KOLMOGOROFF-NAGUMO MEAN OVER THE AFFINE SYMPLECTIC GROUP OF MATRICES

SIMONE FIORI*

Abstract. The present work shows that Harris' exponential-mean-log averaging rule over the space of optical transference matrices may be regarded as an instance of the Kolmogoroff-Nagumo averaging rule over the affine symplectic group. As such, Harris' averaging rule may be generalized to a φ-mean-φ^{-1} rule that can be implemented by different φ maps. The present work also shows that the involved maps may be computed in closed form by low-degree polynomial expressions.

Keywords: Kolmogoroff–Nagumo mean; Harris exponential-mean-log; Affine Symplectic group; Hamiltonian matrices.

1. Introduction. The notion of *mean value* of a set of *structured* samples is not univocally defined. Even in the case of two positive real-valued numbers, there exists a variety of methods to define their mean value, each of which leads to a different numerical result. The best known computational rules to average out two positive real-valued numbers are the arithmetic mean, the harmonic mean and the geometric mean. A generalization of such means is the Heinz mean, which turns out to be a combination of the arithmetic and the geometric mean.

The problem of generalizing the notion of mean value over multidimensional data was given an elegant formulation by Fréchet [12] (see, e.g., [8] for a recent review) and by Karcher [18], in terms of minimal Riemannian dispersion. Karcher's empirical mean of a finite set of samples, in particular, is defined as any minimizer of a criterion function given by the sum of squared Riemannian distances between each sample and the sought mean. As such, its uniqueness is not guaranteed, unless the sample-space and the samples themselves obey some geometric restriction. The notion of mean value of structured samples was subjected to extensive investigation and a number of generalizations were devised. For instance, it is known that the arithmetic, harmonic and geometric mean may be given a unifying view by formulating the definition of mean value as an optimization problem based on the Bregman divergence (for a recent review, see, e.g., [7]). Moreover, it is known that most averaging formulas lead to instances of the *Chisini mean*[1] as defined in [4] (see, e.g., [20] for a detailed discussion on its properties). The Chisini mean provides a good illustration of the fact that it is possible to define the mean value of a set of numbers without any particular requirement about, e.g., convexity. In addition, whenever the sample spaces exhibit the structure of a Lie group or of a smooth manifold, *not necessarily metrizable*, the notion of arithmetic mean may be extended to those spaces by formulating the average over the Lie algebra associated with the sample Lie group [11] or over the tangent bundle associated with the sample-manifold [17]. All such generalizations of empirical mean value give rise to iterative algorithms that converge to an approximation of the actual center of mass of a given set of samples.

A special instance of Chisini mean that gives rise to a closed-form expression of the empirical center of mass of a given set of samples is the *Kolmogoroff-Nagumo mean* [20]. The present work discusses the application of the Kolmogoroff-Nagumo mean theory to a computational optics problem.

In computational optics, interest lays upon computing the mean value out of a set of *optical*

*Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Via Brecce Bianche, Ancona I-60131, Italy `E-mail: s.fiori@univpm.it`

The present paper was published as **S. Fiori, *Kolmogoroff-Nagumo mean over the affine symplectic group of matrices*, Applied Mathematics and Computation, Vol. 266, pp. 820 – 837, September 2015**.

[1]Oscar Chisini (March 4, 1889 – April 10, 1967) was an Italian mathematician. In 1929, Chisini founded the Institute of Mathematics at the University of Milan and held the position of chairman of the Institute from the early 1930s until 1959.

transference matrices [14]. Descriptions of optical systems based on transference matrices are quite general, as they allow to describe coaxial spherical systems, astigmatic surfaces, non-coaxial optical systems and systems containing prisms and decentered lenses, as well as compound optical systems [13]. A compound optical system, as for instance, an eyeball, may be described, via a composition law, through the transference matrices of its subsystems [19].

In the contributions [5, 14], van Gool and Harris proposed an *exponential-mean-log* rule to define the notion of average optical system. Such non-iterative rule was recently compared with iterative rules based on Riemannian optimization of appropriate criterion functions [9, 10] and was numerically proven to provide optimal results.

The present work aims at showing that Harris' exponential-mean-log rule appears as an instance of the Kolmogoroff-Nagumo averaging rule over a structured space known as *affine symplectic group*. As such, the exponential-mean-log may be generalized to a φ*-mean-*φ^{-1} rule that can be implemented by different φ maps. In addition, the present contribution shows that, since the space of interest in computational optics is a low-dimensional instance of the general affine symplectic group, the involved maps may be computed in closed form, *exactly* (up to machine precision), by low-degree polynomial expressions.

2. Averaging over the optical transference group.

The present Section recalls the fundamental definitions regarding the group of optical transference matrices (which is built on the affine symplectic group) and sets out the notion of Kolmogoroff-Nagumo mean on the optical transference group.

2.1. Lie groups, Lie algebras and their local diffeomorphisms.

A smooth matrix manifold G is a continuous set of structured matrices. The tangent space of the manifold G at a point $X \in \mathrm{G}$ is denoted by $T_X\mathrm{G}$. A tangent space $T_X\mathrm{G}$ is a vector space spanned by vectors tangent to each possible smooth curve passing through the point X. Namely, denoting by $\Gamma_X : [-a,\ a] \to \mathrm{G}$, with $a > 0$, any smooth curve $\Gamma_X(t)$ such that $\Gamma_X(0) = X$, the tangent space $T_X\mathrm{G}$ is spanned by vectors $\dot{\Gamma}_X(0)$ for all possible Γ_X's (where the over-dot denotes differentiation with respect to the parameter t).

An algebraic matrix group structure $(\mathrm{G}, \mu, \omega, I)$ is made of a set G endowed with a multiplication operation μ, an inverse operation ω and an identity element I, such that for any $X, Y \in \mathrm{G}$, it holds that $\mu(X, Y) \in \mathrm{G}$, $\mu(X, \omega(X)) = \mu(\omega(X), X) = I$, and $\mu(X, I) = \mu(I, X) = X$. Group identity and inverse need to be unique and the group multiplication needs to be associative, namely, $\mu(X, \mu(Y, Z)) = \mu(\mu(X, Y), Z)$ for every $X, Y, Z \in \mathrm{G}$. In general, group multiplication is not commutative, though.

A *Lie group* is an *algebraic group* that also possesses the structure of a *smooth manifold*. Associated with the Lie group G is a *Lie algebra* $T_I\mathrm{G}$, namely, the tangent space at identity. Any local diffeomorphism $\varphi : T_I\mathrm{G} \to \mathrm{G}$ maps an element of the Lie algebra to an element of the Lie group (where a local diffeomorphism denotes a smooth map which is smoothly invertible at least locally). The inverse of a local diffeomorphism φ is denoted by $\varphi^{-1} : \mathrm{G} \to T_I\mathrm{G}$. Every matrix Lie group admits at least a local diffeomorphism which is termed *group exponential*, whose inverse is termed *logarithm*.

A good reference on general differential geometry is [22], while a more recent reference on Lie groups with applications is [3].

2.2. Real symplectic group, real affine symplectic group, optical transference group and their Lie algebras.

The *real symplectic group* is defined as:

$$\mathrm{Sp}(2p, \mathbb{R}) \stackrel{\text{def}}{=} \{S \in \mathbb{R}^{2p\times 2p} | S^T Q S = Q\},\ \text{with}\ Q \stackrel{\text{def}}{=} \begin{bmatrix} 0_p & I_p \\ -I_p & 0_p \end{bmatrix}, \tag{2.1}$$

where the symbol I_p denotes the $p \times p$ identity matrix, the symbol 0_p denotes a whole-zero $p \times p$ matrix, the symbol ^{T} denotes matrix transpose and the matrix Q is a fundamental skew-symmetric matrix. The fundamental skew-symmetric matrix is such that $Q^2 = -I_{2p}$, $Q^{-1} = Q^T = -Q$.

The space $\mathrm{Sp}(2p, \mathbb{R})$ may be regarded as a smooth manifold and as an algebraic group endowed with multiplication and inversion operations, which possesses an identity element. In particular:

- Group multiplication coincides with matrix multiplication: For every $S_1, S_2 \in \mathrm{Sp}(2p, \mathbb{R})$, the product $S_1 S_2$ belongs to $\mathrm{Sp}(2p, \mathbb{R})$.
- Group identity coincides with the identity matrix: On a real symplectic group $\mathrm{Sp}(2p, \mathbb{R})$, the identity with respect to matrix multiplication is the identity matrix I_{2p}.
- Group inversion coincides with matrix inversion: Any symplectic matrix $S \in \mathrm{Sp}(2p, \mathbb{R})$ is such that $\det^2(S) = 1$, where the operator $\det(\cdot)$ denotes determinant. An implication of such property is that any symplectic matrix is invertible. In addition, it holds that that if $S \in \mathrm{Sp}(2p, \mathbb{R})$, then also S^{-1} belongs to $\mathrm{Sp}(2p, \mathbb{R})$.

