

Commutators of skew-symmetric matrices

Anthony M. Bloch* Arieh Iserles[†]

March 25, 2004

Abstract

In this paper we develop a theory for analysing the size of a Lie bracket or commutator in a matrix Lie algebra. Complete details are given for the Lie algebra $\mathfrak{so}(n)$ of skew symmetric matrices.

1 Norms and commutators in $M_n[\mathbb{R}]$ and $\mathfrak{so}(n)$

This paper is concerned with the following question. Let \mathfrak{g} be a *Lie algebra* (Carter, Segal & Macdonald 1995, Humphreys 1978, Varadarajan 1984). Given $X, Y \in \mathfrak{g}$ and a norm $\|\cdot\| : \mathfrak{g} \rightarrow \mathbb{R}_+$, what is the size of $\|[X, Y]\|$ in comparison with $\|X\| \cdot \|Y\|$?

On the face of this, this question has little merit since the elementary inequality

$$\|[X, Y]\| \leq 2\|X\| \cdot \|Y\| \tag{1.1}$$

always holds for $X, Y \in M_n[\mathbb{R}]$, the set of all $n \times n$ real matrices and an arbitrary norm $\|\cdot\|$. Moreover, it is easy to prove that the bound (1.1) can be attained for most norms of practical interest. In particular, this is the case for two types of norms closely associated with a remarkable paper of von Neumann (1937).

We recall that a *symmetric gauge* is a vector norm $|\cdot|$ which is both symmetric and positive. In other words, for every $\mathbf{x} \in \mathbb{R}^n$ it is true that $|\mathbf{x}_\pi| = |\mathbf{x}|$ and $|\|\mathbf{x}\|| = |\mathbf{x}|$, where π is a permutation of $\{1, 2, \dots, n\}$, $\mathbf{x}_\pi^\top = [x_{\pi_1}, x_{\pi_2}, \dots, x_{\pi_n}]$ and $|\mathbf{x}|^\top = [|x_1|, |x_2|, \dots, |x_n|]$. We consider two norms, firstly the *operator norm*

$$|X| = \max_{\mathbf{v} \neq \mathbf{0}} \frac{|X\mathbf{v}|}{|\mathbf{v}|}$$

and secondly the norm

$$\|X\| = |\sigma(X)|, \tag{1.2}$$

where $\sigma(X)$ are the *singular values* of X , arbitrarily ordered. While it is easy to see that (1.2) is a unitary norm (i.e., invariant under multiplication by a unitary matrix), von Neumann proved that *all* unitary norms are of this form.

*Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA, email abloch@umich.edu, supported in part by the National Science Foundation.

[†]Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, England, email A.Iserles@damtp.cam.ac.uk.

We consider just the case $n = 2$, since it can be imbedded in $M_n[\mathbb{R}]$ for any $n \geq 2$. Let

$$X = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad Z = [X, Y] = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}.$$

It is easy to verify that $|X| = |Y| = 1$ and $|Z| = 2$. Moreover, since $\sigma(X) = \sigma(Y) = [1, 1]$ and $\sigma(Z) = [2, 2]$, it is also true that $\|X\|, \|Y\| = |\mathbf{1}|$ and $\|Z\| = 2|\mathbf{1}|$, where $\mathbf{1}^\top = [1, 1]$. In both cases the upper bound in (1.1) is attained.

Yet, there is a basic difference between $M_n[\mathbb{R}]$ and a Lie algebra $\mathfrak{g} \subset M_n[\mathbb{R}]$: while $\dim M_n[\mathbb{R}] = n^2$, the Lie algebra typically has lower dimension. Thus, it makes sense to pose the question whether, *once X and Y are restricted to \mathfrak{g}* , the inequality (1.1) might still be obeyed as an equality or whether 2 might be replaced by a smaller constant for all $X, Y \in \mathfrak{g}$. Thus, given a norm $\|\cdot\|$, we say that the *radius* of a Lie algebra \mathfrak{g} is the least number $\omega(\mathfrak{g}) \in [0, 2]$ such that

$$\|[X, Y]\| \leq \omega(\mathfrak{g})\|X\| \cdot \|Y\|, \quad X, Y \in \mathfrak{g}.$$

In other words,

$$\omega(\mathfrak{g}) = \max \left\{ \frac{\|[X, Y]\|}{\|X\| \cdot \|Y\|} : X, Y \in \mathfrak{g}, X, Y \neq O \right\}, \quad (1.3)$$

the operator norm of the commutator.

It is important when defining ω to keep in mind which underlying norm we are using. In the following we shall denote by $\|v\|_p$, $v \in \mathbb{R}^n$ the vector p -norm and by

$$\|X\|_p = \max \left\{ \frac{\|Xv\|_p}{\|v\|_p} ; v \neq 0 \right\}$$

the corresponding operator norm as above. In the case $p = 2$ we shall call this the Euclidean norm.

We denote by

$$\|X\|_F = \left(\sum_{k,l=1}^n X_{kl}^2 \right)^{\frac{1}{2}}$$

the Frobenius norm.

We recall also the important facts to be used below, namely that $\|X\|_2$ is equal to the magnitude of the largest singular value of X while $\|X\|_F = \|\sigma(X)\|_2$.

When the context is not clear we will label the ω by a subscript denoting which norm is being used.

