Lie algebra representations, nilpotent matrices, and the C–numerical range

G. Dirr∗, U. Helmke∗ and M. Kleinsteuber∗

June 4, 2005

Abstract

In this paper we survey some applications of the representation theory of Lie algebras to Linear Algebra. This includes the derivation of the Jordan form for tensor products of invertible matrices, the study of normal form problems for nilpotent matrices, as well as the derivation of explicit formulas for the C-numerical radius of certain nilpotent block-shift matrices, arising in quantum mechanics and nuclear magnetic resonance (NMR-) spectroscopy. In this latter case, a conjecture concerning an explicit formula for the C-numerical radius is stated. We show the existence of unitary transformations that realize the prospective maxima. Our approach depends on the Clebsch-Gordan decomposition for unitary representations of su2(C).

1 Introduction

Techniques and ideas from the theory of Lie groups and Lie algebras have long played a prominent role in mathematics and physics alike, with widespread applications to topics such as harmonic analysis, number theory, differential equations and dynamical systems, Yang-Mills gauge theory, quantum mechanics and geometric classical mechanics. In Linear Algebra, Lie theoretic tools have been applied to a considerably lesser extent, despite of their natural and often recognized role in solving classification and normal form problems of matrices. This concerns for instance the analysis of matrix eigenvalue algorithms, where a deeper understanding of methods such as

∗Mathematical Institute, University of Würzburg, Am Hubland, 97074 Würzburg, Germany, {dirr,helmke,kleinsteuber}@mathematik.uni-wuerzburg.de; partial support by the Australian-German DAAD project D/0243869 (U.H.) and by the Marie–Curie PhD programme Control Training Site (M.K.) is gratefully acknowledged.
Jacobi-type or $QR$-algorithms for structured matrices is gained by realizing, that the underlying matrix factorizations are just special cases of more generally defined factorizations for arbitrary semi-simple Lie groups; cf. [5, 7, 15, 18, 29, 30]. Similarly, versal or mini-versal deformations of similarity orbits of matrices are best understood in a Lie algebraic context, as has been observed first by Arnol'd [2]. For applications of these techniques to control theory and Linear Algebra see e.g. [6] and [28]. We also refer to the recent work of Fulton [9] on eigenvalue inequalities for sums of Hermitian matrices, where representation theoretic methods play an important role. For an overview on applications of Lie theoretic tools to the numerical integration of differential equations on Lie groups we recommend [19].

In this survey paper we will not discuss these topics any further but rather focus on the task of considering applications of Lie algebra representations to Linear Algebra. We will illustrate the power of representation theoretic methods by discussing their role in three different areas:

- Jordan canonical forms for tensor products of matrices.
- Geometry and parametrization of unitary orbits of nilpotent matrices.
- The calculation of the $C$-numerical range of nilpotent matrices.

In fact, one of the most elegant applications of representation theoretic ideas in Linear Algebra is connected with the task of obtaining explicit decomposition formulas for the Jordan canonical form of the Kronecker product of two matrices. This classical problem has been known to be closely related to the Clebsch-Gordan decomposition of Lie algebra representations for $\mathfrak{sl}_2(\mathbb{C})$, see [10, 23]. Therefore, after having presented some preliminary material on Lie algebra representations, we explain this connection in Section 2.1. Using the Clebsch-Gordan decomposition for Lie group representations of the special linear group $\text{SL}_2(\mathbb{C})$, we derive an explicit formula for the Jordan form of the Kronecker product of two invertible Jordan blocks; this is in fact a special case of a more general formula for the Jordan form of Kronecker products of arbitrary, not necessarily invertible matrices; see e.g. [21] and the references therein.

The classification of Lie algebra representations of $\mathfrak{sl}_2(\mathbb{C})$ is closely related to the task of parameterizing nilpotent similarity orbits. Thus we discuss this connection in Section 2.2. By utilizing earlier results by Kostant and Sekiguchi we show that there is a bijective correspondence between equivalence classes of Lie algebra representations of $\mathfrak{sl}_2(\mathbb{C})$ and similarity orbits of nilpotent matrices. The so-called Kostant-Sekiguchi correspondence [25] refines this even further to a bijective correspondence between similarity orbits of real nilpotent matrices and those of complex symmetric nilpotent matrices. There are other interesting connections between unitary similarity orbits of nilpotent matrices and the classification task for unitary Lie algebra representations of $\mathfrak{sl}_2(\mathbb{C})$ which are also discussed in Section 2.2.
As a subject of interest in its own, we focus in Section 3 on the optimization task of finding representations that are as close as possible to the equivalence class of a given one. This makes contact with computing normal forms for the simultaneous similarity action on $N$-tuples of matrices, as well as with the $C$-numerical range of nilpotent matrices. We develop a general representation theoretic framework for investigating such problems by studying least squares matching problems for representations of arbitrary Lie algebras. If representations of Abelian Lie algebras are considered, we obtain the classical simultaneous similarity problem. However, other choices of Lie algebras lead to new types of simultaneous classification problems. For unitary representations of $\mathfrak{su}_2(\mathbb{C})$ we introduce the concept of a relative numerical range that measures the distance between the two representations. Using a result of [22], we prove that the relative numerical range is a circular disc in the complex plane, centered at the origin.

Finally, in the last section, we discuss an application to quantum mechanics, i.e. the maximization of the transfer function of $N + 1$ weakly coupled spin-$\frac{1}{2}$-systems arising in NMR–spectroscopy and quantum computing; see e.g. [12] for a more detailed description of the physical background. Explicit formulas for the global maximal values of the transfer are apparently unknown, except for trivial cases. Our approach consists in reformulating the task as a least squares matching problem for two unitary representations of $\mathfrak{su}_2(\mathbb{C})$. This then leads us to explicit formulas for the conjectured maxima in terms of the Clebsch-Gordan decomposition. It is shown that unitary transformations exist that achieve these values.

2 Representations and Normal Form Problems

2.1 Direct Sums and Tensor Product Decompositions

The purpose of this section is to show how one can employ the Clebsch-Gordan formula for representations of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ to determine the Jordan structure of a tensor product of two invertible matrices. We begin by summarizing some basic facts and definitions on Lie algebras and representation theory. For further background material we refer to the textbooks by [11], [17], and [20].

Let $\mathfrak{g}$ be a finite dimensional real or complex Lie algebra with Lie bracket $[\cdot, \cdot]$, and let $\mathfrak{gl}_N(\mathbb{C}) := \mathbb{C}^{N \times N}$ denote the set of all complex $(N \times N)$-matrices. As it is known, $\mathfrak{gl}_N(\mathbb{C})$ endowed with the usual commutator operation $[A, B] = AB - BA$ forms a Lie algebra which is isomorphic to the set End($\mathbb{C}^N$) of all linear maps from $\mathbb{C}^N$ to
Natural subalgebras of $\mathfrak{gl}_N(\mathbb{C})$ that will play a major role in the sequel are

$$\mathfrak{sl}_N(\mathbb{C}) := \{ x \in \mathbb{C}^{N \times N} \mid \text{tr} \, x = 0 \},$$

$$\mathfrak{u}_N(\mathbb{C}) := \{ x \in \mathbb{C}^{N \times N} \mid x^\dagger = -x \},$$

$$\mathfrak{su}_N(\mathbb{C}) := \{ x \in \mathbb{C}^{N \times N} \mid \text{tr} \, x = 0, x^\dagger = -x \},$$

where $(\cdot)^\dagger$ denotes conjugate transpose. A linear map $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}_N(\mathbb{C}) \cong \text{End}(\mathbb{C}^N)$ is called a representation of $\mathfrak{g}$ in $\mathfrak{gl}_N(\mathbb{C})$ if

$$\rho([x, y]) = [\rho(x), \rho(y)]$$

for all $x, y \in \mathfrak{g}$. Basic facts about representations are that the pre-image of an ideal of $\mathfrak{gl}_N(\mathbb{C})$ is an ideal in $\mathfrak{g}$, and that the image of a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra in $\mathfrak{gl}_N(\mathbb{C})$. Furthermore, the following result will prove to be useful.

**Proposition 1.** Let $\rho$ be a representation of a semisimple $\mathfrak{g}$. Then $\rho(\mathfrak{g}) \subset \mathfrak{sl}_N(\mathbb{C})$.

**Proof.** Cf. [17], Section 6.3. □

Two representations $\rho_i : \mathfrak{g} \rightarrow \mathfrak{gl}_N(\mathbb{C})$, $i = 1, 2$, are equivalent, if there exists a $T \in \text{GL}_N(\mathbb{C}) := \{ X \in \mathbb{C}^{N \times N} \mid \text{det} \, X \neq 0 \}$ such that

$$\rho_2(x) = T\rho_1(x)T^{-1}$$

for all $x \in \mathfrak{g}$. They are called unitary equivalent if $T$ can be chosen unitary, i.e.

$$T \in \text{U}_N(\mathbb{C}) := \{ U \in \mathbb{C}^{N \times N} \mid U^\dagger U = \mathbb{I}_N \}.$$

Here $\mathbb{I}_N$ denotes the $N \times N$ identity matrix. In the sequel, we will focus mainly on representations of $\mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{su}_2(\mathbb{C})$. In this context we use the following terminology. A representation $\tau : \mathfrak{su}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C})$ is called unitary, if its image $\tau(\mathfrak{su}_2(\mathbb{C}))$ is contained in $\mathfrak{su}_N(\mathbb{C})$. In abuse of the above definition we also call a representation $\sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C})$ unitary, if its restriction $\sigma|_{\mathfrak{su}_2(\mathbb{C})}$ is unitary.

Let $\mathfrak{g}$ be a real Lie algebra with complexification $\mathfrak{g}^\mathbb{C}$ and let $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}_N(\mathbb{C})$ be any representation. Then the representation $\rho^\mathbb{C} : \mathfrak{g}^\mathbb{C} \rightarrow \mathfrak{gl}_N(\mathbb{C})$ defined by

$$\rho^\mathbb{C}(x + iy) := \rho(x) + i\rho(y)$$

is called the complexification of $\rho$. A representation $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}_N(\mathbb{C})$, $N > 0$, is irreducible, if $\{0\}$ and $\mathbb{C}^N$ are the only subspaces which are invariant under all linear
transformations $\rho(x), x \in \mathfrak{g}$. The class of Lie algebra representations is closed under direct sum and tensor product operations, which are defined as follows. The direct sum of two representations $\rho_i : \mathfrak{g} \to \mathfrak{gl}_{N_i}(\mathbb{C}), i = 1, 2,$ of $\mathfrak{g}$ is the representation $\rho_1 \oplus \rho_2$ defined by

$$\rho_1 \oplus \rho_2 : \mathfrak{g} \longrightarrow \mathfrak{gl}_{N_1+N_2}(\mathbb{C}) \cong \text{End}(\mathbb{C}^{N_1} \oplus \mathbb{C}^{N_2})$$

$$x \longmapsto \begin{bmatrix} \rho_1(x) & 0 \\ 0 & \rho_2(x) \end{bmatrix}.$$  

For the direct sum of $n$ identical representations $\rho$ we shorten notation by writing

$$n \star \rho := \underbrace{\rho \oplus \ldots \oplus \rho}_{n\text{-times}}.$$

**Theorem 2** (Weyl). Let $\mathfrak{g}$ be a semisimple Lie algebra. Then every representation $\rho : \mathfrak{g} \longrightarrow \mathfrak{gl}_N(\mathbb{C})$ is equivalent to a direct sum of irreducible representations, i.e

$$\rho \cong \rho_1 \oplus \ldots \oplus \rho_r,$$

where $\rho_i : \mathfrak{g} \longrightarrow \mathfrak{sl}_{n_i}(\mathbb{C})$ are irreducible for $i = 1, \ldots, r$.

