

Structure Preserving Algorithm for the Logarithm of Symplectic Matrices

Bahar Arslan*

Abstract

The current algorithms use either the full form or the Schur decomposition of the matrix in the inverse scaling and squaring method to compute the matrix logarithm. The inverse scaling and squaring method consists of two main calculations: taking a square root and evaluating the Padé approximants. In this work, we suggest using the structure preserving iteration as an alternative to Denman-Beavers iteration for taking a square root. Numerical experiments show that while using the structure preserving square root iteration in the inverse scaling and squaring method preserves the Hamiltonian structure of matrix logarithm, Denman-Beavers iteration and Schur decomposition cause a structure loss.

Keywords: Matrix functions; matrix logarithm; symplectic matrix; Hamiltonian matrix; inverse scaling and squaring method

AMS Subject Classification (2020): Primary: 00A00 ; Secondary: 00B00; 00C00; 00D00; 00E00; 00F00.

1. Introduction

The matrix logarithm is not only important for being the inverse function of the matrix exponential, it has also many applications. It has been used by engineers in the continuization process. They compute the logarithm of matrices in converting a discrete process into a continuous one [17, 18]. It has also applications to the stability of differential equations [14, 16]. The growing interest in computing structured matrix functions stems from the fact that predicting and preserving the structure of matrices can help us to explain the results physically and geometrically. The logarithm of structured matrices has applications in the control mechanical systems [4, 15] and in the optometry [6]. Structured matrix logarithm is also used for generalizing Bézier curves to non-Euclidean spaces. Crouch's algorithm, which generalizes De Casteljaou algorithm to find polynomial splines on Riemannian manifolds, requires the computation of matrix logarithm when this manifold is a Lie group of matrices [4]. The theory of splines on Lie groups has applications in robotics path planning and air traffic control.

This paper focuses on computing the logarithm of a real symplectic matrix A with the spectrum $\rho(A)$ such that $\rho(A) \cap \mathbb{R}^- = \emptyset$, for which $W = \log A$ is Hamiltonian. In the computation of matrix logarithm, we use the inverse scaling and squaring method proposed by Kenney and Laub [12] and which is based on the relation

$$\log(A) = 2^k \log(A^{1/2^k}).$$

There are two important calculations in the inverse scaling and squaring method. The first one is taking a square root of a symplectic matrix and the second is the Padé approximation [7]. Moreover, the inverse scaling and squaring method can be applied to A directly or it can be used with the Schur decomposition of A . However, we show that the latter case does not preserve the structure of the symplectic matrices. The aim of this paper is to propose using the structure preserving iteration for the square root in the inverse scaling and squaring method. We analyse this approach in terms of structure error, accuracy and computational cost. Numerical experiments assess the advantage of this approach and suggest using the structure preserving iteration in the inverse scaling and squaring method for the logarithm of symplectic matrices to preserve the Hamiltonian structure.

The paper is organized as follows. Section 2 begins with the definition of symplectic matrices and the matrix logarithm. We also review the inverse scaling and squaring method in this section. In Section 3, we propose our algorithm using the structure preserving square root iteration in the inverse scaling and squaring method. Section 4 presents the numerical findings and analyses our approach in terms of structure error, accuracy and cost. Finally, Section 5 gives a brief summary and critique of the findings.

2. Logarithm of symplectic matrices

2.1 Symplectic matrices

Let \mathbb{K} denote the field \mathbb{R} or \mathbb{C} . Consider a scalar product $\langle \cdot, \cdot \rangle_M$ defined by any nonsingular matrix M , for $x, y \in \mathbb{K}^n$,

$$\langle x, y \rangle_M = \begin{cases} x^T M y, & \text{for real or complex bilinear forms,} \\ x^* M y, & \text{for sesquilinear forms.} \end{cases}$$

For any matrix $A \in \mathbb{K}^{n \times n}$, there exists a unique operator $A^* \in \mathbb{K}^{n \times n}$, called the adjoint of A with respect to the scalar product, such that

$$\langle Ax, y \rangle_M = \langle x, A^* y \rangle_M, \quad \forall x, y \in \mathbb{K}^n.$$