Trivially, the Lie algebra \mathfrak{g} is commutative if and only if $\omega(\mathfrak{g}) = 0$, but this observation is devoid of any insight. More interestingly, consider $\mathfrak{so}(3)$ and the Euclidean norm. Letting

$$X = \begin{bmatrix} 0 & x_1 & x_2 \\ -x_1 & 0 & x_3 \\ -x_2 & -x_3 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & y_1 & y_2 \\ -y_1 & 0 & y_3 \\ -y_2 & -y_3 & 0 \end{bmatrix}$$

and observing that in $\mathfrak{so}(n)$ the Euclidean norm coincides with the spectral radius, we commence by noting that

$$\|X\| = \|\mathbf{x}\|, \quad \|Y\| = \|\mathbf{y}\|.$$

Moreover, if $Z = [X, Y]$ then, by an easy direct calculation,

$$\|z\|^2 = \|x\| \cdot \|y\| - (\mathbf{x}^\top \mathbf{y})^2. \quad (1.4)$$

Therefore

$$\|[X, Y]\| = [\|X\|^2 \|Y\|^2 - (\mathbf{x}^\top \mathbf{y})^2]^{\frac{1}{2}} \leq \|X\| \cdot \|Y\|,$$

with the upper bound holding as an equality when x is orthogonal to y . We thus deduce that $\omega(\mathfrak{so}(3)) = 1$.

Remark 1: There is a natural Lie algebra homomorphism between $\mathfrak{so}(3)$ and \mathbb{R}^3 endowed with the cross product. The above computation may be repeated with this in mind and Req:1.4 is a standard vector identity.

Remark 2: It is also of interest to repeat the above computation for the Frobenius norm. One determines immediately that $\omega_F(\mathfrak{so}(3)) = 1/\sqrt{2}$. However for Lie algebraic reasons that will become apparent below it is more natural to scale the Frobenius norm by a factor of $\sqrt{2}$. With this scaling we also have $\omega_F(\mathfrak{so}(3)) = 1$. Strikingly this result does not hold for n larger than 3.

In this paper we determine $\omega(\mathfrak{so}(n))$ for all $n \geq 3$ ($\mathfrak{so}(2)$ is a commutative algebra, hence $\omega(\mathfrak{so}(2)) = 0$) with respect to the (scaled) Frobenius norm. Specifically, we prove that $\omega_F(\mathfrak{so}(n)) = \sqrt{2}$ for $n \geq 4$ (with the above-mentioned scaling). Note that $\|X\|_F = -\langle X, X \rangle$, where $\langle \cdot, \cdot \rangle$ is a multiple of the *Killing form* in $\mathfrak{so}(n)$, hence it has deeper Lie-algebraic significance. The Killing form evaluated on a pair of $n \times n$ skew-symmetric matrices A, B is actually $(n-2)\text{trace } AB$ (see (Kobayashi & Nomizu 1969)). (Of course for a non-compact Lie algebra the Killing form does not provide a norm since it is not definite.) Another reason why the Frobenius norm is of interest is that the radius of $\mathfrak{so}(n)$, $n \geq 4$, is absolutely trivial for most other norms of interest.

Thus, again, let $|\cdot|$ be a symmetric gauge and

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

therefore

$$[X, Y] = \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}.$$

Note that for every $\mathbf{v} \in \mathbb{R}^4$ positivity and symmetry of the symmetric gauge imply that

$$|X\mathbf{v}| = \left| \begin{bmatrix} v_2 \\ -v_1 \\ v_4 \\ -v_3 \end{bmatrix} \right| = |\mathbf{v}|$$

and, similarly, $|Y\mathbf{v}| = |\mathbf{v}|$ and $|[X, Y]\mathbf{v}| = 2|\mathbf{v}|$. Therefore, in the underlying operator norm $|X|, |Y| = 1$ and $|[X, Y]| = 2$. Consequently $\omega(\mathfrak{so}(4)) = 2$ and this can be extended to all Lie algebras $\mathfrak{so}(n)$, $n \geq 4$, since they form a flag. This example cannot, however, be

extended to unitary norms (1.2), unless $\|\mathbf{1}\| = 1$. Note that the latter condition holds when $\|\cdot\| = \|\cdot\|_\infty$ (the ∞ -Schatten norm (Horn & Johnson 1994)), but in that instance $\|\cdot\| = \|\cdot\|_2$ and we are back to the area covered earlier in this paragraph. On the other hand, $\|\cdot\| = \|\cdot\|_2$, whence $\|\mathbf{1}\| = 2$, results in $\|\cdot\| = \|\cdot\|_F$. This, of course, does not necessarily mean that $\omega(\mathfrak{so}(4)) < 2$ in the Frobenius norm.

In Section 2 we discuss the structure of the commutator operator, considered as a linear transformation from $\mathfrak{so}(n)$ to itself. We prove that, subject to an appropriate representation of $\mathfrak{so}(n)$, the commutator matrix in the $\frac{1}{2}(n-1)n$ linear space can be read explicitly from a certain directed graph and investigate its eigenstructure. Section 3 is devoted to the proof of the main result of this paper, namely that, once we use the Frobenius norm, $\omega(\mathfrak{so}(n)) = \sqrt{2}$ for all $n \geq 4$.

The subject matter of this paper is motivated by a raft of issues arising from *geometric numerical integration*. The simplest such problem is the convergence of the sum

$$f(t; X, Y) = \sum_{m=0}^{\infty} a_m t^m \operatorname{ad}_X^m Y, \quad X, Y \in \mathfrak{g},$$

where $\{a_m\}_{m \in \mathbb{Z}_+}$ is a given sequence and ad_X is the *adjoint operator* of the Lie algebra \mathfrak{g} ,

$$\operatorname{ad}_X^0 Y = Y, \quad \operatorname{ad}_X^m Y = [X, \operatorname{ad}_X^{m-1} Y], \quad m \in \mathbb{N}.$$

It is trivial to deduce from the triangle inequality that

$$\|f(t; X, Y)\| \leq \|Y\| \sum_{m=0}^{\infty} a_m [t\omega(\mathfrak{g})\|X\|]^m,$$

thereby relating the convergence of F to the domain of analyticity of the generating function $\sum_{m=0}^{\infty} a_m z^m$. Similar and more complicated problems abound in the analysis of Lie-group methods (Iserles, Munthe-Kaas, Nørsett & Zanna 2000). The benefit of smaller $\omega(\mathfrak{g})$ is clear.

