Computing Exponentials of Real Matrices Diagonalizable Over $\mathbb{C}$

Jean Gallier

Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
jean@saul.cis.upenn.edu

January 24, 2008

Abstract. In this note, we consider the problem of computing the exponential of a real matrix. It is shown that if A is a real $n \times n$ matrix and A can be diagonalized over $\mathbb{C}$, then there is a formula for computing e^A involving only real matrices. When A is a skew symmetric matrix, the formula reduces to the generalization of Rodrigues's formula given in Gallier and Xu [1].

1 Introduction

In this note, we consider the problem of computing the exponential of a real matrix. As much as possible, we would like to find a closed-form formula only involving real matrices. This is possible in some cases, for instance, if the matrix A has real eigenvalues and is diagonalizable. This is also possible when A is a 3×3 skew symmetric matrix, by the well–known formula due to Rodrigues (see Marsden and Ratiu [4], McCarthy [5], or Murray, Li and Sastry [7]). In a recent paper [1], Gallier and Xu prove that there is a natural generalization of Rodrigues's formula applying to any $n \times n$ skew symmetric matrix. In this paper, we show that if A is a real $n \times n$ matrix and if A can be diagonalized over $\mathbb{C}$, then there is a formula for computing e^A involving only real matrices. When A is a skew symmetric matrix, the formula reduces to the formula given in Gallier and Xu [1]. The key point is that A can be decomposed as

$$A = \sum_{i=1}^{m} \left(-\lambda_i V_i^2 + \mu_i V_i\right) + \sum_{i=1}^{p} \lambda_{m+i} W_i,$$

where the $\lambda_j \pm i\mu_j$'s and λ_{m+i}'s are the complex and real eigenvalues of A respectively, and the matrices V_i and W_j satisfy certain conditions so that

$$e^A = I_n + \sum_{i=1}^{m} \left(e^{\lambda_i} \sin \mu_i \, V_i + (1 - e^{\lambda_i} \cos \mu_i)\, V_i^2\right) + \sum_{i=1}^{p} (e^{\lambda_{m+i}} - 1)\, W_i.$$

Furthermore, the matrices V_i and W_j are unique.

The general problem of computing the exponential of a matrix is discussed in Moler and Van Loan [6]. However, more general types of matrices are considered, and Moler and Van Loan's investigations have basically no bearing on the results of this paper.

We begin with a special case that turns out to be crucial for the general case.

2 Preliminary Results

Let J be the matrix

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Obviously,

$$J^2 = -I_2.$$

Thus, J behaves like the complex number i, and this is basically why the following lemma holds.

Lemma 2.1 *Given any real 2×2 matrix A of the form*

$$A = \begin{pmatrix} \lambda & -\mu \\ \mu & \lambda \end{pmatrix},$$

we have

$$e^A = e^\lambda\left(\cos\mu\, I_2 + \sin\mu\, J\right) = e^\lambda\begin{pmatrix}\cos\mu & -\sin\mu\\ \sin\mu & \cos\mu\end{pmatrix}.$$

Proof. The matrix A can be written as

$$A = \lambda\, I_2 + \mu\, J,$$

where

$$J = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$

satisfies the identity $J^2 = -I_2$. It is easily checked by induction that

$$A^k = \frac{1}{2}\left((\lambda+i\mu)^k + (\lambda-i\mu)^k\right) I_2 + \frac{1}{2i}\left((\lambda+i\mu)^k - (\lambda-i\mu)^k\right) J$$

for all $k \geq 0$. Thus,

$$\begin{aligned}
e^A &= \sum_{k\geq 0}\frac{A^k}{k!}\\
&= \frac{1}{2}\sum_{k\geq 0}\frac{1}{k!}\left((\lambda+i\mu)^k + (\lambda-i\mu)^k\right) I_2 + \frac{1}{2i}\sum_{k\geq 0}\frac{1}{k!}\left((\lambda+i\mu)^k - (\lambda-i\mu)^k\right) J\\
&= \frac{1}{2}\left(e^{\lambda+i\mu} + e^{\lambda-i\mu}\right) I_2 + \frac{1}{2i}\left(e^{\lambda+i\mu} - e^{\lambda-i\mu}\right) J\\
&= e^\lambda\left(\cos\mu\, I_2 + \sin\mu\, J\right).
\end{aligned}$$

□

The proof of Lemma 2.1 yields the following useful result.