A^* can be written explicitly

$$A^* = \begin{cases} M^{-1} A^T M, & \text{for real or complex bilinear forms,} \\ M^{-1} A^* M, & \text{for sesquilinear forms.} \end{cases}$$

Symplectic matrices belong to the automorphism group which is characterized by the adjoint matrix as

$$\mathbb{G} = \{A \in \mathbb{K}^{n \times n} : A^* = A^{-1}\}.$$

So for $M = J$ the matrix $A \in \mathbb{K}^{2n \times 2n}$ is symplectic if $A^T J A = J$. The permutation matrix $J \in \mathbb{R}^{2n \times 2n}$ is given as

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

where I_n is the identity matrix of order n .

2.2 Inverse scaling and squaring method

For a given $A \in \mathbb{K}^{n \times n}$ a logarithm of A is any matrix W such that $e^W = A$. We assume that A has no eigenvalues on \mathbb{R}^- so that the existence of a unique principal logarithm is assured as shown in the following theorem.

Theorem 2.1. [7, Thm. 1.31] *Let $A \in \mathbb{K}^{n \times n}$ have no eigenvalues on \mathbb{R}^- . There is a unique logarithm W of A all of whose eigenvalues lie in the strip $\{z : -\pi < \text{Im}(z) < \pi\}$. We refer to W as the principal logarithm of A and write $W = \log A$. If A is real, then its principal logarithm is real.*

For $M = J$, a matrix W is called Hamiltonian if $W^* = -W$, which implies $W^T = J W J$. The function logarithm maps a symplectic matrix to a Hamiltonian matrix as proved in Theorem 2.2.

Theorem 2.2. [3, Thm. 2.1] *If $A \in \mathbb{K}^{n \times n}$ is a symplectic matrix and $\rho(A) \cap \mathbb{R}^- = \emptyset$, then $\log A = W$ is Hamiltonian.*

Proof. Since A is symplectic it satisfies

$$A^T = -J A^{-1} J = J^{-1} A^{-1} J.$$

Then,

$$\log(A^T) = J^{-1} \log A^{-1} J.$$

As stated in [7, Thm.1.13 and Thm.11.2], $\log(A^T) = (\log A)^T$ and $\log A^{-1} = -\log A$. Therefore,

$$\begin{aligned} (\log A)^T &= -J^{-1} \log A J \\ &= J \log A J. \end{aligned}$$

Hence, it shows that $\log A$ is Hamiltonian. \square

In the computation of matrix logarithm we use the inverse scaling and squaring method. The basic idea of the inverse scaling and squaring method is to take the repetitive square root of A so the result is close to the identity matrix and then use the m -th order Padé approximant r_m . We summarise the method in the following algorithm.

Algorithm 2.3. [12] Given $A \in \mathbb{K}^{n \times n}$ with no eigenvalues on \mathbb{R}^- this algorithm employs the inverse scaling and squaring method to compute $W = \log A$.

- 1 Bring A close to an identity matrix by taking k repetitive square root of A
- 2 Decide the order of $r_m(A^{1/2^k} - I)$ by minimising the cost and maximising the accuracy
- 3 Approximate $\log(A^{1/2^k})$ by using $r_m(A^{1/2^k} - I) \approx \log(A^{1/2^k})$
- 4 Rescale to obtain $W \approx 2^k r_m(A^{1/2^k} - I)$

The question we need to deal with is whether the square root function and the Padé approximation preserve the structure or not. After taking the square root of any matrix in the automorphism group it stays in the automorphism group, which is proved in the following theorem.

Theorem 2.4. [13] *Let A be a matrix that has a principal square root $A^{1/2}$. If A is symplectic, then $A^{1/2}$ is symplectic.*

Proof. If A is symplectic, then $A^* = A^{-1}$. We have the equality

$$(A^*)^{1/2} = (A^{-1})^{1/2} \Rightarrow (A^{1/2})^* = (A^{1/2})^{-1}.$$

\square

Theorem 2.5. [9, Thm. 6.2] *Let \mathbb{G} be any automorphism group and $A \in \mathbb{G}$. If A has no eigenvalues on \mathbb{R}^- , then the iteration*

$$\begin{aligned} Y_{k+1} &= \frac{1}{2}(Y_k + Y_k^{-*}) \\ &= \frac{1}{2}(Y_k + M^{-1}Y_k^{-T}M) \end{aligned} \quad (2.1)$$

with starting matrix $Y_1 = \frac{1}{2}(I + A)$, is well defined and Y_k converges quadratically to $A^{1/2}$.