The norm of a bracket is also important in determining the maximum allowable step size in certain minimization problems on adjoint orbits, see the work of Brockett (1993).

A related problem of interest in analysing certain systems of differential equations is that of finding a bound on the norm of the bracket $[X, N]$ where N is fixed and X varies over the adjoint orbit of a group. This problem is discussed in (Brockett 1994). In that setting for $\mathfrak{so}(n)$ one has to solve the problem of maximizing $\|[X, N]\|$ over all $X = \theta^T \Lambda \theta$ for N, Λ fixed in $\mathfrak{so}(n)$ and θ in the group $\operatorname{SO}(n)$.

2 The reduced commutator matrix in $\mathfrak{so}(n)$

2.1 The reduced commutator matrix

Let $\mathfrak{g} \subseteq M_n[\mathbb{R}]$ be an m -dimensional matrix Lie algebra, $1 \leq m \leq n^2$. An obvious means to explore the norm of the commutator in \mathfrak{g} is by means of the *natural embedding* $\theta : \mathfrak{g} \rightarrow \mathbb{R}^{n^2}$ that ‘stretches’ a matrix X into a vector, e.g. by letting $\theta_{(l-1)n+k}(X) = X_{k,l}$, $k, l = 1, 2, \dots, n$ (columnwise ordering). Since commutation is a linear transformation, it follows that for every $X \in \mathfrak{g}$ there exists a matrix $\tilde{C}_X \in M_{n^2}[\mathbb{R}]$ such that

$$\theta([X, Y]) = \tilde{C}_X \theta(Y), \quad Y \in \mathfrak{g}.$$

It is known that

$$\sigma(\tilde{C}_X) = \{\lambda_k - \lambda_l : \lambda_k, \lambda_l \in \sigma(X), k, l = 1, 2, \dots, n\}$$

(Hille 1969), and this provides a useful tool to explore commutators in a classical setting. Yet, this line of reasoning disregards the fact that \mathfrak{g} is a Lie algebra, typically of much smaller dimension than that of $M_n[\mathbb{R}]$. Thus, in place of θ , we propose a *restricted embedding* $\nu : \mathfrak{g} \rightarrow \mathbb{R}^m$. Let $\mathcal{Q} = \{Q_1, Q_2, \dots, Q_m\}$ be a basis of \mathfrak{g} . We define an isomorphism ν from \mathfrak{g} on \mathbb{R}^m through

$$\mathfrak{g} \ni X = \sum_{k=1}^m x_k Q_k \quad \Leftrightarrow \quad \nu(X) = \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}.$$

Remark: We note that this is just a vector space isomorphism and not in general a Lie algebra homomorphism and there is in general no natural cross-product operation in \mathbb{R}^m .

The *restricted commutator matrix* $C_X \in M_m[\mathbb{R}]$ is then defined by the identity

$$\nu([X, Y]) = C_X \nu(Y), \quad Y \in \mathfrak{g}.$$

Spectral information on C_X is no longer readily and explicitly available, yet the procedure has the great virtue of reducing the dimension and allowing for a more natural incorporation of Lie-algebraic information. Specifically, let $\{c_{k,l}^j\}_{k,l,j=1,2,\dots,m}$ be the *structure constants* of \mathfrak{g} with respect to \mathcal{Q} ,

$$[Q_k, Q_l] = \sum_{j=1}^m c_{k,l}^j Q_j.$$

It is an elementary exercise that

$$(C_X)_{j,l} = \sum_{k=1}^m x_k c_{k,l}^j, \quad j, l = 1, 2, \dots, m. \quad (2.1)$$

Defining the *radius* $\omega(\mathfrak{g})$ of the Lie algebra as

$$\max \left\{ \frac{\|[X, Y]\|}{\|X\| \cdot \|Y\|} : X, Y \in \mathfrak{g}, X, Y \neq O \right\},$$

where $\|\cdot\|$ is a given norm induced by the vector norm on \mathbb{R}^m , i.e. $\|X\| = \|\nu(X)\|$, we observe that

Theorem 1

$$\omega(\mathfrak{g}) = \max \left\{ \frac{\|C_X\|}{\|X\|} : X \in \mathfrak{g}, X \neq O \right\}, \quad (2.2)$$

where the Lie algebra norm is that induced by the map ν .

