**MIT 6.7230 - Algebraic techniques and semidefinite optimization** May 3, 2024

# Lecture 19

*Lecturer: Pablo A. Parrilo* *Scribe: ???*

In this lecture we study techniques to exploit symmetries that may be present in semidefinite programming problems, particularly those arising from sum of squares decompositions [GP04]. For this, we present the basic elements of the representation theory of finite groups. There are many possible applications of these ideas in different fields; for the case of Markov chains, see [BDPX05]. The celebrated Delsarte linear programming upper bound for codes (and generalizations by Levenshtein, McEliece, etc., [DL98]) can be understood as a natural symmetry reduction of the SDP relaxations based on the Lovász theta function; see e.g. [Sch79].

## 1 Groups and their representations

The representation theory of finite groups is a classical topic; good descriptions are given in [FS92, Ser77]. We concentrate here on the case of finite groups; extensions to compact groups are relatively straightforward.

**Definition 1.** *A* group *consists of a set $G$ and a binary operation "$\cdot$" defined on $G$, for which the following conditions are satisfied:*

1. *Associative: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, for all $a, b, c \in G$.*
2. *Identity: There exist $1 \in G$ such that $a \cdot 1 = 1 \cdot a = a$, for all $a \in G$.*
3. *Inverse: Given $a \in G$, there exists $b \in G$ such that $a \cdot b = b \cdot a = 1$.*

We consider a finite group $G$, and an $n$-dimensional vector space $V$. We also consider the associated (infinite) group $GL(V)$ of nonsingular linear transformations of $V$, which we can interpret as the set of invertible $n \times n$ matrices. A *linear representation* of the group $G$ is a homomorphism $\rho : G \to GL(V)$. In other words, we have a mapping from the group into linear transformations of $V$, that respects the group structure, i.e.

$$\rho(st) = \rho(s)\rho(t) \quad \forall s, t \in G.$$

The *dimension* of the representation $\rho$ is the dimension of $V$.

**Example 2.** *Let $\rho(g) = 1$ for all $g \in G$. This is the* trivial *representation of the group. It is a one-dimensional representation.*

**Example 3.** *For a more interesting example, consider the symmetric group $S_n$, and the "natural" representation $\rho : S_n \to GL(\mathbb{C}^n)$, where $\rho(g)$ is a permutation matrix. For instance, for the group of permutations of two elements, $S_2 = \{e, g\}$, where $g^2 = e$, we have*

$$\rho(e) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \rho(g) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The representation given in Example 3 has an interesting property. The set of matrices $\{\rho(e), \rho(g)\}$ have common invariant subspaces (other than the trivial ones, namely $\{\mathbf{0}\}$ and $\mathbb{C}^2$). Indeed, we can easily verify that the (orthogonal) one-dimensional subspaces given by $(t, t)$ and $(t, -t)$ are invariant under the action of these matrices. Therefore, the restriction of $\rho$ to those subspaces also gives representations of the group $G$. In this case, the one corresponding to the subspace $(t, t)$ is "equivalent" (in a well-defined sense) to the trivial representation described in Example 2. The other subspace $(t, -t)$ gives the one-dimensional *alternating* representation of $S_2$, namely $\rho_A(e) = 1, \rho_A(g) = -1$. Thus, the representation $\rho$ decomposes as $\rho = \rho_T \oplus \rho_A$, a direct sum of the trivial and the alternating representations.

As we will see, the same ideas extend to arbitrary finite groups.

**Definition 4.** *An* irreducible representation *of a group is a linear representation with no nontrivial invariant subspaces.*

**Definition 5.** *Two representations $\rho_1 : G \to GL(V_1)$ and $\rho_2 : G \to GL(V_2)$ are* equivalent *if there exists a vector space isomorphism $T : V_1 \to V_2$ such that*

$$\rho_1(g) = T^{-1}\rho_2(g)T \qquad \forall g \in G.$$

**Theorem 6.** *Every finite group $G$ has a finite number of nonequivalent complex irreducible representations $\rho_i$, of dimension $d_i$. The relation $\sum_i d_i^2 = |G|$ holds.*

**Example 7.** *Consider the group $S_3$ (permutations in three elements). This group is generated by the two permutations $s : 123 \to 213$ and $c : 123 \to 312$ ("swap" and "cycle"), and has six elements $\{e, s, c, c^2, cs, sc\}$. Notice that $c^3 = e, s^2 = e$, and $s = csc$.*