Since a symplectic matrix belongs to an automorphism group \mathbb{G} the advantage of using iteration (2.1) is that it will preserve the symplectic structure and the result will lie in the group to approximately machine precision. With this iteration, we preserve the symplectic structure and when we evaluate the Padé approximant, we obtain the Hamiltonian structure which is proved in the following theorem.

Theorem 2.6. [5] *Let $r_m(X)$ be the diagonal Padé approximants to $\log(I+X)$, $m = 0, 1, \dots$. Let $W = \log A$ and $X = A - I$ with $\rho(X) < 1$. If A is symplectic, then $r_m(A - I)$ is Hamiltonian.*

Proof. We will use the homographic invariance [2, Thm. 1.5.2] under the argument transformations for this proof. Since $f(x) = \log x$ does not have a power series we take $f(x) = \log(1+x)$. By using the equality $\log(1+x) = -\log\left(1 + \frac{-x}{x+1}\right)$ and [2, Thm. 1.5.2] we get $r_m(x) = -r_m(-x/(x+1))$.

For the matrix case this formula yields $r_m(X) = -r_m(-X(X+I)^{-1})$. If A is a symplectic matrix, then $A^{-1} = -JA^T J$. Thus, we can write

$$\begin{aligned} r_m(A - I) &= -r_m(A^{-1} - I) \\ &= -r_m(-JA^T J - I) \\ &= -r_m(-J(A^T - I)J). \end{aligned}$$

We obtain $r_m(X) = -r_m(-J(X^T)J) = Jr_m(X^T)J = Jr_m(X)^T J = -J^{-1}r_m(X)^T J$ which indicates that $r_m(X)$ is Hamiltonian. \square

We state in the next theorem that the error in matrix Padé approximation is less than the error in scalar Padé approximation at the norm of the matrix, which is used in the inverse scaling and squaring method to decide the order of Padé approximation.

Theorem 2.7. [11] For $\|A - I\| < 1$ and any subordinate matrix norm,

$$\|r_m(A - I) - \log A\| \leq |r_m(-\|A - I\|) - \log(1 - \|A - I\|)|. \quad (2.2)$$

Table 1. Maximal values θ_m of $\|A - I\|$ ensure that the bound $\|r_m(A - I) - \log A\|$ does not exceed $u = 2^{-53}$ [7, Table 11.1].

m	1	2	3	4	5	6	7	8	9
θ_m	1.10e-5	1.82e-3	1.62e-2	5.39e-2	1.14e-1	1.87e-1	2.64e-1	3.40e-1	4.11e-1
m	10	11	12	13	14	15	16	32	64
θ_m	4.75e-1	5.31e-1	5.81e-1	6.24e-1	6.62e-1	6.95e-1	7.24e-1	9.17e-1	9.78e-1

The maximal values θ_m of $\|A - I\|$ such that the error bound $\|r_m(A - I) - \log A\|$ does not exceed $u = 2^{-53} \approx 1.1 \times 10^{-16}$ are given in Table 1.

3. Using the structure preserving square root iteration

We adapt the algorithm [7, Alg. 11.10] by using the structure preserving iteration to take a square root. Iteration (2.1) is used to compute the square root of symplectic matrix and it exploits the symplecticity in each iteration. Let it_j be the number of iterations required in each square root. If M was a full matrix, then the operation count would include the inverse of M and the matrix multiplication. However, since $M = J$ is a permutation of $\text{diag}(\pm 1)$ multiplication by J^{-1} is trivial and the cost of each iteration is one matrix inversion per iteration which is $2n^3$ flops. Evaluating the partial fraction form of the Padé approximation with the order m costs $\frac{8}{3}mn^3$ flops. In iteration (2.1) the number of iterations required to take a square root of A typically changes from 16 on the first iterations to 4 for the last few iterations. So the cost of taking a square root of symplectic matrix A at the last few iterations is $8n^3$ flops. It is worth only taking one more square root if it reduces the order of Padé approximation by at least 3. That decrease in the order of Padé approximation can only be obtained when $\|A^{1/2^s} - I\| > \theta_{16}$, where θ_{16} is the value given in Table 1. Taking a square root of A approximately reduces the distance of $A^{1/2^k}$ to the identity matrix by a half. This is easy to see since