*Proof.* Cf. [17], Section 6.3. $\square$

The tensor product of two representations $\rho_i, i = 1, 2,$ of $\mathfrak{g}$ is the representation defined by

$$\rho_1 \otimes \rho_2 : \mathfrak{g} \longrightarrow \mathfrak{gl}_{N_1N_2}(\mathbb{C}) \cong \text{End}(\mathbb{C}^{N_1} \otimes \mathbb{C}^{N_2})$$

$$x \longmapsto \rho_1(x) \otimes I_{N_2} + I_{N_1} \otimes \rho_2(x),$$

where $\otimes$, applied to matrices, denotes the usual matrix Kronecker product. Note that $(\rho_1 \otimes \rho_2)^C = \rho_1^C \otimes \rho_2^C$ and $(\rho_1 \otimes \rho_2)^C = \rho_1^C \otimes \rho_2^C$. Analogously to the direct sum we write

$$\rho^n := \underbrace{\rho \otimes \ldots \otimes \rho}_{n\text{-times}}$$

for the tensor product of $n$ identical representations $\rho$. The operations of direct sums and tensor products are compatible in the sense that the tensor product

$$\rho \otimes (\rho_1 \oplus \rho_2)$$

is equivalent to the direct sum of tensor products

$$(\rho \otimes \rho_1) \oplus (\rho \otimes \rho_2).$$

Next, we investigate $\mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{su}_2(\mathbb{C})$ in more detail. They are the simplest examples of semisimple Lie algebras and their representations will be of central importance.
for our subsequent analysis and applications. A standard basis for the complex 3-
dimensional Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ is

$$E := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad F := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad H := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$  \hspace{1cm} (1)

with commutator relations

$$[H, E] = -2E, \quad [H, F] = 2F, \quad \text{and} \quad [F, E] = H. \hspace{1cm} (2)$$

Analogously, for the real 3-dimensional Lie algebra $\mathfrak{su}_2(\mathbb{C})$ we define a standard basis via

$$X := \frac{1}{2} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \quad Y := \frac{1}{2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad Z := \frac{1}{2} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$  \hspace{1cm} (3)

Here the basis elements satisfy the identities

$$[X, Y] = Z, \quad [Z, X] = Y, \quad \text{and} \quad [Y, Z] = X. \hspace{1cm} (4)$$

The irreducible representations of these Lie algebras are well known and can be completely characterized. In preparation of the results below we introduce the following definition. Denote by $J_\nu$, $\nu \in \frac{1}{2} \mathbb{N}$, the nilpotent matrix

$$J_\nu := \begin{bmatrix} 0 \\ \vdots \\ d_\nu \\ \vdots \\ \vdots \\ \vdots \\ d_{-\nu+1} \end{bmatrix} \in \mathfrak{sl}_{2\nu+1}(\mathbb{C}) \hspace{1cm} (5)$$

with $d_\mu := \sqrt{\nu + \mu}(\nu - \mu + 1)$, $\mu = \nu, \nu - 1, \ldots, -\nu + 1$. The matrix $J_\nu$ is a scaled version of the canonical nilpotent Jordan block of size $2\nu + 1$. This implies the following lemma whose straightforward proof via the Jordan normal form is omitted.

**Lemma 3.** Every nilpotent matrix $A \in \mathfrak{sl}_N(\mathbb{C})$ is similar to a block diagonal matrix of the form

$$J_A = \begin{bmatrix} J_{\nu_1} & & \\ & \ddots & \\ & & J_{\nu_r} \end{bmatrix},$$

with $J_{\nu_i}$, $i = 1, \ldots, r$, defined as in Eq. (5).

We refer to $J_A$ as the CGJ-form (Clebsch-Gordan-Jordan) of $A$ and call $J_\nu$ a CGJ-block.

**Proposition 4.** Let $J_\nu$ be defined as in Eq. (5). Then the following holds
(a) Every irreducible representation \( \sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_{2\nu+1}(\mathbb{C}) \) is equivalent to the standard irreducible representation \( \sigma_\nu \) defined by

\[
\sigma_\nu(E) := J_\nu, \quad \sigma_\nu(F) := J_\nu^\dagger, \\
\sigma_\nu(H) := 2 \begin{bmatrix}
\nu & 0 & 0 \\
0 & \nu - 1 & 0 \\
0 & \ddots & \ddots \\
\ddots & \ddots & 0 \\
0 & 0 & -\nu
\end{bmatrix}.
\]

(b) Every irreducible representation \( \tau : \mathfrak{su}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_{2\nu+1}(\mathbb{C}) \) is equivalent to the standard irreducible representation \( \tau_\nu \) defined by

\[
\tau_\nu(X) := \frac{i}{2} (J_\nu + J_\nu^\dagger), \quad \tau_\nu(Y) := \frac{1}{2} (J_\nu - J_\nu^\dagger) \quad \text{and} \\
\tau_\nu(Z) := i \begin{bmatrix}
\nu & 0 & 0 \\
0 & \nu - 1 & 0 \\
\ddots & \ddots & \ddots \\
\ddots & \ddots & 0 \\
0 & 0 & -\nu
\end{bmatrix}.
\]

Moreover, if \( \tau \) is unitary, then it is unitarily equivalent to \( \tau_\nu \).

Proof. (a) The first part follows by [14], Ch. 4, §1 or [9], Ch. 11 together with Lemma 3. (b) Cf. [11], Part I, Ch. I, Sec. 2. \( \square \)

Note that \( \tau_\nu^C = \sigma_\nu \) for all \( \nu \in \frac{1}{2} \mathbb{N} \). This result together with Proposition 2 immediately yields the following corollary.

**Corollary 5.** (a) Every representation \( \sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C}) \) is equivalent to the direct sum

\[
\sigma \cong c_1 \star \sigma_{\nu_1} \oplus \ldots \oplus c_r \star \sigma_{\nu_r}, \quad c_i \in \mathbb{N}
\]

where the \( \sigma_{\nu_i} \) denote the corresponding standard irreducible representations.

(b) Every representation \( \tau : \mathfrak{su}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C}) \) is equivalent to the direct sum

\[
\tau \cong c_1 \star \tau_{\nu_1} \oplus \ldots \oplus c_r \star \tau_{\nu_r}, \quad c_i \in \mathbb{N}
\]

where the \( \tau_{\nu_i} \) denote the corresponding standard irreducible representations. Moreover, if \( \tau \) is unitary, then the equivalence is unitary.
We refer to the integers $c_i$ appearing in the above direct sum decomposition as the Clebsch-Gordan multiplicities of $\sigma$ and $\tau$. Note that if $\sigma = \tau^{c}$, then the coefficients in (a) and (b) coincide. Moreover, they are uniquely determined and invariant under equivalence transformations on $\sigma$ and $\tau$ and thus characterize completely the equivalence type of a representation. For the tensor product of two irreducible representations of either $\mathfrak{sl}_2(\mathbb{C})$ or $\mathfrak{su}_2(\mathbb{C})$ one has the following classical result.

**Theorem 6** (Clebsch-Gordan Decomposition). Let $\sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_{2\mu+1}(\mathbb{C})$ and $\sigma' : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_{2\nu+1}(\mathbb{C})$ be two irreducible representations. Then, up to $GL_N(\mathbb{C})$-equivalence, there is a unique direct sum decomposition

$$\sigma \otimes \sigma' \cong \sigma_{\mu+\nu} \oplus \sigma_{\mu+\nu-1} \oplus \ldots \oplus \sigma_{|\mu-\nu|}.$$  

Similarly, for irreducible representations $\tau : \mathfrak{su}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_{2\mu+1}(\mathbb{C})$ and $\tau' : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_{2\nu+1}(\mathbb{C})$ there is a unique, up to $GL_N(\mathbb{C})$-equivalence, direct sum decomposition

$$\tau \otimes \tau' \cong \tau_{\mu+\nu} \oplus \tau_{\mu+\nu-1} \oplus \ldots \oplus \tau_{|\mu-\nu|}.$$  

Moreover, if $\tau$ and $\tau'$ are unitary, then the equivalence is unitary.

**Proof.** (a) The proof is a consequence of Proposition 4 (a) and [9], Ch. 11. (b) Part (b) follows by Proposition 4 (b) and [11], Part I, Ch. I, Sec. 4. □

As a simple application of the above circle of ideas we derive a formula for the Jordan structure of tensor products of complex invertible matrices. Decomposition formulas for the tensor product of Jordan matrices are of course not new, with early contributions going back to the work of Aitkin, Roth, Littlewood, Marcus, Robinson, Brualdi, and others. The result below is a special case of a more general formula for tensor products of Jordan forms of arbitrary, not necessarily invertible, matrices; cf. e.g. [21] for a simple proof, and an extension to infinite dimensions.

In this context it is actually more convenient to work with Lie group representations rather than Lie algebra representations. Recall, that a complex representation of a Lie group $G$ is a homomorphism $\theta : G \rightarrow GL_N(\mathbb{C})$, i.e. $\theta(gh) = \theta(g)\theta(h), \theta(e) = I_N$. The study of Lie group representations is quite analogous to that of Lie algebra representations, with similar definitions and concepts; cf. [13]. For example, the direct sum of two Lie group representations $\theta_1 : G \rightarrow GL_{N_1}(\mathbb{C}), i = 1, 2$, is the representation

$$\theta_1 \oplus \theta_2 : G \rightarrow GL_{N_1+N_2}(\mathbb{C})$$

$$x \mapsto \begin{bmatrix} \theta_1(x) & 0 \\ 0 & \theta_2(x) \end{bmatrix}.$$

Analogously, the tensor product of $\theta_1, \theta_2$ is defined as

$$\theta_1 \otimes \theta_2 : G \rightarrow GL_{N_1N_2}(\mathbb{C})$$

$$x \mapsto \theta_1(x) \otimes \theta_2(x).$$
Note, that the finite-dimensional irreducible representations of

\[ SL_2(\mathbb{C}) := \{ X \in \mathbb{C}^{2 \times 2} \mid \det X = 1 \} \]

are equivalent to \( \kappa_{\mu} : SL_2(\mathbb{C}) \to GL_{2\mu+1}(\mathbb{C}), g \mapsto \kappa_{\mu}(g), \mu \in \frac{1}{2}\mathbb{N} \). Here \( \mathbb{C}^{2\mu+1} \) is identified with the vector space of complex homogeneous polynomials

\[ \Phi(x, y) = \sum_{i=0}^{2\mu} c_i x^i y^{2\mu-i} \]

of degree \( 2\mu \). Then for any \( A \in SL_2(\mathbb{C}) \), \( \kappa_{\mu}(A) \) acts on \( \Phi \) as

\[ \kappa_{\mu}(A) \Phi(x, y) = \Phi((x, y)A^{-1}) \]

and this action on complex homogeneous polynomials defines the irreducible representation \( \kappa_{\mu} \), cf. [14], Ch. III, §2. For group representations of \( SL_2(\mathbb{C}) \) the following group variant of the Clebsch-Gordan decomposition holds, which is easily deduced from the above Lie algebra version, cf. [14], Ch. IV §1.

**Theorem 7** (Clebsch-Gordan Decomposition). Let \( \kappa_{\mu} \) and \( \kappa_{\nu} \) be two irreducible representations of \( SL_2(\mathbb{C}) \). There exists a \( GL_N(\mathbb{C}) \)-equivalence of \( \kappa_{\mu} \otimes \kappa_{\nu} \) with the direct sum decomposition

\[ \kappa_{\mu+\nu} \oplus \kappa_{\mu+\nu-1} \oplus \ldots \oplus \kappa_{|\mu-\nu|}. \]  

(10)

From this basic result, we can obtain an explicit formula for the Jordan structure of a tensor product of two invertible Jordan blocks and hence a formula for the Jordan structure of the tensor product of two arbitrary invertible matrices.

