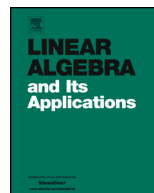




ELSEVIER

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

A Lie algebra view of matrix splittings

Michele Benzi*, Milo Viviani

Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56126, Pisa, Italy

ARTICLE INFO

Article history:

Received 19 March 2025

Received in revised form 29 October 2025

Accepted 30 October 2025

Available online xxxx

Submitted by V. Mehrmann

Dedicated to Daniel Szyld on the occasion of his 70th birthday

MSC:

65F10

15A23

22-8

Keywords:

Matrix decompositions

Matrix splittings

Lie groups

Lie algebras

Jordan algebras

ABSTRACT

In this paper we use some basic facts from the theory of (matrix) Lie groups and algebras to show that many of the classical matrix splittings used to construct stationary iterative methods and preconditioners for Krylov subspace methods can be interpreted as linearizations of matrix factorizations. Moreover, we show that new matrix splittings are obtained when we specialize these splittings to some of the classical matrix groups and their Lie and Jordan algebras. As an example, we derive structured generalizations of the HSS (Hermitian and skew-Hermitian splitting) iteration, and provide sufficient conditions for their convergence.

© 2025 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many classical algorithms of numerical linear algebra are based on matrix decompositions. These decompositions can be either multiplicative or additive; following standard

* Corresponding author.

E-mail addresses: michele.benzi@sns.it (M. Benzi), milo.viviani@sns.it (M. Viviani).

<https://doi.org/10.1016/j.laa.2025.10.035>

0024-3795/© 2025 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

terminology, herein we will refer to the former type as *matrix factorizations*, and to the latter one as *matrix splittings*.

Matrix factorizations are at the basis of many of the most efficient algorithms for solving linear systems and for computing eigenvalues (or singular values) of matrices. The LU, Cholesky, and QR (or LQ) factorizations are among the best known matrix decompositions [19]. Additive decompositions, or splittings, of the form $A = B + C$ (frequently written in the form $A = B - C$) are used to define stationary iterative methods, including the classical Jacobi and Gauss-Seidel iterations [38], and to construct preconditioners used in conjunction with Krylov subspace methods [35]. Splitting methods are also widely used in the numerical solution of systems of ODEs, particularly on manifolds, and semi-discretized PDEs [29].

It has long been recognized that many matrix factorizations can be interpreted in terms of decompositions of a matrix group as a product of subgroups.¹ Matrix groups endowed with additional (topological or differentiable) structure are fundamental objects in many areas of mathematics, from Lie theory to numerical analysis [25], and the connection between matrix factorizations and different representations of a matrix group as a product of subgroups is fairly standard, though not widely used in numerical linear algebra.

In this paper we show that just as many matrix factorizations can be understood in terms of (multiplicative) decompositions of a matrix Lie group as a product of subgroups, certain matrix splittings can similarly be understood in terms of (additive) decompositions of the corresponding Lie algebra (the tangent space of the Lie group at the identity). In some cases the Lie algebra of the group splits as a sum of Lie subalgebras which corresponds to subgroups of the matrix group. In others, as we will see, the Lie algebra splits as the sum of a Lie subalgebra and a Jordan algebra. This second situation arises when one of factors in the matrix factorization comes from a subgroup while the other comes from a vector space of matrices. Moreover, many additive splittings can be obtained as the linearization of an underlying multiplicative factorization of the underlying matrix group. Hence, many classical matrix splittings appear in a new light: they are *infinitesimal counterparts* of well-known matrix factorizations. To the best of our knowledge this observation, while simple, has not been explicitly made before in the numerical analysis literature.

To informally describe what this paper is about, we recall that passing from a matrix Lie group G to the tangent space at the identity (i.e., linearizing G near I_n) results in a vector space \mathfrak{g} of matrices which, endowed with the commutator product $[A, B] = AB - BA$, becomes a Lie algebra. Consider now a smooth path of matrices $t \mapsto A(t)$ in G with $A(0) = I_n$ and a smooth factorization

$$A(t) = B(t)C(t), \tag{1}$$

¹ One of the first papers drawing the attention of numerical analysts to the relation of matrix factorizations to group decompositions is [16].

where $B(t)$ and $C(t)$ are smooth paths entirely lying in submanifolds $H, K \subset G$ for $|t|$ small and such that $B(0) = C(0) = I_n$. Differentiating (1) for $t = 0$ yields

$$A'(0) = B'(0)C(0) + B(0)C'(0) = B'(0) + C'(0), \quad (2)$$

and we see that the tangent vector $A'(0)$ to G at the identity, an element of the Lie algebra \mathfrak{g} of G , decomposes as the sum of two matrices $B'(0)$ and $C'(0)$, one tangent to H and the other tangent to K at the identity. In other words, the Lie algebra \mathfrak{g} splits as the sum of two vector subspaces, which in turn may be endowed with additional algebraic structure. For instance, if H and K are matrix subgroups of G , the corresponding tangent spaces \mathfrak{h} and \mathfrak{k} are themselves Lie algebras; if $H \cap K = \{I_n\}$, then $\mathfrak{h} \cap \mathfrak{k} = \{0\}$ and we obtain that \mathfrak{g} , as a vector space, admits the direct sum decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}$.

When $G = GL(n, \mathbb{R})$, the general linear group, the corresponding Lie algebra \mathfrak{g} , denoted with \mathfrak{gl}_n , is just the total matrix space $\mathbb{R}^{n \times n}$ endowed with the commutator product. Any multiplicative decomposition of G into a product of submanifolds leads to a corresponding additive splitting of \mathfrak{gl}_n , and the structure of the matrices B and C in the splitting $A = B + C$ will depend on the choice of H and K . For instance, linearization of the *polar factorization* of $GL(n, \mathbb{R})$ leads to the *symmetric and skew-symmetric splitting* of a generic $n \times n$ matrix, an instance of the *Cartan decomposition*; see subsection 3.1. Furthermore, this splitting, widely used in numerical linear algebra (e.g., [6,8,11,15,20,40]), corresponds to a decomposition of \mathfrak{gl}_n into subspaces that are mutually orthogonal with respect to the Frobenius inner product. One of these subspaces carries a Lie algebra structure, the other one a Jordan algebra structure. This provides us with a canonical example of what we call here a *Lie-Jordan splitting*; see Section 4.

In this paper we consider several different group decompositions and derive the corresponding matrix splittings. In some cases, instead of considering a global factorization of G , it will be convenient to localize the discussion near the identity I_n , i.e., to consider a decomposition of an open neighborhood of the identity matrix. Since the linearization \mathfrak{g} of G consists of all the tangent vectors at the identity, this still leads to an additive splitting for every matrix in \mathfrak{g} .

Instead of starting from a multiplicative group decomposition in order to obtain a class of additive matrix splittings via linearization, it is also possible to proceed in the reverse order: starting from a matrix splitting, we can try to determine the corresponding matrix factorization, or group decomposition. This passage from local to global can be interpreted as integration, as opposed to differentiation, and just as indefinite integration leads to a non-unique solution (up to additive constants), there may be more than one group factorization leading to the same splitting of the tangent space (Lie algebra). Another way to see this is to recall that while the Lie algebra of a Lie group is uniquely determined, being the tangent space to the Lie group at the identity, there can be more than one Lie group with the same Lie algebra: for instance, the orthogonal group $O(n, \mathbb{R})$ and the special orthogonal group $SO(n, \mathbb{R})$ both have \mathfrak{so}_n , the algebra of skew-symmetric matrices, as their Lie algebra. In the following, sometimes we will start from a matrix

factorization and derive the corresponding matrix splitting, while in others we start from the splitting, regarded as an additive decomposition of a matrix Lie algebra, and then discuss one or more group decompositions, i.e., matrix factorizations, the linearization of which results in the additive splitting.

Classical stationary iterative methods, like Jacobi and Gauss-Seidel, are based on the splittings of A in terms of the diagonal, strictly lower, and strictly upper triangular parts [38]. As we will see, these classical splittings are infinitesimal counterparts, or linearizations, of the LU or (LDU) factorization. Here again we find that linearization leads to a splitting of the Lie algebra of all $n \times n$ matrices into a direct sum of subspaces, which in the case of Gauss-Seidel is an orthogonal decomposition relative to the Frobenius inner product.

We also examine other matrix splittings, for example those corresponding to the QR factorization (a special case of the *Iwasawa decomposition* of a Lie group, see [22,26]), as well as others which generalize the symmetric and skew-symmetric splitting.

The two main contributions of the paper are the following. First, we provide a link (of a geometric nature) between some of the classical matrix decompositions and the corresponding matrix splittings, which has been hitherto overlooked in the numerical linear algebra literature; and second, we present new matrix splittings, in particular some generalizations of the well-known HSS method [6], which may prove useful in the construction of iterative solvers and preconditioners for block linear systems. Some of these splittings may also find application, for example, in developing structure-preserving ODE or eigenvalue solvers.

The remainder of the paper is organized as follows. In Section 2 we provide basic background information on Lie groups, Lie algebras and Jordan algebras, primarily restricted to the classical (matrix) groups and their algebras. In Section 3 we survey several types of matrix splittings and interpret them as linearizations of well-known matrix factorizations. In Section 4 we discuss the class of Lie–Jordan splittings and in Section 5 we establish a general convergence criterion for what we call the J -HSS method (a generalization of the standard HSS iteration [6]), and provide examples of this method. Section 6 contains brief discussions of two alternating-type iterations and the corresponding splittings. Some conclusions and suggestions for future work are given in Section 7.

2. Preliminaries and background

In this Section we recall some basic notions about matrix Lie groups, Lie algebras and Jordan algebras, with the aim of introducing notations and making the paper as self-contained as possible. For ease of exposition, in this paper we limit ourselves to real matrices, but much of the material in the paper can be extended to the complex case. Our treatment is elementary and informal; we refer the reader to, e.g., [22,26,27,32,34] for thorough and general accounts of the theory of Lie groups and algebras, differentiable manifolds, and Jordan algebras.

Let $GL(n, \mathbb{R})$ be the general linear group of degree n over \mathbb{R} , i.e., the set of all real invertible $n \times n$ matrices endowed with the standard matrix product. We note that $GL(n, \mathbb{R})$ is both a group and a differentiable manifold (of dimension n^2), and that the mappings $(A, B) \mapsto AB$ and $A \mapsto A^{-1}$ are smooth. These facts can be condensed in the statement that $GL(n, \mathbb{R})$ is a (matrix) Lie group.

The tangent space to $GL(n, \mathbb{R})$ at the identity I_n coincides with the whole space $\mathbb{R}^{n \times n}$. To see this, take $A \in \mathbb{R}^{n \times n}$, consider (for example) the smooth path $t \mapsto A(t) \in GL(n, \mathbb{R})$ defined by $A(t) = \exp(tA)$, and observe that $A(0) = I_n$, $A'(0) = A$. Hence, every $A \in \mathbb{R}^{n \times n}$ is a tangent vector to $GL(n, \mathbb{R})$ at the identity.

With respect to the usual matrix operations, $\mathbb{R}^{n \times n}$ is an associative, non-commutative algebra with unit. However, we can also introduce two other types of products in $\mathbb{R}^{n \times n}$: the commutator product $[A, B] = AB - BA$, and the Jordan product $A \bullet B = \frac{1}{2}(AB + BA)$. The commutator product does not admit an identity element, is anticommutative ($[B, A] = -[A, B]$), nilpotent ($[A, A] = 0$) and is non-associative; however, it satisfies a weak version of associativity, the *Jacobi identity*:

$$[[A, B], C] + [[C, A], B] + [[B, C], A] = 0 \quad \forall A, B, C, \in \mathbb{R}^{n \times n}.$$

The Jordan product admits I_n as identity element, is commutative, satisfies $A \bullet A = A^2$, and is non-associative; it satisfies a weak version of associativity, the *Jordan identity*:

$$(A \bullet B) \bullet (A \bullet A) = A \bullet (B \bullet (A \bullet A)) \quad \forall A, B \in \mathbb{R}^{n \times n}.$$

We express these properties by saying that $\mathbb{R}^{n \times n}$ is a *Lie algebra* with respect to the commutator product, and a *Jordan algebra* with respect to the Jordan product. We will denote the Lie algebra of all real $n \times n$ matrices with the commutator product by \mathfrak{gl}_n , and the Jordan algebra of all real $n \times n$ matrices with the commutator product by $\mathbb{M}(n, \mathbb{R})$.

We will be dealing with vector subspaces of $\mathbb{R}^{n \times n}$ which are also Lie subalgebras or Jordan subalgebras. Prominent Lie subalgebras of \mathfrak{gl}_n are \mathfrak{sl}_n , the Lie algebra of all $n \times n$ matrices with zero trace, and \mathfrak{so}_n , the Lie algebra of all $n \times n$ skew-symmetric matrices: note that $\mathfrak{so}_n \subset \mathfrak{sl}_n \subset \mathfrak{gl}_n$. The Lie algebra \mathfrak{sl}_n is the tangent space at the identity of the *special linear group* $SL(n, \mathbb{R})$, consisting of all the real $n \times n$ matrices with unit determinant, while \mathfrak{so}_n is the Lie algebra of the *special orthogonal group* $SO(n, \mathbb{R})$, which consists of all the real $n \times n$ orthogonal matrices with unit determinant. This group is the connected component (containing the identity) of the group $O(n, \mathbb{R})$ of all the real $n \times n$ orthogonal matrices, therefore \mathfrak{so}_n is also the Lie algebra (tangent space) of $O(n, \mathbb{R})$. The subgroups $O(n, \mathbb{R})$ and $SO(n, \mathbb{R})$ are topologically closed subgroups of the Lie group $GL(n, \mathbb{R})$ and therefore they are themselves Lie groups. The Lie algebra of a Lie group is often thought of as the *linearization* of the Lie group near the identity.