*The group $S_3$ has three irreducible representations, two one-dimensional, and one two-dimensional (so $1^2 + 1^2 + 2^2 = |S_3| = 6$). These are:*

$$\begin{aligned} \rho_T(s) &= 1, & \rho_T(c) &= 1 \\ \rho_A(s) &= -1, & \rho_A(c) &= 1 \\ \rho_S(s) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, & \rho_S(c) &= \begin{bmatrix} \omega & 0 \\ 0 & \omega^2 \end{bmatrix} \end{aligned}$$

*where $\omega = e^{\frac{2\pi}{3}i}$ is a cube root of 1. Notice that it is enough to specify a representation on the generators of the group.*

**Example 8.** *Consider the cyclic group $C_n$ (cyclic permutations of $n$ elements). This group has a single generator $c$, which satisfies $c^n = e$. Explicitly, the group elements are $\{e, c, c^2, \ldots, c^{n-1}\}$.*

*The group $C_n$ has $n$ (complex) irreducible representations $\rho_0, \ldots, \rho_{n-1}$, given by $\rho_k(c) = \omega_k$, where $\omega_k = e^{k\frac{2\pi}{n}j}$ for $k = 0, \ldots, n-1$. All representations are one-dimensional, satisfying $1^2 + \cdots + 1^2 = |C_n| = n$.*

## 2 Symmetry and convexity

Optimization problems that invariant under the action of a symmetry group appear quite often in applications. Whenever these problems are convex, many simplifications are possible (in fact, even for nonconvex problems it is possible to exploit symmetry).

A key property of symmetric *convex* sets is the fact that the "group average" $\frac{1}{|G|}\sum_{g \in G} \sigma(g)x$ always belongs to the set.

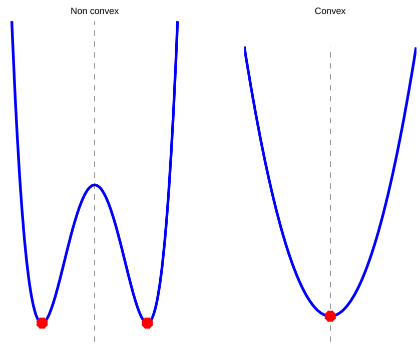


**Figure 1**: Two symmetric optimization problems, one non-convex and the other convex. For the latter, optimal solutions always lie on the fixed-point subspace.

Therefore, in convex optimization we can always restrict the solution to the fixed-point subspace

$$\mathcal{F} := \{x|\sigma(g)x = x, \quad \forall g \in G\}.$$

In other words, for convex problems, no "symmetry-breaking" is ever necessary.

As another interpretation, that will prove useful later, the "natural" decision variables of a symmetric optimization problem are the *orbits*, not the points themselves. Thus, we may look for solutions in the quotient space.

### 2.1 Invariant SDPs

We consider a general SDP, described in geometric form. If $\mathcal{L}$ is an affine subspace of $\mathcal{S}^n$, and $C, X \in \mathcal{S}^n$, an SDP is given by:

$$\min \langle C, X\rangle \qquad \text{s.t.} \quad X \in \mathcal{X} := \mathcal{L} \cap \mathcal{S}^n_+.$$

**Definition 9.** *Given a finite group $G$, and associated representation $\sigma : G \to GL(\mathcal{S}^n)$, a $\sigma$-*invariant SDP *is one where both the feasible set and the cost function are invariant under the group action, i.e.,*

$$\langle C, X\rangle = \langle C, \sigma(g)X\rangle, \qquad \forall g \in G, \qquad\qquad X \in \mathcal{X} \Rightarrow \sigma(g)X \in \mathcal{X} \qquad \forall g \in G$$

**Example 10.** *Consider the SDP given by*

$$\min a + c, \qquad s.t. \quad \begin{bmatrix} a & b \\ b & c \end{bmatrix} \succeq 0,$$

*which is invariant under the $Z_2$ action:*

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{bmatrix} \to \begin{bmatrix} X_{22} & -X_{12} \\ -X_{12} & X_{11} \end{bmatrix}.$$

Usually in SDP, the group acts on $\mathcal{S}^n$ through a congruence transformation, i.e., $\sigma(g)M = \rho(g)^T M \rho(g)$, where $\rho$ is a representation of $G$ on $\mathbb{C}^n$. In this case, the restriction to the fixed-point subspace takes the form:

$$\sigma(g)M = M \qquad \Longrightarrow \qquad \rho(g)M - M\rho(g) = 0, \quad \forall g \in G. \tag{1}$$

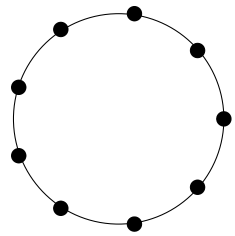

**Figure 2**: The cyclic graph $C_n$ in $n$ vertices (here, $n = 9$).

The Schur lemma of representation theory exactly characterizes the matrices that commute with a group action.