Proof We have

$$\|[X, Y]\| = \|\nu([X, Y])\| = \|C_X \nu(Y)\|.$$

Hence

$$\begin{aligned} \omega(\mathfrak{g}) &= \max_{X \in \mathfrak{g} \setminus \{0\}} \max_{Y \in \mathfrak{g} \setminus \{0\}} \frac{\|C_X \nu(Y)\|}{\|\nu(X)\| \cdot \|\nu(Y)\|} \\ &= \max_{X \in \mathfrak{g} \setminus \{0\}} \frac{\|C_X\|}{\|\nu(X)\|}. \end{aligned}$$

□

We conclude this section by addressing the question of multiple representations. Suppose thus that we have two bases of \mathfrak{g} , $\mathcal{Q} = \{Q_1, Q_2, \dots, Q_m\}$ and $\mathcal{P} = \{P_1, P_2, \dots, P_m\}$, say. Set

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_m], \quad Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_m],$$

where $\mathbf{p}_k = \nu(P_k)$, $\mathbf{q}_k = \nu(Q_k)$, $k = 1, 2, \dots, m$. Then

$$\mathfrak{g} \ni X = \sum_{k=1}^m x_k P_k = \sum_{k=1}^m \tilde{x}_k Q_k$$

implies at once that $\mathbf{x} = Q \tilde{\mathbf{x}}$. Therefore $C_X^Q = Q^{-1} C_X^P Q$, where C_X^P and C_X^Q are reduced commutators with respect to the two bases. In particular, if Q is an orthogonal matrix and the bases are orthogonally similar then the radius of \mathfrak{g} does not depend on the choice of the basis.

Example: We compute $\omega(\mathfrak{so}(3))$ using this formalism.

If

$$X = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$$

then we easily compute

$$C_X = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}.$$

Now, $\nu(X) = [x, y, z]^T$, and hence $\|\nu(X)\|_2 = (a^2 + b^2 + c^2)^{1/2}$. On the other hand, $\sigma(C_X) = \{0, \pm(a^2 + b^2 + c^2)^{1/2}\}$, hence $\|C_X\|_2 = (a^2 + b^2 + c^2)^{1/2}$ and we obtain $\omega(\mathfrak{so}(3)) = 1$.

2.2 The reduced commutator matrix in $\mathfrak{so}(n)$ and directed graphs

We denote by $E_{k,l} \in M_n[\mathbb{R}]$ the matrix whose (k, l) th component is +1 and otherwise is zero, $k, l = 1, 2, \dots, n$, and choose the basis

$$\mathcal{Q} = \{Q_{k,l} = E_{k,l} - E_{l,k} : 1 \leq k < l \leq n\}$$

of $\mathfrak{so}(n)$. The restricted embedding ν takes each $Q_{k,l}$ to $\mathbf{e}_{\mu(k,l)} \in \mathbb{R}^m$, where $m = \frac{1}{2}(n-1)n$ and μ is an arbitrary isomorphism mapping pairs $\mathcal{I} = \{(k, l) : 1 \leq k < l \leq n\}$, into $\{1, 2, \dots, m\}$. However, it is more convenient to discuss restricted commutator matrices in

the formalism of the $Q_{k,l}$, bypassing ν altogether. Thus, we index the structure constants and the entries of the restricted commutator matrix by pairs $(i, j) \in \mathcal{I}$.

For ease of notation we let $Q_{k,l} = -Q_{l,k}$ for $k > l$ and $Q_{k,k} = 0$. Since

$$[Q_{k,l}, Q_{r,s}] = \delta_{l,r}Q_{k,s} - \delta_{k,r}Q_{l,s} - \delta_{l,s}Q_{k,r} + \delta_{k,s}Q_{l,r},$$

the structure constants are

$$c_{(k,l),(r,s)}^{(i,j)} = \begin{cases} -1, & k = r, l \neq s, i = l, j = s, \\ +1, & k \neq r, l = r, i = k, j = s, \\ -1, & k \neq r, l = s, i = k, j = r, \\ +1, & k = s, l \neq r, i = l, j = r, \\ 0, & \text{otherwise,} \end{cases} \quad (i, j) \in \mathcal{I}. \quad (2.3)$$

In other words, most structure constants vanish: not surprising, given that our basis is consistent with the root space decomposition of $\mathfrak{so}(n)$, hence likely to lend itself to a sparse set of structure constants. Specifically, the nonzero structure constants are precisely

$$\begin{aligned} c_{(k,l),(k,s)}^{(l,s)} &= -1, & l < s, & & c_{(k,l),(k,s)}^{(l,s)} &= +1, & l > s, \\ c_{(k,l),(l,s)}^{(k,s)} &= +1, & k < s, & & c_{(k,l),(l,s)}^{(k,s)} &= -1, & k > s, \\ c_{(k,l),(r,l)}^{(k,r)} &= -1, & k < r, & & c_{(k,l),(r,l)}^{(k,r)} &= +1, & k > r, \\ c_{(k,l),(r,k)}^{(l,r)} &= +1, & l < r, & & c_{(k,l),(r,k)}^{(l,r)} &= -1, & l > r, \end{aligned}$$

Given

$$\mathfrak{so}(n) \ni X = \sum_{k=1}^{n-1} \sum_{l=k+1}^n x_{k,l} Q_{k,l},$$

(2.1) implies that

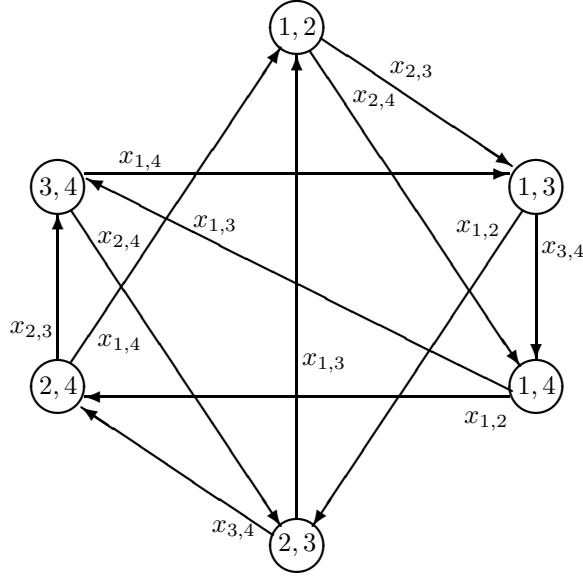
$$(C_X)_{(k,l),(i,j)} = \sum_{(r,s) \in \mathcal{I}} x_{r,s} c_{(r,s),(k,l)}^{(i,j)} = - \sum_{(r,s) \in \mathcal{I}} c_{(k,l),(i,j)}^{(r,s)}, \quad (k,l), (i,j) \in \mathcal{I}.$$