Therefore, the space $\mathrm{Sp}(2p, \mathbb{R})$ has the structure of a Lie group.

The Lie algebra $\mathrm{Ham}(2p, \mathbb{R})$ associated with the Lie group $\mathrm{Sp}(2p, \mathbb{R})$, has the structure:

$$\mathrm{Ham}(2p, \mathbb{R}) = \{H \in \mathbb{R}^{2p \times 2p} | H^T Q + QH = 0_{2p}\}, \tag{2.2}$$

namely, it coincides with the space of $2p \times 2p$ Hamiltonian matrices.

Since the late sixties, there has been a burst of applications of symplectic techniques to mathematics and physics, and even to engineering of medical sciences (magnetic resonance imaging is a typical example) [6].

Applications to computational optics involve a Lie group that is built on the symplectic group and that is termed *real affine symplectic group*, defined as:

$$\mathrm{ASp}(2p, \mathbb{R}) \stackrel{\mathrm{def}}{=} \left\{ \left[\begin{array}{cc} S & \delta \\ 0 & 1 \end{array} \right] \middle| \, S \in \mathrm{Sp}(2p, \mathbb{R}),\ \delta \in \mathbb{R}^{2p} \right\}. \tag{2.3}$$

The space $\mathrm{ASp}(2p, \mathbb{R})$ may be endowed with the structure of an algebraic group by matrix multiplication and inversion. In fact, if $T_1, T_2 \in \mathrm{ASp}(2p, \mathbb{R})$, it holds that:

$$\mu(T_1, T_2) \stackrel{\mathrm{def}}{=} T_1 T_2 = \left[\begin{array}{cc} S_1 S_2 & S_1 \delta_2 + \delta_1 \\ 0 & 1 \end{array} \right], \tag{2.4}$$

therefore $T_1 T_2 \in \mathrm{ASp}(2p, \mathbb{R})$. The group identity with respect to matrix multiplication in $\mathrm{ASp}(2p, \mathbb{R})$ is given by the identity matrix I_{2p+1}. Given the general form (2.3) of an optical transference matrix $T \in \mathrm{ASp}(2p, \mathbb{R})$, it is readily seen that $\det^2(T) = \det^2(S)$, therefore, any transference matrix is invertible. Moreover, if $T \in \mathrm{ASp}(2p, \mathbb{R})$, its inverse reads:

$$\omega(T) \stackrel{\mathrm{def}}{=} T^{-1} = \left[\begin{array}{cc} S^{-1} & -S^{-1}\delta \\ 0 & 1 \end{array} \right], \tag{2.5}$$

therefore $T^{-1} \in \mathrm{ASp}(2p, \mathbb{R})$. The space $\mathrm{ASp}(2p, \mathbb{R})$ is a subgroup of the general linear group $\mathrm{Gl}(2p + 1, \mathbb{R})$.

The Lie algebra associated with the affine symplectic group reads:

$$\mathrm{EHam}(2p, \mathbb{R}) \stackrel{\mathrm{def}}{=} \left\{ \left[\begin{array}{cc} H & v \\ 0 & 0 \end{array} \right] \middle| \, H \in \mathrm{Ham}(2p, \mathbb{R}),\ v \in \mathbb{R}^{2p} \right\}, \tag{2.6}$$

and is termed *extended Hamiltonian algebra*.

Optical transference matrices provide a way of tracing a ray through a compound optical system [19]. The group of optical transference matrices coincides with the affine symplectic group $\mathrm{Sp}(4, \mathbb{R})$

and will be hereafter denoted as Opt for short. In the language of computational ophthalmology, the real symplectic sumbatrix S is termed *dioptric matrix* and represents the optical power of the system, while the vector δ accounts for the amount of decentration of the optical system (namely, the amount of displacement, horizontal and/or vertical, of the centration point of the optical system from the standard optical center position).

The Lie algebra associated with the Lie group Opt will be termed *log-optical algebra* and denoted as LOpt for short, which coincides with the extended Hamiltonian algebra $\mathrm{EHam}(4,\mathbb{R})$.

Hereafter, for the sake of notation conciseness, the identity matrices are denoted by the symbol I and the zero matrices/vectors are denoted by symbol 0, wherever such simplified notation does not give rise to any confusion.

2.3. Harris 'exponential-mean-log' and Kolmogoroff-Nagumo mean. In the paper [14], Harris proposed an 'exponential-mean-log' scheme to compute the average $\bar{T}_{\mathrm{H}} \in \mathrm{Opt}$ of a set of N optical-system transference matrices $T_n \in \mathrm{Opt}$, with $n = 1, \ldots, N$.

In order to express Harris' averaging rule, it is necessary to define the notion of *matrix exponential* and *matrix logarithm*. Given a square matrix $X \in \mathbb{R}^{m\times m}$, its *matrix exponential* is defined as:

$$\exp(X) \stackrel{\mathrm{def}}{=} I + \sum_{k=1}^{\infty} \frac{X^k}{k!}. \tag{2.7}$$

The matrix exponential function is, hence, everywhere analytic. The *matrix logarithm* [23] of a square matrix $X \in \mathbb{R}^{m\times m}$ is defined as:

$$\log(X) \stackrel{\mathrm{def}}{=} -\sum_{k=1}^{\infty} \frac{(I-X)^k}{k}. \tag{2.8}$$

The matrix logarithm is, therefore, analytic for $\|X - I\| < 1$. The matrix logarithm is the inverse of the matrix exponential.

On the basis of the above definitions, Harris' averaging scheme may be expressed as:

$$\bar{T}_{\mathrm{H}} \stackrel{\mathrm{def}}{=} \exp\left(\frac{1}{N}\sum_{n=1}^{N} \log T_n\right). \tag{2.9}$$

The above averaging scheme is consistent, namely, it returns an optical transference matrix as result. Its consistency was proven directly in [15] by calculating explicitly the exponential and the logarithm. In the present context, it may be observed that the consistency is guaranteed by the fact that the spaces of interest are a Lie group and its associated Lie algebra and that the functions exp / log are local diffeomorphisms between such spaces.

Harris' mean value may be regarded as a special case of a *Kolmogoroff-Nagumo* (or *quasi-arithmetic*) mean of a set of affine-symplectic-group samples. The Kolmogoroff-Nagumo mean value may be expressed through the φ-mean-φ^{-1} rule:

$$\bar{T}_{\mathrm{KN}} \stackrel{\mathrm{def}}{=} \varphi\left(\frac{1}{N}\sum_{n=1}^{N} \varphi^{-1}(T_n)\right), \tag{2.10}$$

for a continuous strictly monotonic function φ (for a general review of the Kolmogoroff-Nagumo mean, see, e.g., [20]). In the present context, it must hold $\varphi : \mathrm{LOpt} \to \mathrm{Opt}$ and, consequently, $\varphi^{-1} : \mathrm{Opt} \to \mathrm{LOpt}$, at least locally. The rule (2.10) clearly generalizes the exponential-mean-log rule (2.9) that arises by setting $\varphi(T) = \exp(T)$. A possible alternative choice for the function φ is the *Cayley transform*. Given a square matrix $X \in \mathbb{R}^{m\times m}$, its Cayley transform is defined as:

$$\mathrm{cay}(X) \stackrel{\mathrm{def}}{=} (I + \rho X)(I - \rho X)^{-1}, \tag{2.11}$$

for every $\rho \neq 0$, which is well-defined as long as $\det(I - \rho X) \neq 0$. In the present contribution, by Cayley transform it is meant the definition (2.11) with $\rho = 1$. For a recent account on the numerical properties of the Cayley transform, see [16]. A noticeable property of the Cayley transform is that it coincides with its inverse, namely, $\mathrm{cay}(\mathrm{cay}(X)) = X$. The inverse matrix function in the expression (2.11) is locally analytic, in fact, it holds that:

$$(I - X)^{-1} = I + \sum_{k=1}^{\infty} X^k, \tag{2.12}$$

for $\|X\| < 1$. As a consequence, the Cayley transform is locally analytic. Setting $\varphi(T) \stackrel{\mathrm{def}}{=} \mathrm{cay}(-T)$ in the definition (2.10), with $T \in \mathrm{Opt}$, yields an alternative to the exponential-mean-log averaging rule, which may be termed *Cayley-mean-Cayley* averaging rule. Such alternative to Harris' computational scheme (2.9) reads:

$$\bar{T}_{\mathrm{C}} \stackrel{\mathrm{def}}{=} \mathrm{cay}\left(-\frac{1}{N}\sum_{n=1}^{N} \mathrm{cay}(-T_n)\right). \tag{2.13}$$

As a further advancement with respect to Harris' work, on the basis of Lie-group theory, it pays to define the *Lie-group variance* of the set of optical transference matrices $\{T_n\}_{n=1}^N$ as:

$$\sigma^2 \stackrel{\mathrm{def}}{=} \frac{1}{N}\sum_{n=1}^{N} \|\log(\bar{T}_{\mathrm{KN}}^{-1} T_n)\|_{\mathrm{F}}^2, \tag{2.14}$$

where the symbol $\|\cdot\|_{\mathrm{F}}$ denotes a matrix Frobenius norm. The Lie-group empirical variance measures, in a manner that is fully compatible with the chosen metrization of the Lie group of the optical-transference matrices, the spread of a sample set around its empirical mean matrix.