$$(I - A^{1/2^{k+1}})(I + A^{1/2^{k+1}}) = I - A^{1/2^k},$$

and $A^{1/2^k} \rightarrow I$ as $k \rightarrow \infty$, it gives

$$\|I - A^{1/2^{k+1}}\| \approx \frac{1}{2}\|I - A^{1/2^k}\|. \quad (3.1)$$

When $\|A^{1/2^s} - I\| \leq \theta_{16}$ is obtained, in order to compare the cost of the Padé approximation and the cost of the square root iteration, we check the inequality

$$\frac{8}{3}(m_1 - m_2)n^3 \leq 2n^3 it_j \quad \Rightarrow \quad \frac{4}{3}(m_1 - m_2) \leq it_j \quad (3.2)$$

by assuming the same number of iterations is required. In equation (3.2) m_1 and m_2 are the order of Padé approximants before and after the extra square root, respectively. Since the cost of taking 2 more extra square roots exceeds the cost of evaluating the Padé approximant we limit it by taking $p = 2$. That is, only one extra square root is taken if it is required. By using the cost checking (3.2), we present the modified algorithm of the inverse scaling and squaring method using the structure preserving square root iteration (2.1).

Algorithm 3.1. Given a symplectic matrix $A \in \mathbb{K}^{n \times n}$ with no eigenvalues on \mathbb{R}^- this algorithm computes $W = \log A$ by the inverse scaling and squaring method. It uses the constants θ_m given in Table 1 for the Padé approximation and iteration (2.1) to take a square root of A . This algorithm is intended for IEEE double precision arithmetic.

```

1   $k = 0, p = 0$ 
2  while true
3   $\tau = \|A - I\|_1$ 
4  if  $\tau < \theta_{16}$ 
5   $p = p + 1$ 
6   $m_1 = \min\{i : \tau \leq \theta_i, i = 3:16\}$ 
7   $m_2 = \min\{i : \tau/2 \leq \theta_i, i = 3:16\}$ 
8  if  $4(m_1 - m_2)/3 \leq it_j$  or  $p = 2, m = m_1$ , go to line 13, end
9  end
10  $A \leftarrow A^{1/2}$  by using iteration (2.1)
11  $k = k + 1$ 
12 end
13 Evaluate  $Y = r_m(A - I)$ 
14  $W = 2^k Y$ 

```

Cost: Taking a square root costs $(\sum_{j=1}^k it_j)2n^3$ flops where k is the number of square root and evaluating the partial fraction form of the Padé approximation costs about $\frac{8}{3}mn^3$ flops. It is $(\sum_{j=1}^k it_j)2n^3 + \frac{8}{3}mn^3$ flops in total.

4. Numerical experiments

4.1 Error measure

The appropriate relative measure of departure from Hamiltonian structure can be computed by [1]

$$\text{err}_H(W) = \frac{\|W^* + W\|_2}{\|W\|_2}. \quad (4.1)$$

The relative error for the computed logarithm \widehat{W} is given by

$$\text{rel}_{\text{err}}(\widehat{W}) = \frac{\|\widehat{W} - W\|_2}{\|W\|_2}, \quad (4.2)$$

where $W = \log A$ is the "exact" logarithm. In the numerical tests, W is computed at 100 digit precision and we measure the departure from the Hamiltonian structure and the relative error by equations (4.1) and (4.2), respectively.

4.2 Numerical tests

In the numerical experiments, we test using the structure preserving square root iteration (2.1) in the inverse scaling and squaring method in terms of structure loss, accuracy and computational cost. The experiments are carried out in MATLAB R2020b with $u = 1.1 \times 10^{-16}$. We form the full form of symplectic test matrices $A \in \mathbb{R}^{10 \times 10}$ by using the function `rand_rsypm` from Jagger's MATLAB Toolbox [10].