As a consequence of an important structural result (Schur's lemma), it turns out that every representation can be written in terms of a finite number of primitive blocks, the *irreducible representations* of a group.

**Theorem 11.** *Every group representation $\rho$ decomposes as a direct sum of irreducible representations:*

$$\rho = m_1\vartheta_1 \oplus m_2\vartheta_2 \oplus \cdots \oplus m_N\vartheta_N$$

*where $m_1, \ldots, m_N$ are the multiplicities.*

This decomposition induces an isotypic decomposition of the space

$$\mathbb{C}^n = V_1 \oplus \cdots \oplus V_N, \quad V_i = V_{i1} \oplus \cdots \oplus V_{in_i}.$$

In the symmetry-adapted basis, the matrices in the SDP have a block diagonal form:

$$(I_{m_1} \otimes M_1) \oplus \ldots \oplus (I_{m_N} \otimes M_N)$$

In terms of our symmetry-reduced SDPs, this means that not only the SDP block-diagonalizes, but there is also the possibility that many blocks are identical.

## 2.2 Example: symmetric graphs

Consider the MAXCUT problem on the cycle graph $C_n$ with $n$ vertices (see Figure 2). It is easy to see that the optimal cut has cost equal to $n$ or $n-1$, depending on whether $n$ is even or odd, respectively. What would the SDP relaxation yield in this case? If $A$ is the adjacency matrix of the graph, then the SDP relaxations have essentially the form

$$\begin{array}{rlcrl} \text{minimize} & \operatorname{Tr} AX & \qquad & \text{maximize} & \operatorname{Tr} \Lambda \\ \text{s.t.} & X_{ii} = 1 & & \text{s.t.} & A \succeq \Lambda \\ & X \succeq 0 & & & \Lambda \ \text{diagonal} \end{array} \tag{2}$$

By the symmetry of the graph, the matrix $A$ is *circulant*, i.e., $A_{ij} = a_{i-j \bmod n}$.

We focus now on the dual form. It should be clear that the cyclic symmetry of the graph induces a cyclic symmetry in the SDP, i.e., if $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a feasible solution, then $\tilde{\Lambda} = \operatorname{diag}(\lambda_n, \lambda_1, \lambda_2, \ldots, \lambda_{n-1})$ is also feasible and achieves the same objective value. Thus, by averaging over the cyclic group, we can always restrict $D$ to be a multiple of the identity matrix, i.e., $\Lambda = \lambda I$. Furthermore, the constraint $A \succeq \lambda I$ can be block-diagonalized via the Fourier matrix (i.e., the irreducible representations of the cyclic group), yielding:

$$A \succeq \lambda I \qquad \Leftrightarrow \qquad 2\cos\frac{k\pi}{n} \geq \lambda \qquad k = 0, \dots, n-1.$$

From this, the optimal solution of the relaxation can be directly computed, yielding the exact expressions for the upper bound on the size of the cut

$$mc(C_n) \leq SDP(C_n) = \begin{cases} n & n \text{ even} \\ n\cos^2\frac{\pi}{2n} & n \text{ odd.} \end{cases}$$

Although this example is extremely simple, exactly the same techniques can be applied to much more complicated problems; see for instance [PP04, dKMP$^+$06, Sch05, BV08] for some recent examples.

### 2.3 Example: even polynomials

Another (but illustrative) example of symmetry reduction is the case of SOS decompositions of even polynomials. Consider a polynomial $p(x)$ that is *even*, i.e., it satisfies $p(x) = p(-x)$. Does this symmetry help in making the computations more efficient?

### 2.4 Benefits

In the case of semidefinite programming, there are many benefits to exploiting symmetry:

- Replace one big SDP with smaller, coupled problems.
- Instead of checking if a big matrix is PSD, we use one copy of each repeated block (constraint aggregation).
- Eliminates multiple eigenvalues (numerical difficulties).
- For groups, the coordinate change depends only on the group, and not on the problem data.
- Can be used as a general preprocessing scheme. The coordinate change $T$ is unitary, so well-conditioned.

As we will see in the next section, this approach can be extended to more general algebras that do not necessarily arise from groups.