In the course of the paper we will also deal with groups and algebras of triangular and diagonal matrices. The groups $LT(n, \mathbb{R})$ and $UT(n, \mathbb{R})$ of all invertible lower (respectively, upper) triangular matrices are Lie groups; their Lie algebras are easily seen

to be the (not necessarily invertible) lower and upper triangular matrices, denoted here by \mathfrak{lt}_n and \mathfrak{ut}_n , respectively. Important subgroups of $LT(n, \mathbb{R})$ are $LT_+(n, \mathbb{R})$ (lower triangular matrices with positive diagonal entries), $SLT(n, \mathbb{R}) = SL(n, \mathbb{R}) \cap LT(n, \mathbb{R})$ (lower triangular matrices with unit determinant), and $LT_1(n, \mathbb{R})$ (unit lower triangular matrices). The subgroups $UT_+(n, \mathbb{R})$, $SUT(n, \mathbb{R})$ and $UT_1(n, \mathbb{R})$ of $UT(n, \mathbb{R})$ are defined analogously. The Lie algebra of $LT_+(n, \mathbb{R})$ is easily seen to be again \mathfrak{lt}_n , the Lie algebra of $SLT(n, \mathbb{R})$ is $\mathfrak{slt}_0 = \mathfrak{sl}_n \cap \mathfrak{lt}_n$ (the lower triangular matrices with zero trace), and the Lie algebra of $LT_1(n, \mathbb{R})$ is the nilpotent Lie subalgebra of the strictly lower triangular matrices, $\mathfrak{lt}_0 \subset \mathfrak{lt}_n$. The Lie algebra of upper triangular matrices with zero trace will be denoted by \mathfrak{sut}_0 and the nilpotent Lie algebra of all strictly upper triangular matrices by \mathfrak{ut}_0 .

We will also use \mathfrak{d} and \mathfrak{d}_0 to denote the Lie algebras of all diagonal matrices and diagonal matrices with zero trace, respectively. They are the Lie algebras of the groups of all invertible diagonal matrices (and of its subgroup of all diagonal matrices with positive diagonal entries) and of the group of diagonal matrices with unit determinant, respectively.

The most important example of a Jordan subalgebra of $\mathbb{M}(n, \mathbb{R})$ is provided by the space of all real $n \times n$ symmetric matrices. Indeed, the Jordan product was introduced (by German quantum physicist Pascual Jordan) as a way to “symmetrize” the product of two possibly non-commuting symmetric operators.² We denote this Jordan subalgebra by \mathfrak{p}_n .

The real orthogonal group $O(n, \mathbb{R})$ and its subgroup $SO(n, \mathbb{R})$ are examples of *classical groups*. The group $O(n, \mathbb{R})$ can be characterized as the group of linear transformations Q on \mathbb{R}^n such that $\langle Qx, Qy \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{R}^n$, where $\langle x, y \rangle = x^T y$ denotes the standard (Euclidean) inner product. Other examples of classical groups are the *pseudo-orthogonal group* $O(p, q, \mathbb{R})$ (where $p + q = n$) of all linear transformations on \mathbb{R}^n which preserve the indefinite inner product $x^T I_{p,q} y$, where $I_{p,q} = \text{diag}(I_p, -I_q)$, and the *symplectic group* $Sp(n, \mathbb{R})$ of all linear transformations on \mathbb{R}^{2m} which preserve the symplectic form $x^T \Sigma_n y$, where $n = 2m$ and

$$\Sigma_n = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}.$$

We also have the group of pseudo-orthogonal matrices with unit determinant, $SO(p, q, \mathbb{R})$. (Note that symplectic matrices always have unit determinant.) The Lie algebra of the group $O(p, q, \mathbb{R})$ (and also of $SO(p, q, \mathbb{R})$) consists of all the pseudo-skew-symmetric matrices, while the Lie algebra of $Sp(n, \mathbb{R})$ consists of all the Hamiltonian matrices; these algebras are described in Section 4, where we also describe the corresponding Jordan algebras, namely, the pseudo-symmetric and skew-Hamiltonian matrices, respectively.

² Note that $A \bullet B = AB$ whenever $AB = BA$.

A connected Lie group that does not contain any non-trivial, connected, Abelian normal subgroup is said to be a *semisimple Lie group*. The groups $GL(n, \mathbb{R})$ and $O(n, \mathbb{R})$ are not connected and therefore they are not semisimple. The connected Lie group $SO(n, \mathbb{R})$, on the other hand, is semisimple, as is the symplectic group $Sp(n, \mathbb{R})$.

The corresponding notion for a Lie algebra \mathfrak{g} is defined as follows. If \mathfrak{a} and \mathfrak{b} are vector subspaces of \mathfrak{g} , denote by $[\mathfrak{a}, \mathfrak{b}]$ the subspace generated by the elements of the form $[A, B]$ with $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$. The property of \mathfrak{a} being a Lie subalgebra of \mathfrak{g} can then be expressed as $[\mathfrak{a}, \mathfrak{a}] \subseteq \mathfrak{a}$. The condition $[\mathfrak{a}, \mathfrak{g}] \subseteq \mathfrak{a}$ expresses the property of \mathfrak{a} being a *Lie ideal* of \mathfrak{g} . As an example, the space \mathfrak{lt}_0 of strictly lower triangular matrices is a Lie ideal of the Lie algebra \mathfrak{lt}_n of all lower triangular matrices. Clearly, any Lie ideal is a Lie subalgebra of \mathfrak{g} . A Lie subalgebra \mathfrak{a} is *Abelian* if $[A, B] = 0$ for all $A, B \in \mathfrak{a}$. A Lie algebra that does not contain any non-trivial Abelian ideals is said to be a *semisimple Lie algebra*. A basic result in Lie theory is that *a Lie group is semisimple if and only if its Lie algebra is semisimple*. Hence, \mathfrak{sl}_n is semisimple, while \mathfrak{gl}_n is not: the set of all scalar matrices kI_n ($k \in \mathbb{R}$) is a nontrivial Abelian ideal in \mathfrak{gl}_n . Denoting by \mathfrak{r} this ideal, the Lie algebra \mathfrak{gl}_n decomposes as

$$\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r}. \quad (3)$$

The decomposition (3) is known as the *Levi decomposition* of \mathfrak{gl}_n . More generally, every Lie algebra \mathfrak{g} admits a Levi decomposition, i.e., it can be written as the direct sum of a semisimple Lie subalgebra and the *radical* \mathfrak{r} of \mathfrak{g} (we refer to [22] or [34] for the definition of the radical of a Lie algebra). The importance of the Levi decomposition is that it can be used to determine the structure of Lie groups and algebras by reducing it to the semisimple case, which is much more tractable.

3. Matrix factorizations and matrix splittings

In this Section we discuss several matrix splittings of interest in numerical linear algebra and describe how they correspond to linearizations of well known matrix factorizations.

3.1. Cartan decomposition

One of the most important decompositions of a Lie algebra (and of the corresponding Lie group) is the *Cartan decomposition*; see for example [22,26]. Rather than giving the general definition, here we describe it in the special cases we are mostly interested in.

Let us consider the map $\theta : \mathfrak{gl}_n \rightarrow \mathfrak{gl}_n$ defined by $\theta(X) = -X^T$. Then θ is a Lie algebra automorphism which satisfies $\theta^2 = \text{Id}$, i.e., it is an involution. Hence, θ has two eigenvalues, $+1$ and -1 . The corresponding eigenspaces are, respectively, the Lie subalgebra of all skew-symmetric matrices \mathfrak{so}_n and the subspace \mathfrak{p}_n of all $n \times n$ real symmetric matrices. This results in the decomposition

$$\mathfrak{gl}_n^{\mathbb{R}} = \mathfrak{so}_n \oplus \mathfrak{p}_n, \quad (4)$$

which is a Cartan decomposition of $\mathfrak{gl}_n^{\mathbb{R}}$.³ Note that (4) is simply the statement that any $n \times n$ real matrix A can be written uniquely as the sum of a skew-symmetric and a symmetric matrix:

$$A = \mathcal{S}(A) + \mathcal{H}(A), \quad (5)$$

where $\mathcal{S}(A) := \frac{1}{2}(A - A^T) \in \mathfrak{so}_n$ is the skew-symmetric part of A and $\mathcal{H}(A) := \frac{1}{2}(A + A^T) \in \mathfrak{p}_n$ the symmetric part. However, there is something more to this. First of all, the two subspaces \mathfrak{so}_n and \mathfrak{p}_n are mutually orthogonal with respect to the Frobenius inner product on $\mathbb{R}^{n \times n}$, defined as $\langle A, B \rangle = \text{Tr}(A^T B)$, as is easily checked ($\text{Tr}(\cdot)$ denotes the trace). Hence, the direct sum (4) is actually an orthogonal sum. In other words, \mathfrak{so}_n is the tangent space to $SO(n, \mathbb{R})$ at the identity, while \mathfrak{p}_n is the normal space. Second, while \mathfrak{so}_n is a Lie subalgebra of $\mathfrak{gl}_n^{\mathbb{R}}$ (and indeed of \mathfrak{sl}_n), \mathfrak{p}_n is not a Lie algebra; it is, however, a Jordan algebra with respect to the Jordan product $A \bullet B = \frac{1}{2}(AB + BA)$.

As already mentioned in the Introduction, the symmetric and skew-symmetric splitting is the infinitesimal version of the polar decomposition $A = QP$, where Q is orthogonal and P is symmetric and positive definite; since A is nonsingular, this factorization is unique. However, some care is needed here in order to have a precise statement of this fact. First we restrict our attention to the (semisimple) Lie group of all real $n \times n$ matrices with unit determinant, $SL(n, \mathbb{R})$; its Lie algebra, \mathfrak{sl}_n , consists of all real $n \times n$ matrices with zero trace. The *global Cartan decomposition* of the special linear group is

$$SL(n, \mathbb{R}) = SO(n, \mathbb{R}) \cdot \exp(\mathfrak{p}_0), \quad (6)$$

where $\exp(\mathfrak{p}_0)$ denotes the image under the exponential map of the set of all $n \times n$ real symmetric matrices with zero trace, consisting of all symmetric and positive definite matrices S with $\det S = 1$, since $\det \exp(S) = e^{\text{Tr}(S)}$. At the matrix level, the decomposition (6) is nothing but the polar factorization $A = QP$ of a real nonsingular matrix with unit determinant as the product of a special orthogonal matrix and a symmetric positive definite matrix with unit determinant.

Remark 3.1. More generally, a global Cartan decomposition of a non-compact, semisimple Lie group G with Lie algebra \mathfrak{g} is defined as follows. There is a unique Lie group automorphism Θ with differential at the identity given by θ (the Cartan involution of \mathfrak{g}) and such that $\Theta^2 = \text{Id}$ (the identity automorphism). Moreover, there is a compact

³ Here we follow the example of [17] and commit a slight abuse of language: strictly speaking, the term *Cartan decomposition* refers to the corresponding decomposition of the semisimple Lie algebra $\mathfrak{sl}_n = \mathfrak{so}_n \oplus \mathfrak{p}_0$, where \mathfrak{p}_0 is the subspace of all real symmetric matrices with zero trace. Since \mathfrak{so}_n is the Lie algebra corresponding to both $O(n, \mathbb{R})$ and $SO(n, \mathbb{R})$, this abuse is inconsequential for our purposes. Furthermore, recalling the Levi decomposition $\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r}$ and observing that $\mathfrak{p}_n = \mathfrak{p}_0 \oplus \mathfrak{r}$, we conclude that (4) is equivalent to the “genuine” Cartan decomposition $\mathfrak{sl}_n = \mathfrak{so}_n \oplus \mathfrak{p}_0$.

subgroup $K \subset G$ that is left fixed by Θ . Let \mathfrak{p} be the eigenspace of θ corresponding to the eigenvalue -1 . The mapping

$$K \times \mathfrak{p} \longrightarrow G \quad \text{given by} \quad (k, X) \mapsto k \cdot \exp(X)$$

is a diffeomorphism. The idempotent automorphism Θ is called a global Cartan involution. Writing $P = \exp(\mathfrak{p})$, we have that the product map $K \times P \rightarrow G$ is a diffeomorphism, so that $G = K \cdot P$. If $G = SL(n, \mathbb{R})$, the global Cartan involution is $\Theta : X \mapsto (X^{-1})^T$, hence $K = SO(n, \mathbb{R})$ and $\mathfrak{p} = \mathfrak{p}_0$.

We can now determine the additive counterpart of the polar decomposition. Taking a smooth path $t \mapsto A(t)$ in $SL(n, \mathbb{R})$ and writing $A(t) = Q(t)P(t)$ (smooth paths in $SO(n, \mathbb{R})$ and $\exp(\mathfrak{p}_0)$, respectively, with $Q(0) = P(0) = I_n$) and differentiating for $t = 0$ we find

$$A'(0) = Q'(0) + P'(0), \tag{7}$$

with $Q'(0) \in \mathfrak{so}_n$ and $P'(0)$ symmetric with $\text{Tr}(P'(0)) = 0$, i.e., $Q'(0)$ is the skew-symmetric part of $A'(0) \in \mathfrak{sl}_n$ and $P'(0)$ its symmetric part. We can therefore write the Cartan decomposition of \mathfrak{sl}_n as the infinitesimal counterpart of (6), namely:

$$\mathfrak{sl}_n = \mathfrak{so}_n \oplus \mathfrak{p}_0. \tag{8}$$

This shows that indeed the symmetric and skew-symmetric splitting of a zero-trace matrix is the linearization near the identity of the polar factorization of matrices with unit determinant.

