjugate gradient method for solving consecutive symmetric positive definite linear systems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1279–1299.
[17] Y. A. Erlangga and R. Nabben, *Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 684–699.
[18] Y. A. Erlangga and R. Nabben, *Multilevel projection-based nested Krylov iteration for boundary value problems*, SIAM J. Sci. Comput., 30 (2008), pp. 1572–1595.
[19] Y. A. Erlangga and R. Nabben, *On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian*, Electron. Trans. Numer. Anal., 31 (2008), pp. 403–424.
[20] Y. A. Erlangga and R. Nabben, *Algebraic multilevel Krylov methods*, SIAM J. Sci. Comput., 31 (2009), pp. 3417–3437.
[21] J. Frank and C. Vuik, *On the construction of deflation-based preconditioners*, SIAM J. Sci. Comput., 23 (2001), pp. 442–462 (electronic).
[22] A. Galántai, *Projectors and Projection Methods*, Springer, New York, 2004.
[23] L. García Ramos and R. Nabben, *On optimal algebraic multigrid methods*, submitted.
[24] L. García Ramos and R. Nabben, *On the spectrum of deflated matrices with applications to the deflated shifted Laplace preconditioner for the Helmholtz equation*, SIAM J. Matrix. Anal. Appl., 39 (2018), pp. 262–286.
[25] A. Gaul, *Recycling Krylov Subspace Methods for Sequences of Linear Systems*, Ph.D. thesis, Technische Universität Berlin, 2014.
[26] A. Gaul, M. H. Gutknecht, J. Liesen, and R. Nabben, *A framework for deflated and augmented Krylov subspace methods*, SIAM. J. Matrix Anal. Appl., 34 (2013), pp. 495–518.
[27] L. Giraud, S. Gratton, X. Pinel, and X. Vasseur, *Flexible GMRES with deflated restarting*, SIAM J. Sci. Comput., 32 (2010), pp. 1858–1878.
[28] M. H. Gutknecht, *Spectral deflation in Krylov solvers: A theory of coordinate space based methods*, Electron. Trans. Numer. Anal., 39 (2012), pp. 156–185.
[29] M. H. Gutknecht, *Deflated and augmented Krylov subspace methods: A framework for deflated BiCG and related solvers*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1444–1466.
[30] K. Hayami and M. Sugihara, *A geometric view of Krylov subspace methods on singular systems*, Numer. Linear Algebra Appl., 18 (2011), pp. 449–469.
[31] K. Kahl and H. Rittich, *The deflated conjugate gradient method: Convergence, perturbation and accuracy*, Linear Algebra Appl, (2016), pp. 1–26.
[32] M. E. Kilmer and E. de Sturler, *Recycling subspace information for diffuse optical tomography*, SIAM J. Sci. Comput., 27 (2006), pp. 2140–2166.
[33] A. Klawonn and O. Rheinbach, *Deflation, projector preconditioning, and balancing in iterative substructuring methods: Connections and new results*, SIAM J. Sci. Comput., 34 (2012), pp. A459–A484.
[34] L. Y. Kolotilina, *Twofold deflation preconditioning of linear algebraic systems.* I. *Theory*, J. Math. Sci., 89 (1998), pp. 1652–1689.
[35] J. Liesen and Z. Strakoš, *Krylov Subspace Methods: Principles and Analysis*, Numer. Math. Sci. Comput., Oxford University Press, New York, 2012.
[36] L. Mansfield, *On the use of deflation to improve the convergence of conjugate gradient iteration*, Comm. Appl. Numer. Methods, 4 (1988), pp. 151–156.
[37] L. Mansfield, *On the conjugate gradient solution of the Schur complement system obtained from domain decomposition*, SIAM J. Numer. Anal., 27 (1990), pp. 1612–1620.
[38] L. Mansfield, *Damped Jacobi preconditioning and coarse grid deflation for conjugate gradient iteration on parallel computers*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 1314–1323.

[39] R. B. Morgan, *A restarted GMRES method augmented with eigenvectors*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1154–1171.
[40] R. B. Morgan, *Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1112–1135.
[41] R. B. Morgan, *GMRES with deflated restarting*, SIAM J. Sci. Comput., 24 (2002), pp. 20–37 (electronic).
[42] R. B. Morgan, *Restarted block-GMRES with deflation of eigenvalues*, Appl. Numer. Math., 54 (2005), pp. 222–236.