**Corollary 8.** Let \( \mu, \nu \in \frac{1}{2}\mathbb{N} \) and

\[ J_{\mu}(\alpha) := \begin{bmatrix} \alpha \\ 1 & \ddots \\ & 1 & \ddots \\ & & \ddots & \ddots \\ & & & 1 & \alpha \end{bmatrix} \]  

(11)

denote an \( (2\mu + 1) \times (2\mu + 1) \) Jordan block. Then for complex numbers \( \alpha, \beta \neq 0 \), the Kronecker product \( J_{\mu}(\alpha) \otimes J_{\nu}(\beta) \) is similar to the direct sum

\[ J_{\mu}(\alpha) \otimes J_{\nu}(\beta) \cong \begin{bmatrix} J_{\mu+\nu}(\alpha\beta) \\ J_{\mu+\nu-1}(\alpha\beta) \\ \vdots \\ J_{|\mu-\nu|}(\alpha\beta) \end{bmatrix}. \]  

9
Proof. Since the claim is invariant under multiplication by nonzero complex numbers, we can assume without loss of generality that $\alpha = \beta = 1$. A simple computation shows that the standard irreducible representation $\kappa_\mu : SL_2(\mathbb{C}) \to GL_{2\mu+1}(\mathbb{C})$ maps the $2 \times 2$-matrix
\begin{equation}
\begin{bmatrix}
1 & 0 \\
1 & 1 \\
\end{bmatrix}
\end{equation}
to a lower triangular $(2\mu + 1) \times (2\mu + 1)$-matrix with the nonzero entry 1 on the diagonal. By inspection, this matrix is easily seen to be similar to the Jordan block $J_\mu(1)$. The result now follows from the above group version of the Clebsch-Gordan decomposition and by applying suitable similarity transformations to the diagonal-blocks.

For an early version of this representation theoretic approach via the Clebsch-Gordan decomposition we refer to [10]. More recently, the Clebsch-Gordan decomposition for modules of a polynomial ring in one variable is employed by Martsinkovsky and Vlassov in [23], who obtain an elegant proof of the tensor product formula for general, not necessarily invertible matrices.

## 2.2 Nilpotent Matrices and Representations

There is a surprising connection between the classification of similarity classes of complex nilpotent matrices and representations of $\mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{su}_2(\mathbb{C})$. This will be exploited in the following subsection, when studying unitary similarity orbits of nilpotent matrices. In fact, this study goes back to the work by Jacobson, Morosov and Kostant, and is explained in more detail now. We begin with an elementary lemma.

**Lemma 9.** Let $\sigma : \mathfrak{sl}_2(\mathbb{C}) \longrightarrow \mathfrak{gl}_N(\mathbb{C})$ and $\tau : \mathfrak{su}_2(\mathbb{C}) \longrightarrow \mathfrak{gl}_N(\mathbb{C})$ be two representations. Then

(a) The images of $\sigma$ and $\tau$ are contained in $\mathfrak{sl}_N(\mathbb{C})$.

(b) The matrix $A := \rho(E)$ is nilpotent.

**Proof.** (a) As $\mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{su}_2(\mathbb{C})$ are semisimple, the assertion follows from Proposition 1.

(b) Let $B := \rho(\frac{1}{2}H)$. Then it holds $A = AB - BA$ and therefore
\begin{equation}
A^n = A^k(AB - BA)A^{n-k-1} = A^{k+1}BA^{n-k-1} - A^kBA^{n-k}
\end{equation}
for all $k = 0, \ldots, n - 1$. This yields
\begin{equation}
nA^n = \sum_{k=0}^{n-1} (A^{k+1}BA^{n-k-1} - A^kBA^{n-k}) = -BA^n + A^n B.
\end{equation}
Now let \( \| \cdot \| \) be any submultiplicative norm on \( \mathfrak{gl}_N(\mathbb{C}) \). Using Eq. (14) we obtain
\[
n\|A^n\| = \|A^nB - BA^n\| \leq 2\|B\| \cdot \|A^n\|. \quad (15)
\]
Eq. (15) implies \( A^n = 0 \) for some \( n \in \mathbb{N} \). In fact, if this were not the case, one would obtain a contradiction to the finiteness of \( \|B\| \).

**Lemma 10** (Jacobson-Morosov). For any nilpotent matrix \( A \in \mathfrak{sl}_N(\mathbb{C}) \) there exists a representation \( \sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C}) \) such that
\[
\sigma(E) = A \quad (16)
\]

**Proof.** This is an immediate consequence of the results of the previous subsection. In fact, from Lemma 3 and Proposition 4 we see, that every nilpotent matrix is similar to a direct sum of standard irreducible representations of \( \mathfrak{sl}_2(\mathbb{C}) \), evaluated at \( E \). See also [20], Ch. III, Theorem 17 for another proof.

Now the beautiful fact is that arbitrary nilpotent matrices can not only be defined via representations, but they are even completely classified by the equivalence type of the associated representations. This is a result of Kostant.

**Theorem 11** (Kostant). Two representations \( \sigma, \tilde{\sigma} : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C}) \) are equivalent if and only if the nilpotent matrices \( \sigma(E) \) and \( \tilde{\sigma}(E) \) are similar.

**Proof.** If \( \sigma(E) \) and \( \tilde{\sigma}(E) \) are similar, they possess in particular the same CGJ-form. Thus only the sufficiency of the condition needs to be proven. Without loss of generality, we can assume that \( \sigma(E) = \tilde{\sigma}(E) = J_A \) is in CGJ-form. By Weyl’s theorem, \( \sigma \) decomposes into the direct sum of irreducible representations, where the equivalence type of these irreducible summands is uniquely determined by \( \sigma \). As the direct sum of irreducible representations maps \( E \) always onto a uniquely determined CGJ-form \( J_A \), with the multiplicities of the components bijectively corresponding to the sizes of the CGJ-blocks of \( J_A \), it follows that the equivalence type of \( \sigma \) is in turn uniquely determined by \( J_A = \sigma(E) \). The same reasoning applies to \( \tilde{\sigma} \) and the result follows.

**Corollary 12.** Let \( \sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_N(\mathbb{C}) \) be a representation with
\[
\sigma(E) = A = TJ_AT^{-1}.
\]
Then \( \sigma \) is equivalent to the unique representation defined by
\[
\sigma_{J_A}(E) := J_A, \quad \sigma_{J_A}(F) := J_A, \quad \sigma_{J_A}(H) := [J_A, J_A^\dagger]. \quad (17)
\]

**Proof.** Both \( \sigma \) and \( \sigma_{J_A} \) are representations such that \( \sigma(E) \) and \( \sigma_{J_A}(E) \) are similar. Thus the result follows from Theorem 11.
A similar bijective correspondence holds between unitary similarity classes of nilpotent matrices and equivalence classes of unitary representations of \( \mathfrak{su}_2(\mathbb{C}) \). This is shown in the next result.

**Proposition 13.** Let \( A \in \mathfrak{sl}_N(\mathbb{C}) \) be nilpotent and denote by \( J_A \) its CGJ-form. The following statements are equivalent.

(a) The matrix \( A \) is unitarily similar to \( J_A \), i.e. there exists a unitary \( U \in U_N(\mathbb{C}) \) such that

\[
A = U J_A U^\dagger.
\]

(b) There exists a unitary representation \( \sigma : \mathfrak{sl}_2(\mathbb{C}) \longrightarrow \mathfrak{sl}_N(\mathbb{C}) \) such that \( \sigma(E) = A \).

(c) There exists a representation \( \tau : \mathfrak{su}_2(\mathbb{C}) \longrightarrow \mathfrak{su}_N(\mathbb{C}) \) such that

\[
\tau(X) = \frac{i}{2}(A + A^\dagger), \quad \text{and} \quad \tau(Y) = \frac{1}{2}(A - A^\dagger).
\]

If representations \( \sigma \) and \( \tau \) as in (b) and (c) exist, then they are unique. In particular, \( \sigma \) is given by \( \sigma = U \sigma J_A U^\dagger \).

**Proof.** (a) \( \Rightarrow \) (b): Let \( \sigma : \mathfrak{sl}_2(\mathbb{C}) \longrightarrow \mathfrak{sl}_N(\mathbb{C}) \) be given by

\[
\sigma(E) := A, \quad \sigma(F) := A^\dagger, \quad \sigma(H) := [A, A^\dagger]. \tag{18}
\]

Then a straightforward calculation shows that \( \sigma \) defines a unitary representation of \( \mathfrak{sl}_2(\mathbb{C}) \).

(b) \( \Rightarrow \) (a): Let \( A = T J_A T^{-1} \). Then by Theorem 11 there exists \( S \in GL_N(\mathbb{C}) \) such that \( \sigma = S \sigma J_A S^{-1} \). This particularly yields \( A = SJ_A S^{-1} \) and

\[
i S J_A J_A^\dagger S^{-1} = i \sigma([E, F]) = \sigma(iH) = \sigma(Z) \in \mathfrak{su}_N(\mathbb{C}).
\]

Hence \( S \) can be chosen unitary, cf. [11], Part I, Ch. I, Sec. 2.

(a) \( \Rightarrow \) (c): Define \( \tau \) as the restriction of (18) to \( \mathfrak{su}_2(\mathbb{C}) \).

(c) \( \Rightarrow \) (b): The complexification of \( \tau \) satisfies (b).

Assume now that \( \sigma, \tilde{\sigma} : \mathfrak{sl}_2 \longrightarrow \mathfrak{sl}_N(\mathbb{C}) \) are representations satisfying (b). Then by part (a) there exist unitary transformations \( S_1, S_2 \) such that

\[
\sigma = S_1 \sigma J_A S_1^\dagger, \quad \text{and} \quad \tilde{\sigma} = S_2 \sigma J_A S_2^\dagger.
\]

In particular this yields

\[
\sigma(F) = S_1 J_A^\dagger S_1^\dagger = \sigma(E) = \tilde{\sigma}(E) = S_2 J_A^\dagger S_2^\dagger = \tilde{\sigma}(F),
\]

and therefore \( \sigma = \tilde{\sigma} \). The uniqueness of \( \tau \) in (c) is obvious. \( \square \)
Remark 14. If in Proposition 13 (c) the condition $\tau(Y) = \frac{1}{2}(A - A^\dagger)$ is omitted, then in general the representation $\tau$ is no longer unique. However, there should be only finitely many of them, which are all equivalent among each other.

The above result implies the following bijective correspondence between unitary representations of $\mathfrak{su}_2(\mathbb{C})$ and $GL_N(\mathbb{C})$-orbits of complex nilpotent matrices.

Corollary 15. Let $\mathcal{O}$ denote the $GL_N(\mathbb{C})$-similarity orbit of a nilpotent matrix $A \in \mathfrak{sl}_N(\mathbb{C})$ with CGJ-form $J_A$.

(a) There exists a unique representation $\tau_{J_A} : \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{su}_N(\mathbb{C})$ with

$$\tau_{J_A}(Y) - i\tau_{J_A}(X) = J_A.$$

(b) Any representation $\tau : \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{su}_N(\mathbb{C})$ with

$$\tau(Y) - i\tau(X) \in \mathcal{O}$$

is unitarily equivalent to $\tau_{J_A}$.

Proof. The existence and uniqueness of $\tau_{J_A}$ follows immediately from part (b) of Proposition 13. Moreover, if $\tau : \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{su}_N(\mathbb{C})$ is any representation with

$$\widetilde{A} = \tau(Y) - i\tau(X) \in \mathcal{O},$$

then Proposition 13 also implies that $\tau$ is unitarily equivalent to $\tau_{J_A}$. \hfill \Box

Note, in view of Remark 14, that the uniqueness part of statement (a) in Corollary 15 is false, if we require $\tau_{J_A}(X) = \frac{1}{2}(J_A + J_A^\dagger)$ instead of the above identity.