At the Lie group level, the global Cartan decomposition also takes the name of *KAK decomposition*: there exist a compact subgroup \mathcal{K} and an Abelian subgroup \mathcal{A} of G such that

$$G = \mathcal{K} \cdot \mathcal{A} \cdot \mathcal{K}, \tag{9}$$

so that every $g \in G$ can be written as the product $g = k_1 \cdot a \cdot k_2$ with $k_1, k_2 \in \mathcal{K}$ and $a \in \mathcal{A}$; see [17] for additional information. When $G = SL(n, \mathbb{R})$, $\mathcal{K} = SO(n, \mathbb{R})$ and \mathcal{A} is the subgroup of diagonal matrices with unit determinant we obtain the foremost example of KAK decomposition, namely, the singular value decomposition (SVD):

$$A = U\Sigma V^T,$$

with U and V real orthogonal $n \times n$ matrices and Σ the diagonal matrix with the singular values (in non-increasing order) down the main diagonal. Rewriting the SVD as $A = (UV^T) \cdot (V\Sigma V^T)$ we recover the polar factorization $A = QP$ with $Q = UV^T \in SO(n, \mathbb{R})$ and $P = V\Sigma V^T = (A^T A)^{\frac{1}{2}}$ (symmetric positive definite with $\det P = 1$).

Conversely, starting from $A = QP$ with $Q \in SO(n, \mathbb{R})$ and $P = VDV^T$ positive definite (where $V \in SO(n, \mathbb{R})$ and D is diagonal positive definite with $\det D = 1$) we readily obtain $A = U\Sigma V^T$ with $U = QV \in SO(n, \mathbb{R})$ and $\Sigma = D$. It must be noted, however, that the infinitesimal variation of the SVD near the identity does not yield the same tangent space decomposition as the linearization of the polar decomposition, and does not lead to an additive splitting of the type considered in this paper.

So far we have restricted the discussion to matrices with unit determinant. Let now $A \in GL(n, \mathbb{R})$. Since any matrix sufficiently close to the identity I_n has positive determinant, we can assume $\det(A) > 0$. For any such matrix, consider the “trivial” factorization

$$A = \frac{A}{(\det A)^{\frac{1}{n}}} \cdot (\det A)^{\frac{1}{n}} I_n. \tag{10}$$

Hence, any matrix close to I_n can be written as the product of a matrix in $SL(n, \mathbb{R})$ times a matrix of the form αI_n , with $\alpha > 0$. At the group level, this corresponds to the decomposition

$$GL_+(n, \mathbb{R}) = SL(n, \mathbb{R}) \cdot \mathbb{R}_+^*, \tag{11}$$

where $GL_+(n, \mathbb{R})$ denotes the group of all real $n \times n$ matrices with positive determinant and \mathbb{R}_+^* stands for the group of all scalar matrices of the form αI_n with $\alpha > 0$. At the Lie algebra level this amounts to the addition of the algebra of *all* scalar matrices to the algebra \mathfrak{sl}_n , corresponding to the splitting $A = A_0 + \text{Tr}(A)I_n$, where $A_0 := A - \text{Tr}(A)I_n \in \mathfrak{sl}_n$. That is, we make use again of the Levi decomposition

$$\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r}, \tag{12}$$

where \mathfrak{r} is the subalgebra of all real scalar matrices, including the zero matrix, isomorphic to \mathbb{R} ; note that this is the tangent space, at the identity matrix, to the group \mathbb{R}_+^* (as well as to the group \mathbb{R}^* of invertible scalar matrices). Using (8) we obtain the decomposition

$$\mathfrak{gl}_n = \mathfrak{so}_n \oplus \mathfrak{p}_0 \oplus \mathfrak{r}, \tag{13}$$

which is precisely (4) since $\mathfrak{p}_0 \oplus \mathfrak{r} = \mathfrak{p}_n$, the space of all $n \times n$ symmetric matrices.

The Lie algebra decomposition (13) is the infinitesimal counterpart of the group (polar) decomposition

$$GL_+(n, \mathbb{R}) = SO(n, \mathbb{R}) \cdot \exp(\mathfrak{p}_0) \cdot \mathbb{R}_+^* = SO(n, \mathbb{R}) \cdot \exp(\mathfrak{p}_n). \tag{14}$$

At the matrix level the decomposition can be written as $A = (\det A)^{-\frac{1}{n}} \cdot A \cdot (\det A)^{\frac{1}{n}} = Q \cdot P$. Here $Q \in SO(n, \mathbb{R})$ and $P = (\det A)^{\frac{1}{n}} \cdot P_1$ with P_1 such that $(\det A)^{-\frac{1}{n}} \cdot A = Q \cdot P_1$ (polar factorization); note that P_1 is symmetric positive definite, with $\det P_1 = 1$ (i.e., P_1 is the exponential of a symmetric matrix with zero trace), while $\det P = \det A$.

As is well known, the symmetric and skew-symmetric splitting (5) is the basis for different iterative methods for solving nonsingular linear systems $Ax = b$, including both stationary methods [6,11] and non-stationary/preconditioned iterations [15,20,40]. In Section 4 we take a look at analogous Lie algebra decompositions of \mathfrak{gl}_n and describe the corresponding matrix splittings.

3.2. The Zappa-Szép decomposition

We begin this subsection by recalling a purely algebraic decomposition of a group as product of two of its subgroups, known as the Zappa-Szép product, see [36,41].⁴

Definition 3.2. Let G be a group with identity element e and let H, K be subgroups of G . If $G = HK$ and $H \cap K = \{e\}$, then G is said to be an (internal) Zappa-Szép product of H and K .

Equivalently, G is the Zappa-Szép product of H and K if for every $g \in G$ there exist unique $h \in H$ and $k \in K$ such that $g = hk$. We also observe that if $G = HK$ is a Zappa-Szép product, then also $G = KH$ is. We stress that in this product neither H nor K need to be normal subgroups of G , hence this product is more general than the semidirect product of subgroups. Note that the (global) Cartan decomposition of $SL(n, \mathbb{R})$ (or $GL(n, \mathbb{R})$) is not a Zappa-Szép product since the factor $\exp(\mathfrak{p})$ is not a subgroup of $SL(n, \mathbb{R})$ or $GL(n, \mathbb{R})$.

In the case of (matrix) Lie groups, the Zappa-Szép decomposition induces a splitting of the corresponding Lie algebra as a direct sum of subalgebras, as shown by the following result. Although straightforward, we haven't been able to find this result in the literature and thus we include a proof.

Proposition 3.3. Let the Lie group G be the Zappa-Szép product of Lie subgroups $H \subset G$ and $K \subset G$, and let $\mathfrak{g}, \mathfrak{h}$ and \mathfrak{k} be the corresponding Lie algebras. Then

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}.$$

Proof. First we note that \mathfrak{h} and \mathfrak{k} are Lie subalgebras of \mathfrak{g} and therefore $\mathfrak{h} + \mathfrak{k} \subseteq \mathfrak{g}$. To show that equality holds and that $\mathfrak{h} \cap \mathfrak{k} = \{0\}$ we use a dimensionality argument. The map $(h, k) \mapsto g = hk$ is evidently a diffeomorphism of $H \times K$ onto G (as differentiable manifolds), implying that

$$\dim(G) = \dim(H \times K) = \dim(H) + \dim(K).$$

It follows that the corresponding Lie algebras must also satisfy

⁴ This decomposition is also known under a few other names, including Zappa-Rédei-Szép product, general product, knit product, exact factorization or bicrossed product.

$$\dim(\mathfrak{g}) = \dim(\mathfrak{h}) + \dim(\mathfrak{k})$$

and therefore $\mathfrak{h} \cap \mathfrak{k} = \{0\}$, showing that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}$. \square

Remark 3.4. It is worth mentioning that while \mathfrak{h} and \mathfrak{k} are Lie subalgebras of \mathfrak{g} , the direct sum in Proposition 3.3 need not be a direct sum of Lie algebras, unless the condition $[\mathfrak{h}, \mathfrak{k}] = 0$ holds. In general, it is only a direct sum of vector subspaces. An example of a direct sum decomposition of a Lie algebra is the Levi decomposition $\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r}$, since $[\mathfrak{sl}_n, \mathfrak{r}] = 0$. In order to distinguish between a direct sum decomposition of a Lie algebra and the decomposition of a Lie algebra as a direct sum of vector subspaces, which may or may not be Lie subalgebras, the symbol \oplus_L is sometimes used instead of \oplus .

An example of Zappa-Szép product is given by the QR factorization [19]. Any nonsingular matrix A can be factorized as $A = QR$ with Q orthogonal and R upper triangular with positive diagonal entries, and the factors are unique. Letting $G = GL(n, \mathbb{R})$, $H = O(n, \mathbb{R})$ and $K = UT_+(n, \mathbb{R})$, the group of upper triangular matrices with positive diagonal entries, we have that $G = HK$, a Zappa-Szép product since $H \cap K = \{I_n\}$. Moreover, in this case G is a Lie group and H, K are Lie subgroups. The corresponding Lie algebras are, respectively, $\mathfrak{gl}_n, \mathfrak{so}_n$ and \mathfrak{ut}_n (the algebra of all $n \times n$ upper triangular matrices). The foregoing Proposition 3.3 implies that

$$\mathfrak{gl}_n = \mathfrak{so}_n \oplus \mathfrak{ut}_n. \tag{15}$$

Equation (15) simply expresses the fact that *any real square matrix A can be written in one and only one way as the sum of a skew-symmetric matrix and an upper triangular matrix*. This *skew-symmetric and upper triangular splitting* is the basis for various iterative methods and preconditioners for solving nonsymmetric linear systems, such as the STS and MSTs iterations (see, e.g., [5,28,39]). The splitting takes the form

$$A = (L_0 - L_0^T) + (D + U_0 + L_0^T) \tag{16}$$

where $A = L_0 + D + U_0$ is the usual splitting of a matrix into its strictly lower triangular, diagonal, and strictly upper triangular parts. Of course, one may also consider instead the *skew-symmetric and lower triangular splitting*,

$$A = (U_0 - U_0^T) + (D + L_0 + U_0^T), \tag{17}$$

corresponding to the direct sum decomposition

$$\mathfrak{gl}_n = \mathfrak{so}_n \oplus \mathfrak{lt}_n, \tag{18}$$

which in turns arises from linearization of the Zappa-Szép decomposition

$$GL(n, \mathbb{R}) = LT_+(n, \mathbb{R}) \cdot O(n, \mathbb{R}),$$

which is sometimes called the LQ factorization ($A = LQ$ with L lower triangular with positive diagonal entries and Q orthogonal).

Remark 3.5. Similar to the case of the Cartan decomposition, we can also derive (15) without using Proposition 3.3, as follows. We first deal with the case of a matrix with unit determinant. Consider a smooth matrix path $t \mapsto A(t) = Q_+(t) \cdot R_1(t)$ in $SL(n, \mathbb{R})$ with $t \in \mathbb{R}$ in an interval around zero, $Q_+(t)$ a path in $SO(n, \mathbb{R})$ and $R_1(t)$ a path in $SUT_n(n, \mathbb{R})$ (the subgroup of upper triangular matrices with unit determinant), with $Q_+(0) = R_1(0) = I_n$. Differentiating for $t = 0$ we obtain

$$A'(0) = Q'_+(0)R_1(0) + Q_+(0)R'_1(0) = Q'_+(0) + R'_1(0).$$

This shows that any matrix in the Lie algebra \mathfrak{sl}_n can be written, uniquely, as the sum of a skew-symmetric matrix, $Q'_+(0) \in \mathfrak{so}_n$, and an upper triangular matrix with zero trace, $R'_1(0) \in \mathfrak{sut}_0$, hence

$$\mathfrak{sl}_n = \mathfrak{so}_n \oplus \mathfrak{sut}_0. \tag{19}$$

To obtain the general case we make use again of the Levi decomposition (3) and observe that

$$\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r} = \mathfrak{so}_n \oplus \mathfrak{sut}_0 \oplus \mathfrak{r},$$

where \mathfrak{r} is the Lie ideal of \mathfrak{gl}_n consisting of all the scalar matrices. Observing that

$$\mathfrak{sut}_0 \oplus \mathfrak{r} = \mathfrak{ut}_n,$$

the Lie algebra of all $n \times n$ upper triangular matrices, we obtain the direct sum decomposition (15). This is the infinitesimal counterpart of the (global) QR factorization of a nonsingular matrix.

Remark 3.6. It is worth noting that the vector subspaces \mathfrak{ut}_n and \mathfrak{lt}_n , besides being Lie algebras with respect to the commutator product, are also Jordan algebras with respect to the Jordan product. This is not surprising in view of the fact that every associative algebra is automatically a Lie algebra under the commutator product and a Jordan algebra under the Jordan product [32], and both \mathfrak{ut}_n and \mathfrak{lt}_n are associative algebras under the standard matrix product. We explicitly note, however, that the direct sum decompositions (15) and (18) of \mathfrak{gl}_n are not orthogonal with respect to the Frobenius inner product.

A further, simple example of Zappa-Szép decomposition is the decomposition of the group $GL_+(n, \mathbb{R})$ given by (10)-(11). As already pointed out, linearizing this decomposition near the identity I_n yields the Lie algebra splitting

$$\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r},$$

which is precisely the Levi decomposition (12) of \mathfrak{gl}_n , corresponding to the matrix splitting $A = (A - \text{Tr}(A)I_n) + \text{Tr}(A)I_n$.