Lemma 2.2 *If U and V are $n\times n$ matrices such that*

$$U^2 = U,\quad UV = VU = V,\quad \textit{and}\quad V^2 = -U,$$

then for any matrix of the form

$$A = \lambda\, U + \mu\, V,$$

we have

$$e^A = I_n + (e^\lambda\cos\mu - 1)\, U + e^\lambda \sin\mu\, V.$$

Proof. The proof of Lemma 2.1 can be adapted as follows. We can still prove by induction that

$$A^k = \frac{1}{2}\left((\lambda+i\mu)^k + (\lambda-i\mu)^k\right) U + \frac{1}{2i}\left((\lambda+i\mu)^k - (\lambda-i\mu)^k\right) V$$

for all $k \geq 1$ (but not for $k = 0$). Thus,

$$
\begin{aligned}
e^A &= I_n + \sum_{k \geq 1} \frac{A^k}{k!} \\
&= I_n + \frac{1}{2} \sum_{k \geq 1} \frac{1}{k!} \left((\lambda + i\mu)^k + (\lambda - i\mu)^k \right) U + \frac{1}{2i} \sum_{k \geq 1} \frac{1}{k!} \left((\lambda + i\mu)^k - (\lambda - i\mu)^k \right) V \\
&= I_n + \frac{1}{2} \left(e^{\lambda + i\mu} - 1 + e^{\lambda - i\mu} - 1 \right) U + \frac{1}{2i} \left(e^{\lambda + i\mu} - 1 - (e^{\lambda - i\mu} - 1) \right) V \\
&= I_n + (e^\lambda \cos \mu - 1)\, U + e^\lambda \sin \mu \, V.
\end{aligned}
$$

□

We are now ready to present a formula for the exponential of a real matrix that can be diagonalized over $\mathbb{C}$.

3 The Formula

Let A be a real $n \times n$ matrix that can be diagonalized over $\mathbb{C}$. The complex eigenvalues of A come in conjugate pairs $\lambda \pm i\mu$ (with $\lambda, \mu \in \mathbb{R}$). It is also obvious that if $u + iv$ is an eigenvector for $\lambda + i\mu$, where u, v are real vectors, then $u - iv$ is an eigenvector for $\lambda - i\mu$. If $\mu \neq 0$, then $\lambda + i\mu \neq \lambda - i\mu$ and thus $u + iv$ and $u - iv$ are linearly independent (over $\mathbb{C}$), which implies that u and v are linearly independent. Since

$$
A(u) + iA(v) = A(u + iv) = (\lambda + i\mu)(u + iv) = \lambda u - \mu v + i(\mu u + \lambda v),
$$

we get

$$
\begin{aligned}
Au &= \lambda u - \mu v, \\
Av &= \mu u + \lambda v.
\end{aligned}
$$

Since A can be diagonalized over $\mathbb{C}$, there is a basis consisting of (complex) eigenvectors, and using the above remarks, we can find a real basis with respect to which A is a block diagonal matrix whose blocks are one-dimensional or of the form

$$
\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix},
$$

where $\lambda, \mu \in \mathbb{R}$. This is summarized in the following Lemma, where it is more convenient to replace μ by $-\mu_1$. This lemma is also a direct consequence of the real Jordan form, see Horn and Johnson [3].

Lemma 3.1 *Let A be a real $n \times n$ matrix that can be diagonalized over $\mathbb{C}$. Then, A can be written as $A = PDP^{-1}$, where P and D are real matrices with D a block diagonal matrix of the form*

$$\begin{pmatrix} D_1 & & \dots & \\ & D_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & D_p \end{pmatrix},$$

such that each block D_j is either a one-dimensional matrix (i.e., a real scalar) or a two-dimensional matrix of the form

$$D_j = \begin{pmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{pmatrix},$$

where $\lambda_j, \mu_j \in \mathbb{R}$.

The two-dimensional blocks correspond to pairs of conjugate eigenvalues $\lambda_j + i\mu_j$ and $\lambda_j - i\mu_j$, and the one-dimensional blocks to the real eigenvalues of A.

Using Lemma 3.1, we can prove that a certain decomposition of A exists. The existence of this decomposition immediately yields a formula for e^A. The uniqueness of the matrices involved in the decomposition will be proved in the next section.