In order to summarize our results, we need some further notation. Let $\text{Rep}(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C}))$ denote the algebraic variety of all Lie algebra representations of $\mathfrak{g}$ in $\mathfrak{gl}_N(\mathbb{C})$. For $\mathfrak{g} = \mathfrak{su}_2(\mathbb{C})$ consider the evaluation map defined as

$$\text{ev} : \text{Rep}(\mathfrak{su}_2(\mathbb{C}), \mathfrak{su}_N(\mathbb{C})) \to \mathfrak{sl}_N(\mathbb{C}), \quad \tau \mapsto \tau(Y) - i\tau(X). \quad (19)$$

Note that the image of $\text{ev}$ consists of complex nilpotent matrices and thus is not contained in $\mathfrak{su}_N(\mathbb{C})$. Moreover, $\text{ev}$ is equivariant in the sense that for any unitary transformation $U \in U_N(\mathbb{C})$ one has

$$\text{ev}(U\tau U^\dagger) = U\text{ev}(\tau)U^\dagger.$$

Now, the following theorem is an immediate consequence of the previous result.

Theorem 16. (i) The map $\text{ev} : \text{Rep}(\mathfrak{su}_2(\mathbb{C}), \mathfrak{su}_N(\mathbb{C})) \to \mathfrak{sl}_N(\mathbb{C}), \tau \mapsto \tau(Y) - i\tau(X)$ is injective.
(ii) The map \( ev \) defines a bijection between the sets of unitary equivalence classes of representations \( \tau : \mathfrak{su}_2(\mathbb{C}) \rightarrow \mathfrak{su}_N(\mathbb{C}) \), and unitary orbits of nilpotent matrices in CGJ-form.

It is interesting to see that Theorem 16 can also be derived from more general geometric facts on Lie group orbits. In fact, basic results on the geometry of orbits of Lie algebra representations and nilpotent matrices are already known for some while, mainly in connection with the so-called Kostant–Sekiguchi correspondence; cf. e.g. [26] for a nice exposition on these ideas. Instead of formulating this correspondence in full generality, we restrict ourselves to the special case of \( \mathfrak{sl}_2(\mathbb{C}) \) and \( \mathfrak{su}_2(\mathbb{C}) \) representations. In this situation the Kostant-Sekiguchi correspondence establishes a bijection between certain classes of nilpotent matrices. The following theorem is obtained by specializing the more general theory outlined in [26] to the situation at hand, and goes back to the seminal work by Sekiguchi [25] and Kostant. We define

\[
O_N(\mathbb{C}) := \{ X \in \mathbb{C}^{N \times N} \mid XX^\top = I_N \},
\]

where \( (\cdot)^\top \) denotes transpose.

**Theorem 17 (Kostant-Sekiguchi).**

(a) There is a bijection between \( GL_N(\mathbb{R}) \)-similarity orbits of real nilpotent \((N \times N)\)-matrices, and \( O_N(\mathbb{C}) \)-similarity orbits of complex symmetric, nilpotent \((N \times N)\)-matrices.

(b) There is a bijection between \( GL_N(\mathbb{C}) \)-similarity classes of complex nilpotent \((N \times N)\)-matrices, and \( U_N(\mathbb{C}) \)-similarity classes of complex nilpotent \((N \times N)\)-matrices \( A \), satisfying

\[
[[A, A^\dagger], A] = -2A, \quad [[A, A^\dagger], A^\dagger] = 2A^\dagger.
\]

The above theorem is easily deduced from a more general result in [24], specialized to the situation at hand. We present in the following the formulation of [26], where \( || \cdot ||_F \) denotes the Frobenius norm, i.e.

\[
||A||_F := (\text{tr} \ A A^\dagger)^{\frac{1}{2}}.
\]

**Theorem 18 (Ness).** Let \( \mathcal{O} \) denote the complex similarity orbit of the nilpotent \((N \times N)\)-matrix \( J \) in CGJ-form. Let \( \mathcal{H}_N \) denote the set of complex Hermitian \((N \times N)\)-matrices and

\[
\mu : \mathcal{O} \rightarrow \mathcal{H}_N, \quad \mu(A) = -\frac{[A, A^\dagger]}{||A||_F^2}
\]

the so-called moment map. An element \( A \in \mathcal{O} \) is a critical point of \( ||\mu||_F^2 \) if and only if there exists a real number \( a < 0 \) with

\[
[[A, A^\dagger], A] = -aA, \quad [[A, A^\dagger], A^\dagger] = aA^\dagger.
\]

The set of critical points is nonempty and consists of a single \( U_N(\mathbb{C}) \times \mathbb{R}^+ \) orbit. The function \( ||\mu||_F^2 \) assumes its minimum value exactly at the critical set.
By choosing $a = 2$ in the theorem by Ness (this can always be achieved w.l.o.g. by an appropriate normalization of $A$) we immediately obtain part (b) of the Kostant-Sekiguchi correspondence. For the first part, see [26]. To clarify the algebraic matrix equations appearing in (b) we recall that a triple $(\tilde{H}, \tilde{E}, \tilde{F})$ of complex matrices with trace zero is called a KS-triple (Kostant-Sekiguchi), if

$$
\tilde{H}^\dagger = \tilde{H}, \quad \tilde{E}^\dagger = \tilde{F}
$$

holds. By inspection it is easily seen that $A \in \mathcal{O}$ satisfies the two matrix equations in (b) for $a = -2$, if and only if $(\tilde{H}, \tilde{E}, \tilde{F})$ defined by

$$
\tilde{H} := [A, A^\dagger], \quad \tilde{E} := A, \quad \tilde{F} := A^\dagger
$$

is a KS-triple satisfying

$$
[[\tilde{H}, \tilde{E}], A] = 2\tilde{E}, \quad [[\tilde{H}, \tilde{F}], A] = -2\tilde{F}, \quad [\tilde{E}, \tilde{F}] = \tilde{H}.
$$

Thus $\tilde{H}, \tilde{E},$ and $\tilde{F}$ define a unitary Lie algebra representation of $\mathfrak{sl}_2(\mathbb{C})$, and conversely any unitary representation of $\mathfrak{su}_2(\mathbb{C})$ can be obtained in this way. Therefore the Kostant-Sekiguchi correspondence (b) just describes the bijective correspondence between nilpotent orbits, characterized in Theorem 16. Moreover, this leads to the following result that characterizes the image set of the evaluation map $ev$.

**Corollary 19.** The image of the evaluation map

$$
ev : \text{Rep}(\mathfrak{su}_2(\mathbb{C}), \mathfrak{sl}_N(\mathbb{C})) \longrightarrow \mathfrak{sl}_N(\mathbb{C}), \quad \tau \mapsto \tau(Y) - i\tau(X).
$$

consists precisely of all complex nilpotent matrices $A$, satisfying

$$
[[A, A^\dagger], A] = -2A, \quad [[A, A^\dagger], A^\dagger] = 2A^\dagger.
$$

### 2.3 The Stabilizer and Unitary Orbit of a Representation

Let $\rho : g \longrightarrow \mathfrak{gl}_N(\mathbb{C})$ be a representation of a Lie algebra $g$. We define a group action of a subgroup $G \subset GL_N(\mathbb{C})$ on the space of linear maps $L(g, \mathfrak{gl}_N(\mathbb{C}))$ via

$$G \times L(g, \mathfrak{gl}_N(\mathbb{C})) \longrightarrow L(g, \mathfrak{gl}_N(\mathbb{C})), \quad T.\rho = T \rho(\cdot) T^{-1}.
$$

The **stabilizer** of $\rho$ in $G$ is the subgroup of $G$ defined by

$$
\text{Stab}_G(\rho) = \{T \in G \mid T.\rho = \rho\}.
$$

**Lemma 20.** Let $\rho = c_1 \ast \rho_1 \oplus ... \oplus c_r \ast \rho_r : g \longrightarrow \mathfrak{gl}_N(\mathbb{C})$ be a direct sum where $\rho_i : g \longrightarrow \mathfrak{gl}_N(\mathbb{C}), i = 1, ..., r$ are inequivalent irreducible representations. Then for the stabilizer of $\rho$ in $G \subset GL_N(\mathbb{C})$ it holds

$$
\text{Stab}_G(\rho) = \{T \in G \mid T = \text{diag}(S_1, ..., S_r) \text{ is block diagonal}\},
$$

15
where the \( S_i \) have the structure \( S_i = A_i \otimes I_{N_i}, A_i \in \mathbb{C}^{c_i \times c_i} \). In particular, we obtain the dimension formulas

\[
(a) \quad \dim \text{Stab}_{GL_N(\mathbb{C})}(\rho) = 2 \sum_{i=1}^{r} c_i^2, \quad (b) \quad \dim \text{Stab}_{U_N(\mathbb{C})}(\rho) = \sum_{i=1}^{r} c_i^2.
\]

Proof. We provide the matrix \( T \in G \) with the block structure

\[
T = (T_{ij}), \quad i, j = 1, \ldots, \sum c_i
\]

corresponding to the direct sum decomposition of \( \rho \). The \((i, j)\)-block of the matrix equation

\[
T \rho(x) = \rho(x) T
\]

yields \( \rho_i(x)T_{ij} = T_{ij}\rho_j(x) \) for all \( x \in \mathfrak{g} \). If \( \rho_i = \rho_j \), this implies \( T_{ij} = \lambda_{ij} I_{N_i} \) because of the irreducibility of the \( \rho_i \) and Schur’s Lemma [17]. Now let \( \rho_i \not\cong \rho_j \). If \( N_i > N_j \), the equation

\[
\rho_i T_{ij} = T_{ij} \rho_j
\]

implies that the linear span of \( T_{ij} \) is an invariant subspace of \( \rho_i \). Thus the irreducibility of \( \rho_i \) yields \( T_{ij} = 0 \). Analogously, if \( N_i < N_j \), the kernel of \( T_{ij} \) is an invariant subspace of \( \rho_j \), and hence \( T_{ij} = 0 \). Consider now the case \( N_i = N_j \). Then \( T_{ij} \) cannot be invertible because of the inequivalence of \( \rho_i \) and \( \rho_j \). But then its nontrivial kernel is an invariant subspace of \( \rho_j \) and again \( T_{ij} = 0 \). Thus the stabilizer has the required structure of Eq. (24). The formula for the dimensions is obtained by taking into account that \( T \) has to be in the subgroup \( G \). In particular, in case (a) one has \( A_i \in GL_{c_i}(\mathbb{C}) \) and hence \( 2c_i^2 \) free parameters whereas in (b) the \( A_i \) have to be unitary which admits \( c_i^2 \) free parameters. \( \square \)

For the unitary orbit of a representation \( \rho \) that decomposes into a direct sum of irreducible ones we have the following result.

**Proposition 21.** Let \( \rho \) be equivalent to the representation defined in Lemma 20 and let

\[
\mathcal{O}(\rho) := \{ U.\rho \mid U \in U_N(\mathbb{C}) \}
\]

be the unitary orbit of \( \rho \). Then \( \mathcal{O}(\rho) \) is a compact submanifold of \( L(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C})) \) and its dimension is

\[
\dim \mathcal{O}(\rho) = N^2 - \sum_{i=1}^{r} c_i^2.
\]

**Proof.** The proof is a direct consequence of Lemma 20 (b) and [4], Ch. III, §1.8. \( \square \)

The last result of this subsections yields a relation between the unitary stabilizer of a representation \( \tau \) of \( \mathfrak{su}_2(\mathbb{C}) \) and the unitary stabilizer of the nilpotent matrix \( \tau^C(E) \).