We observe that $C_X \in \mathfrak{so}(m)$ and that it is a very sparse matrix. specifically, for any $(k,l) \in \mathcal{I}$ the only nonzero entries are

$$\begin{aligned} (C_X)_{(k,l),(l,j)} &= x_{k,j}, & j &= l+1, l+2, \dots, n, \\ (C_X)_{(k,l),(k,j)} &= x_{j,l}, & j &= k+1, k+2, \dots, n, j \neq l, \\ (C_X)_{(k,l),(i,l)} &= x_{i,k}, & i &= 1, 2, \dots, l-1, i \neq k, \\ (C_X)_{(k,l),(i,k)} &= x_{l,i}, & i &= 1, 2, \dots, k-1. \end{aligned} \quad (2.4)$$

Altogether, just $(n-2)(n-1)$, out of $\frac{1}{2}(n-1)^2 n^2$, entries of C_X are nonzero.

The elements of C_X lend themselves to a very convenient representation in terms of *labelled digraphs*. Any matrix $A \in M_m[\mathbb{R}]$ can be represented by a digraph with m vertices, adopting the convention that, once $A_{k,l} \neq 0$, then there is a directed edge from vertex k to vertex l with the label $A_{k,l}$. As an example, let us examine the digraph corresponding to C_X for $n = 4$ (hence $m = 6$):

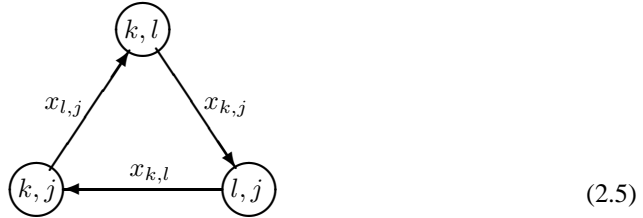


We commence by noting that the graph is 4-regular (Chartrand & Lesniak 1986): each vertex is of degree 4. Moreover, two of the edges at each vertex commence and two terminate there.

A generalization for all $n \geq 3$ is clear from (2.4). For every $1 \leq k < l < j \leq n$ we have

$$(C_X)_{(k,l),(l,j)} = x_{k,j}, \quad (C_X)_{(l,j),(k,j)} = x_{k,l}, \quad (C_X)_{(k,j),(k,l)} = x_{l,j}.$$

In the notation of labelled digraphs this corresponds to the 3-cycle



Needless to say, (2.5) can be read ‘backwards’: an arrow from (l, j) to (k, j) with a label $x_{k,l}$ is the same as an arrow from (k, j) to (l, j) with the label $-x_{k,l}$.

Lemma 2 For all $n \geq 3$ the directed graph of C_X is the sum of all $\binom{n}{3}$ 3-cycles (2.5) for all $1 \leq k < l < j \leq n$. It is r -regular, where $r = 2(n - 2)$.

Proof The first statement of the lemma follows at once from our analysis. Because of symmetry, clearly the graph must be r -regular for some $r \geq 1$. Therefore, the sum of all the degrees of all the vertices is mr . Since each 3-cycle (2.5) accounts for exactly six degrees and $m = \frac{1}{2}(n - 1)n$, we have

$$r = \frac{6\binom{n}{3}}{\frac{1}{2}(n - 1)n} = 2(n - 2).$$

□

Given $X, Y \in \mathfrak{so}(n)$, $n \geq 3$, we can reconstruct the representation of $[X, Y]$ in the basis \mathcal{Q} directly from the digraph in Lemma 2. Since

$$[X, Y] = \sum_{(k,l) \in \mathcal{I}} \sum_{(r,s) \in \mathcal{I}} x_{k,l} y_{r,s} (\delta_{l,r} Q_{k,s} - \delta_{k,r} Q_{l,s} - \delta_{l,s} Q_{k,r} + \delta_{k,s} Q_{l,r}),$$

we have the following contributions to the $(i, j) \in \mathcal{I}$ component of the commutator,

$$\begin{aligned} (k, s) = (i, j), \quad r = l &\Rightarrow +x_{i,l} y_{l,j}, \quad i+1 \leq l \leq j-1, \\ (l, s) = (i, j), \quad r = k &\Rightarrow -x_{k,i} y_{k,j}, \quad 1 \leq k \leq \min\{i, j\} - 1, \\ (k, r) = (i, j), \quad s = l &\Rightarrow -x_{i,l} y_{j,l}, \quad \max\{i, j\} + 1 \leq l \leq n, \\ (l, r) = (i, j), \quad s = k &\Rightarrow +x_{k,i} y_{j,k}, \quad j+1 \leq k \leq i-1, \\ (k, s) = (j, i), \quad r = l &\Rightarrow -x_{j,l} y_{l,i}, \quad j+1 \leq l \leq i-1, \\ (l, s) = (j, i), \quad r = k &\Rightarrow +x_{k,j} y_{k,i}, \quad 1 \leq k \leq \min\{i, j\} - 1, \\ (k, r) = (j, i), \quad s = l &\Rightarrow +x_{j,l} y_{i,l}, \quad \max\{i, j\} + 1 \leq l \leq n, \\ (l, r) = (j, i), \quad s = k &\Rightarrow -x_{k,j} y_{i,k}, \quad i+1 \leq k \leq n. \end{aligned}$$

For $n = 4$ and $(i, j) = (1, 3)$ just four terms survive from the above list,

$$[X, Y]_{1,3} = x_{1,2} y_{2,3} - x_{1,4} y_{3,4} + x_{3,4} y_{1,4} - x_{2,3} y_{1,2}.$$

Examine now the digraph for $n = 4$: $(1, 3)$ is connected to $(2, 3)$ with label $x_{1,2}$ and outgoing arrow, to $(3, 4)$ with label $x_{1,4}$ and incoming arrow etc. In general, it is easy to confirm the following general rule for the reconstruction of the commutator in our basis.