Lemma 3.2 *Let A be a real $n \times n$ matrix that can be diagonalized over $\mathbb{C}$ and assume that its set of nonnull eigenvalues is*

$$\{\lambda_1 + i\mu_1, \lambda_1 - i\mu_1, \dots, \lambda_m + i\mu_m, \lambda_m - i\mu_m, \lambda_{m+1}, \dots, \lambda_{m+p}\},$$

with $\lambda_j, \mu_j, \lambda_{m+k} \in \mathbb{R}$ and $\mu_j \neq 0$. Then, there are real matrices V_i, W_k such that

$$\begin{aligned} V_i^3 &= -V_i, \\ V_i V_j &= V_j V_i = 0, \ (i \neq j) \\ V_i W_k &= W_k V_i = 0, \ (i \neq k) \\ W_k W_l &= W_l W_k = 0, \ (k \neq l) \\ W_k^2 &= W_k, \end{aligned}$$

and

$$A = \sum_{i=1}^{m} \left(-\lambda_i V_i^2 + \mu_i V_i\right) + \sum_{i=1}^{p} \lambda_{m+i} W_i.$$

Furthermore,

$$e^A = I_n + \sum_{i=1}^{m} \left(e^{\lambda_i} \sin \mu_i \, V_i + (1 - e^{\lambda_i} \cos \mu_i)\, V_i^2\right) + \sum_{i=1}^{p} (e^{\lambda_{m+i}} - 1)\, W_i.$$

Proof. Let $A = PDP^{-1}$, as in Lemma 3.1. For every $\lambda_j + i\mu_j$, let E_j be the matrix obtained from D by deleting all the blocks corresponding to eigenvalues distinct from $\lambda_j + i\mu_j$. Then,

$$E_j = \lambda_j F_j + \mu_j G_j,$$

where $\lambda_j F_j$ is the diagonal part of E_j and $\mu_j G_j$ is obtained by deleting the diagonal entries from E_j. For every λ_{m+k}, let $\lambda_{m+k} F_{m+k}$ be the matrix obtained from D by deleting all the blocks corresponding to eigenvalues distinct from λ_{m+k}. Observe that

$$G_j^2 = -F_j \quad \text{and} \quad F_{m+k}^2 = F_{m+k}.$$

Then, let

$$U_j = PF_jP^{-1}, \quad V_j = PG_jP^{-1}, \quad \text{and} \quad W_k = PF_{m+k}P^{-1}.$$

Clearly,

$$\begin{aligned}
V_i^3 &= -V_i, \\
V_iV_j &= V_jV_i = 0, \ (i \neq j) \\
V_iW_k &= W_kV_i = 0, \ (i \neq k) \\
W_kW_l &= W_lW_k = 0, \ (k \neq l) \\
W_k^2 &= W_k, \\
V_i^2 &= -U_i,
\end{aligned}$$

and

$$A = \sum_{i=1}^{m} (\lambda_i U_i + \mu_i V_i) + \sum_{i=1}^{p} \lambda_{m+i} W_i = \sum_{i=1}^{m} \left(-\lambda_i V_i^2 + \mu_i V_i\right) + \sum_{i=1}^{p} \lambda_{m+i} W_i.$$

Since $U_i = -V_i^2$, from $V_i^3 = -V_i$, we get

$$U_i^2 = U_i, \quad U_iV_i = V_iU_i = V_i.$$

Since the V_i's commute, the W_k's commute, and the V_i's and the W_k's commute, we get

$$e^A = \prod_{i=1}^{m} e^{\lambda_i U_i + \mu_i V_i} \prod_{i=1}^{p} e^{\lambda_{m+i} W_i}.$$

However, by Lemma 2.2,

$$e^{\lambda_i U_i + \mu_i V_i} = I_n + (e^{\lambda_i} \cos \mu_i - 1)\, U_i + e^{\lambda_i} \sin \mu_i \, V_i,$$

and since $U_i = -V_i^2$, we get

$$e^{\lambda_i U_i + \mu_i V_i} = I_n + e^{\lambda_i} \sin \mu_i \, V_i + (1 - e^{\lambda_i} \cos \mu_i)\, V_i^2.$$