**Proposition 22.** Let \( \tau : \mathfrak{su}_2(\mathbb{C}) \rightarrow \mathfrak{su}_N(\mathbb{C}) \) be a representation, denote by \( \sigma \) its complexification and let \( A := \sigma(E) \). Then it holds

\[
\text{Stab}_{U_N(\mathbb{C})}(\tau) = \text{Stab}_{U_N(\mathbb{C})}(A).
\]

16
Proof. For any subgroup \( G \subset GL_N(\mathbb{C}) \) it holds \( \text{Stab}_G(\tau) = \text{Stab}_G(\sigma) \). On the other hand, denoting
\[
A := \sigma(E), \quad B := \sigma(F),
\]
one has \( \text{Stab}_G(\sigma) = \text{Stab}_G(A) \cap \text{Stab}_G(B) \). By assumption, \( \tau \) is a representation in \( SU_N(\mathbb{C}) \) and hence Proposition 13 yields \( B = A^\dagger \). Therefore \( \text{Stab}_{U_N(\mathbb{C})}(A) = \text{Stab}_{U_N(\mathbb{C})}(A^\dagger) \) and we are done.

3 Least Squares Matching of Representations

In this section we discuss the main objective of this paper, i.e. the analysis and computation of representations, that are in some sense as close as possible to a given one. In general, i.e. for arbitrary Lie algebra representations, it is of course difficult to solve this problem, but we consider an interesting case at the end of the paper. We begin by constructing metrics on the set of all representations of a given Lie algebra \( g \) in \( \mathfrak{gl}_N(\mathbb{C}) \). Let \( \Omega \) be any subset of \( g \) and denote by
\[
\langle \Omega \rangle_L := \bigcap \{ \mathfrak{h} \subset g \mid \mathfrak{h} \text{ is subalgebra and } \Omega \subset \mathfrak{h} \} \tag{29}
\]
the smallest subalgebra generated by \( \Omega \). Although for most parts of the theory it is not necessary, we assume for simplicity that \( \Omega = \{ \omega_1, ..., \omega_r \} \) is finite. Furthermore, let \( || \cdot || \) be a norm on \( \mathfrak{gl}_N(\mathbb{C}) \) and let \( | \cdot | \) be a monotone norm on \( \mathbb{R}^r \), i.e.
\[
0 \leq x_i \leq y_i \quad \text{for } i = 1, ..., r \implies |x| \leq |y|. \tag{30}
\]
For any two representations \( \rho_i : g \to \mathfrak{gl}_N(\mathbb{C}) \), \( i = 1, 2 \), we define
\[
\delta_\Omega(\rho_1, \rho_2) := \left| \left( \left| \rho_1(\omega_1) - \rho_2(\omega_1) \right|, ..., \left| \rho_1(\omega_r) - \rho_2(\omega_r) \right| \right) \right|^\top. \tag{31}
\]

Proposition 23. (a) The map \( \delta_\Omega \) is induced by a semi-norm \( || \cdot ||_\Omega \) on the space \( L(g, \mathfrak{gl}_N(\mathbb{C})) \) of all linear maps from \( g \) to \( \mathfrak{gl}_N(\mathbb{C}) \).

(b) The semi-norm in (a) is a norm on \( L(g, \mathfrak{gl}_N(\mathbb{C})) \) if and only if the linear span of \( \Omega \) is \( g \).

(c) If \( \Omega \) generates \( g \), i.e. \( \langle \omega_1, ..., \omega_r \rangle_L = g \), then \( \delta_\Omega \) defines a metric on the set \( \text{Rep}(g, \mathfrak{gl}_N(\mathbb{C})) \) of all representations of \( g \) in \( \mathfrak{gl}_N(\mathbb{C}) \).

Proof. (a) Define
\[
||| \cdot |||_\Omega := \left| \left( \left| \phi(\omega_1) \right|, ..., \left| \phi(\omega_r) \right| \right) \right|^\top \quad \tag{32}
\]
for all \( \phi \in L(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C})) \). Obviously \( ||| \cdot |||_\Omega \) satisfies \( |||\phi|||_\Omega \geq 0 \) and \( |||\lambda \phi|||_\Omega = |\lambda| \ |||\phi|||_\Omega \) for all \( \phi \in L(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C})) \) and \( \lambda \in \mathbb{K} \). Moreover, by the triangle inequality and the monotony it follows

\[
|||\phi + \psi|||_\Omega = \left( |||\phi(\omega_1)|||, ..., |||\phi(\omega_r)||| \right)^\top 
\leq \left( |||\phi(\omega_1)||| + |||\psi(\omega_1)|||, ..., |||\phi(\omega_r)||| + |||\psi(\omega_r)||| \right)^\top 
\leq |||\phi|||_\Omega + |||\psi|||_\Omega
\]

(33)

Hence \( ||| \cdot |||_\Omega \) is a semi-norm on \( L(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C})) \).

(b) Obviously,

\[
|||\phi|||_\Omega = 0 \iff \phi(\omega_i) = 0 \text{ for all } i = 1, ..., r,
\]

which is equivalent to the fact that the restriction of \( \phi \) to the linear span of \( \Omega \) vanishes. Therefore \( ||| \cdot |||_\Omega \) defines a norm on \( L(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C})) \) if and only if the linear span of \( \Omega \) is \( \mathfrak{g} \).

(c) Let \( \rho_1 \) and \( \rho_2 \) be two representations of \( \mathfrak{g} \). Then we have

\[
\delta_\Omega(\rho_1, \rho_2) = 0 \iff \rho_1(\omega_i) = \rho_2(\omega_i) \text{ for all } i = 1, ..., r.
\]

This implies

\[
\rho_1([\omega_i, \omega_j]) = \rho_1(\omega_i)\rho_1(\omega_j) - \rho_1(\omega_j)\rho_1(\omega_i) = \rho_2(\omega_i)\rho_2(\omega_j) - \rho_2(\omega_j)\rho_2(\omega_i) = \rho_2([\omega_i, \omega_j])
\]

(34)

for all \( i, j = 1, ..., r \) and hence

\[
\delta_\Omega(\rho_1, \rho_2) = 0 \iff \rho_1|_{\langle \omega_1, ..., \omega_r \rangle_L} = \rho_2|_{\langle \omega_1, ..., \omega_r \rangle_L}.
\]

Therefore \( \delta_\Omega \) yields a metric on \( \text{Rep}(\mathfrak{g}, \mathfrak{gl}_N(\mathbb{C})) \) if \( \langle \omega_1, ..., \omega_r \rangle_L = \mathfrak{g} \).

Note that the ”only-if”-part of statement (c) in Proposition 23 is in general not true as the following example shows.

**Example** Let \( \mathfrak{g} := \left\{ \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix} \mid \alpha, \beta, \gamma \in \mathbb{C} \right\} \) and \( \Omega := \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \). Then for any representation \( \rho \) of \( \mathfrak{g} \) in \( \mathfrak{gl}_1(\mathbb{C}) = \mathbb{C} \) we have

\[
\rho\left( \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) = \rho\left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) 
= \rho\left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) \rho\left( \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) - \rho\left( \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) \rho\left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = 0.
\]

(35)
Hence $\delta_\Omega$ is a metric on $\text{Rep}(g, \mathbb{C})$, but $\Omega$ does not generate the Lie algebra $g$.

For the remainder of this section we concentrate on the following choice of norms in Eq. (31), which are of particular interest for later applications. Let $\| \cdot \|$ be the Frobenius norm on $\mathfrak{gl}_N(\mathbb{C})$ and let $| \cdot |$ be the Euclidian norm on $\mathbb{R}^r$, i.e. $|x| = (\sum_{i=1}^{r} x_i^2)^{\frac{1}{2}}$. With these choices, we write $d_\Omega$ instead of $\delta_\Omega$ and obtain the explicit formula

$$d_\Omega(\rho_1, \rho_2) = \left( \sum_{i=1}^{r} \|\rho_1(\omega_i) - \rho_2(\omega_i)\|_F^2 \right)^{\frac{1}{2}}. \quad (36)$$

Now, with the above notation we can state what we call the general least squares matching problem of two representations. Given two representations $\rho_i : g \to \mathfrak{gl}_N(\mathbb{C})$, $i = 1, 2$, find matrices $S_0, T_0 \in \text{GL}_N(\mathbb{C})$ such that

$$d_\Omega(S_0 \rho_1 S_0^{-1}, T_0 \rho_2 T_0^{-1}) = \inf_{S,T \in \text{GL}_N(\mathbb{C})} d_\Omega(S \rho_1 S^{-1}, T \rho_2 T^{-1}). \quad (37)$$

In general, however, such an optimal $(S_0, T_0)$ might not exist.

**Example.** Let $\rho_1, \rho_2 : \mathbb{R} \to \mathfrak{gl}_2(\mathbb{C})$ be defined by

$$\rho_1(t) := \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}, \quad \rho_2(t) := \begin{bmatrix} t & 1 \\ 0 & t \end{bmatrix}$$

and choose $\Omega = \{1\}$ as a generating set of the Lie algebra $\mathbb{R}$. Then we have

$$\inf_{S,T \in \text{GL}_2(\mathbb{C})} d_\Omega(S \rho_1 S^{-1}, T \rho_2 T^{-1}) = 0.$$ 

However, there is no $(S_0, T_0)$ such that

$$T_0 \rho_2(1) T_0^{-1} = T_0 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} T_0^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = S_0 \rho_1(1) S_0^{-1}.$$ 

Note that the aim of the least square matching problem is not to match the images of the representations rather than the representations themselves. Hence, even if the images of two representations coincide, their minimal distance need not necessarily be equal to zero.

In many applications it is more meaningful to restrict the class of admissible coordinate transformations to a suitable subgroup $G$ of $\text{GL}_N(\mathbb{C})$, where the minimum is known to exist. A natural choice in quantum mechanics is the unitary group $G = U_N(\mathbb{C})$. In this case, the compactness of $U_N(\mathbb{C})$ guarantees the existence of a minimum.

**Unitary least squares matching Problem.** For two given representations $\rho_i : \mathfrak{g} \to \mathfrak{gl}_N(\mathbb{C})$, $i = 1, 2$, find matrices $U_0, V_0 \in U_N(\mathbb{C})$ such that

$$d_\Omega(U_0 \rho_1 U_0^\dagger, V_0 \rho_2 V_0^\dagger) = \min_{U,V \in U_N(\mathbb{C})} d_\Omega(U \rho_1 U^\dagger, V \rho_2 V^\dagger).$$
Lemma 24. Let $\rho_i : g \to gl_N(\mathbb{C})$, $i = 1, 2$, be two representations of the Lie algebra $g$ and let $\Omega = \{\omega_1, \ldots, \omega_r\}$ be a finite set of generators of $g$.