3.3. The Iwasawa decomposition

The QR factorization also occurs as a special case of the *Iwasawa decomposition*; see for example [22,26], also referred to as the ‘KAN’ decomposition. If G is a connected semisimple Lie group, there exists a compact subgroup $\mathcal{K} \subset G$, an Abelian subgroup $\mathcal{A} \subset G$, and a nilpotent subgroup $\mathcal{N} \subset G$ such that

$$G = \mathcal{K} \cdot \mathcal{A} \cdot \mathcal{N}. \tag{20}$$

Consider the case where $G = SL(n, \mathbb{R})$. Let $\mathcal{K} = SO(n, \mathbb{R})$, let \mathcal{A} be the subgroup of all diagonal matrices with positive diagonal entries and unit determinant, and $\mathcal{N} = UT_1(n, \mathbb{R})$ (the subgroup of all $n \times n$ unit triangular matrices). These are all connected Lie subgroups of $SL(n, \mathbb{R})$ with \mathcal{K} compact, \mathcal{A} Abelian, and \mathcal{N} nilpotent. We can take a smooth path $t \mapsto A(t)$ in $SL(n, \mathbb{R})$ passing through the origin of the form

$$A(t) = Q_+(t) \cdot D_+(t) \cdot U(t), \tag{21}$$

with $Q_+(t) \in \mathcal{K}$, $D_+(t) \in \mathcal{A}$, $U(t) \in \mathcal{N}$ and $Q_+(0) = D_+(0) = U(0) = I_n$. This is just the QR factorization written in QDR form, where the ‘D’ factor is diagonal and the ‘R’ factor is unit upper triangular. Differentiating (21) at $t = 0$ we obtain the splitting

$$A'(0) = Q'_+(0) + D'_+(0) + U'(0)$$

with $Q'_+(0) \in \mathfrak{so}_n$, $D'_+(0)$ a diagonal matrix, and $U'(0)$ strictly upper triangular. Hence, the Iwasawa decomposition of the group $SL(n, \mathbb{R})$ is just the QDR factorization, and its linearization near the identity yields the Iwasawa decomposition of the semisimple Lie algebra \mathfrak{sl}_n :

$$\mathfrak{sl}_n = \mathfrak{so}_n \oplus \mathfrak{d}_0 \oplus \mathfrak{ut}_0, \tag{22}$$

where \mathfrak{d}_0 denotes the subalgebra of diagonal matrices with zero trace and \mathfrak{ut}_0 that of strictly upper triangular matrices. Noting that $\mathfrak{d}_0 \oplus \mathfrak{ut}_0 = \mathfrak{sut}_0$, the Lie algebra of all upper triangular matrices with zero trace, we recover the splitting (19).

For the general case ($G = GL(n, \mathbb{R})$, with Lie algebra \mathfrak{gl}_n) we follow the usual script; using the Levi decomposition (3) we obtain the direct sum decomposition

$$\mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathfrak{r} = \mathfrak{so}_n \oplus \mathfrak{d}_0 \oplus \mathfrak{ut}_0 \oplus \mathfrak{r} = \mathfrak{so}_n \oplus \mathfrak{ut}_n, \tag{23}$$

i.e., the splitting of an arbitrary square matrix as the sum of a skew-symmetric matrix and an upper triangular matrix, both uniquely determined: if $A = L_0 + D + U_0$, then $A = (L_0 - L_0^T) + (D + U_0 + L_0^T)$, exactly as in subsection 3.2. This splitting is the basis for the method first studied in [5]. We will come back to this method in Section 6.1.

3.4. The LU decomposition and the Gauss-Seidel and Jacobi splittings

Perhaps the best known factorization of a square matrix is the triangular factorization, or LU decomposition. It is a basic fact that if all the leading principal minors of $A \in GL(n, \mathbb{R})$ are nonzero, there exist a unit lower triangular matrix L and an upper triangular matrix U such that $A = LU$, and this decomposition is unique. This is sometimes called the *Doolittle factorization* of A ; in alternative, we can write $A = LU$ with L lower triangular and U unit upper triangular (this is known as the *Crout factorization*).

It is sometimes convenient to express the factorization in LDU form:

$$A = LDU, \quad (24)$$

where L and U are unit triangular and D is the diagonal matrix of pivots. Then $A = L \cdot (DU)$ is the Doolittle factorization, and $A = (LD) \cdot U$ is Crout.

It is also well known that if $A \in GL(n, \mathbb{R})$ does not have the LU factorization, there exists a permutation matrix P such that PA has the LU (or LDU) factorization or, equivalently,

$$A = P^T LU. \quad (25)$$

However, this factorization is no longer unique in general. It is also well known that when implementing the LU factorization on a computer a permutation matrix P , representing row interchanges (partial pivoting), is typically incorporated as a means to promote numerical stability even though the matrix A may have the LU factorization. In this paper, however, we leave finite precision and numerical stability issues aside.

While the LU (or LDU) factorization is a staple of numerical linear algebra, it is much less used in the context of Lie theory. In its place, the *Bruhat decomposition* plays an important role. The Bruhat decomposition of a nonsingular matrix has the form

$$A = U_1 P U_2, \quad (26)$$

with U_1, U_2 invertible upper triangular matrices and P a permutation matrix.⁵ At the group level, the Bruhat decomposition is

$$GL(n, \mathbb{R}) = UT(n, \mathbb{R}) \cdot S_n \cdot UT(n, \mathbb{R}), \quad (27)$$

⁵ Some authors define the Bruhat decomposition using lower triangular instead of upper triangular factors. In practice, this makes no difference.

where $UT(n, \mathbb{R})$ is the group of all invertible upper triangular matrices and S_n denotes the group of all $n \times n$ permutation matrices (isomorphic to the symmetric group on n objects). A variant is the *modified Bruhat decomposition*, defined in [37]:

$$A = L \Pi U,$$

with L lower triangular, Π a permutation matrix, and U upper triangular.⁶

Unfortunately, none of these triangular factorizations is a Zappa-Szép decomposition of $GL(n, \mathbb{R})$. Depending on the decomposition, either existence or uniqueness of the decomposition is not satisfied. Specifically, existence may fail for the LU factorization, and uniqueness for its permuted versions. The set of invertible matrices for which the LDU factorization exists (and is unique), however, is dense in $GL(n, \mathbb{R})$ and moreover, the LDU factorization exists and is unique for all matrices $A \in GL(n, \mathbb{R})$ that are sufficiently close to the identity I_n ; in other words, there exists $\varepsilon > 0$ such that if $A \in B(I_n, \varepsilon)$, the open ball of radius ε centered at I_n , then $A = LU$ (or $A = LDU$) exists and is unique.⁷ Indeed, the determinant is a continuous function, hence a matrix sufficiently close to I_n will have all its leading principal minors different from zero, guaranteeing the existence and uniqueness of the LU (or LDU) factorization. Therefore, we can linearize the LDU factorization near the identity matrix to determine its infinitesimal counterpart.

To this end, consider for sufficiently small $|t|$ a smooth path of invertible matrices $t \mapsto A(t) \in B(I_n, \varepsilon)$ of the form

$$A(t) = L(t)U(t) \quad \text{with} \quad L(0) = I_n, \quad U(0) = I_n, \tag{28}$$

where $L(t) \in LT_1(n, \mathbb{R})$ and $U(t) \in UT(n, \mathbb{R})$. Differentiating (34) at $t = 0$ we obtain

$$A'(0) = L'(0) + U'(0) \tag{29}$$

with $L'(0)$ strictly lower triangular and $U'(0)$ upper triangular. At the Lie algebra level we have the direct sum decomposition

$$\mathfrak{gl}_n = \mathfrak{lt}_0 \oplus \mathfrak{ut}_n. \tag{30}$$

Letting $A'(0) = A$, $L'(0) = L_0$ and $U'(0) = U = D + U_0$ with D diagonal and U_0 strictly upper triangular, we see that the lower and upper triangular splitting

$$A = L_0 + U \tag{31}$$

⁶ We point out that the modified Bruhat decomposition is sometimes referred to as the *Gelfand-Naimark decomposition*, see [24]; however, no reference to Gelfand or Naimark is given in [24].

⁷ Here the topology is the one defined by the operator norm induced by the vector 2-norm, but other matrix norms could be used as well, since all norms on $\mathbb{R}^{n \times n}$ induce the same topology.

is the linearization of the Doolittle form of the LU factorization. In the case of the Crout form we have again the smooth factorization (28) but now with $L(t) \in LT(n, \mathbb{R})$ (lower triangular) and $U(t) \in UT_1(n, \mathbb{R})$ (unit upper triangular). In this case, linearization near the identity yields the direct sum decomposition

$$\mathfrak{gl}_n = \mathfrak{lt}_n \oplus \mathfrak{ut}_0. \tag{32}$$

In matrix terms this corresponds to the splitting

$$A = L + U_0, \tag{33}$$

where $L = L_0 + D$.

We observe that both the Lie algebra splittings (30) and (32) are orthogonal direct sums with respect to the Frobenius inner product, since

$$\mathfrak{lt}_0 = (\mathfrak{ut}_n)^\perp \quad \text{and} \quad \mathfrak{ut}_0 = (\mathfrak{lt}_n)^\perp$$

as vector subspaces of $\mathbb{R}^{n \times n}$.

In numerical linear algebra, when D is nonsingular the splitting (33) corresponds to the Gauss-Seidel method, whereas (31) corresponds to the Gauss-Seidel method applied in reverse order. It is an interesting coincidence that linearization of the LU factorization underlying Gaussian elimination leads to the Gauss-Seidel splittings. If $t \mapsto A(t) = A(t)^T$, then for $|t|$ small enough $A(t)$ will be positive definite, the LDU factorization becomes the root-free Cholesky factorization $A = LDL^T$, and linearizing at $t = 0$ yields the symmetric matrix splitting $A'(0) = L_0 + D_0 + L_0^T$.

We can also write

$$A(t) = L(t)D(t)U(t) \tag{34}$$

with $D(t)$ diagonal such that $D(0) = I_n$, $L(t)$ unit lower triangular and $U(t)$ unit upper triangular, and obtain (31) or (33) by differentiating (34) at $t = 0$. The Lie algebra splitting can be then written as the orthogonal direct sum

$$\mathfrak{gl}_n = \mathfrak{lt}_0 \oplus \mathfrak{d} \oplus \mathfrak{ut}_0. \tag{35}$$

Writing this as $\mathfrak{gl}_n = (\mathfrak{lt}_0 \oplus \mathfrak{d}) \oplus \mathfrak{ut}_0$ or as $\mathfrak{gl}_n = \mathfrak{lt}_0 \oplus (\mathfrak{d} \oplus \mathfrak{ut}_0)$ leads to the two types of Gauss-Seidel splittings.

We note incidentally that the decomposition of the matrix group $UT(n, \mathbb{R})$ as the product of the subgroup of (invertible) diagonal matrices and that of the unit upper triangular ones and the decomposition of the matrix group $LT(n, \mathbb{R})$ as the product of the subgroup of the unit lower triangular matrices and that of (invertible) diagonal ones are both examples of Zappa-Szép products.

The splitting associated to the Jacobi stationary iteration can be written as

$$A = D_0 + (L_0 + U_0) = D_0 - (A - D_0). \quad (36)$$

If we introduce the $n(n-1)$ -dimensional vector subspace

$$\mathfrak{w}_n = \{C = (c_{ij}) \in \mathbb{R}^{n \times n} \mid c_{ii} = 0 \ \forall i = 1, \dots, n\} \quad (37)$$

consisting of all the real $n \times n$ matrices with zero diagonal entries we have that

$$\mathfrak{w}_n = \mathfrak{lt}_0 \oplus \mathfrak{ut}_0$$

and therefore the Jacobi splitting corresponds to the orthogonal direct sum

$$\mathfrak{gl}_n = \mathfrak{d} \oplus \mathfrak{w}_n. \quad (38)$$

Note that \mathfrak{w}_n is not a Lie algebra with respect to the commutator product. It is the tangent space at the identity of the matrix manifold

$$W_n = LT_1(n, \mathbb{R}) \cdot UT_1(n, \mathbb{R}),$$

product of the Lie subgroups of $SL(n, \mathbb{R})$ consisting of all unit lower and unit upper triangular matrices. Note that W_n is not a group, consistent with the fact that \mathfrak{w}_n is not an algebra.

3.5. Spectral decompositions

In linear algebra there are a number of canonical forms of matrices that can be cast as multiplicative matrix decompositions, such as the Jordan form $A = X(D+N)X^{-1}$ (with D diagonal and N a particular nilpotent matrix such that $DN = ND$) and the Schur form $A = URU^T$ with U unitary and R upper triangular. These reduce to $A = XDX^{-1}$ and $A = UDU^T$ in the diagonalizable and normal case, respectively. We refer to these as ‘spectral decompositions.’

We will not consider infinitesimal counterparts of these decompositions in this paper. The main reason is that, generically, a matrix $A \in GL(n, \mathbb{R})$ has complex eigenvalues, and thus in order to interpret these spectral decompositions as factorizations of a group into subgroups (even if only locally) would imply working in the complex group $GL(n, \mathbb{C})$ and its Lie algebra, which we do not consider in this paper. Restricting the analysis to classes of matrices with real eigenvalues does not solve the problem, since these matrix classes do not, in general, have a Lie group structure, and moreover linearizing near the identity cannot result in a global direct sum decomposition of the tangent space of $GL(n, \mathbb{R})$ (or $SL(n, \mathbb{R})$) into subspaces, since any neighborhood of the identity matrix I_n in $GL(n, \mathbb{R})$ contains matrices with non-real eigenvalues.

A possibility would be to replace the complex Schur form of A with its real Schur form. The real Schur decomposition is $A = QRQ^T$ with $Q \in O(n, \mathbb{R})$ and $R \in \mathbb{R}^{n \times n}$ block upper triangular with 2×2 diagonal blocks corresponding to complex conjugate eigenvalues and 1×1 blocks corresponding to the real eigenvalues (when present). Doing this requires extending our treatment from point to block factorizations and block splittings; this can also be done, in principle, for the LU and QR factorization. We leave these generalizations for future work.

4. Lie-Jordan splittings

As already mentioned, the splitting of a matrix into its symmetric and skew-symmetric parts is the basis for various iterative algorithms and preconditioning techniques. A prominent example of this is the well-known *Hermitian and skew-Hermitian splitting* (HSS) method:

$$\begin{cases} (\alpha I_n + \mathcal{H}(A))x^{(k+\frac{1}{2})} = (\alpha I_n - \mathcal{S}(A))x^{(k)} + b \\ (\alpha I_n + \mathcal{S}(A))x^{(k+1)} = (\alpha I_n - \mathcal{H}(A))x^{(k+\frac{1}{2})} + b \end{cases} \quad (k = 0, 1, 2, \dots), \quad (39)$$

first proposed in [6], where it is shown that if $\mathcal{H}(A)$ is positive definite, the HSS method converges to the unique solution of the linear system $Ax = b$, for any initial guess and $\alpha > 0$. In [11], the HSS method was extended to generalized saddle point problems in which the Hermitian part of the coefficient matrix of the system is possibly singular, and the use of HSS as a preconditioner was also proposed and investigated. Since then, hundreds of papers have been written on this method and its many variants; see, e.g., the treatment in [8] and references therein.