For the sake of consistency, note that both the scalar-to-scalar functions φ used to define Kolmogoroff-Nagumo means, namely, $\varphi(x) = \exp(x)$ and $\varphi(x) = \frac{1+x}{1-x}$, are strictly monotonic as $\varphi'(x) > 0$.

3. Matrix functions on the real symplectic Lie group and its Lie algebra. A popular way to compute the value of an analytic function $f : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ of a squared matrix $X \in \mathbb{R}^{m\times m}$ consists in computing its Jordan canonical form $X = B\Delta B^{-1}$, where the matrix Δ contains (non-diagonal, in general) Jordan blocks and B denotes a similarity matrix. The matrix Δ shows, along its main diagonal, the eigenvalues of the matrix X. (The paper [2] gives a physical interpretation, in terms of elementary optical subsystems, of the Jordan blocks of a symplectic matrix.) The numerical advantage of a Jordan canonical representation is that, for k integer:

$$(B\Delta B^{-1})^k = B\Delta^k B^{-1}, \tag{3.1}$$

hence, it follows that the function $f(X)$ simplifies to:

$$f(B\Delta B^{-1}) = Bf(\Delta)B^{-1}, \tag{3.2}$$

where $f(\Delta)$ is easier to evaluate than $f(X)$. The disadvantage of such procedure is that it requires the computation of both the Jordan-blocks-matrix Δ *and* of the similarity matrix B, which may be assimilated to the problem of computing both the (non-distinct) eigenvalues and the eigenvectors of a linear operator.

The key idea of the present contribution is that *functions of finite-dimensional matrices may be evaluated through exact polynomial formulas of finite order*, where *such computation only requires the evaluation of the eigenvalues of the matrix in argument.* In the present context, interest lays upon the matrix exponential (exp), the matrix logarithm (log) and the Cayley transform (cay) applied to 4×4 real-valued matrices, which may be expressed exactly as third-order matrix polynomials.

Let X denote again a $m \times m$ real-valued matrix. The *secular function* (or *characteristic polynomial*) associated with the matrix X is $P_X(z) \stackrel{\text{def}}{=} \det(zI - X)$, where $z \in \mathbb{C}$ denotes a scalar complex-valued variable. The function $P_X(z)$ is a monic polynomial of degree m:

$$P_X(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0, \tag{3.3}$$

whose roots, namely, the solutions of the characteristic equation $P_X(z) = 0$, are the eigenvalues of the matrix X. The *Cayley-Hamilton theorem* states that each matrix satisfies its own characteristic equation, namely, that:

$$P_X(X) \stackrel{\text{def}}{=} X^m + a_{m-1}X^{m-1} + \cdots + a_1 X + a_0 I = 0. \tag{3.4}$$

By definition, any analytic function $f : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ may be expanded as a polynomial (either of finite or infinite degree). Moreover, to any matrix-to-matrix function $f : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$, it may be associated a scalar-to-scalar function by replacing the matrix argument with a scalar argument. Hence, the function $f(X)$ may be thought of as a polynomial $f(z)$ in the variable $z \in \mathbb{C}$. The polynomial $f(z)$ may be written as:

$$f(z) = Q(z)P_X(z) + R(z), \tag{3.5}$$

where $Q(z)$ is found by a *long division* and the remainder polynomial $R(z)$ is of degree $m - 1$ or less, namely:

$$R(z) = b_{m-1}z^{m-1} + b_{m-2}z^{m-2} + \cdots + b_2 z^2 + b_1 z + b_0. \tag{3.6}$$

The corresponding relationship in the matrix variable X reads:

$$f(X) = Q(X)P_X(X) + R(X). \tag{3.7}$$

Since $P_X(X) = 0$, it holds that $f(X) = R(X)$. Therefore, the matrix function $f(X)$ may be evaluated, *exactly*, as a polynomial of degree $m - 1$ or less, namely:

$$f(X) = b_{m-1}X^{m-1} + \cdots + b_2 X^2 + b_1 X + b_0 I. \tag{3.8}$$

It is worth recalling that, given a 4×4 real-valued matrix X, the coefficients of its characteristic polynomial (3.3) may be computed through the Newton's identities:

$$\begin{cases} a_3 = -\mathrm{tr}(X), \\ a_2 = \frac{1}{2}(\mathrm{tr}^2(X) - \mathrm{tr}(X^2)), \\ a_1 = -\frac{1}{6}(\mathrm{tr}^3(X) - 3\mathrm{tr}(X^2)\mathrm{tr}(X) + 2\mathrm{tr}(X^3)), \\ a_0 = \det(X). \end{cases} \tag{3.9}$$

In the present Section, the following matrix functions are evaluated:

- Exponential (exp) of a 4×4 Hamiltonian matrix.
- Logarithm (log) of a 4×4 symplectic matrix.
- Cayley transform (cay) of a 4×4 Hamiltonian matrix.
- Cayley transform (cay) of a 4×4 symplectic matrix.

Such cases are covered separately in the following subsections.

3.1. Exponential of a fourth-order Hamiltonian matrix. An account of numerically convenient and accurate methods to compute exponential of general-dimension Hamiltonian matrices is given in [1], while an account of the closely-related problem of the explicit computation of the exponential of a 4×4 skew-symmetric matrix was given in [21]. Let $H \in \mathrm{Ham}(4, \mathbb{R})$ denote a 4×4 Hamiltonian matrix. According to the Cayley-Hamilton theorem, every analytic function of H may be computed as a polynomial of degree 3 or less. In particular, the matrix exponential $f(H) = \exp(H)$ may be evaluated as a polynomial:

$$\exp(H) = b_3 H^3 + b_2 H^2 + b_1 H + b_0 I. \tag{3.10}$$

Note that the coefficients b_i's depend on the eigenvalues of the matrix H, hence, each $b_i = b_i(H)$.

Any Hamiltonian matrix H may be written as:

$$H = \begin{bmatrix} A & B \\ C & -A^T \end{bmatrix}, \; B^T = B, \; C^T = C, \; A \text{ arbitrary}. \tag{3.11}$$

Therefore, it holds that:

$$\begin{aligned} \mathrm{tr}(H) &= \mathrm{tr}(A - A^T) = 0, \\ \mathrm{tr}(H^3) &= 3\mathrm{tr}(ABC - A^T CB) = 0. \end{aligned}$$

The coefficients of the secular function $P_H(z)$, expressed through the Newton's identities (3.9), hence, simplify to:

$$\begin{cases} a_3 = 0, \\ a_2 = -\frac{1}{2}\mathrm{tr}(H^2), \\ a_1 = 0, \\ a_0 = \det(H). \end{cases} \tag{3.12}$$

and the characteristic polynomial of the matrix H reads $p_H(z) = z^4 + a_2 z^2 + a_0$. The four eigenvalues of the matrix H take the closed-form expression:

$$\pm \frac{1}{2} \sqrt{\mathrm{tr}(H^2) \pm \sqrt{\mathrm{tr}^2(H^2) - 16 \det(H)}}. \tag{3.13}$$

The four eigenvalues of an Hamiltonian matrix come in complex-valued antipodal pairs, namely $\lambda_1, -\lambda_1, \lambda_2, -\lambda_2 \in \mathbb{C}$. In order to determine the coefficients b_i's of the polynomial expansion (3.10), it is necessary to distinguish between the cases that the complex-valued numbers λ_1 and λ_2 are distinct or coincident.

3.1.1. Exponential of an Hamiltonian matrix: Eigenvalues of multiplicity 1. The only such a case is $(\lambda_1, -\lambda_1, \lambda_2, -\lambda_2)$ with $\lambda_1, \lambda_2 \in \mathbb{C} - \{0\}$ and $\lambda_1^2 \neq \lambda_2^2$. In order to determine the coefficients b_i of the expansion (3.10), recall that the equation (3.5), evaluated by replacing the variable z with each eigenvalue of the matrix H, reads $\exp(z) = b_3 z^3 + b_2 z^2 + b_1 z + b_0$. Since the eigenvalues are, by hypothesis, all distinct, such equation, written for each eigenvalue separately, gives rise to the consistent linear system:

$$\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 \\ 1 & -\lambda_1 & \lambda_1^2 & -\lambda_1^3 \\ 1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 \\ 1 & -\lambda_2 & \lambda_2^2 & -\lambda_2^3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \exp \lambda_1 \\ \exp(-\lambda_1) \\ \exp \lambda_2 \\ \exp(-\lambda_2) \end{bmatrix}. \tag{3.14}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = \frac{\lambda_2^2 \cosh(\lambda_1) - \lambda_1^2 \cosh(\lambda_2)}{\lambda_2^2 - \lambda_1^2}, \\ b_1 = \frac{\lambda_2^3 \sinh(\lambda_1) - \lambda_1^3 \sinh(\lambda_2)}{\lambda_1 \lambda_2 (\lambda_2^2 - \lambda_1^2)}, \\ b_2 = \frac{\cosh(\lambda_1) - \cosh(\lambda_2)}{\lambda_2^2 - \lambda_1^2}, \\ b_3 = \frac{\lambda_1 \sinh(\lambda_2) - \lambda_2 \sinh(\lambda_1)}{\lambda_1 \lambda_2 (\lambda_2^2 - \lambda_1^2)}. \end{cases} \tag{3.15}$$

The solution coefficients b_i would clearly become undefined if either $\lambda_1 = 0$ or $\lambda_2 = 0$ or if $\lambda_1 = \lambda_2$ or $\lambda_1 = -\lambda_2$.