Let \widehat{W} represent the computed logarithm and $W = \log A$ represent the "exact" logarithm obtained by using MATLAB's Symbolic Math Toolbox, where A is diagonalized in 100 digit precision as $A = VDV^{-1}$ and computed by $\log A = V \log DV^{-1}$.

In the experiments, the relative error of the logarithm of the symplectic matrices $\text{rel}_{\text{err}}(\widehat{W})$, the structure error $\text{err}_H(W)$ for the Hamiltonian structure and the computational cost are presented. All the results are plotted with the condition number of the symplectic matrices computed by $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$. The legend labels in figures are described as follows:

1. `full_preserve`: \widehat{W} is computed by Algorithm 3.1 using the full form of A with the structure preserving square root iteration (2.1).
2. `full_DB`: \widehat{W} is computed using the full form of A with the scaled product Denman-Beavers iteration [7, Alg. 11.10].
3. `Schur`: \widehat{W} is computed using the Schur decomposition of A [7, Alg. 11.9] with a square root algorithm [7, Alg. 6.3].

Figure 1 provides the comparison of the structure error. We can see that while using the scaled product Denman-Beavers iteration with the full form of the matrix and Schur decomposition cause a structure loss for the ill-conditioned matrices, using Algorithm 3.1 preserves the Hamiltonian structure. We compare the accuracy of the approaches in Figure 2. As shown, while for the well-conditioned matrices, i.e., $\kappa_2(A) \approx 1$ the methods give the good estimate to the "exact" logarithm, we obtain less accurate results for the badly conditioned matrices. Since we obtain almost the same accuracy from three different approaches, we cannot say one is superior to other in terms of accuracy. Figure 3 reveals that using the iterative methods to compute the square root is computationally expensive. However, reducing the matrix to an upper triangular matrix with the Schur decomposition and using a square root algorithm [7, Alg. 6.3] is relatively cheaper than other approaches.

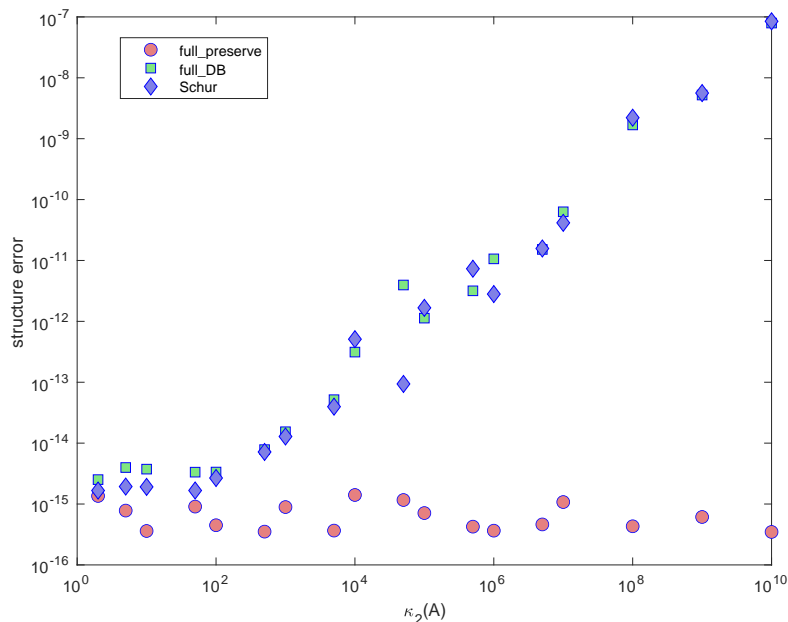


Figure 1. Comparison of the structure error for full_preserve, full_DB and Schur.