### 2.5 Sum of squares

In the case of SDPs arising from sum of squares decompositions, a parallel theory can be developed by considering the symmetry-induced decomposition of the full polynomial ring $\mathbb{R}[x]$. Since a complete explanation involves some elements of invariant theory, we omit the details here; see [GP04] for the full story.

**Example 12.** *Consider the (non-convex) quartic trivariate polynomial:*

$$p(x, y, z) = x^4 + y^4 + z^4 - 4xyz + x + y + z.$$

*This polynomial is invariant under all permutations of $\{x, y, z\}$ (the full symmetric group $S_3$). The global minimum of $p$ is $p_\star \approx -2.1129$, and is achieved at the orbit of global minimizers:*

$$(0.988, -1.102, -1.102)\,,\ (-1.102, 0.988, -1.102)\,,\ (-1.102, -1.102, 0.988).$$

*For this polynomial, it holds that $p_{sos} = p_\star$.*

*We show now how to compute $p_{sos}$ by exploiting symmetry. Since $p(x, y, z)$ has $n = 3$ variables, degree $2d = 4$, and a full Newton polytope, its standard sos formulation is indexed by all $\binom{n+d}{d} = \binom{5}{2} = 10$ monomials of degree 2, i.e.,*

$$p(x,y,z)-\gamma = \begin{bmatrix} 1 \\ x \\ y \\ z \\ x^2 \\ y^2 \\ z^2 \\ yz \\ xz \\ xy \end{bmatrix}^T \begin{bmatrix} q_{00} & q_{01} & q_{02} & q_{03} & q_{04} & q_{05} & q_{06} & q_{07} & q_{08} & q_{09} \\ q_{01} & q_{11} & q_{12} & q_{13} & q_{14} & q_{15} & q_{16} & q_{17} & q_{18} & q_{19} \\ q_{02} & q_{12} & q_{22} & q_{23} & q_{24} & q_{25} & q_{26} & q_{27} & q_{28} & q_{29} \\ q_{03} & q_{13} & q_{23} & q_{33} & q_{34} & q_{35} & q_{36} & q_{37} & q_{38} & q_{39} \\ q_{04} & q_{14} & q_{24} & q_{34} & q_{44} & q_{45} & q_{46} & q_{47} & q_{48} & q_{49} \\ q_{05} & q_{15} & q_{25} & q_{35} & q_{45} & q_{55} & q_{56} & q_{57} & q_{58} & q_{59} \\ q_{06} & q_{16} & q_{26} & q_{36} & q_{46} & q_{56} & q_{66} & q_{67} & q_{68} & q_{69} \\ q_{07} & q_{17} & q_{27} & q_{37} & q_{47} & q_{57} & q_{67} & q_{77} & q_{78} & q_{79} \\ q_{08} & q_{18} & q_{28} & q_{38} & q_{48} & q_{58} & q_{68} & q_{78} & q_{88} & q_{89} \\ q_{09} & q_{19} & q_{28} & q_{39} & q_{49} & q_{59} & q_{69} & q_{79} & q_{89} & q_{99} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ z \\ x^2 \\ y^2 \\ z^2 \\ yz \\ xz \\ xy \end{bmatrix},$$

*where the matrix $Q$ above will be constrained to be positive semidefinite. Recall that $p$ is invariant under all permutation of the variables (the full symmetric group $S_3$). Thus, we can constrain the matrix $Q$ to be in the fixed-point subspace, i.e., it should satisfy $Q = \rho(g)^T Q \rho(g)$, where $g \in G$ and $\rho : G \to GL(\mathbb{R}^{10})$ is the induced representation on the vector of monomials that arises from permuting the variables $(x, y, z)$. Solving the equations that define the fixed-point subspace, we find that the Gram matrix must have the structure*

$$\widehat{Q} = \begin{bmatrix} r_0 & r_1 & r_1 & r_1 & r_2 & r_2 & r_2 & r_3 & r_3 & r_3 \\ r_1 & r_4 & r_5 & r_5 & r_6 & r_7 & r_7 & r_8 & r_9 & r_9 \\ r_1 & r_5 & r_4 & r_5 & r_7 & r_6 & r_7 & r_9 & r_8 & r_9 \\ r_1 & r_5 & r_5 & r_4 & r_7 & r_7 & r_6 & r_9 & r_9 & r_8 \\ r_2 & r_6 & r_7 & r_7 & r_{10} & r_{11} & r_{11} & r_{12} & r_{13} & r_{13} \\ r_2 & r_7 & r_6 & r_7 & r_{11} & r_{10} & r_{11} & r_{13} & r_{12} & r_{13} \\ r_2 & r_7 & r_7 & r_6 & r_{11} & r_{11} & r_{10} & r_{13} & r_{13} & r_{12} \\ r_3 & r_8 & r_9 & r_9 & r_{12} & r_{13} & r_{13} & r_{14} & r_{15} & r_{15} \\ r_3 & r_9 & r_8 & r_9 & r_{13} & r_{12} & r_{13} & r_{15} & r_{14} & r_{15} \\ r_3 & r_9 & r_9 & r_8 & r_{13} & r_{13} & r_{12} & r_{15} & r_{15} & r_{14} \end{bmatrix}. \tag{3}$$