As we saw in subsection 3.1, the splitting $A = \mathcal{S}(A) + \mathcal{H}(A)$ corresponding to the direct sum decomposition (4) is a splitting of the total matrix space $\mathbb{R}^{n \times n}$ as the orthogonal sum of a Lie algebra and a Jordan algebra. This is not the only example of such a splitting; as we saw in subsection 3.4, the Gauss-Seidel splittings (30) and (32) are also of this type. We call such a splitting a *Lie-Jordan splitting*, see Definition 4.1 below. The skew-symmetric and triangular splitting and the Jacobi splitting, on the other hand, are not Lie-Jordan splittings.

Definition 4.1. We say that a matrix $A \in \mathfrak{gl}_n = \mathbb{R}^{n \times n}$ admits a (non-trivial) Lie-Jordan splitting if it can be written as $A = B + C$ (with $B \neq 0$, $C \neq 0$) where $B \in \mathfrak{g}$, $C \in \mathbb{J}$, with \mathfrak{g} a Lie subalgebra and \mathbb{J} a Jordan subalgebra such that $\mathbb{R}^{n \times n} = \mathfrak{g} \oplus \mathbb{J}$ (direct sum of vector subspaces) and $\mathfrak{g} \perp \mathbb{J}$ with respect to the Frobenius inner product.

In the next two subsections we will describe two additional examples of Lie-Jordan splittings. Although not under this name, these Lie algebra splittings (and the corresponding group-level decompositions) are well known in Lie theory; a recent treatment

can be found in Section 3 of [17], see also [18, Ch. 2]. In Section 5 we show how these splittings lead to generalizations of the HSS iterative method for solving linear systems.

A key feature of the symmetric and skew-symmetric direct sum decomposition of $\mathbb{R}^{n \times n}$ is that the Lie algebra component \mathfrak{so}_n (the skew-symmetric matrices) is the tangent space at the identity of the orthogonal group $O(n, \mathbb{R})$, which is the group of linear transformations on \mathbb{R}^n that preserve the standard (Euclidean) inner product. The Jordan algebra component (the space of symmetric matrices) is the orthogonal complement of the tangent space to $O(n, \mathbb{R})$ with respect to the Frobenius inner product $\langle A, B \rangle_F = \text{Tr}(A^T B)$. In other words, it is the normal space to $O(n, \mathbb{R})$ at the identity.

A natural question, then, is to ask what happens if the standard inner product in \mathbb{R}^n , $\langle x, y \rangle = x^T y$, is replaced with a different bilinear form, associated with a matrix $J \in \mathbb{R}^{n \times n}$:

$$b(x, y) = x^T J y$$

(which reduces to the standard inner product when $J = I_n$). Notice that, in principle, J need not be symmetric positive definite; that is, J may not define an inner product. Recalling now that the classical groups $O(p, q, \mathbb{R})$ and $Sp(n, \mathbb{R})$, like $O(n, \mathbb{R})$, can all be defined as the sets of matrices Q such that $b(Qx, Qy) = b(x, y)$ for all $x, y \in \mathbb{R}^n$ for a suitable bilinear form b , it is natural to begin by examining the cases of the pseudo-orthogonal group $O(p, q, \mathbb{R})$ and the symplectic group $Sp(n, \mathbb{R})$. We do this in the next two subsections.

For future reference, we recall that the J -adjoint of a matrix $A \in \mathbb{R}^{n \times n}$ is defined as $A^* = J^{-1} A^T J$, and that a matrix is J -symmetric if $A = A^*$ and J -skew-symmetric if $A = -A^*$.

4.1. The pseudo-symmetric and pseudo-skew-symmetric splitting

Let p and q be positive integers and let $n = p + q$. Consider the bilinear form associated with the matrix

$$J = I_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}. \tag{40}$$

The group leaving $I_{p,q}$ invariant is the *pseudo-orthogonal* (or *indefinite*) group $O(p, q, \mathbb{R})$. It can also be characterized as the group of all matrices A satisfying the equivalent conditions $I_{p,q} A^T I_{p,q} = A^{-1}$, $A^T I_{p,q} A = I_{p,q}$, or $A = I_{p,q} (A^T)^{-1} I_{p,q}$.

The Lie algebra (tangent space at the identity) of the group $O(p, q, \mathbb{R})$ consists of the (p, q) -pseudo-skew-symmetric matrices, which have the following structure (see, e.g., [1]):

$$A = \begin{bmatrix} F & G^T \\ G & H \end{bmatrix}, \quad F = -F^T \in \mathbb{R}^{p \times p}, \quad H = -H^T \in \mathbb{R}^{q \times q}, \quad G \in \mathbb{R}^{q \times p}.$$

Note that this is precisely the set of all matrices $A \in \mathbb{R}^{n \times n}$ which satisfy the condition

$$A^T I_{p,q} + I_{p,q} A = 0$$

(a linearized version of the condition $A^T I_{p,q} A = I_{p,q}$). This set of matrices is clearly a vector subspace of $\mathbb{R}^{n \times n}$, is closed under the commutator product, and it is a Lie subalgebra of dimension $\frac{p(p-1)}{2} + \frac{q(q-1)}{2} + pq = \frac{n(n-1)}{2}$ of \mathfrak{gl}_n . We shall denote this Lie algebra by $\mathfrak{so}_{p,q}$. Next, consider the set of $n \times n$ matrices of the form

$$A = \begin{bmatrix} F & G^T \\ -G & H \end{bmatrix}, \quad F = F^T \in \mathbb{R}^{p \times p}, \quad H = H^T \in \mathbb{R}^{q \times q}, \quad G \in \mathbb{R}^{q \times p}.$$

This is the set of (p, q) -pseudo-symmetric matrices. It is readily verified that the matrices in this set are precisely those that satisfy the condition

$$A^T I_{p,q} - I_{p,q} A = 0.$$

This set of matrices is again a vector subspace of $\mathbb{R}^{n \times n}$, is closed under the Jordan product, and it is a Jordan algebra, which we denote by $\mathfrak{p}_{p,q}$. Its dimension is $\frac{p(p+1)}{2} + \frac{q(q+1)}{2} + pq = \frac{n(n+1)}{2}$, hence as a subspace of $\mathfrak{gl}_n = \mathbb{R}^{n \times n}$ it is complementary to $\mathfrak{so}_{p,q}$:

$$\mathfrak{gl}_n = \mathbb{R}^{n \times n} = \mathfrak{so}_{p,q} \oplus \mathfrak{p}_{p,q}. \tag{41}$$

The decomposition (41) states that any matrix $\mathcal{A} \in \mathbb{R}^{(p+q) \times (p+q)}$ can be written in one and only one way as the sum of a matrix $H \in \mathfrak{p}_{p,q}$ and a matrix $S \in \mathfrak{so}_{p,q}$. The splitting is given by

$$\mathcal{A} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(A + A^T) & \frac{1}{2}(B - C^T) \\ -\frac{1}{2}(B^T - C) & \frac{1}{2}(D + D^T) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}(A - A^T) & \frac{1}{2}(B + C^T) \\ \frac{1}{2}(B^T + C) & \frac{1}{2}(D - D^T) \end{bmatrix} = H + S. \tag{42}$$

We call the splitting (42) the *pseudo-symmetric and pseudo-skew-symmetric splitting*. It is easily checked that $H \in \mathfrak{p}_{p,q}$, $S \in \mathfrak{so}_{p,q}$, and $\text{Tr}(H^T S) = 0$. Indeed, (41) is an orthogonal direct sum decomposition, and (42) is an example of a Lie-Jordan splitting. It depends, of course, on the choice of p and q (where $p + q = n$). In the case $p = n$ and $q = 0$, we have $O(p, q, \mathbb{R}) = O(n, \mathbb{R})$, $\mathcal{A} = A \in \mathbb{R}^{n \times n}$, and (41)-(42) reduce to the usual symmetric and skew-symmetric splitting (4)-(5).

In Section 3.1 we saw how the symmetric and skew-symmetric splitting can be interpreted as the infinitesimal version of the polar factorization. It is thus natural to ask if the pseudo-symmetric and pseudo-skew-symmetric splitting corresponds to a kind of polar factorization. In terms of Lie algebras and Lie groups, noting that (41) is a decomposition of \mathfrak{gl}_n into eigenspaces corresponding to the ± 1 eigenvalues of the involution

$\theta(X) = -I_{p,q}A^T I_{p,q}$, it is natural to ask if there exists an associated global decomposition of $GL(n, \mathbb{R})$, of which (41) is the infinitesimal counterpart. This *indefinite polar factorization*, if it exists, must take the form

$$A = QP \quad \text{where } A \in GL(n, \mathbb{R}), Q \in O(p, q, \mathbb{R}), P \in \exp(\mathfrak{p}_{p,q}). \quad (43)$$

The idempotent diffeomorphism Θ leaving $O(p, q, \mathbb{R})$ fixed is given by $\Theta(X) = I_{p,q}(X^T)^{-1}I_{p,q}$.

As it turns out, the decomposition (43) exists under certain conditions on A . These conditions have been established in [23], where the *generalized polar decomposition* is investigated; see also [30]. In particular, specialization of Theorem 4.1 in [23] to the pseudo-orthogonal group implies that $A \in GL(n, \mathbb{R})$ has the factorization (43) if and only if the pseudo-symmetric matrix $A^*A = I_{p,q}A^T I_{p,q}A$ has no eigenvalues on the negative real axis.⁸ Hence, for a “generic” $A \in GL(n, \mathbb{R})$ the generalized polar decomposition (43) exists, in the sense that the set of matrices for which the decomposition exists is dense in $GL(n, \mathbb{R})$. Moreover, it is easily seen that as A approaches the identity matrix in $GL(n, \mathbb{R})$, $I_{p,q}A^T I_{p,q}A$ also tends to the identity and therefore has no eigenvalues on the negative real axis. In other words, there exists an open neighborhood of the identity matrix in $GL(n, \mathbb{R})$ such that every matrix in this neighborhood has the generalized polar decomposition (43). We can therefore linearize this factorization near the identity to obtain the splitting (42). The situation is similar to the one we encountered when discussing the LDU factorization, which always exists in a sufficiently small neighborhood of the identity.

Remark 4.2. It is natural to introduce a “pseudo-HSS iteration”, i.e., an iterative method of the form (39) but based on the pseudo-symmetric and pseudo-skew-symmetric splitting (42), and to seek convergence conditions for it. We do this in Section 5.

4.2. The Hamiltonian and skew-Hamiltonian splitting

Let $n = 2m$ (even) and consider the anti-symmetric bilinear form associated with the matrix

$$J = \Sigma_n = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}. \quad (44)$$

The group leaving Σ_n invariant is the *symplectic group* $Sp(n, \mathbb{R})$. These are the matrices A that satisfy the equivalent conditions $\Sigma_n A^T \Sigma_n = -A^{-1}$, $A^T \Sigma_n A = \Sigma_n$, or $A = -\Sigma_n(A^T)^{-1}\Sigma_n$.

⁸ Note that if either $p = 0$ or $q = 0$, then $A^*A = I_{p,q}A^T I_{p,q}A = A^T A$, which is positive definite and therefore satisfies the condition for the existence of the (standard) polar factorization, for any A .

The Lie algebra of the group $Sp(n, \mathbb{R})$ consists of the *Hamiltonian matrices*, which have the following structure:

$$A = \begin{bmatrix} F & G \\ H & -F^T \end{bmatrix}, \quad F \in \mathbb{R}^{m \times m}, \quad G = G^T \in \mathbb{R}^{m \times m}, \quad H = H^T \in \mathbb{R}^{m \times m},$$

see [1]. Note that this is precisely the set of all matrices $A \in \mathbb{R}^{n \times n}$ which satisfy the condition

$$A^T \Sigma_n + \Sigma_n A = 0$$

(a linearized version of the condition $A^T \Sigma_n A = \Sigma_n$). This set of matrices is clearly a vector subspace of $\mathbb{R}^{n \times n}$, is closed under the commutator product, and it is a Lie subalgebra of dimension $m(m + 1) + m^2 = 2m^2 + m$ of \mathfrak{gl}_n . We shall denote this Lie algebra by \mathfrak{sp}_n . Next, consider the set of $n \times n$ matrices of the form

$$A = \begin{bmatrix} F & G \\ H & F^T \end{bmatrix}, \quad F \in \mathbb{R}^{m \times m}, \quad G = -G^T \in \mathbb{R}^{m \times m}, \quad H = -H^T \in \mathbb{R}^{m \times m}.$$

A matrix of this type is said to be *skew-Hamiltonian*. It is readily verified that the skew-Hamiltonian matrices are precisely those that satisfy the condition

$$A^T \Sigma_n - \Sigma_n A = 0.$$

This set is again a vector subspace of $\mathbb{R}^{n \times n}$, and it forms a Jordan algebra with respect to the Jordan product. We denote it by \mathfrak{j}_n . Its dimension is $m(m - 1) + m^2 = 2m^2 - m$, hence as a subspace of $\mathfrak{gl}_n = \mathbb{R}^{n \times n}$ it is complementary to \mathfrak{sp}_n :

$$\mathfrak{gl}_n = \mathbb{R}^{n \times n} = \mathfrak{sp}_n \oplus \mathfrak{j}_n. \tag{45}$$

The decomposition (45) states that any matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ with $n = 2m$ can be written in one and only one way as the sum of a matrix $S \in \mathfrak{sp}_n$ and a matrix $H \in \mathfrak{j}_n$. The splitting is given by

$$\mathcal{A} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(A - D^T) & \frac{1}{2}(B + B^T) \\ \frac{1}{2}(C + C^T) & -\frac{1}{2}(A^T - D) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}(A + D^T) & \frac{1}{2}(B - B^T) \\ \frac{1}{2}(C - C^T) & \frac{1}{2}(A^T + D) \end{bmatrix} = S + H. \tag{46}$$

We call the splitting (46) the *Hamiltonian and skew-Hamiltonian splitting*. It is only defined for n even. It is easily checked that $S \in \mathfrak{sp}_n$ and $H \in \mathfrak{j}_n$, and that $\text{Tr}(S^T H) = 0$. Indeed, (45) is an orthogonal direct sum decomposition, and (46) is another example of a Lie-Jordan splitting.