[43] R. Nabben and J. Schramm, *Improved error bounds for the deflated multi-preconditioned CG method*, submitted.
[44] R. Nabben and C. Vuik, *A comparison of deflation and coarse grid correction applied to porous media flow*, SIAM J. Numer. Anal., 42 (2004), pp. 1631–1647.
[45] R. Nabben and C. Vuik, *A comparison of deflation and the balancing preconditioner*, SIAM J. Sci. Comput., 27 (2006), pp. 1742–1759.
[46] R. A. Nicolaides, *Deflation of conjugate gradients with applications to boundary value problems*, SIAM J. Numer. Anal., 24 (1987), pp. 355–365.
[47] Y. Notay, *Algebraic analysis of two-grid methods: The nonsymmetric case*, Numer. Linear Algebra Appl., 17 (2010), pp. 73–96.
[48] Y. Notay, *Algebraic theory of two-grid methods*, Numer. Math. Theory Methods Appl., 8 (2015), pp. 168–198.
[49] Y. Notay, *Analysis of Two-Grid Methods: The Nonnormal Case*, Université Libre de Bruxelles, Brussels, 2018.
[50] M. L. Parks, E. de Sturler, G. Mackey, D. D. Johnson, and S. Maiti, *Recycling Krylov subspaces for sequences of linear systems*, SIAM J. Sci. Comput., 28 (2006), pp. 1651–1674.
[51] Y. Saad, *Analysis of augmented Krylov subspace methods*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 435–449.
[52] Y. Saad, *Iterative methods for sparse linear systems*, 2nd ed., SIAM, Philadelphia, 2003.
[53] Y. Saad, M. Yeung, J. Erhel, and F. Guyomarc'h, *A deflated version of the conjugate gradient algorithm*, SIAM J. Sci. Comput., 21 (2000), pp. 1909–1926.
[54] A. H. Sheikh, D. Lahaye, L. García Ramos, R. Nabben, and C. Vuik, *Accelerating the shifted Laplace preconditioner for the Helmholtz equation by multilevel deflation*, J. Comput. Phys., 322 (2016), pp. 473–490.
[55] A. H. Sheikh, D. Lahaye, and C. Vuik, *On the convergence of shifted Laplace preconditioner combined with multilevel deflation*, Numer. Linear Algebra Appl., 20 (2013), pp. 645–662.
[56] J. A. Sifuentes, M. Embree, and R. B. Morgan, *GMRES convergence for perturbed coefficient matrices, with application to approximate deflation preconditioning*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1066–1088.
[57] K. M. Soodhalter, *Block Krylov subspace recycling for shifted systems with unrelated right-hand sides*, SIAM J. Sci. Comput., 38 (2016), pp. A302–A324.
[58] K. M. Soodhalter, D. B. Szyld, and F. Xue, *Krylov subspace recycling for sequences of shifted linear systems*, Appl. Numer. Math., 81 (2014), pp. 105–118.
[59] N. Spillane, *An adaptive multipreconditioned conjugate gradient algorithm*, SIAM J. Sci. Comput., 38 (2016), pp. A1896–A1918.
[60] D.-L. Sun, T.-Z. Huang, B. Carpentieri, and Y.-F. Jing, *Flexible and deflated variants of the block shifted GMRES method*, J. Comput. Appl. Math., 345 (2019), pp. 168–183.
[61] D.-L. Sun, T.-Z. Huang, Y.-F. Jing, and B. Carpentieri, *A block GMRES method with deflated restarting for solving linear systems with multiple shifts and multiple right-hand sides*, Numer. Linear Algebra Appl., 25 (2018), p. e2148.
[62] J. M. Tang, S. P. MacLachlan, R. Nabben, and C. Vuik, *A comparison of two-level preconditioners based on multigrid and deflation*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1715–1739.
[63] J. M. Tang, R. Nabben, C. Vuik, and Y. A. Erlangga, *Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods*, J. Sci. Comput., 39 (2009), pp. 340–370.
[64] S. Wang, E. de Sturler, and G. H. Paulino, *Large-scale topology optimization using preconditioned Krylov subspace methods with recycling*, Internat. J. Numer. Methods Engrg., 69 (2007), pp. 2441–2468.
[65] Y.-F. Xiang, Y.-F. Jing, and T.-Z. Huang, *A new projected variant of the deflated block conjugate gradient method*, J. Sci. Comput., 80 (2019), pp. 1116–1138.