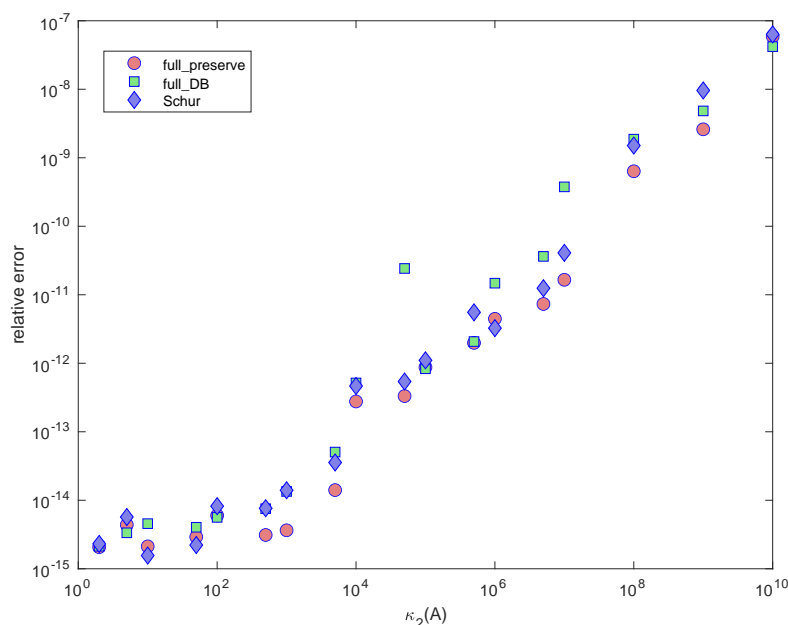


Figure 2. Comparison of the relative error for full_preserve, full_DB and Schur.

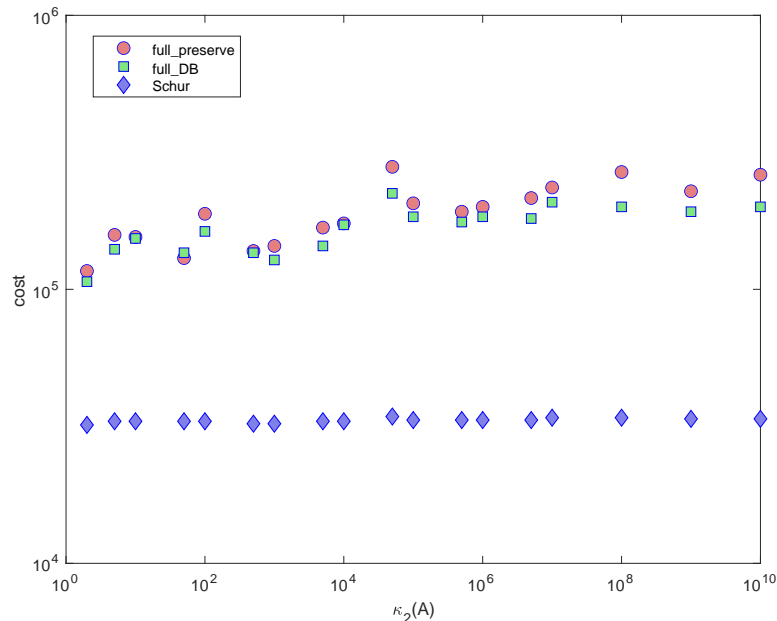


Figure 3. Comparison of the computational cost of full_preserve, full_DB and Schur.

5. Conclusion

We proposed using the structure preserving square root iteration in the inverse scaling and squaring method to compute the logarithm of symplectic matrices and we compared it with the algorithms using either the full form of the matrix with the scaled product Denman-Beavers iteration or the Schur decomposition in terms of the structure loss, the accuracy and the computational cost. The findings show that the best structure is obtained by using the structure preserving square root iteration (2.1) instead of using scaled product Denman-Beavers iteration or the Schur decomposition. While there is not much big difference in their accuracy, however, using the iteration (2.1) with the full form of the symplectic matrices is computationally expensive.

Acknowledgments

We would like to thank the referees for the feedbacks which helped improving the quality of the paper.

Acknowledgment.

We thank the reviewer for their insightful comments and suggestions that helped us improve the paper.