Lemma 3 *Let $n \geq 3$. Then, for every $(i, j) \in \mathcal{I}$ the element $[X, Y]_{i,j}$ is the sum of terms of the form $\pm x_{k,l} y_{r,s}$ over all the $2(n-2)$ edges adjoining the vertices (i, j) and (r, s) with the weight $x_{k,l}$, and with the sign being $+1$ if the arrow is outgoing from (i, j) , -1 otherwise.*

Lemma 3 becomes very useful when the matrix Y is sparse, since the algorithm therein lends itself handily to the exploitation of structure and sparsity.

3 The radius of $\mathfrak{so}(n)$ for $n \geq 4$

3.1 The eigenstructure of $\mathfrak{so}(n)$ in \mathbb{R}^m

The evaluation of the Frobenius norm of a commutator comes as something of an anticlimax, since the spectrum of the *restricted* commutator operator can be evaluated with relative ease. We have already noted that the eigenvalues of the *full* commutator operator, acting in \mathbb{R}^{n^2} , are $\{i(\lambda_k - \lambda_l) : k, l = 1, 2, \dots, n\}$, where $\sigma(X) = \{i\lambda_1, i\lambda_2, \dots, i\lambda_n\}$. Our contention is that $m = \frac{1}{2}(n-1)n$ of these eigenvalues survive intact once we consider the restricted commutator.

To this end, we commence by revisiting the classical analysis of the eigenstructure of the full commutator. Thus, suppose that $X \in M_n[\mathbb{R}]$ has a full set of eigenvectors, therefore $X = VDV^{-1}$, where $D = \text{diag } \lambda$. For every $k, l = 1, 2, \dots, n$, $k \neq l$, we set $E_{k,l} \in M_n[\mathbb{R}]$ as a zero matrix, except for a unit element at the (k, l) entry. Therefore

$$V^{-1}[X, E_{k,l}]V = DE_{k,l} - E_{k,l}D = (\lambda_k - \lambda_l)E_{k,l}$$

and $[X, W_{k,l}] = (\lambda_k - \lambda_l)W_{k,l}$, where $W_{k,l} = VE_{k,l}V^{-1}$, $k, l = 1, 2, \dots, n$. However, if X resides in a Lie algebra \mathfrak{g} , we cannot expect $W_{k,l}$ to belong to \mathfrak{g} : If $\mathfrak{g} = \mathfrak{so}(n)$ then this is in general false.

Suppose that $X \in \mathfrak{so}(n)$ and assume that $n = 2N$ – the case of an odd n will be addressed briefly in the sequel. We set $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then there exists a matrix $Q \in \text{SO}(n)$ such that

$$QXQ^T = A = \begin{bmatrix} A_1 & O & \cdots & O \\ O & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & A_N \end{bmatrix},$$

where $A_k = \alpha_k J$, $k = 1, 2, \dots, N$. Note that the eigenvalues of X are $\pm i\alpha_k$, $k = 1, 2, \dots, N$.

Choose $1 \leq k < l \leq N$ and let $V \in \mathfrak{so}(n, \mathbb{C})$ be a zero matrix, except that

$$\begin{bmatrix} V_{2k-1,2l-1} & V_{2k-1,2l} \\ V_{2k,2l-1} & V_{2k,2l} \end{bmatrix} = U, \quad \begin{bmatrix} V_{2l-1,2k-1} & V_{2l-1,2k} \\ V_{2l,2k-1} & V_{2l,2k} \end{bmatrix} = -U,$$

where $U = \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}$. Letting $Z = [A, V]$, we observe that all the entries of Z vanish, except for

$$\begin{bmatrix} Z_{2k-1,2l-1} & Z_{2k-1,2l} \\ Z_{2k,2l-1} & Z_{2k,2l} \end{bmatrix} = A_k U - U A_l, \quad \begin{bmatrix} Z_{2l-1,2k-1} & Z_{2l-1,2k} \\ Z_{2l,2k-1} & Z_{2l,2k} \end{bmatrix} = U A_k - A_l U.$$

Assume that $\gamma \in \mathbb{C}$ and $\mathbf{u} \neq \mathbf{0}$ are an eigenvalue and an eigenvector, respectively, of the matrix

$$\begin{bmatrix} 0 & \alpha_l & \alpha_k & 0 \\ -\alpha_l & 0 & 0 & \alpha_k \\ -\alpha_k & 0 & 0 & \alpha_l \\ 0 & -\alpha_k & -\alpha_l & 0 \end{bmatrix}.$$

Then $A_k U - U A_l = \gamma U$ and it follows that $\nu(V)$ is an eigenvector of C_A , corresponding to the eigenvalue γ . This results for each $k < l$ in *four* eigenvalue/eigenvector pairs,

$$\begin{aligned} \gamma &= i(\alpha_k + \alpha_l), & U &= \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}; \\ \gamma &= i(-\alpha_k + \alpha_l), & U &= \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}; \\ \gamma &= i(\alpha_k - \alpha_l), & U &= \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}; \\ \gamma &= i(-\alpha_k - \alpha_l), & U &= \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}. \end{aligned}$$

Altogether, this results in $\frac{1}{2}(N-1)N = \frac{1}{2}(n-1)n - \frac{1}{2}n$ eigenvalues of C_A .