3.1.2. Exponential of an Hamiltonian matrix: Eigenvalues of multiplicity 2. For eigenvalues of multiplicity 2, $P_H(z) = 0$ as well as $\dot{P}_H(z) = 0$ when z is set to each of the eigenvalues[2]. Deriving term by term the relationship (3.5), gives:

$$\dot{f}(z) = \dot{P}_H(z)Q(z) + P_H(z)\dot{Q}(z) + \dot{R}(z). \tag{3.16}$$

When z is set to each of the eigenvalues of the matrix H, the first two terms on the right-hand side vanish to zero, therefore it holds that $\dot{f}(z) = \dot{R}(z)$.

Exponential of an Hamiltonian matrix: Case $(\lambda, -\lambda, \lambda, -\lambda)$ **with** $\lambda \in \mathbb{C} - \{0\}$**.** Denoting by λ the common value of λ_1 and λ_2, the combination of the conditions $\exp(z) = b_3 z^3 + b_2 z^2 + b_1 z + b_0$ and of the condition on first order derivatives, namely $\exp(z) = 3b_3 z^2 + 2b_2 z + b_1$, gives the linear system:

$$\begin{bmatrix} 1 & \lambda & \lambda^2 & \lambda^3 \\ 1 & -\lambda & \lambda^2 & -\lambda^3 \\ 0 & 1 & 2\lambda & 3\lambda^2 \\ 0 & 1 & -2\lambda & 3\lambda^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \exp\lambda \\ \exp(-\lambda) \\ \exp\lambda \\ \exp(-\lambda) \end{bmatrix}. \tag{3.17}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = \cosh\lambda - \frac{\lambda}{2}\sinh\lambda, \\ b_1 = \frac{3}{2}\frac{\sinh\lambda}{\lambda} - \frac{1}{2}\cosh\lambda, \\ b_2 = \frac{1}{2\lambda}\sinh\lambda, \\ b_3 = \frac{1}{2\lambda^2}\left(\cosh\lambda - \frac{\sinh\lambda}{\lambda}\right). \end{cases} \tag{3.18}$$

The solution coefficients would clearly become undefined if $\lambda = 0$.

Exponential of an Hamiltonian matrix: Case $(\lambda, -\lambda, 0, 0)$ **with** $\lambda \in \mathbb{C} - \{0\}$**.** The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & \lambda & \lambda^2 & \lambda^3 \\ 1 & -\lambda & \lambda^2 & -\lambda^3 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \exp\lambda \\ \exp(-\lambda) \\ 1 \\ 1 \end{bmatrix}. \tag{3.19}$$

[2] Note that the function $f(z)$ is analytic in a domain containing the eigenvalues, hence, its derivatives of any order are well-defined.

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = 1, \\ b_1 = 1, \\ b_2 = \frac{\cosh\lambda - 1}{\lambda^2}, \\ b_3 = \frac{\sinh\lambda - \lambda}{\lambda^3}. \end{cases} \quad (3.20)$$

The problem would clearly become ill-posed if $\lambda = 0$.

3.1.3. Exponential of an Hamiltonian matrix: Eigenvalues of multiplicity 4. The only such a case is $(0, 0, 0, 0)$, which is a trivial case that leads to the solution $\exp(H) = I$.

3.2. Logarithm of a fourth-order symplectic matrix. Let $S \in \mathrm{Sp}(4, \mathbb{R})$ denote a 4×4 symplectic matrix. According to the Cayley-Hamilton theorem, every analytic function of S may be computed as a polynomial of degree 3 or less. In particular, the matrix logarithm $f(S) = \log(S)$ may be evaluated as a polynomial:

$$\log(S) = b_3 S^3 + b_2 S^2 + b_1 S + b_0 I. \quad (3.21)$$

The coefficients a_i's of the characteristic polynomial $P_S(z)$ of a real-valued symplectic matrix S are real-valued, therefore, if λ is a root, also $\bar{\lambda}$ is a root (the over-bar denotes complex conjugation). Moreover, if λ is a root, also $\frac{1}{\lambda}$ is a root. In addition, it holds that $a_0 = \det(S) = 1$. Therefore, the polynomial $P_S(z)$ is *monic palindromic*:

$$P_S(z) = z^4 + a_1 z^3 + a_2 z^2 + a_1 z + 1, \quad (3.22)$$

and its roots are computed as:

$$\frac{\zeta \pm \sqrt{\zeta^2 - 4}}{2}, \text{ where } \zeta = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2 + 8}}{2}. \quad (3.23)$$

For a quick account on the eigenvalue structure of a real symplectic matrix, see [2]. When all eigenvalues are of multiplicity 1, the following cases may occur:

1. Case of complex-valued quartet: There are four distinct complex-valued roots $\lambda \in \mathbb{C}$, $\bar{\lambda}$, λ^{-1}, $\bar{\lambda}^{-1}$.
2. Case of a real-valued duet and a unimodular duet: There are two distinct real-valued roots $\lambda_1 \in \mathbb{R}$, λ_1^{-1} and two distinct unimodular roots $\lambda_2 \in \mathbb{C}$, $\bar{\lambda}_2$ with $|\lambda_2| = 1$ (the symbol $|\cdot|$ denotes the modulus of a complex-valued number).
3. Case of two real-valued duets: There are four distinct real-valued roots $\lambda_1 \in \mathbb{R}$, λ_1^{-1}, $\lambda_2 \in \mathbb{R}$, λ_2^{-1}.
4. Case of two unimodular duets: There are four distinct real-valued roots $\lambda_1 \in \mathbb{C}$, $\bar{\lambda}_1$ with $|\lambda_1| = 1$ and $\lambda_2 \in \mathbb{C}$, $\bar{\lambda}_2$ with $|\lambda_2| = 1$.

The eigenvalues may come with multiplicity 2 and with multiplicity 4 as well. All the cases are covered in the following Subsections.

3.2.1. Logarithm of a symplectic matrix: Eigenvalues of multiplicity 1. There are four cases of quadruples of eigenvalues with multiplicity 1. In order to determine the coefficients b_i's of the expansion (3.21), recall that the equation (3.5), evaluated by replacing the variable z with each eigenvalue of the matrix S, reads $\log(z) = b_3 z^3 + b_2 z^2 + b_1 z + b_0$.

Logarithm of a symplectic matrix: Case $(\lambda, \bar{\lambda}, \lambda^{-1}, \bar{\lambda}^{-1})$, with $\lambda \in \mathbb{C} - \{0\}$. In order for the four eigenvalues to be distinct from each other, it must hold $|\lambda| \neq 1$ and $\mathrm{Im}\{\lambda\} \neq 0$ (where $\mathrm{Im}\{\cdot\}$ denotes imaginary part). The equation to determine the coefficients b_i's gives rise to the consistent linear system:

$$\begin{bmatrix} 1 & \lambda & \lambda^2 & \lambda^3 \\ 1 & \bar{\lambda} & \bar{\lambda}^2 & \bar{\lambda}^3 \\ 1 & \lambda^{-1} & \lambda^{-2} & \lambda^{-3} \\ 1 & \bar{\lambda}^{-1} & \bar{\lambda}^{-2} & \bar{\lambda}^{-3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \log \lambda \\ \log \bar{\lambda} \\ -\log \lambda \\ -\log \bar{\lambda} \end{bmatrix}. \tag{3.24}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = 2\Re\left\{\frac{(\lambda^4+1)\bar{\lambda}}{(\lambda^2-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)} \log \lambda\right\}, \\ b_1 = 2\Re\left\{\frac{(\bar{\lambda}^4+1)(\lambda^2+1)+|\lambda|^2(\bar{\lambda}^2+1)}{(\bar{\lambda}^2-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)} \log \bar{\lambda}\right\}, \\ b_2 = 2\Re\left\{\frac{(\lambda^3+\bar{\lambda})(|\lambda|^2+1)+\lambda(|\lambda|^4+1)}{(\lambda^2-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)} \log \lambda\right\}, \\ b_3 = 2|\lambda|^2\Re\left\{\frac{\bar{\lambda}^2+1}{(\bar{\lambda}^2-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)} \log \bar{\lambda}\right\}. \end{cases} \tag{3.25}$$

The problem would clearly be ill-posed if the quantity λ would be unimodular or real-valued (including 0).