We note that (45) is a decomposition of \mathfrak{gl}_n into eigenspaces corresponding to the eigenvalues ± 1 of the involution $\theta(X) = \Sigma_n X^T \Sigma_n$, with the idempotent diffeomorphism Θ leaving $Sp(n, \mathbb{R})$ fixed being given by $\Theta(X) = -\Sigma_n (X^T)^{-1} \Sigma_n$. As in the previous subsection, specializing Theorem 4.1 in [23] to the symplectic group implies that a matrix $A \in GL(n, \mathbb{R})$ (n even) has a “symplectic” polar factorization of the form

$$A = WZ \quad \text{where } W \in Sp(n, \mathbb{R}), Z \in \exp(\mathfrak{j}_n) \quad (47)$$

if and only if the skew-Hamiltonian matrix $-\Sigma_n A^T \Sigma_n A$ has no eigenvalues on the negative real axis. Again, this is a condition that holds generically. It is also easy to see that as A approaches the identity matrix in $GL(n, \mathbb{R})$, so does $-\Sigma_n A^T \Sigma_n A$, hence there is a neighborhood of the identity in $GL(n, \mathbb{R})$ such that every matrix in this neighborhood has the symplectic polar factorization (47). Therefore, the splitting (46) is obtained as the linearization near the identity of the factorization (47).

Remark 4.3. As in the case of the pseudo-symmetric and pseudo-skew-symmetric splitting, we can introduce an HSS-like iteration, i.e., an iterative method of the form (39) but based on the Hamiltonian and skew-Hamiltonian splitting (46), and to seek convergence conditions for it. We do this in Section 5.

4.3. Quadratic algebra decomposition

The symmetric and skew-symmetric, pseudo-symmetric and pseudo-skew-symmetric, and the Hamiltonian and skew-Hamiltonian splittings are all instances of what we called Lie-Jordan splittings; the first one corresponds to the Cartan decomposition of \mathfrak{gl}_n , the other two to decompositions that are not Cartan but only “Cartan-like,” all well known in Lie theory. Moreover, their global variants are all types of polar factorizations which are known in numerical linear algebra. In this subsection we discuss one possible unified treatment of these splittings, showing that they are all special cases of a general construction. To this end we recall the definition of a J -quadratic Lie algebra and the corresponding notion of a J -quadratic Jordan algebra, slightly more general than the definition in [14].

Here we use the notation \mathfrak{gl}_n for the space of all $n \times n$ matrices $\mathbb{R}^{n \times n}$ endowed with the Lie algebra structure given by the commutator product, and the notation $\mathbb{M}(n, \mathbb{R})$ for the same space endowed with the Jordan algebra structure given by the Jordan product.

Let J be an invertible matrix in $\mathbb{R}^{n \times n}$ such that $J = J^T$ or and $J = -J^T$.

Definition 4.4. Given J as above, a Lie subalgebra \mathfrak{g} of \mathfrak{gl}_n is called a J -quadratic Lie algebra if

$$W \in \mathfrak{g} \iff W^T J + JW = 0. \quad (48)$$

Related to a J -quadratic Lie algebra, we have the notion of J -quadratic Jordan algebra.

Definition 4.5. Given J as above, a Jordan subalgebra \mathbb{J} of $\mathbb{M}(n, \mathbb{R})$ is called a J -quadratic Jordan algebra if

$$W \in \mathbb{J} \iff W^T J - JW = 0. \tag{49}$$

Remark 4.6. We observe that up to a change of coordinates, any J -quadratic Lie algebra in the hypothesis above is isomorphic to either \mathfrak{so}_n , $\mathfrak{so}_{p,q}$ or \mathfrak{sp}_n .

5. The J -HSS iteration

In this Section we obtain a convergence result for a generalization of the HSS method for solving linear systems which we call the J -HSS iteration; it can be described as an *alternating Lie-Jordan iteration*.

5.1. A general convergence result

Let us consider the linear system of equations in \mathbb{R}^n :

$$Ax = b, \tag{50}$$

where $A \in GL(n, \mathbb{R})$. For any J invertible symmetric or skew-symmetric, we can define an involution on $\mathbb{R}^{n \times n}$ as $A \mapsto -JA^T J^{-1}$. This map is an isometry with respect to the positive definite form $\langle A, B \rangle = \text{Tr}(A^T J^T J B)$. By Theorem 3.14 in [2], we have the following splitting:

$$\mathbb{R}^{n \times n} = \mathfrak{g} \oplus \mathbb{J}, \tag{51}$$

where \mathfrak{g} is the J -quadratic Lie algebra, corresponding to the eigenspace of the involution with eigenvalue 1, and \mathbb{J} the J -quadratic Jordan algebra, corresponding to the eigenspace of the involution with eigenvalue -1 . More explicitly, we can define the J -symmetric part of A by $H = \frac{1}{2}(A + JA^T J^{-1}) \in \mathbb{J}$ and its J -skew-symmetric part by $S = \frac{1}{2}(A - JA^T J^{-1}) \in \mathfrak{g}$.

Consider the following alternating (two-step) scheme, which we refer to as the J -HSS iteration:

$$\begin{cases} (H + \alpha J^{-1})x^{k+\frac{1}{2}} &= (\alpha J^{-1} - S)x^k + b \\ (S + \alpha J^{-1})x^{k+1} &= (\alpha J^{-1} - H)x^{k+\frac{1}{2}} + b \end{cases} \quad (k = 0, 1, 2, \dots). \tag{52}$$

Proposition 5.1. Let $A, J \in \mathbb{R}^{n \times n}$ be invertible. Write $A = H + S$ where $H \in \mathbb{J}$ and $S \in \mathfrak{g}$, where \mathbb{J} and \mathfrak{g} are J -quadratic Jordan and Lie algebras (resp.) such that (51) holds.

- (i) If $J = J^T$, and the symmetric matrix HJ is positive definite, then the iteration (52) is well-defined and converges to the unique solution of $Ax = b$ for all $\alpha > 0$ and $x^0 \in \mathbb{R}^n$.
- (ii) If $J = -J^T$, and the symmetric matrix SJ is positive definite, then the iteration (52) is well-defined and converges to the unique solution of $Ax = b$ for all $\alpha > 0$ and $x^0 \in \mathbb{R}^n$. In this case, n is necessarily even.

Proof. The linear system $Ax = b$ can be rewritten as

$$AJy = b, \quad x = Jy.$$

Then

$$AJ = (H + S)J = HJ + SJ. \tag{53}$$

If $J = J^T$, then $HJ = \frac{1}{2}(AJ + JA^T)$ is symmetric and $SJ = \frac{1}{2}(AJ - JA^T)$ is skew-symmetric, hence the splitting (53) is just the standard HSS splitting of AJ . Moreover, $H + \alpha J^{-1} = (HJ + \alpha I)J^{-1}$, hence $H + \alpha J^{-1}$ is invertible (as the product of invertible matrices) when HJ is positive definite; moreover, $S + \alpha J^{-1} = (SJ + \alpha I)J^{-1}$ is invertible for all $\alpha \neq 0$, since SJ is skew-symmetric, and the iteration is well-defined. By Theorem 2.2 in [6], the sequence $y^k = J^{-1}x^k$ with x^k defined by iteration (52) converges, for all $\alpha > 0$ and $y^0 = J^{-1}x^0$, to the unique solution y^* of $AJy = b$, hence x^k converges to the unique solution x^* of $Ax = b$. This proves part (i).

To prove (ii), note that when $J = -J^T$ then HJ and SJ in (53) are, respectively, skew-symmetric and symmetric. The remainder of the proof is analogous to the one of part (i). Note that in this case n must be even, since J must be nonsingular. \square

It is instructive to also prove Proposition 5.1 directly, without relying on Theorem 2.2 in [6]. Let $J = J^T$ and assume HJ is positive definite. Eliminating $x^{k+\frac{1}{2}}$ from (52) we can rewrite the J -HSS iteration as $x^{k+1} = T_\alpha x^k + c$, where

$$\begin{aligned} T_\alpha &= (\alpha I + JS)^{-1}(\alpha I - JH)(JH + \alpha I)^{-1}(\alpha I - JS) \\ &= (\alpha I + SJ)^{-1}(\alpha I - HJ)(HJ + \alpha I)^{-1}(\alpha I - SJ) \end{aligned}$$

and

$$c = 2\alpha(\alpha I + JS)^{-1}(JH + \alpha I)^{-1}Jb = 2\alpha J(SJ + \alpha I)^{-1}(HJ + \alpha I)^{-1}b. \tag{54}$$

We notice that T_α is similar to

$$\widehat{T}_\alpha = (\alpha I - HJ)(HJ + \alpha I)^{-1}(\alpha I - SJ)(\alpha I + SJ)^{-1},$$

hence the spectral radius of T_α is equal to that of \widehat{T}_α . Since $(SJ - \alpha I)(SJ + \alpha I)^{-1}$ is orthogonal as the Cayley transform of a skew-symmetric matrix, we get

$$\varrho(T_\alpha) = \varrho(\widehat{T}_\alpha) \leq \|\widehat{T}_\alpha\|_2 \leq \|(\alpha I - HJ)(\alpha I + HJ)^{-1}\|_2 = \max_{1 \leq i \leq n} \frac{|\alpha - \lambda_i(HJ)|}{|\alpha + \lambda_i(HJ)|} < 1,$$

where $\lambda_i(HJ)$ denotes the i -th eigenvalue of the symmetric matrix HJ . Hence the iteration converges, for all $\alpha > 0$ and all x^0 . Moreover, the upper bound $\max_{1 \leq i \leq n} \frac{|\alpha - \lambda_i(HJ)|}{|\alpha + \lambda_i(HJ)|}$ is minimized by

$$\alpha_* = \sqrt{\lambda_{\min}(HJ)\lambda_{\max}(HJ)}.$$

The case $J = -J^T$ is analogous, interchanging the roles of HJ and SJ .

Remark 5.2. When $J = I_n$, the J -HSS iteration reduces to the standard HSS iteration (39) introduced in [6]. Also note that the J -HSS iteration is mathematically equivalent to the standard HSS iteration applied to the linear system $AJy = b$, $y = J^{-1}x$, when the splitting $A = H + S$ is a Lie-Jordan splitting with $S \in \mathfrak{g}$ and $H \in \mathfrak{J}$, where \mathfrak{g} and \mathfrak{J} are, respectively, J -quadratic Lie and Jordan algebras.

Remark 5.3. The first half-step of the J -HSS iteration requires solving linear systems with coefficient matrix $H + \alpha J^{-1}$. Writing $H + \alpha J^{-1} = (HJ + \alpha I)J^{-1}$, this reduces to solving a system with symmetric positive definite coefficient matrix $HJ + \alpha I$ followed by multiplication by J . Since J is usually of very simple form, the cost of the first half-step is essentially identical to that for the standard HSS method. Likewise, the second half-step requires solving systems with coefficient matrix $SJ + \alpha I$, where SJ is skew-symmetric, followed by a matrix-vector product with J , again essentially identical to the standard HSS iteration in terms of cost.

Remark 5.4. The single-step iteration $x^{k+1} = T_\alpha x^k + c$ can also be written as

$$x^{k+1} = (I - P_\alpha^{-1}A)x^k + P_\alpha^{-1}b$$

with

$$P_\alpha^{-1} = 2\alpha J(SJ + \alpha I)^{-1}(HJ + \alpha I)^{-1}b, \quad \text{or} \quad P_\alpha = \frac{1}{2\alpha}(HJ + \alpha I)(SJ + \alpha I)J^{-1}, \tag{55}$$

see for instance [13]. We refer to the matrix P_α as the J -HSS preconditioner. Its application requires the solution of two shifted linear systems, one of which is symmetric positive definite and the other shifted skew-symmetric, and a matrix-vector product with J . The scalar factor $\frac{1}{2\alpha}$ is immaterial for preconditioning, and it can be ignored. Since the matrix J is usually of very simple form, the cost is essentially the same as for the standard HSS

preconditioner. When used to accelerate the convergence of Krylov subspace methods, exact solves are usually replaced with inexact ones (see, e.g., [7,11]).

Remark 5.5. We mention that in [21], the author has proposed two variants of the HSS scheme that bear some resemblance to the J -HSS method, but should not be confused with it. The first variant uses the (standard) HSS splitting $A = H + S$ and considers the iteration obtained by replacing shifts of the form αI with αJ where $J = \Sigma_n$, with n even. In the second variant the symmetric matrix H is shifted by αI in the first half step, while the skew-symmetric matrix S is shifted by αJ in the second half step (here, again, $J = \Sigma_n$). For these iterations to be well-defined, restrictions must be imposed on the parameter α or on the matrix S ; moreover, convergence conditions appear to be difficult to establish, see [21].

In the next two subsections we discuss the two choices of the matrix J from subsections 4.1 and 4.2, which provide ‘canonical’ examples of the two cases covered in Proposition 5.1.