Funding

There is no funding for this work.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

References

- [1] Arslan, B.: Ph.D. Thesis: Functions of Structured Matrices. University of Manchester, Manchester, England, (2017).
- [2] Baker, G. A and Graves-Morris, P.: Padé Approximants. Cambridge University Press, New York, NY 10011-4211, USA, (1996).
- [3] Cardoso, J. R., Leite, S. F.: *On computing real logarithms for matrices in the Lie group of special Euclidean motions in \mathbb{R}^n* . Centro de Matemática da Universidade de Coimbra. (1999).
- [4] Crouch, P., Kun, G., Leite, F. S.: *The De Casteljau Algorithm on Lie Groups and spheres*. Journal of Dynamical and Control Systems. **5** (3), 397–429 (1999). <https://doi.org/10.1023/A:1021770717822>
- [5] Dieci, L.: *Considerations on computing real logarithms of matrices, Hamiltonian logarithms, and skew-symmetric logarithms*. Linear Algebra and its Applications. **244**, 35-54 (1996). [https://doi.org/10.1016/0024-3795\(94\)00206-1](https://doi.org/10.1016/0024-3795(94)00206-1)
- [6] Harris, W. F., Cardoso, J. R.: *The exponential-mean-log-transference as a possible representation of the optical character of an average eye..* Ophthalmic Physiol Opt. **26** (4), 380-383 (2006). <https://doi.org/10.1111/j.1475-1313.2006.00365.x>
- [7] Higham, N. J.: Functions of Matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, (2008).
- [8] Higham, N. J., Mackey, D. S., Mackey, N., Tisseur, F.: *Computing the polar decomposition and the matrix sign decomposition in matrix groups*. SIAM Journal on Matrix Analysis and Applications. **25**(4), 1178-1192 (2004). <https://doi.org/10.1137/S0895479803426644>
- [9] Higham, N. J., Mackey, D. S., Mackey, N., Tisseur, F.: *Functions preserving matrix groups and iterations for the matrix square root*. SIAM Journal on Matrix Analysis and Applications. **26**(3), 849-877 (2005). <https://doi.org/10.1137/S0895479804442218>
- [10] Jagger, D. P.: M.Sc. Thesis: MATLAB toolbox for classical matrix groups. University of Manchester, Manchester, England, (2003).
- [11] Kenney, C. S., Laub, A. J.: *Padé error estimates for the logarithm of a matrix*. International Journal of Control. **50** (3), 707-730 (1989). <https://doi.org/10.1080/00207178908953392>
- [12] Kenney, C. S., Laub, A. J.: *Condition estimates for matrix functions*. SIAM Journal on Matrix Analysis and Applications. **10** (2), 191-209 (1989). <https://doi.org/10.1137/0610014>
- [13] Mackey, D. S., Mackey, N., Tisseur, F. : *Structured factorizations in scalar product spaces*. SIAM Journal on Matrix Analysis and Applications. **27**, 821-850 (2006). <https://doi.org/10.1137/040619363>
- [14] Morgan, K.: *Linear differential equations with periodic coefficients*. International Journal for Numerical Methods in Engineering. **10** (4), 976–976 (1976). <https://doi.org/10.1002/nme.1620100430>
- [15] Park, F. C., Ravani, B.: *Bézier Curves on Riemannian manifolds and Lie groups with kinematic applications*. ASME, Journal of Mechanical Design . **117**, 36-40 (1995). <https://doi.org/10.1115/1.2826114>
- [16] Sibuya, Y.: *Note on real matrices and linear dynamical systems with periodic coefficients*. Journal of Mathematical Analysis and Applications. **1** (3), 363 - 372 (1960). [http://doi.org/10.1016/0022-247X\(60\)90010-X](http://doi.org/10.1016/0022-247X(60)90010-X)
- [17] Sinha, N. K., Lastman, G. J.: *Transformation algorithm for identification of continuous-time multivariable systems from discrete data*. Electronics Letters. **17** (21), 779-780 (1981). <http://doi.org/10.1049/el:19810546>
- [18] Sinha, N. K., Lastman, G. J.: *Infinite series for logarithm of matrix, applied to identification of linear continuous-time multivariable systems from discrete-time models*. Electronics Letters. **27** (16), 1468-1470 (1991). <http://doi.org/10.1049/el:19910919>

Affiliations

BAHAR ARSLAN

ADDRESS: Bursa Technical University, Mathematics Department, Bursa, Turkey.

E-MAIL: bahar.arslan@btu.edu.tr

ORCID ID:0000-0002-7750-4325