Logarithm of a symplectic matrix: Case of two distinct unimodular duets. It is convenient to write the eigenvalues quadruple as $(e^{\iota\theta}, e^{-\iota\theta}, e^{\iota\psi}, e^{-\iota\psi})$, with $\theta, \psi \in (0, 2\pi)$, $\theta \neq \psi$ and $\theta + \psi \neq 2\pi$ (where the symbol ι denotes the *imaginary unit*, namely, $\iota^2 = -1$). Since the eigenvalues are, by hypothesis, all distinct, the equation $b_3 z^3 + b_2 z^2 + b_1 z + b_0 = \log(z)$, written for each eigenvalue separately, gives rise to the consistent linear system:

$$\begin{bmatrix} 1 & e^{\iota\theta} & e^{2\iota\theta} & e^{3\iota\theta} \\ 1 & e^{-\iota\theta} & e^{-2\iota\theta} & e^{-3\iota\theta} \\ 1 & e^{\iota\psi} & e^{2\iota\psi} & e^{3\iota\psi} \\ 1 & e^{-\iota\psi} & e^{-2\iota\psi} & e^{-3\iota\psi} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \iota\theta \\ -\iota\theta \\ \iota\psi \\ -\iota\psi \end{bmatrix}. \tag{3.26}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = \frac{\psi(\sin(\psi-\theta)-\sin(3\psi+\theta)-\sin(2\psi-2\theta)+\sin(2\psi+2\theta)+\sin(3\psi-\theta)-\sin(\psi+\theta))}{D(\theta,\psi)} - \\ \qquad \frac{\theta(\sin(\psi-\theta)+\sin(\psi-3\theta)+\sin(\psi+3\theta)-\sin(2\psi-2\theta)-\sin(2\psi+2\theta)+\sin(\psi+\theta))}{D(\theta,\psi)}, \\ b_1 = \frac{\psi(\sin(2\psi-3\theta)-\sin(2\psi+3\theta)-\sin(3\psi-2\theta)+\sin(3\psi+2\theta)+2\sin(\theta))}{D(\theta,\psi)} + \\ \qquad \frac{\theta(\sin(2\psi-3\theta)+\sin(2\psi+3\theta)-\sin(3\psi-2\theta)-\sin(3\psi+2\theta)+2\sin(\psi))}{D(\theta,\psi)}, \\ b_2 = \frac{-\psi(\sin(\psi-3\theta)-\sin(\psi+3\theta)+\sin(3\psi+\theta)+2\sin(2\theta)-\sin(3\psi-\theta))}{D(\theta,\psi)} - \\ \qquad \frac{\theta(\sin(\psi-3\theta)+\sin(\psi+3\theta)-\sin(3\psi+\theta)+2\sin(2\psi)-\sin(3\psi-\theta))}{D(\theta,\psi)}, \\ b_3 = \frac{\psi(\sin(\psi-2\theta)-\sin(\psi+2\theta)+\sin(2\psi+\theta)-\sin(2\psi-\theta)+2\sin(\theta))}{D(\theta,\psi)} + \\ \qquad \frac{\theta(\sin(\psi-2\theta)+\sin(\psi+2\theta)-\sin(2\psi+\theta)-\sin(2\psi-\theta)+2\sin(\psi))}{D(\theta,\psi)}. \end{cases} \tag{3.27}$$

where:

$$D(\theta,\psi)\stackrel{\text{def}}{=}2\cos(\psi-\theta)+\cos(\psi-3\theta)-\cos(\psi+3\theta)-\cos(3\psi+\theta)-2\cos(2\psi-2\theta)+ \\ 2\cos(2\psi+2\theta)+\cos(3\psi-\theta)-2\cos(\psi+\theta). \quad (3.28)$$

Note, for instance, that $D(\theta,0)=D(0,\psi)=D(\theta,\theta)=0$, hence, the problem of determining the coefficients $b_i's$ would be ill-posed for $\theta=0$ or $\psi=0$ or $\theta=\psi$.

Logarithm of a symplectic matrix: Case of two distinct real-valued eigenvalues and two distinct unimodular eigenvalues. It is convenient to write the eigenvalues as r, r^{-1}, $e^{\iota\theta}$, $e^{-\iota\theta}$, with $r\in\mathbb{R}-\{0,\pm1\}$, $\theta\in\mathbb{R}-\pi\mathbb{Z}$. The linear system of equations to determine the coefficients b_i's read:

$$\begin{bmatrix} 1 & r & r^2 & r^3 \\ 1 & r^{-1} & r^{-2} & r^{-3} \\ 1 & e^{\iota\theta} & e^{2\iota\theta} & e^{3\iota\theta} \\ 1 & e^{-\iota\theta} & e^{-2\iota\theta} & e^{-3\iota\theta} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \log r \\ -\log r \\ \iota\theta \\ -\iota\theta \end{bmatrix}. \quad (3.29)$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{(1+r^4)\log r}{(1-r^2)|1-re^{\iota\theta}|^2}+\frac{r\theta\cos 2\theta}{|1-re^{\iota\theta}|^2\sin\theta}, \\ b_1 = \frac{r(r^2+1)+2(r^4+1)\cos\theta}{(r^2-1)|1-re^{\iota\theta}|^2}\log r-\frac{(r^2+1)\cos 2\theta+r\cos\theta}{|1-re^{\iota\theta}|^2\sin\theta}\theta, \\ b_2 = \frac{r^4+2r(r^2+1)\cos\theta}{(1-r^2)|1-re^{\iota\theta}|^2}\log r+\frac{(r^2+1)\cos\theta+r\cos 2\theta}{|1-re^{\iota\theta}|^2\sin\theta}\theta, \\ b_3 = \frac{r(r^2+1)\log r}{(r^2-1)|1-re^{\iota\theta}|^2}-\frac{r\theta\cos\theta}{|1-re^{\iota\theta}|^2\sin\theta}. \end{cases} \quad (3.30)$$

The problem is clearly ill-posed if $r=\pm1$ or if $r=0$. In addition, note that, for $\theta=0$, the algebraic system (3.29) is inconsistent, since the matrix of the known-terms on the left-hand side is not invertible. Consequently, the solution (3.30) shows a singularity for $\theta=0$. Curiously enough, such singularity is removable. For example, think of b_0 as a function $b_0(r,\theta)$: It holds that:

$$\lim_{\theta\to0} b_0(r,\theta)=\frac{(1+r^4)\log r}{(1-r^2)^2}+\frac{r}{(1-r^2)},$$

where the limit $\lim_{\theta\to0}\frac{\theta}{\sin\theta}=1$ was made use of. The solution obtained by removing the singularity does not seem to correspond to the solution of the corresponding case $(r,r^{-1},1,1)$ (see below) and its meaning remains an open question.

Logarithm of a symplectic matrix: Case of four distinct real-valued eigenvalues. It is convenient to write the eigenvalues as x, x^{-1}, y, y^{-1}, with $x,y\in\mathbb{R}-\{0,\pm1\}$ and $x\neq y$. The linear system of equations to determine the coefficients b_i's read:

$$\begin{bmatrix} 1 & x & x^2 & x^3 \\ 1 & x^{-1} & x^{-2} & x^{-3} \\ 1 & y & y^2 & y^3 \\ 1 & y^{-1} & y^{-2} & y^{-3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \log x \\ -\log x \\ \log y \\ -\log y \end{bmatrix}. \quad (3.31)$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{y(x^4+1)\log x}{(x^2-1)(x-y)(1-xy)} - \frac{x(y^4+1)\log y}{(y^2-1)(x-y)(1-xy)}, \\ b_1 = \frac{y^4(x^2+1)+y^3x+x^2+xy+1}{(y^2-1)(x-y)(1-xy)}\log y - \frac{x^4(y^2+1)+x^3y+y^2+xy+1}{(x^2-1)(x-y)(1-xy)}\log x, \\ b_2 = \frac{x^3(y^2+xy+1)+xy^2+y+x}{(x^2-1)(x-y)(1-xy)}\log x - \frac{y^3(x^2+xy+1)+x^2y+x+y}{(y^2-1)(x-y)(1-xy)}\log y, \\ b_3 = \frac{xy(y^2+1)\log y}{(y^2-1)(x-y)(1-xy)} - \frac{xy(x^2+1)\log x}{(x^2-1)(x-y)(1-xy)}. \end{cases} \tag{3.32}$$

The problem is clearly ill-posed if $x = \pm 1$ or $y = \pm 1$ or $x = y$.