5.2. Pseudo-HSS method

When $J = I_{p,q}$ (see (40)), we refer to the J -HSS iteration as the *pseudo-HSS* iteration. Here we assume that $p \neq 0, q \neq 0$, for otherwise we revert to the standard HSS iteration. As shown in subsection 4.1, every $A \in \mathbb{R}^{n \times n}$ (with $n = p + q$) can be split as

$$A = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(A + A^T) & \frac{1}{2}(B - C^T) \\ -\frac{1}{2}(B^T - C) & \frac{1}{2}(D + D^T) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}(A - A^T) & \frac{1}{2}(B + C^T) \\ \frac{1}{2}(B^T + C) & \frac{1}{2}(D - D^T) \end{bmatrix} = H + S.$$

We have

$$HJ = \begin{bmatrix} \frac{1}{2}(A + A^T) & -\frac{1}{2}(B - C^T) \\ -\frac{1}{2}(B^T - C) & -\frac{1}{2}(D + D^T) \end{bmatrix}, \tag{56}$$

a symmetric matrix. For this matrix to be positive definite, it is necessary that both $\frac{1}{2}(A + A^T)$ and $-\frac{1}{2}(D + D^T)$ be positive definite. A necessary and sufficient condition is that both $\frac{1}{2}(A + A^T)$ and the Schur complement matrix

$$-\frac{1}{2}(D + D^T) - \frac{1}{2}(B^T - C)(A + A^T)^{-1}(B - C)^T \tag{57}$$

be positive definite. In general, this condition is not easy to check. An exception is when $C = B^T$, in which case

$$HJ = \begin{bmatrix} \frac{1}{2}(A + A^T) & 0 \\ 0 & -\frac{1}{2}(D + D^T) \end{bmatrix},$$

a block diagonal matrix, which is positive definite when both A and $-D$ are positive real. In this case we also have

$$SJ = \begin{bmatrix} \frac{1}{2}(A - A^T) & -B \\ B^T & -\frac{1}{2}(D - D^T) \end{bmatrix}.$$

Hence, when $B = C^T$ the two half-steps of the pseudo-HSS iteration require solving shifted linear systems with coefficient matrices given by

$$HJ + \alpha I = \begin{bmatrix} \frac{1}{2}(A + A^T) + \alpha I & 0 \\ 0 & \alpha I - \frac{1}{2}(D + D^T) \end{bmatrix}, \quad (58)$$

and

$$SJ + \alpha I = \begin{bmatrix} \frac{1}{2}(A - A^T) + \alpha I & -B \\ B^T & \alpha I - \frac{1}{2}(D - D^T) \end{bmatrix}. \quad (59)$$

If in addition $D = D^T$, as is frequently the case in the solution of saddle point problems (see [12]), we have

$$HJ + \alpha I = \begin{bmatrix} \frac{1}{2}(A + A^T) + \alpha I & 0 \\ 0 & \alpha I - D \end{bmatrix}, \quad (60)$$

a symmetric positive definite, block diagonal matrix for all $\alpha > 0$, and

$$SJ + \alpha I = \begin{bmatrix} \frac{1}{2}(A - A^T) + \alpha I & -B \\ B^T & \alpha I \end{bmatrix}. \quad (61)$$

In this special case, the resulting iterative method is essentially the one introduced in [11]. Systems with coefficient matrix given by (60) can be solved by any solver for symmetric positive definite systems. The shifted skew-symmetric systems with coefficient matrix (61) can be reduced, via Schur complement reduction, to solving $p \times p$ systems with matrix $\frac{1}{2}(A - A^T) + \alpha I + \frac{1}{\alpha}BB^T$, for which several specialized solvers exist; see, for instance, [20] and [10].

We note that while the partitioning of n as $n = p + q$ is in principle arbitrary, in practice the choice of p and q may be induced by a natural partitioning of the variables, as in saddle point problems; for example, in fluid flow problems p could be the number of velocity degrees of freedom and q the number of pressure degrees of freedom. In other cases the choice of p and q may be dictated by computational considerations.

5.3. Hamiltonian-HSS method

When n is even and $J = \Sigma_n$, we refer to the J -HSS iteration as the *Hamiltonian-HSS* iteration. As shown in subsection 4.2, every $\mathcal{A} \in \mathbb{R}^{n \times n}$ can be split as

$$\mathcal{A} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(A - D^T) & \frac{1}{2}(B + B^T) \\ \frac{1}{2}(C + C^T) & -\frac{1}{2}(A^T - D) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}(A + D^T) & \frac{1}{2}(B - B^T) \\ \frac{1}{2}(C - C^T) & \frac{1}{2}(A^T + D) \end{bmatrix} = S + H,$$

with $S \in \mathfrak{sp}_n$ (Hamiltonian) and $H \in \mathfrak{j}_n$ (skew-Hamiltonian).

We have

$$HJ = \begin{bmatrix} -\frac{1}{2}(B - B^T) & \frac{1}{2}(A + D^T) \\ -\frac{1}{2}(A^T + D) & \frac{1}{2}(C - C^T) \end{bmatrix},$$

a skew-symmetric matrix, and

$$SJ = \begin{bmatrix} -\frac{1}{2}(B + B^T) & \frac{1}{2}(A - D^T) \\ \frac{1}{2}(A^T - D) & \frac{1}{2}(C + C^T) \end{bmatrix},$$

a symmetric matrix, which must be positive definite for the Hamiltonian-HSS method to converge for all $\alpha > 0$. In particular, a necessary condition is that B and C have, respectively, negative definite and positive definite symmetric part. In the special case in which $B = B^T$, $C = C^T$ and $A = D^T$ these matrices reduce to

$$HJ = \begin{bmatrix} 0 & A \\ -A & 0 \end{bmatrix},$$

and

$$SJ = \begin{bmatrix} -B & 0 \\ 0 & C \end{bmatrix},$$

and the Hamiltonian-HSS scheme reduces to a special case of the standard HSS method, which converges for all $\alpha > 0$ if B is negative and C is positive definite. When $C = -B$ and $A = D$ we have

$$\mathcal{A} = \begin{bmatrix} A & B \\ -B & A \end{bmatrix},$$

which is one of the real forms of the complex matrix $C = A + iB$. In this case we have

$$AJ = \begin{bmatrix} -B & A \\ -A & -B \end{bmatrix} = HJ + SJ,$$

where

$$HJ = \begin{bmatrix} -\frac{1}{2}(B - B^T) & \frac{1}{2}(A + A^T) \\ -\frac{1}{2}(A + A^T) & -\frac{1}{2}(B - B^T) \end{bmatrix},$$

a skew-symmetric Hamiltonian matrix, and

$$SJ = \begin{bmatrix} -\frac{1}{2}(B + B^T) & \frac{1}{2}(A - A^T) \\ -\frac{1}{2}(A - A^T) & -\frac{1}{2}(B + B^T) \end{bmatrix}$$

a symmetric skew-Hamiltonian matrix. If in addition $A = A^T$, SJ becomes block diagonal and the Hamiltonian-HSS iteration converges if $-B$ is positive real, a condition often satisfied in applications; see, e.g., [3,4,9].

6. Further discussion of alternating iterations

In this Section we give additional details on the triangular and skew-symmetric splitting iteration (following [5], see also [8]) and we discuss a further class of splitting methods obtained by linearizing the Kronecker product decomposition. The latter class includes the well-known Alternating Direction Implicit (ADI) method as a special case.

6.1. Convergence of the triangular and skew-symmetric splitting iteration

Similar to the HSS method, the authors of [5] have proposed an alternating iteration based on the triangular and skew-symmetric splitting of A . If $A = L_0 + D + U_0$ is the usual splitting of A into its strictly lower triangular, diagonal and strictly upper triangular parts (or their block analogues), we can write $A = U + S$ with $U = U_0 + D + L_0^T$ (upper triangular) and $S = L_0 - L_0^T$ (skew-symmetric). One can then define the following alternating iteration:

$$\begin{cases} (\alpha I_n + U)x^{(k+\frac{1}{2})} = (\alpha I_n - S)x^{(k)} + b \\ (\alpha I_n + S)x^{(k+1)} = (\alpha I_n - U)x^{(k+\frac{1}{2})} + b \end{cases} \quad (k = 0, 1, 2, \dots), \tag{62}$$

where $\alpha > 0$. Note that the first subsystem in (62) is triangular, and therefore easy to solve, while the coefficient matrix subsystem is a shifted skew-symmetric matrix, as in the HSS method. The convergence analysis of this method is very similar to that of the HSS iteration. In particular, the method is unconditionally convergent if A is positive definite. Indeed, the iteration matrix T_α of this algorithm is similar to

$$\hat{T}_\alpha = (\alpha I_n - U)(\alpha I_n + U)^{-1}(\alpha I_n - S)(\alpha I_n + S)^{-1}$$

so that for all $\alpha > 0$ we have

$$\rho(T_\alpha) = \rho(\hat{T}_\alpha) \leq \|(\alpha I_n - U)(\alpha I_n + U)^{-1}\|_2 \|(\alpha I_n - S)(\alpha I_n + S)^{-1}\|_2 < 1.$$

The contractivity of $(\alpha I_n - U)(\alpha I_n + U)^{-1}$ follows from the fact that U is positive definite, since

$$\frac{1}{2}(U + U^T) = \frac{1}{2}(U_0 + D + L_0^T + U_0^T + D + L_0) = \frac{1}{2}(A + A^T).$$

Kellogg’s Lemma (see, e.g., [31]) insures that $\|(\alpha I_n - U)(\alpha I_n + U)^{-1}\|_2 < 1$; see also [5].

As already observed in Sections 3.2-3.3, the underlying Lie algebra decomposition

$$\mathfrak{gl}_n = \mathfrak{so}_n \oplus \mathfrak{ut}_n = \mathfrak{so}_n \oplus \mathfrak{d}_0 \oplus \mathfrak{ut}_0$$

is the linearization of the QR (or QDR) factorization of the general linear group $GL(n, \mathbb{R})$, a special case of the Zappa-Szép (or of the Iwasawa) decomposition. Unlike the Lie algebra decompositions considered in the previous Section, this is *not* a quadratic algebra decomposition.

6.2. ADI-like splittings

Let now $A, B \in \mathbb{R}^{n \times n}$ and consider the $n^2 \times n^2$ matrix $M = A \otimes I_n + I_n \otimes B$, the Kronecker sum of A and B . The assumption that A and B are of the same size is not necessary, and is made here only for notational simplicity. To solve the linear system $Mx = b$, let $\alpha > 0$ and consider the alternating iteration

$$\begin{cases} (\alpha I_{n^2} + A \otimes I_n)x^{(k+\frac{1}{2})} = (\alpha I_{n^2} - I_n \otimes B)x^{(k)} + b \\ (\alpha I_{n^2} + I_n \otimes B)x^{(k+1)} = (\alpha I_{n^2} - A \otimes I_n)x^{(k+\frac{1}{2})} + b \end{cases} \quad (k = 0, 1, 2, \dots). \quad (63)$$

Note that the second of the two subsystems in (63) is block diagonal (with all diagonal blocks equal to B), while the first one can be symmetrically permuted to a system of the same type. The classical Alternating Direction Implicit (ADI) method, see for instance [31,33,38], is of this form. When applied to the standard five-point difference approximation of the Poisson equation on a rectangular region (with an equal number of mesh points in each direction), $A = B$ is just the tridiagonal approximation of the 1D Laplacian and M is the discrete 2D Laplacian. As is well-known, for this problem the ADI method is unconditionally convergent.

More generally, it is straightforward to check that iteration (63) is unconditionally convergent to the solution of $Mx = b$ if A and B are positive semidefinite (not necessarily symmetric) with at least one of them positive definite, in which case M is also necessarily positive definite. Indeed, the iteration matrix of (63) is similar to

$$(\alpha I_{n^2} - A \otimes I_n)(\alpha I_{n^2} + A \otimes I_n)^{-1}(\alpha I_{n^2} - I_n \otimes B)(\alpha I_{n^2} + I_n \otimes B)^{-1}$$

and under the stated assumptions we have, for all $\alpha > 0$, that

$$\|(\alpha I_{n^2} - A \otimes I_n)(\alpha I_{n^2} + A \otimes I_n)^{-1}\|_2 \leq 1 \quad \text{and} \quad \|(\alpha I_{n^2} - I_n \otimes B)(\alpha I_{n^2} + I_n \otimes B)^{-1}\|_2 \leq 1,$$

with at least one of the two norms being < 1 , since $A \otimes I_n$ and $I_n \otimes B$ inherit the positive (semi)definiteness properties of A and B .

The alternating iteration (63) is obtained from the natural splitting

$$M = (A \otimes I_n) + (I_n \otimes B), \tag{64}$$

and it is natural to ask whether this splitting corresponds to a linearization of a Lie group decomposition. Not surprisingly, the answer is affirmative and the corresponding group factorization is the Kronecker (or tensor) product decomposition. This decomposition differs significantly from the ones previously considered in this paper since it is not a global decomposition of the group $GL(n^2, \mathbb{R})$ but only of a particular subgroup of it, namely,

$$G = GL(n, \mathbb{R}) \otimes GL(n, \mathbb{R}) = \{A \otimes B \mid A, B \in GL(n, \mathbb{R})\}. \tag{65}$$

This is to be expected, since not every $n^2 \times n^2$ real matrix is of the form (64). The fact that G is a subgroup of $GL(n^2, \mathbb{R})$ follows immediately from the properties of the Kronecker (or tensor) product of matrices. The identity element is $I_{n^2} = I_n \otimes I_n$ and the inverse of $A \otimes B$ is $A^{-1} \otimes B^{-1}$. Moreover, $(A \otimes B) \cdot (C \otimes D) = (AB) \otimes (CD)$. The dimension of G as a Lie group is $2n^2 - 1$. To see this, observe that the group G can also be written as

$$G = GL(n, \mathbb{R}) \otimes GL(n, \mathbb{R}) / \mathbb{R}^* \tag{66}$$

Indeed, any matrix of the form $A \otimes B$ with $A \in GL(n, \mathbb{R})$ and $B \in GL(n, \mathbb{R})$ can be written equivalently as $A_1 \otimes B_1$ with $A_1 \in GL(n, \mathbb{R})$ and B_1 a matrix of determinant ± 1 , owing to the identity

$$A \otimes B = (tA) \otimes (t^{-1}B), \quad t \in \mathbb{R}^* .$$

Note that the ordering chosen for the factors is arbitrary (we could have equally well picked A with $\det A = \pm 1$ and $B \in GL(n, \mathbb{R})$). Since $GL(n, \mathbb{R})$ has dimension n^2 and $GL(n, \mathbb{R}) / \mathbb{R}^*$ has dimension $n^2 - 1$, G has dimension $2n^2 - 1$.