(a) For all $U, V \in U_N(\mathbb{C})$ it holds
$$d_\Omega(U\rho_1U^\dagger, V\rho_2V^\dagger) = d_\Omega(\rho_1, U^\dagger V\rho_2V^\dagger U).$$

In particular $(U_0, V_0)$ solves the unitary least squares matching problem if and only if $U_0^\dagger V_0$ minimizes the distance function
$$U_N(\mathbb{C}) \ni U \mapsto d_\Omega(\rho_1, U\rho_2U^\dagger).$$

(b) The minima of the distance function (38) coincide with the maxima of the associated trace function
$$U \mapsto \sum_{i=1}^r \Re \tr(\rho_1(\omega_i)^\dagger U\rho_2(\omega_i)U^\dagger).$$

Proof. (a) By definition of $d_\Omega$,
$$d_\Omega(U\rho_1U^\dagger, V\rho_2V^\dagger) = \left(\sum_{i=1}^r \left\| U\rho_1(\omega_i)U^\dagger - V\rho_2(\omega_i)V^\dagger \right\|^2_F \right)^{\frac{1}{2}}.$$

The Frobenius norm is invariant under unitary transformations, and thus
$$\left\| U\rho_1(\omega_i)U^\dagger - V\rho_2(\omega_i)V^\dagger \right\|^2_F = \left(\sum_{i=1}^r \left| U\rho_1(\omega_i)U^\dagger - V\rho_2(\omega_i)V^\dagger \right|^2 \right)^{\frac{1}{2}}$$
$$= \left| U\rho_1(\omega_i)U^\dagger \right|^2_F + \left| V\rho_2(\omega_i)V^\dagger \right|^2_F - 2 \Re \tr \left( (U\rho_1(\omega_i)U^\dagger)V\rho_2(\omega_i)V^\dagger \right)$$
$$= \left| U\rho_1(\omega_i)U^\dagger \right|^2_F + \left| V\rho_2(\omega_i)V^\dagger \right|^2_F - 2 \Re \tr \left( U\rho_1(\omega_i)U^\dagger V\rho_2(\omega_i)V^\dagger \right)$$
$$= \left| U\rho_1(\omega_i)U^\dagger - V\rho_2(\omega_i) \right|^2_F$$
for $i = 1, \ldots, r$. Hence assertion (a) follows.

(b) The same argument as in (a) shows
$$d_\Omega(\rho_1, U\rho_2U^\dagger)^2 = K - 2 \sum_{i=1}^r \Re \tr(\rho_1(\omega_i)^\dagger U\rho_2(\omega_i)U^\dagger),$$
where the constant $K$ is given as
$$K = \sum_{i=1}^r \left| \rho_1(\omega_i) \right|^2_F + \sum_{i=1}^r \left| \rho_2(\omega_i) \right|^2_F.$$

Thus the minima of the distance function (38) coincide with the maxima of the associated trace function. \qed
Proposition 25 (Critical Point Condition). A unitary transformation $U_0 \in U_N(\mathbb{C})$ is a critical point of the associated trace function (39) if and only if

$$\sum_{i=1}^{r} \left[ \rho_1(\omega_i)^\dagger, U_0^\dagger \rho_2(\omega_i) U_0^\dagger \right]$$

is Hermitian.

Proof. Let $\Psi \in u_N(\mathbb{C})$ be an arbitrary Lie algebra element and identify the tangent space of $U_0 \in U_N(\mathbb{C})$ in the natural way with $u_N(\mathbb{C})$, i.e. $T_{U_0}U_N(\mathbb{C}) = u_N(\mathbb{C}) \cdot U_0$. Thus the derivation of the trace function $f$ in $U_0$ is given by

$$Df(U_0)\Psi = \frac{d}{dt} \sum_{i=1}^{r} \text{Re tr} \left( \rho_1(\omega_i)^\dagger \exp(t\Psi) U_0^\dagger \rho_2(\omega_i) U_0^\dagger \exp(-t\Psi) \right) \bigg|_{t=0}$$

$$= \sum_{i=1}^{r} \text{Re tr} \left( \rho_1(\omega_i)^\dagger [\Psi, U_0^\dagger \rho_2(\omega_i) U_0^\dagger] \right)$$

$$= - \sum_{i=1}^{r} \text{Re tr} \left( [\rho_1(\omega_i)^\dagger, U_0^\dagger \rho_2(\omega_i)] \right)$$

$$= - \text{Re tr} \left( \Psi \left( \sum_{i=1}^{r} [\rho_1(\omega_i)^\dagger, U_0^\dagger \rho_2(\omega_i)] \right) \right).$$

(41)

Hence a necessary and sufficient condition for the derivative to vanish is that

$$\sum_{i=1}^{r} \left[ \rho_1(\omega_i)^\dagger, U_0^\dagger \rho_2(\omega_i) U_0^\dagger \right]$$

is Hermitian.

Concerning subsequent applications in quantum mechanics, the following proposition relates the unitary least squares matching problem of unitary $su_2(\mathbb{C})$-representations and their complexifications to $sl_2(\mathbb{C})$.

Proposition 26. Let $\tau, \tilde{\tau} : su_2(\mathbb{C}) \to su_N(\mathbb{C})$, be two representations and let $\sigma$ and $\tilde{\sigma}$ be their complexifications. Then the corresponding distance functions and their associated trace functions coincide, i.e.

$$d_E(\sigma, U\tilde{\sigma}U^\dagger) = d_{\{X,Y\}}(\tau, U\tilde{\tau}U^\dagger)$$

$$\text{Re tr} \sigma(E)^\dagger U\tilde{\sigma}(E) U^\dagger = \text{tr} \tau(X)^\dagger U\tilde{\tau}(X) U^\dagger + \text{tr} \tau(Y)^\dagger U\tilde{\tau}(Y) U^\dagger$$

for all $U \in U_N(\mathbb{C})$. 

21
Proof. Using the identity $E = Y - iX$ one has

$$
d_E(\sigma, U\bar{\sigma}U^\dagger)^2 = ||\sigma(E) - U\bar{\sigma}(E)U^\dagger||_F^2 =
||\tau(Y) - i\tau(X) - U\bar{\tau}(Y)U^\dagger + iU\bar{\tau}(X)U^\dagger||_F^2 =
||\tau(X) - U\bar{\tau}(X)U^\dagger||_F^2 + ||\tau(Y) - U\bar{\tau}(Y)U^\dagger||_F^2 -
2\Re \text{tr} \left( i(\tau(X) - U\bar{\tau}(X)U^\dagger)(\tau(Y) - U\bar{\tau}(Y)U^\dagger) \right).
$$

(42)

The last term vanishes because, $\tau$ and $\bar{\tau}$ are representations in $\mathfrak{su}_N(\mathbb{C})$, and hence $U\bar{\tau}(X)U^\dagger$ and $U\bar{\tau}(Y)U^\dagger$ are skew-Hermitian for all $U \in U_N(\mathbb{C})$. The second identity follows by a similar calculation.

To put these elementary definitions and results in a somewhat broader perspective, we introduce the following concept of the relative numerical range of two representations. Although it would be possible, we do not attempt to define this concept in full generality. Here we rather focus on the case induced by the metric $d_\Omega$.

**Definition 27.** Let $\rho, \bar{\rho} : \mathfrak{g} \to \mathfrak{gl}_N(\mathbb{C})$ denote two arbitrary representations and let $\Omega = \{\omega_1, \ldots, \omega_r\}$ be a subset of $\mathfrak{g}$. The $\Omega$-relative numerical range of $\rho$ and $\bar{\rho}$ is the subset of the complex plane

$$W_\Omega(\rho, \bar{\rho}) := \left\{ \sum_{i=1}^r \text{tr}(\rho(\omega_i)U\bar{\rho}(\omega_i)U^\dagger) \mid U \in U_N(\mathbb{C}) \right\}.$$

Of course, if $\Omega = \{\omega\}$, then this is just the usual $\rho(\omega)$-numerical range of $\bar{\rho}(\omega)$ and thus there seems to be nothing worth mentioning about this special case. The interesting fact, however, is that, due to the special structure of unitary $\mathfrak{sl}_2(\mathbb{C})$- and $\mathfrak{su}_2(\mathbb{C})$-representations, the relative numerical ranges $W_E(\sigma, \bar{\sigma})$ and $W_{\{X,Y\}}(\tau, \bar{\tau})$ are always discs in $\mathbb{C}$ or $\mathbb{R}$, respectively.

**Theorem 28.** Let $\tau, \bar{\tau} : \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{su}_N(\mathbb{C})$, be two representations and let $\sigma$ and $\bar{\sigma}$ be their complexifications. Then it holds

(a) $\Re W_E(\sigma, \bar{\sigma}) = W_{\{X,Y\}}(\tau, \bar{\tau})$.

(b) The relative numerical range $W_E(\sigma, \bar{\sigma})$ is a disc in the complex plane, centered at the origin.

(c) The relative numerical range $W_{\{X,Y\}}(\tau, \bar{\tau})$ is an interval in $\mathbb{R}$, centered at the origin.

**Proof.** (a) The first part is an immediate consequence of Proposition 26. (b) By Proposition 13, the nilpotent matrices $\sigma(E), \bar{\sigma}(E)$ are unitarily equivalent to their
CGJ-form. By Theorem 2.1 (e) in [22] this shows that $\sigma(E), \tilde{\sigma}(E)$ are block-shift nilpotent operators. The result now follows from Corollary 2.2 in [22]. (c) The last part follows from (a) and (b).

The class of block-shift matrices introduced in [22] is rather special and does not include arbitrary nilpotent matrices. For example, the nilpotent matrix

$$N = \begin{bmatrix}
0 & 0 & 0 \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\end{bmatrix},$$

(43)
is not block-shift, but it is of the form $N = \sigma(E)$ for a suitable representation $\sigma : \mathfrak{sl}_2(\mathbb{C}) \rightarrow \mathfrak{sl}_3(\mathbb{C})$. Thus the above theorem cannot be extended to arbitrary pairs of representations of $\mathfrak{sl}_2(\mathbb{C})$.

A challenge is of course to compute the radius of the disc $W(\tau, \tilde{\tau})$, but no explicit formula is available. Since the Clebsch-Gordan multiplicities are a complete invariant for the representations $\tau, \tilde{\tau}$, it seems reasonable to expect that one can express the radius in terms of them. In the last section we will consider an example arising from physics, which leads to a conjecture on a formula for the radius via the Clebsch-Gordan decomposition. The general case, however remains open, even with regard to formulating a reasonable conjecture.

To illustrate the use of representation theory in computing normal forms of matrices, we briefly discuss the problem of simultaneous diagonalization of Hermitian matrices. This is closely related to representations of Abelian Lie algebras. In the next section, we will study an example for matching representations of semisimple Lie algebras.

Consider two $r$-tuples of commuting Hermitian $(N \times N)$-matrices

$$(X_1, \ldots, X_r), \quad X_i = X_i^\dagger, \quad [X_i, X_j] = 0, \quad i, j = 1, \ldots, r$$

$$(Y_1, \ldots, Y_r), \quad Y_i = Y_i^\dagger, \quad [Y_i, Y_j] = 0, \quad i, j = 1, \ldots, r$$

(44)

and define representations $\rho_i : \mathbb{C}^r \rightarrow \mathfrak{gl}_n(\mathbb{C}), i = 1, 2$, of the Abelian Lie algebra $\mathbb{C}^r$ via

$$\rho_1(e_i) = X_i, \quad \rho_2(e_i) = Y_i, \quad i = 1, \ldots, r,$$

where $e_i$ denotes the $i$-th standard basis vector of $\mathbb{C}^n$. Since $X_1, \ldots, X_r$ and $Y_1, \ldots, Y_r$, respectively, are commuting Hermitian matrices, there exist unitary transformations $V_0$ and $W_0$ which simultaneously diagonalize all $X_i$ and $Y_i$, respectively, i.e.

$$V_0 X_i V_0^\dagger = \Lambda_i, \quad \Lambda_i \text{ real and diagonal, } \quad i = 1, \ldots, r$$

$$W_0 Y_i W_0^\dagger = \Sigma_i, \quad \Sigma_i \text{ real and diagonal, } \quad i = 1, \ldots, r.$$ 

(45)

The unitary transformations $V_0$ and $W_0$ are closely related to the critical points of the trace function

$$U \mapsto \sum_{i=1}^r \text{tr}(\rho_1(e_i)U\rho_2(e_i)U^\dagger)$$

(46)
and hence to the corresponding unitary least squares matching problem of $\rho_1$ and $\rho_2$. Note that the trace function (46) is a generalization of the function discussed in [3] and the following theorem partially extends the results therein.