3.2.2. Logarithm of a symplectic matrix: Eigenvalues of multiplicity 2. There are six possible cases that cover the occurrence of eigenvalues with multiplicity 2. In order to build up a linear system of equations to calculate the coefficients b_i, it is necessary to make use of both equations $\log(z) = b_3z^3 + b_2z^2 + b_1z + b_0$ and and of its hand-by-hand first-order derivative with respect to the parameter z, namely, $\frac{1}{z} = 3b_3z^2 + 2b_2z + b_1$.

Logarithm of a symplectic matrix: Case of eigenvalues (x, x^{-1}, x, x^{-1}) **with** $x \in \mathbb{R} - \{0, \pm 1\}$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & x & x^2 & x^3 \\ 1 & x^{-1} & x^{-2} & x^{-3} \\ 0 & x & 2x^2 & 3x^3 \\ 0 & x^{-1} & 2x^{-2} & 3x^{-3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \log x \\ -\log x \\ 1 \\ 1 \end{bmatrix}. \tag{3.33}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{3x^2-1-x^4(x^2-3)}{(x^2-1)^3}\log x - \frac{x^4+1}{(x^2-1)^2}, \\ b_1 = \frac{x^4(x^2+2)+2x^2+1}{x(x^2-1)^2} - \frac{12x^3\log x}{(x^2-1)^3}, \\ b_2 = \frac{2x^2(3x^2+3)\log x}{(x^2-1)^3} - \frac{2(x^4+x^2+1)}{(x^2-1)^2}, \\ b_3 = \frac{x(x^2+1)}{(x^2-1)^2} - \frac{4x^3\log x}{(x^2-1)^3}. \end{cases} \tag{3.34}$$

The possibility that $x = \pm 1$ was clearly excluded because it would correspond to eigenvalues of multiplicity 4. Note, in fact, that the obtained solutions are singular for $x = \pm 1$.

Logarithm of a symplectic matrix: Case of eigenvalues $(x, x^{-1}, 1, 1)$ **with** $x \in \mathbb{R} - \{0, \pm 1\}$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & x & x^2 & x^3 \\ 1 & x^{-1} & x^{-2} & x^{-3} \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \log x \\ -\log x \\ 0 \\ 1 \end{bmatrix}. \tag{3.35}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{x}{(x-1)^2} - \frac{(x^4+1)\log x}{(x-1)^3(x+1)}, \\ b_1 = \frac{x^3(2x+1)+x+2}{(x-1)^3(x+1)} \log x - \frac{x^2+x+1}{(x-1)^2}, \\ b_2 = \frac{x^2+x+1}{(x-1)^2} - \frac{x^3(x+2)+2x+1}{(x-1)^3(x+1)} \log x, \\ b_3 = \frac{x(x^2+1)}{(x-1)^3(x+1)} \log x - \frac{x}{(x-1)^2}. \end{cases} \tag{3.36}$$

The possibility that $x = 1$ was clearly excluded because it would correspond to eigenvalues of multiplicity 4. Note, in fact, that the obtained solutions are singular for $x = 1$. The case $x = -1$ corresponds to two eigenvalues with multiplicity 2 and cannot be dealt with through the same structure of the resolving linear system adopted in the present subsection. It will be covered separately in a subsequent subsection.

Logarithm of a symplectic matrix: Case of eigenvalues $(x, x^{-1}, -1, -1)$ **with** $x \in \mathbb{R} - \{0, \pm 1\}$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & x & x^2 & x^3 \\ 1 & x^{-1} & x^{-2} & x^{-3} \\ 1 & -1 & 1 & -1 \\ 0 & -1 & 2 & -3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \log x \\ -\log x \\ \pi\iota \\ 1 \end{bmatrix}. \tag{3.37}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{(2\pi\iota-1)x}{(x+1)^2} - \frac{(x^4+1)\log x}{(x-1)(x+1)^3}, \\ b_1 = \frac{x^2-x+1-\pi\iota(2x^2-x+2)}{(x+1)^2} + \frac{x-2-x^3(2x-1)}{(x-1)(x+1)^3} \log x, \\ b_2 = \frac{x^2-x+1-\pi\iota(x-1)^2}{(x+1)^2} + \frac{2x-1-x^3(x-2)}{(x-1)(x+1)^3} \log x, \\ b_3 = \frac{x(x^2+1)\log x}{(x-1)(x+1)^3} + \frac{(\pi\iota-1)x}{(x+1)^2}. \end{cases} \tag{3.38}$$

Note that the coefficients b_i's are complex-valued, in this case, due to the presence of negative eigenvalues. The corresponding logarithm is complex-valued, which would be inconsistent with the definition of an Hamiltonian matrix (2.2). Such case is, therefore, to be excluded from the present context. The case $x = 1$ corresponds to two eigenvalues with multiplicity 2 and cannot be dealt with through the same structure of the resolving linear system adopted in the present subsection. It will be covered separately in a subsequent subsection.

Logarithm of a symplectic matrix: Case of eigenvalues $(e^{\iota\theta}, e^{-\iota\theta}, 1, 1)$ **with** $\theta \in \mathbb{R} - \pi\mathbb{Z}$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & e^{\iota\theta} & e^{2\iota\theta} & e^{3\iota\theta} \\ 1 & e^{-\iota\theta} & e^{-2\iota\theta} & e^{-3\iota\theta} \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \iota\theta \\ -\iota\theta \\ 0 \\ 1 \end{bmatrix}. \tag{3.39}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{\theta \cos 2\theta}{8\sin^3(\theta/2)\cos(\theta/2)} - \frac{1}{4\sin^2(\theta/2)}, \\ b_1 = -\frac{2\cos 2\theta + \cos\theta}{8\sin^3(\theta/2)\cos(\theta/2)}\theta + \frac{1+2\cos\theta}{4\sin^2(\theta/2)}, \\ b_2 = \frac{\cos 2\theta + 2\cos\theta}{8\sin^3(\theta/2)\cos(\theta/2)}\theta - \frac{1+2\cos\theta}{4\sin^2(\theta/2)}, \\ b_3 = \frac{1}{4\sin^2(\theta/2)} - \frac{\theta\cos\theta}{8\sin^3(\theta/2)\cos(\theta/2)}. \end{cases} \quad (3.40)$$

Logarithm of a symplectic matrix: Case of eigenvalues $(e^{\iota\theta}, e^{-\iota\theta}, -1, -1)$ **with** $\theta \in \mathbb{R} - \pi\mathbb{Z}$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & e^{\iota\theta} & e^{2\iota\theta} & e^{3\iota\theta} \\ 1 & e^{\iota\theta} & e^{-2\iota\theta} & e^{-3\iota\theta} \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 2 & -3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \iota\theta \\ -\iota\theta \\ \pi\iota \\ 1 \end{bmatrix}. \quad (3.41)$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{2\pi\iota - 1}{4\cos^2(\theta/2)} - \frac{\theta\cos 2\theta}{8\cos^3(\theta/2)\sin(\theta/2)}, \\ b_1 = \frac{\theta(\cos\theta - 2\cos 2\theta)}{8\cos^3(\theta/2)\sin(\theta/2)} - \frac{2(2\pi\iota - 1)\cos\theta - \pi\iota + 1}{4\cos^2(\theta/2)}, \\ b_2 = \frac{2\cos\theta - 1 + 4\pi\iota\sin^2(\theta/2)}{4\cos^2(\theta/2)} + \frac{2\cos\theta - \cos 2\theta}{8\cos^3(\theta/2)\sin(\theta/2)}\theta, \\ b_3 = \frac{\pi\iota - 1}{4\cos^2(\theta/2)} + \frac{\theta\cos\theta}{8\cos^3(\theta/2)\sin(\theta/2)}. \end{cases} \quad (3.42)$$

Even in this case, the coefficients b_i's are complex-valued, due to the presence of negative eigenvalues. The corresponding logarithm is complex-valued, which would be inconsistent with the definition of an Hamiltonian matrix (2.2). Such case is to be excluded from the present context.

Case of eigenvalues $(1, 1, -1, -1)$**.** The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & -1 & 2 & -3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \pi\iota \\ 1 \\ 1 \end{bmatrix}. \quad (3.43)$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{\pi\iota - 1}{2}, \\ b_1 = -\frac{3\pi\iota}{4}, \\ b_2 = \frac{1}{2}, \\ b_3 = \frac{\pi\iota}{4}. \end{cases} \quad (3.44)$$

Even in this case, the coefficients b_i's are complex-valued, due to the presence of negative eigenvalues. The corresponding logarithm is complex-valued, which would be inconsistent with the definition of an Hamiltonian matrix (2.2). Such case is, therefore, to be excluded from the present context.

3.2.3. Logarithm of a symplectic matrix: Eigenvalues of multiplicity 4. There are two possible cases that cover the occurrence of eigenvalues with multiplicity 4. In order to build up a linear system of equations to calculate the coefficients b_i, it is necessary to make use of equation $\log(z) = b_3 z^3 + b_2 z^2 + b_1 z + b_0$ along with its term-by-term first-order derivative $\frac{1}{z} = 3b_3 z^2 + 2b_2 z + b_1$, its term-by-term second-order derivative $-\frac{1}{z^2} = 6b_3 z + 2b_2$ and its term-by-term third-order derivative $\frac{2}{z^3} = 6b_3$.