*Notice that the fixed-point subspace is 16-dimensional, as opposed to the $\binom{11}{2} = 55$ degrees of freedom in the original matrix.*

*We can now, however, give a nicer description of this subspace. Consider the coordinate transformation (a symmetry-adapted basis) of the form $X \mapsto T^T X T$, where the orthogonal matrix $T$ is given by*

$$T = \mathit{BlockDiag}(1, R, R, R) \cdot \Pi, \qquad R = \begin{bmatrix} \alpha & \alpha & \alpha \\ \alpha & \beta & \gamma \\ \alpha & \gamma & \beta \end{bmatrix},$$

*where $\alpha = 1/\sqrt{3}$, $\beta = (3-\sqrt{3})/6$, $\gamma = -(3+\sqrt{3})/6$, and $\Pi$ is the permutation matrix satisfying $\Pi^T[x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9] = [x_0, x_1, x_4, x_7, x_2, x_5, x_8, x_3, x_6, x_9]$. It can be verified that* *under this tranformation, the matrix in* (3) *takes now the form*

$$T^T \widehat{Q} T = BlockDiag(Q_1, Q_2, Q_2),$$

*where*

$$Q_1 = \begin{bmatrix} r_0 & \sqrt{3}r_1 & \sqrt{3}r_2 & \sqrt{3}r_3 \\ \sqrt{3}r_1 & r_4 + 2r_5 & r_6 + 2r_7 & r_8 + 2r_9 \\ \sqrt{3}r_2 & r_6 + 2r_7 & r_{10} + 2r_{11} & r_{12} + 2r_{13} \\ \sqrt{3}r_3 & r_8 + 2r_9 & r_{12} + 2r_{13} & r_{14} + 2r_{15} \end{bmatrix},$$

$$Q_2 = \begin{bmatrix} r_4 - r_5 & r_6 - r_7 & r_8 - r_9 \\ r_6 - r_7 & r_{10} - r_{11} & r_{12} - r_{13} \\ r_8 - r_9 & r_{12} - r_{13} & r_{14} - r_{15} \end{bmatrix}.$$

*Notice that the* $10 \times 10$ *matrix has split in three blocks, one of size* $4 \times 4$*, and two identical blocks of size* $3 \times 3$*. Also, all entries are otherwise linearly independent (in fact, we have the dimension count* $\binom{5}{2} + \binom{4}{2} = 10 + 6 = 16$*, the number of free parameters in* (3)*).*

*Since* $\widehat{Q} \succeq 0$ *if and only if* $T^T \widehat{Q} T \succeq 0$*, this implies that instead of solving a semidefinite programming problem with a positivity constraint on a* $10 \times 10$ *matrix, we have now a* $4 \times 4$ *and a* $3 \times 3$ *matrix instead (clearly, we only need one copy of the two identical* $3 \times 3$ *blocks), which is a lot simpler.*

## 3 Algebra decomposition

An alternative (and somewhat more general) approach can be obtained by focusing instead on the *associative algebra* generated by the matrices in a semidefinite program.

**Definition 13.** *An* associative algebra $\mathcal{A}$ *over* $\mathbb{C}$ *is a vector space with a* $\mathbb{C}$*-bilinear operation* $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ *that satisfies*

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z, \qquad \forall x, y, z \in \mathcal{A}.$$

In general, associative algebras do not need to be commutative (i.e., $x \cdot y = y \cdot x$). However, that is an important special case, with many interesting properties. Important examples of finite dimensional associative algebras are:

- Full matrix algebra $\mathbb{C}^{n \times n}$, standard product.
- The subalgebra of square matrices with equal row and column sums.
- The $n$-dimensional algebra generated by a single $n \times n$ matrix.
- The group algebra: formal $\mathbb{C}$-linear combination of group elements.
- Polynomial multiplication modulo a zero dimensional ideal.
- The Bose-Mesner algebra of an association scheme.