**Theorem 29.** (a) Let $\rho_1, \rho_2$ and $V_0, W_0$ be defined as above and let $\Omega = \{e_1, ..., e_r\}$.

Then $U_0 := V_0^\dagger W_0$ is a critical point of the trace function (46).

(b) Let $V_0, W_0$ be defined as above and denote by $\Pi \subset U_N(\mathbb{C})$ the finite subgroup of all permutation matrices. If there exists $P \in \Pi$ such that $\Lambda_i$ and $P\Sigma_i P^\dagger$ are simultaneously ordered such that

$$\text{tr } \Lambda_i P\Sigma_i P^\dagger = \max_{Q \in \Pi} \text{tr } \Lambda_i Q \Sigma_i Q^\dagger$$

(47)

for all $i = 1, ..., r$, then the maximum of the trace function (46) and hence the minimum of the associated distance function, is given by

$$\max_{U \in U_N(\mathbb{C})} \sum_{i=1}^r \text{tr}(\rho_1(e_i)U \rho_2 U^\dagger) = \sum_{i=1}^r \text{tr}(\rho_1(e_i)V_0^\dagger P W_0 \rho_2 W_0^\dagger P^\dagger V_0),$$

$$\min_{U \in U_N(\mathbb{C})} d_\Omega(\rho_1, U \rho_2 U^\dagger) = d_\Omega(\rho_1, V_0^\dagger P W_0 \rho_2 W_0^\dagger P^\dagger V_0).$$

(48)

**Proof.** (a) The first part immediately follows from the critical point condition of Proposition 25, as

$$\left[\rho_1(e_i)^\dagger, U_0 \rho_2(e_i)U_0^\dagger\right] = V_0^\dagger \left[V_0 \rho_1(e_i)^\dagger V_0^\dagger, W_0 \rho_2(e_i)W_0^\dagger\right] V_0 = V_0^\dagger \left[\Lambda_i, \Sigma_i\right] V_0 = 0$$

for all $i = 1, ..., r$.

(b) By [3], each summand in Eq. (46) satisfies the identity

$$\max_{U \in U_N(\mathbb{C})} \text{tr}(\rho_1(e_i)U \rho_2(e_i)U^\dagger) = \max_{Q \in \Pi} \text{tr}(\Lambda_i Q \Sigma_i Q^\dagger).$$

Therefore we have

$$\sum_{i=1}^r \text{tr}(\rho_1(e_i)V_0^\dagger P W_0 \rho_2 W_0^\dagger P^\dagger V_0) \leq \max_{U \in U_N(\mathbb{C})} \sum_{i=1}^r \text{tr}(\rho_1(e_i)U \rho_2 U^\dagger)$$

$$\leq \sum_{i=1}^r \max_{U \in U_N(\mathbb{C})} \text{tr}(\rho_1(e_i)U \rho_2 U^\dagger) = \sum_{i=1}^r \text{tr}(\rho_1(e_i)V_0^\dagger P W_0 \rho_2 W_0^\dagger P^\dagger V_0).$$

(49)

We believe that part (b) of the previous theorem is valid even without any assumption on the ordering condition (47).
Conjecture. Let $V_0, W_0, \Omega$ and $\Pi$ be defined as above. The maximum of the trace function (46) and hence the minimum of the associated distance function is given by

$$
\max_{U \in U_N(C)} \sum_{i=1}^r \text{tr}(\rho_1(e_i)U\rho_2U^\dagger) = \max_{P \in \Pi} \sum_{i=1}^r \text{tr}(\rho_1(e_i)V_0^\dagger PW_0\rho_2W_0^\dagger P^\dagger V_0),
$$

$$
\min_{U \in U_N(C)} d_\Omega(\rho_1, U\rho_2U^\dagger) = \min_{P \in \Pi} d_\Omega(\rho_1, V_0^\dagger PW_0\rho_2W_0^\dagger P^\dagger V_0).
$$

(50)

It is challenging to explore the more general situation of representations of nilpotent, rather than Abelian Lie algebras. This might lead to interesting new matrix classification problems that would on the one hand be more general than simultaneous diagonalization of commuting matrices, but might allow on the other hand more specific classification results than for solvable Lie algebras. Moreover, the algorithmic aspects of these problems seem not to be explored at all.

4 An Application to NMR-Spectroscopy

In this section, motivated by applications in NMR spectroscopy and quantum computing, cf. [12] and [27], we consider the maximization task for the so-called transfer function given by

$$
f_n : U_{2n+1}(\mathbb{C}) \longrightarrow \mathbb{R}, \quad f_n(U) = \text{Re} \text{tr}(C_n^\dagger UA_nU^\dagger) \quad (51)
$$

where for any $n \in \mathbb{N}$ the nilpotent matrices $C_n, A_n \in \mathfrak{gl}_{2n+1}(\mathbb{C})$ are recursively defined as follows.

$$
A_n := \begin{bmatrix} N_n & 0 \\ 0 & N_n \end{bmatrix}, \quad \text{with } N_n := \begin{bmatrix} N_{n-1} & 0 \\ 0 & N_{n-1} \end{bmatrix}, \quad N_0 := 0,
$$

$$
C_n := \begin{bmatrix} 0 & 0 \\ \mathbb{I}_{2n} & 0 \end{bmatrix}, \quad C_0 := \begin{bmatrix} 0 \\ 0 & 1 \end{bmatrix}.
$$

(52)

Here and in the sequel zero entries denote zero matrices of appropriate size. Thus for $n = 1, 2, 3$ we have

$$
C_1 = \begin{bmatrix} 0 \\ \mathbb{I}_2 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0 & 0 \\ \mathbb{I}_4 & 0 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 0 & 0 \\ \mathbb{I}_8 & 0 \end{bmatrix}
$$

(53)
\[
A_1 = \begin{bmatrix}
0 & 0 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix},
A_2 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0
\end{bmatrix},
A_3 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0
\end{bmatrix}
\]

To see the connection to the previously discussed least squares matching problems, we introduce the following representations of \( \mathfrak{su}_2(\mathbb{C}) \). Let \( 0_{\frac{3}{2}} : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}_2(\mathbb{C}) \) denote the trivial representation of \( \mathfrak{sl}_2(\mathbb{C}) \) in \( \mathfrak{gl}_2(\mathbb{C}) \). Hence \( 0_{\frac{3}{2}} = \sigma_0 \oplus \sigma_0 = 2 \ast \sigma_0 \). Define \( \gamma_n \) and \( \alpha_n \) by

\[
\begin{align*}
\gamma_n, \alpha_n & : \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{su}_{2n+1}(\mathbb{C}) \\
\gamma_n & := \tau_1 \otimes \sigma_{\frac{n}{2}} \\
\alpha_n & := 0_{\frac{3}{2}} \otimes \tau_{\frac{n}{2}}.
\end{align*}
\]

**Lemma 30.** For any \( n \in \mathbb{N} \) it holds \( \gamma_n^C(E) = C_n \), \( \sigma_{\frac{n}{2}}^C(E) = N_n \) and \( \alpha_n^C(E) = A_n \).

**Proof.** It is easily seen, that \( \gamma_n^C(E) = C_n \). We prove the second identity by induction. The assertion is obviously true for \( n = 1 \). Assume, it holds for \( n - 1 \). Then

\[
\sigma_{\frac{n}{2}}^C(E) = \sigma_0^C(E) \otimes \sigma_{\frac{n-1}{2}} \otimes \tau_{\frac{n}{2}}
\]

which proves the induction step. The last statement follows immediately, too. \( \square \)

**Corollary 31.** The \( C_n \)-numerical range of \( A_n \) is a circular disc in the complex plane, centered at the origin.

**Proof.** The proof is a direct consequence of Lemma 30 and Theorem 28. \( \square \)

**Theorem 32.** The transfer function \((51)\) coincides with the trace function

\[
U \mapsto \text{tr} \gamma_n(X)^\dagger U \alpha_n(X) U^\dagger + \text{tr} \gamma_n(Y)^\dagger U \alpha_n(Y) U^\dagger.
\]

**Proof.** The proof is an immediate consequence of Proposition 26 and Lemma 30. \( \square \)
The transfer functions can either be regarded as real valued functions from the unitary orbit of $C_n$: 

$$f_n : \mathcal{O}_{C_n} := \{ UC_n U^\dagger \mid U \in U_{2n+1}(\mathbb{C}) \} \rightarrow \mathbb{R}, \quad \widetilde{C} \mapsto \text{Re} \text{tr}(\widetilde{C} A_n^\dagger)$$

or equivalently as a real valued function from the unitary orbit of $A_n$: 

$$f_n : \mathcal{O}_{A_n} := \{ UA_n U^\dagger \mid U \in U_{2n+1}(\mathbb{C}) \} \rightarrow \mathbb{R}, \quad \widetilde{A} \mapsto \text{Re} \text{tr}(C_n \widetilde{A}^\dagger).$$

For optimization tasks of the transfer functions, it is of interest to know the dimensions of these unitary orbits. As a consequence of Theorem 6, the $n$-fold tensor product of the standard irreducible representation $\tau_n^{1/2}$ of $\mathfrak{su}_2(\mathbb{C})$ and $\sigma_n^{1/2}$ of $\mathfrak{sl}_2(\mathbb{C})$ respectively, decomposes into

$$\tau_n^{1/2} \cong \bigoplus_{\nu=0}^{n/2} c^{(n)}_{\nu} \star \tau_\nu, \quad \text{and} \quad \sigma_n^{1/2} \cong \bigoplus_{\nu=0}^{n/2} c^{(n)}_{\nu} \star \sigma_\nu, \quad c^{(n)}_{\nu} \in \mathbb{N}. \quad (57)$$

The next lemma shows how to compute the Clebsch-Gordan multiplicities $c^{(n)}_{\nu} \in \mathbb{N}$ in Eq. (57) recursively.

**Lemma 33.** For $n \in \mathbb{N}_0$ and $\nu \in \frac{1}{2}\mathbb{N}_0$ it holds the recursive relation

$$c^{(n)}_{\nu} = 1, \quad c^{(n)}_{\nu} + c^{(n)}_{\nu + 1},$$

where $c^{(0)}_{\nu} := 0$ if $\nu < 0$ or $\nu > \frac{n}{2}$.

**Proof.** Proposition 6 yields

$$\tau_n^{1/2} \cong \bigoplus_{\nu=0}^{n/2} \tau_\nu \cong \bigoplus_{\nu=0}^{n/2} \bigoplus_{\nu'=0}^{n/2} (c^{(n)}_{\nu} \star \tau_\nu \otimes \tau_\nu') \cong \bigoplus_{\nu=0}^{n/2} \bigoplus_{\nu'=0}^{n/2} c^{(n)}_{\nu} \star (\tau_\nu \otimes \tau_{\nu'})$$

In Table 1 the Clebsch-Gordan multiplicities $c^{(n)}_{\nu}$ are listed for some small values of $n$ and $\nu$. The following theorem is a consequence of Propositions 20, 22 and Eq. (57).