Since $W_k^2 = W_k$, we get

$$e^{\lambda_{m+i} W_i} = I_n + (e^{\lambda_{m+i}} - 1)\, W_i.$$

Since $V_iV_j = V_jV_i = 0$ for $i \neq j$, $V_iW_k = W_kV_i = 0$ for $i \neq k$, and $W_kW_l = W_lW_k = 0$ for $k \neq l$, we finally get

$$e^A = I_n + \sum_{i=1}^{m} \left(e^{\lambda_i} \sin \mu_i \, V_i + (1 - e^{\lambda_i} \cos \mu_i) \, V_i^2\right) + \sum_{i=1}^{p} (e^{\lambda_{m+i}} - 1) \, W_i.$$

$\square$

Recall that a normal matrix is a matrix A such that

$$AA^* = A^*A$$

It is well-known that normal matrices can be diagonalized in terms of unitary matrices (see Golub and Van Loan [2] or Trefethen and Bau [8]). Then, if A is a normal matrix, the matrix P of Lemma 3.1 can be chosen to be orthogonal, which implies that the matrices V_i are skew symmetric and the matrices W_k are symmetric.

If A is skew symmetric, its eigenvalues are pure imaginary of null. Thus, the formula reduces to

$$e^A = I_n + \sum_{i=1}^{m} \left(\sin \mu_i \, V_i + (1 - \cos \mu_i) \, V_i^2\right),$$

the generalization of Rodrigues's formula given in Gallier and Xu [1].

Next, we prove the uniqueness of the V_i's and W_k's.

4 Uniqueness of the V_i's and W_k's

Lemma 4.1 *Let A be a real $n \times n$ matrix that can be diagonalized over $\mathbb{C}$ and assume that its set of nonnull eigenvalues is*

$$\{\lambda_1 + i\mu_1, \lambda_1 - i\mu_1, \ldots, \lambda_m + i\mu_m, \lambda_m - i\mu_m, \lambda_{m+1}, \ldots, \lambda_{m+p}\},$$

with $\lambda_j, \mu_j, \lambda_{m+k} \in \mathbb{R}$ and $\mu_j \neq 0$. The real matrices V_i, W_k satisfying the conditions

$$\begin{aligned} V_i^3 &= -V_i, \\ V_iV_j &= V_jV_i = 0, \ (i \neq j) \\ V_iW_k &= W_kV_i = 0, \ (i \neq k) \\ W_kW_l &= W_lW_k = 0, \ (k \neq l) \\ W_k^2 &= W_k, \end{aligned}$$

and

$$A = \sum_{i=1}^{m} \left(-\lambda_i V_i^2 + \mu_i V_i\right) + \sum_{i=1}^{p} \lambda_{m+i} W_i$$

as in Lemma 3.2, are unique.

Proof. Letting $U_i = -V_i^2$, using the identities

$$\begin{aligned} V_i^3 &= -V_i, \\ V_iV_j &= V_jV_i = 0,\ (i \neq j) \\ V_iW_k &= W_kV_i = 0,\ (i \neq k) \\ W_kW_l &= W_lW_k = 0,\ (k \neq l) \\ W_k^2 &= W_k, \end{aligned}$$

U_i and V_i satisfy the conditions of Lemma 2.2, and we get the following formula for A^k:

$$\begin{aligned} A^k = &\sum_{j=1}^{m} \left(\frac{1}{2} \left((\lambda_j + i\mu_j)^k + (\lambda_j - i\mu_j)^k \right) U_j + \frac{1}{2i} \left((\lambda_j + i\mu_j)^k - (\lambda_j - i\mu_j)^k \right) V_j \right) \\ &+ \sum_{j=1}^{p} \lambda_{m+j}^k W_i. \end{aligned}$$

Thus, we can form a system of $2m+p$ equations with the $2m+p$ variables $U_1, \ldots, U_m$, $V_1, \ldots, V_m$, $W_1, \ldots, W_p$, where the kth equation is the above equation. We claim that the determinant of this system can be reduced to a Vandermonde determinant and that it is nonnull. Indeed, the determinant of the system is

$$\begin{vmatrix} x_{1\,1} & \ldots & x_{m\,1} & y_{1\,1} & \ldots & y_{m\,1} & \lambda_{m+1} & \ldots & \lambda_{m+p} \\ x_{1\,2} & \ldots & x_{m\,2} & y_{1\,2} & \ldots & y_{m\,2} & \lambda_{m+1}^2 & \ldots & \lambda_{m+p}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1\,2m+p} & \ldots & x_{m\,2m+p} & y_{1\,2m+p} & \ldots & y_{m\,2m+p} & \lambda_{m+1}^{2m+p} & \ldots & \lambda_{m+p}^{2m+p} \end{vmatrix}$$

where

$$\begin{aligned} x_{j\,k} &= \frac{1}{2} \left((\lambda_j + i\mu_j)^k + (\lambda_j - i\mu_j)^k \right), \\ y_{j\,k} &= \frac{1}{2i} \left((\lambda_j + i\mu_j)^k - (\lambda_j - i\mu_j)^k \right). \end{aligned}$$