The remaining $N = \frac{1}{2}n$ eigenvalues of C_A are zero. This is easy to verify by letting, for any $k = 1, 2, \dots, N$, $V \in \mathfrak{so}(n, \mathbb{C})$ be zero, except that

$$\begin{bmatrix} V_{2k-1,2k-1} & V_{2k-1,2k} \\ V_{2k,2k-1} & V_{2k,2k} \end{bmatrix} = J,$$

whence $[A, V] = O$.

Once we have determined $\sigma(C_A)$, we note that $\sigma(C_X) = \sigma(C_A)$ whenever X and A are similar, since $X = QAQ^{-1}$ means that

$$C_X \nu(Y) = \nu(Z) \quad \Leftrightarrow \quad C_A \nu(Q^{-1}YQ) = \nu(Q^{-1}ZQ).$$

Lemma 4 *Suppose that $n = 2N$ and that the eigenvalues of $X \in \mathfrak{so}(n)$ are $\pm i\alpha_k$, $k = 1, 2, \dots, N$. Then the eigenvalues of the restricted commutator C_X are*

$$i(\pm\alpha_k \pm \alpha_l), \quad 1 \leq k < l \leq N, \quad (3.1)$$

as well as a zero eigenvalue of multiplicity N .

We note as an aside that we have just determined that the *centralizer* of $X \in \mathfrak{so}(n)$ is $(n/2)$ -dimensional, as well as presenting its basis.

Lemma 5 *Suppose that $n = 2N + 1$ and that the eigenvalues of $X \in \mathfrak{so}(n)$ are $\pm i\alpha_k$, $k = 1, 2, \dots, N$ and zero. Then the eigenvalues of the restricted commutator C_X are*

$$i(\pm\alpha_k \pm \alpha_l), \quad 1 \leq k < l \leq N, \quad (3.2)$$

$$\pm i\alpha_k, \quad 1 \leq k \leq N, \quad (3.3)$$

as well as a zero eigenvalue of multiplicity N .

Proof Since $X \in \mathfrak{so}(n)$ is necessarily singular, we need to add to A a bottom row and rightmost column of zeros: We denote the new, $(2N + 1) \times (2N + 1)$ matrix by \tilde{A} . All the eigenvectors of C_A , suitably padded by zeros, can be extended to $C_{\tilde{A}}$. Moreover, let $v \in \mathbb{C}^{2N}$ be a nonzero eigenvector of \tilde{A} with an eigenvalue $i\gamma$ and set

$$V = \begin{bmatrix} O & v \\ -v^T & 0 \end{bmatrix}.$$

Then we can easily verify that $[\tilde{A}, V] = i\gamma V$. Hence we recover the eigenvalues (3.3). Altogether we have $N(2N + 1)$ eigenvalues, hence the full spectrum of $C_{\tilde{A}}$. Since the spectrum of the restricted commutator is invariant under similarity transformation, the proof is complete. \square

3.2 The radius of $\mathfrak{so}(n)$

Up to $\sqrt{2}$, measuring $\mathfrak{so}(n)$ in the Frobenius norm is the same as using the Euclidean norm in $\mathbb{R}^{\frac{1}{2}(n-1)n}$, $\|X\|_F = \sqrt{2}\|\nu(X)\|_2$. Moreover, C_X is skew symmetric, therefore normal, and its Euclidean norm coincides with its spectral radius.

Theorem 6 For every $n \geq 4$ it is true that

$$\omega(\mathfrak{so}(n)) = \sqrt{2}. \quad (3.4)$$

Proof We commence with even $n = 2N$ and assume that, without loss of generality,

$$|\alpha_1| \geq |\alpha_2| \geq \cdots \geq |\alpha_N|.$$

Therefore, according to (3.1),

$$\|C_X\|_2 = \rho(C_X) = |\alpha_1| + |\alpha_2|.$$

Since $\|X\|_{\mathbb{F}}^2 = \sum_{\lambda \in \sigma(X)} |\lambda|^2$, we deduce that

$$\begin{aligned} \frac{\|C_X\|_2}{\|\nu(X)\|_2} &= \frac{|\alpha_1| + |\alpha_2|}{\frac{1}{\sqrt{2}}\|X\|_{\mathbb{F}}} = \frac{|\alpha_1| + |\alpha_2|}{\sqrt{\sum_{k=1}^N |\alpha_k|^2}} \\ &\leq \frac{|\alpha_1| + |\alpha_2|}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2}} \leq \sqrt{2}, \end{aligned}$$

with the upper bound attainable when $\alpha_1 = \alpha_2 > 0$, $\alpha_k = 0$, $k \geq 3$, which corresponds to an embedding of $\mathfrak{so}(4)$ in the algebra. Note that the inequality above holds by Young's inequality for $p = 2$, i.e. we have $2|\alpha_1||\alpha_2| \leq |\alpha_1|^2 + |\alpha_2|^2$.

Therefore $\omega(\mathfrak{so}(2N)) = \sqrt{2}$.