We have already encountered some of these, when studying the companion matrix and its generalizations to the multivariate case. A particularly interesting class of algebras (for a variety of reasons) are the *semisimple* algebras.

**Definition 14.** *The* radical *of an associative algebra* $\mathcal{A}$*, denoted rad*$(\mathcal{A})$*, is the intersection of all maximal left ideals of* $\mathcal{A}$*.*

**Definition 15.** *An associative algebra $\mathcal{A}$ is* semisimple *if $Rad(\mathcal{A}) = 0$.*

For a semidefinite programming problem in standard (dual) form

$$\max b^T y \qquad \text{s.t.} \quad A_0 - \sum_{i=1}^{m} A_i y_i \succeq 0,$$

we consider the algebra generated by the $A_i$.

**Theorem 16.** *Let $\{A_0, \ldots, A_m\}$ be given* symmetric *matrices, and $\mathcal{A}$ the generated associative algebra. Then, $\mathcal{A}$ is a semisimple algebra.*

Semisimple algebras have a very nice structure, since they are essentially the direct sum of much simpler algebras.

**Theorem 17** (Wedderburn)**.** *Every finite dimensional semisimple associative algebra over $\mathbb{C}$ can be decomposed as a direct sum*

$$\mathcal{A} = \mathcal{A}_1 \oplus \mathcal{A}_2 \oplus \ldots \oplus \mathcal{A}_k.$$

*Each $\mathcal{A}_i$ is isomorphic to a simple full matrix algebra.*

**Example 18.** *A well-known example is the (commutative) algebra of circulant matrices, i.e., those of the form*

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_4 & a_1 & a_2 & a_3 \\ a_3 & a_4 & a_1 & a_2 \\ a_2 & a_3 & a_4 & a_1 \end{bmatrix}.$$

*Circulant matrices are ubiquitous in many applications, such as signal processing. It is well-known that there exists a* fixed *unitary coordinate change (the $n \times n$ discrete Fourier transform matrix with entries $F_{jk} = \frac{1}{\sqrt{n}}\omega^{jk}$ where $\omega$ is an $n$-root of unity) under which* all *matrices $A$ are diagonal (with distinct scalar blocks). For instance, for the example above with $n = 4$, we have*

$$F = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}, \qquad F^* A F = \operatorname{diag}\left( \begin{bmatrix} a_1 + a_2 + a_3 + a_4 \\ a_1 - i a_2 - a_3 + i a_4 \\ a_1 - a_2 + a_3 - a_4 \\ a_1 + i a_2 - a_3 - i a_4 \end{bmatrix} \right).$$

**Remark 19.** *In general, any associative algebra is the direct sum of its radical and a semisimple algebra. For the $n$-dimensional algebra generated by a single matrix $A \in \mathbb{C}^{n \times n}$, we have that $A = S + N$, where $S$ is diagonalizable, $N$ is nilpotent, and $SN = NS$. Thus, this statement is essentially equivalent to the existence of the Jordan decomposition.*

## References


[BDPX05] S. Boyd, P. Diaconis, P. A. Parrilo, and L. Xiao. Symmetry analysis of reversible Markov chains. *Internet Math.*, 2(1):31–71, 2005.

[BV08] C. Bachoc and F. Vallentin. New upper bounds for kissing numbers from semidefinite programming. *J. Amer. Math. Soc*, 21(3):909–924, 2008.

[dKMP+06] E. de Klerk, J. Maharry, D.V. Pasechnik, R.B. Richter, and G. Salazar. Improved bounds for the crossing numbers of $K_{m,n}$ and $K_n$. *SIAM Journal on Discrete Mathematics*, 20:189, 2006.

[DL98] P. Delsarte and VI Levenshtein. Association schemes and coding theory. *IEEE Transactions on Information Theory*, 44(6):2477–2504, 1998.

[FS92] A. Fässler and E. Stiefel. *Group Theoretical Methods and Their Applications.* Birkhäuser, 1992.

[GP04] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95–128, 2004.

[PP04] P. A. Parrilo and R. Peretz. An inequality for circle packings proved by semidefinite programming. *Discrete and Computational Geometry*, 31(3):357–367, 2004.

[Sch79] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory*, 25(4), 1979.

[Sch05] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Transactions on Information Theory*, 51(8):2859–2866, 2005.

[Ser77] J.-P. Serre. *Linear Representations of Finite Groups.* Springer-Verlag, 1977.