**Theorem 34.** The following formulas for the dimensions of the unitary orbits hold.

(a) $\dim \mathcal{O}_{A_n} = 2^{2n+2} - 2 \sum_{\nu=0}^{n/2} (c^{(n)}_{\nu})^2$ \quad (b) $\dim \mathcal{O}_{C_n} = 3 \cdot 2^{2n}$
\begin{table}[h]
\centering
\begin{tabular}{c|cccccccccc}
\(n\) & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \ldots \\
\hline
0 & 1 & 0 & 1 & 0 & 2 & 0 & 5 & 0 & 14 & 0 & 42 \\
\frac{1}{2} & 0 & 1 & 0 & 2 & 0 & 5 & 0 & 14 & 0 & 42 & 0 \\
1 & 0 & 0 & 1 & 0 & 3 & 0 & 9 & 0 & 28 & 0 & 90 \\
\frac{3}{2} & 0 & 0 & 0 & 1 & 0 & 4 & 0 & 14 & 0 & 48 & 0 \\
2 & 0 & 0 & 0 & 0 & 1 & 0 & 5 & 0 & 20 & 0 & 75 \\
\frac{5}{2} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 6 & 0 & 27 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 7 & 0 & 35 & \ldots \\
\end{tabular}
\caption{The Clebsch-Gordan multiplicities \(c^{(n)}_{\nu}\) in Eq. (57).}
\end{table}

Proof. (a) Lemma 20 together with Eq. (57) yields

\[ \dim \text{Stab}_{U_{2n+1}(\mathbb{C})} = 2 \sum_{\nu=0}^{n/2} (c^{(n)}_{\nu})^2. \]

Furthermore, \(\dim U_{2n+1}(\mathbb{C}) = 2^{2n+2}\) and the claim follows from Proposition 22. Part (b) is proven in the same way. \(\square\)

We now show that the subsequently conjectured maxima of the transfer functions \(f_n\) can be achieved by unitary transformations, by transforming \(\gamma_n\) and \(\alpha_n\) into their standard irreducible sum decomposition. In preparation of this result, we state the following lemma, whose straightforward proof is left to the reader.

**Lemma 35.** There exists a permutation \(P_n \in U_{2n+1}(\mathbb{C})\) such that

\[ P_n^{\dagger} \gamma_n P_n = 0_2 \otimes \sigma_{\frac{1}{2}} \otimes 0_{n-1}^{n-1} = (\sigma_{\frac{1}{2}} \otimes 0_{\frac{1}{2}}^{n-1}) \oplus (\sigma_{\frac{1}{2}} \otimes 0_{\frac{1}{2}}^{n-1}). \]

**Theorem 36.** For \(n = 2l + 1, \ l \in \mathbb{N}\), there exists a unitary transformation \(V_n \in U_{2n}(\mathbb{C})\) such that

\[ \text{tr} \ C_{n-1}^{\dagger} V_n N_n V_n^{\dagger} = \sum_{k=0}^{l} c_{(2k+1)/2}^{(n)} \sum_{j=0}^{k} \sqrt{(2j + 1)(2k + 1 - 2j)}, \]

where the \(c_{(2k+1)/2}^{(n)}\) are the Clebsch-Gordan multiplicities in Eq. (57). Furthermore, let \(P_n\) be as in Lemma 35 and let \(U_n := P_n^{\dagger} (I_2 \otimes V_n)\). Then it holds

\[ \text{tr} \ C_n^{\dagger} U_n A_n U_n^{\dagger} = 2 \sum_{k=0}^{l} c_{(2k+1)/2}^{(n)} \sum_{j=0}^{k} \sqrt{(2j + 1)(2k + 1 - 2j)}. \]
Proof. By Theorem 6 and Eq. (57) there exist unitary transformations $R_n, S_n \in U_{2n}(\mathbb{C})$ such that

$$R_n \gamma_{n-1} C R_n^\dagger = 2^{n-1} \sigma_{\frac{1}{2}}, \quad S_n \sigma_{\frac{1}{2}} S_n^\dagger = \bigoplus_{\nu=0}^{n/2} c_{\nu}^{(n)} \sigma_{\nu}.$$ 

A straightforward calculation shows

$$\text{Re tr} \left( (k+1) \sigma_{\frac{1}{2}} (E) \right)^\dagger \sigma_{n+1} (E) = \sum_{j=0}^{k} \sqrt{(2j+1)(2k+1-2j)}$$

for $k \in \mathbb{N}_0$. Taking into account that $c_{\nu}^{(n)} = 0$ for $n$ even and $\nu$ integer, we obtain

$$\text{Re tr} \left( 2^{n-1} \sigma_{\frac{1}{2}} (E) \right)^\dagger \bigoplus_{\nu=0}^{n/2} c_{\nu}^{(n)} \sigma_{\nu} (E) = \sum_{k=0}^{l} c_{(2k+1)/2}^{(n)} \sum_{j=0}^{k} \sqrt{(2j+1)(2k+1-2j)}.$$ 

Moreover, let $V_n := R_n^\dagger S_n$ and define $U_n := P_n^\dagger (\mathbb{I}_2 \otimes V_n)$. Then the first identity follows immediately by Lemma 30 and the second identity is a consequence of Lemma 35. □

**Proposition 37.** The unitary transformation $U_n \in U_{2n+1}(\mathbb{C})$, chosen as in Theorem 36, is a critical point of the transfer function $f_n$.

Proof. By Proposition 25, it remains to show that the commutator $[\gamma_n C, U_n \alpha_n C] U_n^\dagger$ is Hermitian. Recall that $U_n = P_n^\dagger (\mathbb{I}_2 \otimes V_n)$ where $P_n$ and $V_n = R_n^\dagger S_n$ are chosen such that

$$\tilde{C}_n := (\mathbb{I}_2 \otimes R_n) P_n \gamma_n C (E) P_n^\dagger (\mathbb{I}_2 \otimes R_n)^\dagger \quad \text{and} \quad \tilde{A}_n := (\mathbb{I}_2 \otimes S_n) \alpha_n C (E) (\mathbb{I}_2 \otimes S_n)^\dagger$$

are in CGJ-form. Therefore, the commutator $[\tilde{C}_n, \tilde{A}_n]$ is real and diagonal and hence $[\gamma_n C, U_n \alpha_n C] U_n^\dagger = P_n^\dagger (\mathbb{I}_2 \otimes R_n)^\dagger [\tilde{C}, \tilde{A}](\mathbb{I}_2 \otimes R_n) P_n$ is Hermitian. □

It is known that for $n = 1, 2$, the values given in Theorem 36 coincide with the maximal values of the transfer function (51), cf. [16]. We therefore have the following conjecture:

**Conjecture.** Let $n \in \mathbb{N}$ be odd and $V_n, W_n \in U_{2n+1}(\mathbb{C})$ such that $V_n \gamma_n C (E) V_n^\dagger$ and $W_n \alpha_n C (E) W_n^\dagger$ are in CGJ-form. Then $U_n := V_n^\dagger W_n$ is a maximum of the transfer function and the maximal value is given by Theorem 36.

Note that for $n \in \mathbb{N}$ even the above conjecture is false. Albeit the CGJ-forms of $\gamma_n C$ and $\alpha_n C$ lead to critical values of the transfer function, they are not maximal and topped by the values given in the next theorem.

**Theorem 38** (Doubling Argument). Let $U_n \in U_{2n+1}(\mathbb{C})$ be chosen as in Theorem 36. Then there exists a unitary $U_{n+1} \in U_{2n+2}(\mathbb{C})$ such that

$$\text{tr} C_{n+1}^\dagger U_{n+1} A_{n+1} U_{n+1}^\dagger = 2 \text{tr} C_n^\dagger U_n A_n U_n^\dagger.$$
By Lemma 35, there exists a permutation $P_{n+1} \in U_{2n+2}$ such that

$$P_{n+1} \gamma_{n+1}^C P_{n+1}^\dagger = 0_\frac{1}{2} \otimes \sigma_{\frac{1}{2}} \otimes 0^1_{\frac{1}{2}}.$$  

Moreover, it holds

$$\sigma_{\frac{1}{2}}^{n+1} (E) = (\sigma_{\frac{1}{2}} \otimes \sigma_{\frac{1}{2}}^n) (E) = E \otimes I_{2^n} + I_2 \otimes N_n. \quad (60)$$

Now let $U_n = P_n^\dagger (I_2 \otimes V_n)$ be chosen as in Theorem 36 and define

$$U_{n+1} := P_{n+1}^\dagger (I_2 \otimes U_n).$$

Then it follows by Theorem 36

$$\text{tr} C_{n+1} U_{n+1} A_{n+1} U_{n+1}^\dagger =$$

$$= \text{Re} \text{tr} (\sigma_{\frac{1}{2}} \otimes 0_1^{n+1}) (E) U_{n+1} (0_\frac{1}{2} \otimes \sigma_{\frac{1}{2}}^{n+1}) (E) U_{n+1}^\dagger$$

$$= 2 \text{Re} \text{tr} (E \otimes I_{2^n}) U_n (E \otimes I_{2^n} + I_2 \otimes N_n) U_n^\dagger$$

$$= 2 \text{Re} \text{tr} (I_2 \otimes E \otimes I_{2n-1}) (I_2 \otimes V_n) (E \otimes I_{2n} + I_2 \otimes N_n) (I_2 \otimes V_n)^\dagger$$

$$= 4 \text{Re} \text{tr} (E \otimes I_{2n-1}) V_n N_n V_n^\dagger$$

$$= 4 \text{Re} \text{tr} C_{n-1} V_n N_n V_n^\dagger = 2 \text{tr} C_{n} U_{n} A_n U_{n}^\dagger$$

\[\square\]

The values given in Theorem 38 are also critical.

**Proposition 39.** The unitary transformation $U_{n+1} \in U_{2n+2}(\mathbb{C})$, chosen as in Theorem 38, is a critical point of the transfer function $f_{n+1}$.

**Proof.** It holds

$$[\gamma_{n+1}^C (E), U_{n+1} \alpha_{n+1}^C (E) U_{n+1}^\dagger] =$$

$$= P_{n+1}^\dagger [P_{n+1} \gamma_{n+1}^C (E) P_{n+1}^\dagger, (I_2 \otimes U_n) \alpha_{n+1}^C (E)(I_2 \otimes U_n)^\dagger] P_{n+1}$$

$$= P_{n+1}^\dagger [I_2 \otimes \gamma_n^C (E), (I_2 \otimes U_n)(I_2 \otimes \alpha_n^C (E))(I_2 \otimes U_n)^\dagger] P_{n+1}$$

$$+ P_{n+1}^\dagger [I_2 \otimes \gamma_n^C (E), (I_2 \otimes U_n)(E \otimes I_{2^n})(I_2 \otimes U_n)^\dagger] P_{n+1}.$$  

A straightforward calculation now yields $[I_2 \otimes \gamma_n^C (E), (I_2 \otimes U_n)(E \otimes I_{2^n})(I_2 \otimes U_n)^\dagger] = 0$. Hence we obtain

$$[\gamma_{n+1}^C (E), U_{n+1} \alpha_{n+1}^C (E) U_{n+1}^\dagger] =$$

$$= P_{n+1}^\dagger (I_2 \otimes [\gamma_n^C (E), U_n \otimes \alpha_n^C (E) U_n^\dagger]) P_{n+1}, \quad (63)$$

and the result follows by Proposition 37. \[\square\]

30
Finally, this leads to the achievable values for the transfer function listed in Table 2. Based on extensive numerical simulations with gradient flows maximizing the $C$-numerical radius function we believe that these are indeed maximal.

Of course, it would also be interesting to know if one can also use representation theoretic methods to deduce formulas for all of the critical values of the $C$-numerical radius function. This seems to be a hard problem. For $n = 1$ one can show by brute force arguments that the critical values are exactly equal to $-2, -1, 0, 1, 2$. However, for $n > 1$ the situation becomes much more complicated and we have not been able to characterize the critical values, even in the apparently simple looking case $n = 2$.

Table 2: Achievable values for the transfer function.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$f_n^{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>$4(1 + \sqrt{3})$</td>
</tr>
<tr>
<td>4</td>
<td>$8(1 + \sqrt{3})$</td>
</tr>
<tr>
<td>5</td>
<td>$16(1 + \sqrt{3}) + 4\sqrt{5}$</td>
</tr>
<tr>
<td>6</td>
<td>$32(1 + \sqrt{3}) + 8\sqrt{5}$</td>
</tr>
</tbody>
</table>

References