Logarithm of a symplectic matrix: Case of eigenvalues $(1, 1, 1, 1)$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \end{bmatrix}. \tag{3.45}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = -\frac{11}{6}, \\ b_1 = 3, \\ b_2 = -\frac{3}{2}, \\ b_3 = \frac{1}{3}. \end{cases} \tag{3.46}$$

Namely, for every 4×4 symplectic matrix S whose eigenvalues are $(1, 1, 1, 1)$, it holds that $\log(S) = \frac{1}{3}S^3 - \frac{3}{2}S^2 + 3S - \frac{11}{6}I$.

Logarithm of a symplectic matrix: Case of eigenvalues $(-1, -1, -1, -1)$. The linear system of equations to determine the coefficients b_i's reads:

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \pi\iota \\ -1 \\ -1 \\ -2 \end{bmatrix}. \tag{3.47}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \pi\iota - \frac{11}{6}, \\ b_1 = -3, \\ b_2 = -\frac{3}{2}, \\ b_3 = -\frac{1}{3}. \end{cases} \tag{3.48}$$

The coefficients b_i's are complex-valued, due to the presence of negative eigenvalues. The corresponding logarithm is complex-valued, which would be inconsistent with the definition of an Hamiltonian matrix (2.2). Such a case needs, therefore, to be excluded from the present context.

3.3. Cayley transform of a fourth-order Hamiltonian matrix. Let $H \in \mathrm{Ham}(4,\mathbb{R})$. For the sake of conciseness, only the case of four distinct eigenvalues $(\lambda_1, -\lambda_1, \lambda_2, -\lambda_2)$ with $\lambda_1, \lambda_2 \in \mathbb{C} - \{0\}$ and $\lambda_1^2 \neq \lambda_2^2$ is considered. In order to determine the coefficients b_i's of the expansion of the function $f(H) = \mathrm{cay}(H)$, given by:

$$\mathrm{cay}(H) = b_3 H^3 + b_2 H^2 + b_1 H + b_0 I, \tag{3.49}$$

recall that the equation (3.5), evaluated by replacing the variable z with each eigenvalue of the matrix H, reads $\frac{1+z}{1-z} = b_3 z^3 + b_2 z^2 + b_1 z + b_0$. Since the eigenvalues are, by hypothesis, all distinct, such equation, written for each eigenvalue separately, gives rise to the consistent linear system:

$$\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 \\ 1 & -\lambda_1 & \lambda_1^2 & -\lambda_1^3 \\ 1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 \\ 1 & -\lambda_2 & \lambda_2^2 & -\lambda_2^3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1+\lambda_1}{1-\lambda_1} \\ \frac{1-\lambda_1}{1+\lambda_1} \\ \frac{1+\lambda_2}{1-\lambda_2} \\ \frac{1-\lambda_2}{1+\lambda_2} \end{bmatrix}. \tag{3.50}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = \frac{\lambda_2^2(\lambda_1^2+1)}{(\lambda_1^2-\lambda_2^2)(\lambda_1^2-1)} - \frac{\lambda_1^2(\lambda_2^2+1)}{(\lambda_1^2-\lambda_2^2)(\lambda_2^2-1)}, \\ b_1 = \frac{2\lambda_2^2}{(\lambda_1^2-\lambda_2^2)(\lambda_1^2-1)} + \frac{2\lambda_1^2}{(\lambda_1^2-\lambda_2^2)(\lambda_2^2-1)}, \\ b_2 = \frac{\lambda_2^2+1}{(\lambda_1^2-\lambda_2^2)(\lambda_2^2-1)} - \frac{\lambda_1^2+1}{(\lambda_1^2-\lambda_2^2)(\lambda_1^2-1)}, \\ b_3 = -\frac{2}{(\lambda_1^2-\lambda_2^2)(\lambda_1^2-1)} + \frac{2}{(\lambda_1^2-\lambda_2^2)(\lambda_2^2-1)}. \end{cases} \tag{3.51}$$

3.4. Cayley transform of a fourth-order symplectic matrix. Let $S \in \mathrm{Sp}(4,\mathbb{R})$. The coefficients b_i's of the expansion of the function $f(S) = \mathrm{cay}(S)$ given by:

$$\mathrm{cay}(S) = b_3 S^3 + b_2 S^2 + b_1 S + b_0 I, \tag{3.52}$$

will be determined by proceeding as before. For the sake of conciseness, only the cases of four distinct eigenvalues are considered.

Cayley transform of a symplectic matrix: Case $(\lambda, \bar{\lambda}, \lambda^{-1}, \bar{\lambda}^{-1})$, with $\lambda \in \mathbb{C} - \{0\}$. In order for the eigenvalues to be distinct one from another, it is assumed that $|\lambda| \neq 1$ and that $\mathrm{Im}\{\lambda\} \neq 0$. The equation to determine the coefficients b_i's gives rise to the consistent linear system:

$$\begin{bmatrix} 1 & \lambda & \lambda^2 & \lambda^3 \\ 1 & \bar{\lambda} & \bar{\lambda}^2 & \bar{\lambda}^3 \\ 1 & \lambda^{-1} & \lambda^{-2} & \lambda^{-3} \\ 1 & \bar{\lambda}^{-1} & \bar{\lambda}^{-2} & \bar{\lambda}^{-3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1+\lambda}{1-\lambda} \\ \frac{1+\bar{\lambda}}{1-\bar{\lambda}} \\ \frac{\lambda+1}{\lambda-1} \\ \frac{\bar{\lambda}+1}{\bar{\lambda}-1} \end{bmatrix}. \tag{3.53}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\begin{cases} b_0 = 2\Re\left\{\frac{\lambda(1+\bar{\lambda}^4)(\bar{\lambda}+1)}{(\bar{\lambda}-1)(\bar{\lambda}^2-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)}\right\}, \\ b_1 = 2\Re\left\{\frac{\bar{\lambda}^2+|\lambda|^2+1+\lambda^3(\lambda\bar{\lambda}^2+\bar{\lambda}+\lambda)}{(\lambda^2-1)(\lambda-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)}(\lambda+1)\right\}, \\ b_2 = 2\Re\left\{\frac{\lambda+\bar{\lambda}+\lambda^2\bar{\lambda}+\bar{\lambda}^3(|\lambda|^2+\lambda^2+1)}{(\bar{\lambda}-1)(\bar{\lambda}^2-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)}(\bar{\lambda}+1)\right\}, \\ b_3 = 2\Re\left\{\frac{|\lambda|^2(\lambda^2+1)(\lambda+1)}{(\lambda^2-1)(\lambda-1)(\lambda-\bar{\lambda})(1-|\lambda|^2)}\right\}. \end{cases} \tag{3.54}$$

Cayley transform of a symplectic matrix: Case of two distinct unimodular duets. It is convenient to write the eigenvalues as $(e^{\iota\theta}, e^{-\iota\theta}, e^{\iota\psi}, e^{-\iota\psi})$, with $\theta, \psi \in (0, 2\pi)$, $\theta \neq \psi$ and $\theta + \psi \neq 2\pi$. Since the eigenvalues are, by hypothesis, all distinct, the linear system to compute the coefficients b_i's reads:

$$\begin{bmatrix} 1 & e^{\iota\theta} & e^{2\iota\theta} & e^{3\iota\theta} \\ 1 & e^{-\iota\theta} & e^{-2\iota\theta} & e^{-3\iota\theta} \\ 1 & e^{\iota\psi} & e^{2\iota\psi} & e^{3\iota\psi} \\ 1 & e^{-\iota\psi} & e^{-2\iota\psi} & e^{-3\iota\psi} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1+e^{\iota\theta}}{1-e^{\iota\theta}} \\ \frac{1+e^{-\iota\theta}}{1-e^{-\iota\theta}} \\ \frac{1+e^{\iota\psi}}{1-e^{\iota\psi}} \\ \frac{1+e^{-\iota\psi}}{1-e^{-\iota\psi}} \end{bmatrix}. \tag{3.55}$$

The above linear system is equivalent to:

$$\begin{bmatrix} 1 & \cos(\theta) & \cos(2\theta) & \cos(3\theta) \\ 0 & \sin(\theta) & \sin(2\theta) & \sin(3\theta) \\ 1 & \cos(\psi) & \cos(2\psi) & \cos(3\psi) \\ 0 & \sin(\psi) & \sin(2\psi) & \sin(3\psi) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\cos(\theta/2)}{\sin(\theta/2)} \\ 0 \\ \frac{\cos(\psi/2)}{\sin(\psi/2)} \end{bmatrix}. \tag{3.56}$$