The proof for $n = 2N + 1$ is virtually identical, since

$$\rho(C_X) = \max\left\{\max_{1 \leq k < l \leq N} |\alpha_k| + |\alpha_l|, \max_{k=1,2,\dots,N} |\alpha_k|\right\} = |\alpha_1| + |\alpha_2|,$$

and we again obtain the radius (3.4). \square

Example It is instructive to analyse the special case $\mathfrak{so}(4)$. Using the structure constants one can compute that for

$$X = \begin{bmatrix} 0 & x_1 & x_2 & x_3 \\ -x_1 & 0 & x_4 & x_5 \\ -x_2 & -x_4 & 0 & x_6 \\ -x_3 & -x_5 & -x_6 & 0 \end{bmatrix},$$

we have

$$C_X = \begin{bmatrix} 0 & x_4 & x_5 & -x_2 & -x_3 & 0 \\ -x_4 & 0 & x_6 & x_1 & 0 & -x_3 \\ -x_5 & -x_6 & 0 & 0 & x_1 & x_2 \\ x_2 & -x_1 & 0 & 0 & x_6 & -x_5 \\ x_3 & 0 & -x_1 & -x_6 & 0 & x_4 \\ 0 & x_3 & -x_2 & x_5 & -x_4 & 0 \end{bmatrix}.$$

The 2-norm of C_X may then be computed to be $\|x\|^2 + 2|x_1x_6 - x_2x_5 + x_3x_4|$. Using Lagrange multipliers to maximize this subject to $\|x\| = 1$ yields indeed that $\omega(\mathfrak{so}(4)) \leq \sqrt{2}$.

It is of interest in fact to to characterize all $\mathfrak{so}(4)$ matrices whose restricted commutator has the norm $\sqrt{2}$. These take either the form

$$X = \begin{bmatrix} 0 & a & b & c \\ -a & 0 & c & -b \\ -b & -c & 0 & a \\ -c & b & -a & 0 \end{bmatrix},$$

or

$$X = \begin{bmatrix} 0 & a & b & c \\ -a & 0 & -c & b \\ -b & c & 0 & -a \\ -c & -b & a & 0 \end{bmatrix}$$

for arbitrary $a, b, c \in \mathbb{R}$ which are not all zero.

In each case the spectrum of C_X consists of 4 zero eigenvalues and $\pm i2(a^2 + b^2 + c^2)^{1/2}$. Hence $\|X\|_F = 2(a^2 + b^2 + c^2)^{1/2}$ and thus $\|\nu(X)\|_2 = \sqrt{2}(a^2 + b^2 + c^2)^{1/2}$ and $\|C_X\|_2 = 2(a^2 + b^2 + c^2)^{1/2}$ and it follows that $\|C_X\|_2/\|\nu(X)\|_2 = \sqrt{2}$.

4 Conclusion

We have defined the radius of a Lie algebra and computed its value for $\mathfrak{so}(n)$ and the Frobenius norm. It is of interest to compute the radius for other Lie algebras. We intend to do this in a future publication. In generalizing the work here one needs to distinguish between compact and noncompact Lie algebras (where the Killing form is definite and indefinite respectively) and between real and complex algebras. The compact real form of a complex Lie algebra is natural to look at – for example $\mathfrak{su}(n)$, the compact real form of $\mathfrak{sl}(n, \mathbb{C})$. In the case of $\mathfrak{su}(2)$ one has of course a Lie algebra isomorphism between $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ and \mathbb{R}^3 endowed with the cross product. The map in this case is given by (see e.g.(Marsden & Ratiu 1999))

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \rightarrow X = \frac{1}{2} \begin{bmatrix} -ix_3 & -ix_1 - x_2 \\ -ix_1 + x_2 & ix_3 \end{bmatrix}.$$

Thus our earlier argument for $\mathfrak{so}(3)$ shows that $\omega(\mathfrak{su}(2)) = 1$ with respect to the norm induced by the vector norm on \mathbb{R}^3 .

Acknowledgements

We would like to thank Brad Baxter, Reng-Cang Li, Elizabeth Mansfield, Alexei Shadrin and Mike Shub for useful comments.

References

Brockett, R. (1993), ‘Differential geometry and the design of gradient algorithms’, *Proc. Symposia Pure Mathematics* **54**, 69–92.

- Brockett, R. (1994), 'Differential equations and matrix inequalities on isospectral families', *Linear Algebra Applcs* **203/204**, 189–207.
- Carter, R., Segal, G. & Macdonald, I. (1995), *Lectures on Lie Groups and Lie Algebras*, number 32 in 'London Mathematical Society Student Texts', Cambridge University Press, Cambridge.
- Chartrand, G. & Lesniak, L. (1986), *Graphs and Digraphs*, Wadsworth & Brooks/Cole.
- Hille, E. (1969), *Lectures on Ordinary Differential Equations*, Addison–Wesley, Reading, MA.
- Horn, R. A. & Johnson, C. R. (1994), *Topics in Matrix Analysis*, Cambridge University Press, Cambridge.
- Humphreys, J. E. (1978), *Introduction to Lie Algebras and Representation Theory*, number 9 in 'Graduate Texts in Maths', Springer Verlag, Berlin.
- Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. & Zanna, A. (2000), 'Lie-group methods', *Acta Numerica* **9**, 215–365.
- Kobayahsi, S. & Nomizu, K. (1969), *Foundations of Differential Geometry*, John Wiley, New York.
- Marsden, J. & Ratiu, T. (1999), *Introduction to Mechanics and Symmetry*, Springer Verlag, New York.
- Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and Their Representations*, number 102 in 'Graduate Texts in Mathematics', Springer-Verlag, New York.
- von Neumann, J. (1937), 'Some matrix inequalities and metrization of matrix space', *Tomsk University Review* **1**, 286–300. In *Collected Works*, Pergamon, Oxford, Volume IV (1962), pp. 205–218.