Next, we consider the linearization of the Kronecker product near the identity. Let $t \mapsto A(t)$ and $t \mapsto B(t)$ be smooth paths in $GL(n, \mathbb{R})$ and $SL(n, \mathbb{R})$, respectively, such that $A(0) = I_n = B(0)$. Then clearly $t \mapsto M(t) = A(t) \otimes B(t)$ is a smooth path in G , such that $M(0) = I_{n^2}$, and differentiating at $t = 0$ leads to

$$M'(0) = A'(0) \otimes I_n + I_n \otimes B'(0),$$

with $A'(0) \in \mathfrak{gl}_n$ and $B'(0) \in \mathfrak{sl}_n$. Hence, the tangent space of the Lie group G at the identity I_{n^2} is precisely the Lie algebra of all matrices that are Kronecker sums of two matrices:

$$\mathfrak{g} = \{M \in \mathfrak{gl}_{n^2} \mid M = A \otimes I_n + I_n \otimes B, A \in \mathfrak{gl}_n, B \in \mathfrak{sl}_n\}, \tag{67}$$

the dimension of which (as a vector space) is clearly $2n^2 - 1$, as it should be.⁹ We note in passing that the fact that \mathfrak{g} is the Lie algebra of the Lie group G also follows easily from the well-known identity

$$\exp(A \otimes I_n + I_n \otimes B) = \exp(A) \otimes \exp(B),$$

valid for any A and B . The Lie algebra (67) naturally splits as the orthogonal direct sum of two subspaces:

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2, \tag{68}$$

where

$$\mathfrak{g}_1 = \{A \otimes I_n \mid A \in \mathfrak{gl}_n\}, \quad \mathfrak{g}_2 = \{I_n \otimes B \mid B \in \mathfrak{sl}_n\},$$

the first of dimension n^2 , the second of dimension $n^2 - 1$. The fact that $\mathfrak{g}_1 \perp \mathfrak{g}_2$ follows from the well-known property

$$\text{Tr}(A \otimes B) = \text{Tr}(A) \text{Tr}(B),$$

which in turn implies that the Frobenius inner product of a $A \otimes I_n$ and $I_n \otimes B$ is zero if one of A or B has zero trace.

With a slight abuse of notation we can also write the splitting (68) as

$$\mathfrak{g} = (\mathfrak{gl}_n \otimes I_n) \oplus (I_n \otimes \mathfrak{sl}_n), \tag{69}$$

stating that \mathfrak{g} is the *Kronecker sum* of the Lie algebras \mathfrak{gl}_n and \mathfrak{sl}_n . This Lie algebra splitting is the infinitesimal counterpart of the Lie group decomposition (66).

⁹ This Lie algebra is identical to the Lie algebra of all matrices of the form $M = A \otimes I_n + I_n \otimes B, A \in \mathfrak{gl}_n, B \in \mathfrak{gl}_n$, as the trivial identity

$$A \otimes I_n + I_n \otimes B = (A + \text{Tr}(B)I_n) \otimes I_n + I_n \otimes (B - \text{Tr}(B)I_n)$$

shows.

Table 1

Nomenclature of main decompositions, factorizations and corresponding splittings, with the associated iteration.

Lie group decomp.	Matrix factorization	Lie algebra splitting	Iteration
Global Cartan-like	Polar	J -symm. and J -skew symm.	J -HSS
Zappa-Szép	QR	skew-symm. and upper triangular	STS
Iwasawa (KAN)	QDR	skew-symm., diag. and str. upper tr.	MSTS
(Local) Doolittle	(LD)U	lower tr. and str. upper tr.	Gauss-Seidel
(Local) Crout	L(DU)	upper tr. and str. lower tr.	rev. Gauss-Seidel
(Local) LDU	LDU	diag. and str. lower-upper tr.	Jacobi
Kronecker product	$A \otimes B$	Kronecker sum	ADI-like

Using the representation (65) of G leads to a different splitting of the Lie algebra \mathfrak{g} , namely,

$$\mathfrak{g} = \mathfrak{g}_1 + \mathfrak{g}_3 = (\mathfrak{gl}_n \otimes I_n) + (I_n \otimes \mathfrak{gl}_n), \tag{70}$$

which is no longer a direct sum decomposition since $\mathfrak{g}_1 \cap \mathfrak{g}_3$ consists of all matrices of the form kI_{n^2} with $k \in \mathbb{R}$. Nevertheless, this somewhat redundant representation of the tangent space of G is the one which more closely corresponds to the splitting (64) underlying the ADI-like methods.

7. Conclusions and future work

In this paper we have shown that many (well-known and less well-known) matrix splittings can be interpreted as linearizations of matrix factorizations using basic notions from the theory of Lie groups and their Lie and Jordan algebras. We have also introduced the notion of Lie-Jordan splitting and used it to investigate a generalization of the HSS iteration, which we call the J -HSS iteration, for solving block-structured linear systems. Possible applications of this technique include the development of solvers and preconditioners for saddle-point problems and for complex linear systems in real equivalent form. Further exploration of these techniques could be the subject of future work.

In Table 1 we summarize the main group decompositions and corresponding Lie algebra splittings considered in this paper. Our list of splittings (and the corresponding factorizations) is by no means exhaustive, and many more such relationships remain to be explored; for instance, the recent paper [17] describes fifty-three different matrix factorizations, some of which may correspond to interesting matrix splittings not previously considered. Also, the investigation of block factorizations and the associated block splittings is a possible avenue for future work.

Funding

The authors acknowledge funding from INdAM-GNCS (Project “Metodi basati su matrici e tensori strutturati per problemi di algebra lineare di grandi dimensioni”). The first

author also acknowledges support from MUR Ministero dell'Università e della Ricerca through the PRIN Project 20227PCCKZ (“Low Rank Structures and Numerical Methods in Matrix and Tensor Computations and their Applications”).

Declaration of competing interest

There are no competing interests.

Acknowledgements

This paper is dedicated to Daniel Szyld, who has made major contributions to the theory and application of matrix splittings, Krylov subspace methods, preconditioning, and many other topics. The first author would like to thank Daniel for his warm friendship and for all his advice and support over the last 30 years. We are also grateful to two anonymous reviewers and to Angelo Vistoli for helpful suggestions.

Data availability

No data was used for the research described in the article.

References

- [1] G. Ammar, C. Mehl, V. Mehrmann, Schur-like forms for matrix Lie groups, Lie algebras and Jordan algebras, *Linear Algebra Appl.* 287 (1999) 11–39.
- [2] E. Artin, *Geometric Algebra*, Dover Publications, Mineola, NY, 2015.
- [3] Z.-Z. Bai, M. Benzi, F. Chen, Modified HSS iteration method for a class of complex symmetric linear systems, *Computing* 87 (2010) 93–111.
- [4] Z.-Z. Bai, M. Benzi, F. Chen, Z.-Q. Wang, Preconditioned MHSS iteration methods for a class of block two-by-two linear systems with applications to distributed control problems, *IMA J. Numer. Anal.* 33 (2013) 343–369.
- [5] Z.-Z. Bai, G.H. Golub, L.-Z. Lu, J.-F. Yin, Block triangular and skew-Hermitian splitting methods for positive definite linear systems, *SIAM J. Sci. Comput.* 26 (2005) 844–863.
- [6] Z.-Z. Bai, G.H. Golub, M.K. Ng, Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *SIAM J. Matrix Anal. Appl.* 24 (2003) 603–626.
- [7] Z.-Z. Bai, G.H. Golub, M.K. Ng, On inexact Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *Linear Algebra Appl.* 428 (2008) 413–440.
- [8] Z.-Z. Bai, J.-Y. Pan, *Matrix Analysis and Computations*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2021.
- [9] M. Benzi, D. Bertaccini, Block preconditioning of real-valued iterative algorithms for complex linear systems, *IMA J. Numer. Anal.* 28 (2008) 598–618.
- [10] M. Benzi, C. Faccio, Solving linear systems of the form $(A + \gamma UU^T)\mathbf{x} = \mathbf{b}$ by preconditioned iterative methods, *SIAM J. Sci. Comput.* 46 (2024) S51–S70.
- [11] M. Benzi, G.H. Golub, A preconditioner for generalized saddle point problems, *SIAM J. Matrix Anal. Appl.* 26 (2004) 20–41.
- [12] M. Benzi, G.H. Golub, J. Liesen, Numerical solution of saddle-point problems, *Acta Numer.* 14 (2005) 1–137.
- [13] M. Benzi, D.B. Szyld, Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods, *Numer. Math.* 76 (1997) 309–321.
- [14] A.M. Bloch, P.E. Crouch, J.E. Marsden, A.K. Sanyal, Optimal control and geodesics on quadratic matrix Lie groups, *Found. Comput. Math.* 8 (2008) 469–500.

- [15] P. Concus, G.H. Golub, A generalized conjugate gradient method for nonsymmetric systems of linear equations, in: *Computing Methods in Applied Sciences and Engineering (Second Internat. Sympos., Versailles, 1975)*, Part 1, Institut de Recherche d'Informatique et d'Automatique, Rocquencourt, 1976, pp. 56–65.
- [16] J. Della-Dora, Numerical linear algorithms and group theory, *Linear Algebra Appl.* 10 (1975) 267–283.
- [17] A.S. Edelman, S. Jeong, Fifty three matrix factorizations: a systematic approach, *SIAM J. Matrix Anal. Appl.* 44 (2023) 415–480.
- [18] R. Gilmore, *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists*, Cambridge University Press, 2008.
- [19] G.H. Golub, C.F. Van Loan, *Matrix Computations*, fourth edition, Johns Hopkins University Press, Baltimore, MD, 2013.
- [20] C. Güdücü, J. Liesen, V. Mehrmann, D.B. Szyld, On non-Hermitian positive (semi)definite linear algebraic systems arising from dissipative Hamiltonian PDEs, *SIAM J. Sci. Comput.* 44 (2022) A2871–A2894.
- [21] C. Greif, Structured shifts for skew-symmetric matrices, *Electron. Trans. Numer. Anal.* 55 (2022) 455–468.
- [22] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Graduate Studies in Mathematics, vol. 34, American Mathematical Society, Providence, RI, 2001.
- [23] N.J. Higham, D.S. Mackey, N. Mackey, F. Tisseur, Functions preserving matrix groups and iterations for the matrix square root, *SIAM J. Matrix Anal. Appl.* 31 (2010) 849–877.
- [24] H. Huang, T.-Y. Tam, On Gelfand-Naimark decomposition of a nonsingular matrix, *Linear Multilinear Algebra* 58 (2010) 27–43.
- [25] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, A. Zanna, Lie-group methods, *Acta Numer.* 9 (2000) 215–365.
- [26] A.W. Knap, *Lie Groups Beyond an Introduction*, second edition, Birkhäuser, Boston, MA, 2002.
- [27] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry*, vols. 1 and 2, Wiley, New York, 1969.
- [28] L. Krukier, L.G. Chikina, T.V. Belokon, Triangular skew-symmetric iterative solvers for strongly nonsymmetric positive real linear systems of equations, *Appl. Numer. Math.* 41 (2002) 89–105.
- [29] R.I. MacLachlan, G.R.W. Quispel, Splitting methods, *Acta Numer.* 11 (2002) 341–434.
- [30] D.S. Mackey, N. Mackey, F. Tisseur, Structured factorizations in scalar product spaces, *SIAM J. Matrix Anal. Appl.* 27 (2006) 821–850.
- [31] G.I. Marchuk, *Methods of Numerical Mathematics*, Springer-Verlag, New York, 1984.
- [32] K. McCrimmon, *A Taste of Jordan Algebras*, Springer, New York, 2004.
- [33] D.W. Peaceman, H.H. Rachford Jr., The numerical solution of parabolic and elliptic partial differential equations, *J. Soc. Ind. Appl. Math.* 3 (1955) 28–41.
- [34] M. Postnikov, *Lectures in Geometry, Semester V: Lie Groups and Lie Algebras*, Mir Publishers, Moscow, 1986.
- [35] Y. Saad, *Iterative Methods for Sparse Linear Systems*, second edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2003.
- [36] J. Szép, On the structure of groups which can be represented as the product of two subgroups, *Acta Math. Sci. (Szeged)* 12 (1950) 57–61.
- [37] E. Tyrtyshnikov, Matrix Bruhat decomposition with a remark on the QR (GR) algorithm, *Linear Algebra Appl.* 250 (1997) 61–68.
- [38] R.S. Varga, *Matrix Iterative Analysis*, expanded edition, Springer Series in Computational Mathematics, vol. 27, Springer-Verlag, 2000.
- [39] L. Wang, Z.-Z. Bai, Skew-Hermitian triangular splitting iteration methods for non-Hermitian positive definite linear systems of strong skew-Hermitian part, *BIT Numer. Math.* 44 (2004) 363–386.
- [40] O. Widlund, A Lanczos method for a class of nonsymmetric systems of equations, *SIAM J. Numer. Anal.* 15 (1978) 801–812.
- [41] G. Zappa, Sulla costruzione dei gruppi prodotto di due dati sottogruppi permutabili tra loro, in: *Atti Secondo Congresso Un. Mat. Ital., Bologna, 1940, Cremonese, Rome, 1942*, pp. 119–125.