For every j, $1 \leq j \leq m$, if we add the $(m+j)$th column to the jth column, we get

$$\begin{vmatrix} \lambda_1 + i\mu_1 & \ldots & \lambda_m + i\mu_m & y_{1\,1} & \ldots & y_{m\,1} & \lambda_{m+1} & \ldots & \lambda_{m+p} \\ (\lambda_1 + i\mu_1)^2 & \ldots & (\lambda_m + i\mu_m)^2 & y_{1\,2} & \ldots & y_{m\,2} & \lambda_{m+1}^2 & \ldots & \lambda_{m+p}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (\lambda_1 + i\mu_1)^{2m+p} & \ldots & (\lambda_m + i\mu_m)^{2m+p} & y_{1\,2m+p} & \ldots & y_{m\,2m+p} & \lambda_{m+1}^{2m+p} & \ldots & \lambda_{m+p}^{2m+p} \end{vmatrix}.$$

For every j, $1 \leq j \leq m$, if if we subtract $1/2i$ times the jth column from the $(m+j)$th column, we get

$$(i/2)^m \begin{vmatrix} z_1 & \ldots & z_m & \overline{z_1} & \ldots & \overline{z_m} & \lambda_{m+1} & \ldots & \lambda_{m+p} \\ z_1^2 & \ldots & z_m^2 & \overline{z_1}^2 & \ldots & \overline{z_m}^2 & \lambda_{m+1}^2 & \ldots & \lambda_{m+p}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^{2m+p} & \ldots & z_m^{2m+p} & \overline{z_1}^{2m+p} & \ldots & \overline{z_m}^{2m+p} & \lambda_{m+1}^{2m+p} & \ldots & \lambda_{m+p}^{2m+p} \end{vmatrix}$$

where $z_j = \lambda_j + i\mu_j$ (and $\overline{z_j} = \lambda_j - i\mu_j$). Thus, the determinant of the system if proportional to a Vandermonde determinant, which is nonnull, since

$$\{\lambda_1 + i\mu_1, \lambda_1 - i\mu_1, \ldots, \lambda_m + i\mu_m, \lambda_m - i\mu_m, \lambda_{m+1}, \ldots, \lambda_{m+p}\}$$

is the set of nonnull eigenvalues of A (where $\mu_j \neq 0$). $\square$

5 Conclusion

We have given a formula computing the exponential of a real $n \times n$ matrix A that can be diagonalized over $\mathbb{C}$. This formula reduces to a generalization of Rodrigues's formula given in Gallier and Xu [1] when A is skew symmetric. In Gallier and Xu, a formula for computing exponentials of matrices in the Lie algebra $\mathfrak{se}(n)$ of the Lie group $\mathbf{SE}(n)$ of rigid motions was also given. This suggests that there may be other classes of matrices for which a formula for the exponential can be found. Finding such classes is left as an open problem.

References

[1] Jean Gallier and Dianna Xu. Computing exponentials of skew symmetric matrices and logarithms of orthogonal matrices. *International Journal of Robotics and Automation*, 18:10–20, 2003.

[2] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.

[3] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, first edition, 1990.

[4] Jerrold E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry*. TAM, Vol. 17. Springer Verlag, first edition, 1994.

[5] J.M. McCarthy. *Introduction to Theoretical Kinematics*. MIT Press, first edition, 1990.

[6] Cleve Moler and Charles Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, 20(4):801–836, 1978.

[7] R.M. Murray, Z.X. Li, and S.S. Sastry. *A Mathematical Introduction to Robotics Manipulation*. CRC Press, first edition, 1994.

[8] L.N. Trefethen and D. Bau III. *Numerical Linear Algebra*. SIAM Publications, first edition, 1997.