The solution of the above linear system, in the unknown b_i's, is:

$$\left\{ \begin{array}{rcl} b_0 & = & \frac{\cos(\psi/2)(\sin(\psi-\theta)-\sin(3\psi+\theta)-\sin(2\psi-2\theta)+\sin(2\psi+2\theta)+\sin(3\psi-\theta)-\sin(\psi+\theta))}{\sin(\psi/2)D(\theta,\psi)} - \\ & & \frac{\cos(\theta/2)(\sin(\psi-\theta)+\sin(\psi-3\theta)+\sin(\psi+3\theta)-\sin(2\psi-2\theta)-\sin(2\psi+2\theta)+\sin(\psi+\theta))}{\sin(\theta/2)D(\theta,\psi)}, \\ b_1 & = & \frac{\cos(\psi/2)(\sin(2\psi-3\theta)-\sin(2\psi+3\theta)-\sin(3\psi-2\theta)+\sin(3\psi+2\theta)+2\sin(\theta))}{\sin(\psi/2)D(\theta,\psi)} + \\ & & \frac{\cos(\theta/2)(\sin(2\psi-3\theta)+\sin(2\psi+3\theta)-\sin(3\psi-2\theta)-\sin(3\psi+2\theta)+2\sin(\psi))}{\sin(\theta/2)D(\theta,\psi)}, \\ b_2 & = & -\frac{\cos(\psi/2)(\sin(\psi-3\theta)-\sin(\psi+3\theta)+\sin(3\psi+\theta)+2\sin(2\theta)-\sin(3\psi-\theta))}{\sin(\psi/2)D(\theta,\psi)} - \\ & & \frac{\cos(\theta/2)(\sin(\psi-3\theta)+\sin(\psi+3\theta)-\sin(3\psi+\theta)+2\sin(2\psi)-\sin(3\psi-\theta))}{\sin(\theta/2)D(\theta,\psi)}, \\ b_3 & = & \frac{\cos(\psi/2)(\sin(\psi-2\theta)-\sin(\psi+2\theta)+\sin(2\psi+\theta)-\sin(2\psi-\theta)+2\sin(\theta))}{\sin(\psi/2)D(\theta,\psi)} + \\ & & \frac{\cos(\theta/2)(\sin(\psi-2\theta)+\sin(\psi+2\theta)-\sin(2\psi+\theta)-\sin(2\psi-\theta)+2\sin(\psi))}{\sin(\theta/2)D(\theta,\psi)}, \end{array} \right. \tag{3.57}$$

where:

$$D(\theta,\psi) \stackrel{\text{def}}{=} 2\cos(\psi-\theta) + \cos(\psi-3\theta) - \cos(\psi+3\theta) - \cos(3\psi+\theta) - 2\cos(2\psi-2\theta) + 2\cos(2\psi+2\theta) + \cos(3\psi-\theta) - 2\cos(\psi+\theta). \tag{3.58}$$

Cayley transform of a symplectic matrix: Case of two distinct real-valued eigenvalues and two distinct unimodular eigenvalues. It is convenient to write the eigenvalues as r, r^{-1}, $e^{\iota\theta}$, $e^{-\iota\theta}$, with $r \in \mathbb{R} - \{0, \pm 1\}$, $\theta \in \mathbb{R} - \pi\mathbb{Z}$. The linear system of equations to determine the coefficients b_i's read:

$$\begin{bmatrix} 1 & r & r^2 & r^3 \\ 1 & r^{-1} & r^{-2} & r^{-3} \\ 1 & e^{\iota\theta} & e^{2\iota\theta} & e^{3\iota\theta} \\ 1 & e^{-\iota\theta} & e^{-2\iota\theta} & e^{-3\iota\theta} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1+r}{1-r} \\ \frac{r+1}{r-1} \\ \frac{1+e^{\iota\theta}}{1-e^{\iota\theta}} \\ \frac{1+e^{-\iota\theta}}{1-e^{-\iota\theta}} \end{bmatrix}. \tag{3.59}$$

The above linear system is equivalent to:

$$\begin{bmatrix} 1 & r & r^2 & r^3 \\ 1 & r^{-1} & r^{-2} & r^{-3} \\ 1 & \cos(\theta) & \cos(2\theta) & \cos(3\theta) \\ 0 & \sin(\theta) & \sin(2\theta) & \sin(3\theta) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1+r}{1-r} \\ -\frac{1+r}{1-r} \\ 0 \\ \frac{\cos(\theta/2)}{\sin(\theta/2)} \end{bmatrix}. \tag{3.60}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 &= \frac{r\cos(\theta/2)(\cos(2\theta)-r\cos(3\theta)+r^2\cos(2\theta)-r\cos(\theta))}{\sin(\theta/2)D(r,\theta)} + \frac{(r^4+1)(r+1)(\sin(\theta)r^2-\sin(2\theta)r+\sin(\theta))}{(r^2-1)(r-1)D(r,\theta)}, \\ b_1 &= -\frac{\cos(\theta/2)(\cos(2\theta)-r\cos(3\theta)+r^2\cos(2\theta)-r^3\cos(3\theta)+r^4\cos(2\theta)-r^2)}{\sin(\theta/2)D(r,\theta)} - \\ & \quad \frac{(r^3+1)(r+1)(\sin(2\theta)r^3-\sin(3\theta)r^2+\sin(\theta))}{(r^2-1)(r-1)D(r,\theta)}, \\ b_2 &= \frac{(r^3+1)(r+1)(\sin(2\theta)r^3-\sin(3\theta)r^2+\sin(\theta))}{(r^2-1)(r-1)D(r,\theta)} - \\ & \quad \frac{\cos(\theta/2)(r-\cos(\theta)-r^2\cos(\theta)-r^4\cos(\theta)+r^2\cos(3\theta)+r^3)}{\sin(\theta/2)D(r,\theta)}, \\ b_3 &= \frac{r\cos(\theta/2)(r-\cos(\theta)+r\cos(2\theta)-r^2\cos(\theta))}{\sin(\theta/2)D(r,\theta)} - \frac{r(r^2+1)(r+1)(\sin(\theta)r^2-\sin(2\theta)r+\sin(\theta))}{(r^2-1)(r-1)D(r,\theta)}, \end{cases} \tag{3.61}$$

where:

$$D(r,\theta) = \sin(\theta) - 2r\sin(2\theta) + 3r^2\sin(\theta) + r^4\sin(\theta) + r^2\sin(3\theta) - 2r^3\sin(2\theta). \tag{3.62}$$

Cayley transform of a symplectic matrix: Case of four distinct real-valued eigenvalues. It is convenient to write the eigenvalues as x, x^{-1}, y, y^{-1}, with $x, y \in \mathbb{R} - \{0, -1, 1\}$ and $x \neq y$. The linear system of equations to determine the coefficients b_i's read:

$$\begin{bmatrix} 1 & x & x^2 & x^3 \\ 1 & x^{-1} & x^{-2} & x^{-3} \\ 1 & y & y^2 & y^3 \\ 1 & y^{-1} & y^{-2} & y^{-3} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1+x}{1-x} \\ -\frac{1+x}{1-x} \\ \frac{1+y}{1-y} \\ -\frac{1+y}{1-y} \end{bmatrix}. \tag{3.63}$$

The solution of the above linear system reads:

$$\begin{cases} b_0 = \frac{x(y^4+1)(y+1)}{(y^2-1)(y-1)(-x^2y+xy^2+x-y)} - \frac{y(x^4+1)(x+1)}{(x^2-1)(x-1)(-x^2y+xy^2+x-y)}, \\ b_1 = \frac{y^2+xy+1+x^3(xy^2+y+x)}{(x^2-1)(x-1)(-x^2y+xy^2+x-y)}(x+1) - \frac{x^2+yx+1+y^3(yx^2+x+y)}{(y^2-1)(y-1)(-x^2y+xy^2+x-y)}(y+1), \\ b_2 = \frac{yx^2+x+y+y^3(x^2+yx+1)}{(y^2-1)(y-1)(-x^2y+xy^2+x-y)}(y+1) - \frac{xy^2+y+x+x^3(y^2+xy+1)}{(x^2-1)(x-1)(-x^2y+xy^2+x-y)}(x+1), \\ b_3 = \frac{xy(x^2+1)(x+1)}{(x^2-1)(x-1)(-x^2y+xy^2+x-y)} - \frac{xy(y^2+1)(y+1)}{(y^2-1)(y-1)(-x^2y+xy^2+x-y)}. \end{cases} \tag{3.64}$$

4. Matrix functions on the optical transference Lie group and its Lie algebra. In order to implement the Kolmogoroff-Nagumo mean (2.10) in the full-transference-matrix case (that is, in the presence of decentered optical systems), it is necessary to extend the above calculations to the case of logarithm/exponential and to the case of Cayley transform of matrices belonging to the Lie group Opt and to the Lie algebra LOpt.

4.1. Logarithm and exponential on the optical transference group. Let $L \in \text{LOpt}$. Recall from Subsection 2.3 that:

$$\exp(L) = I + \sum_{k=1}^{\infty} \frac{L^k}{k!}, \tag{4.1}$$

with:

$$L = \begin{bmatrix} H & v \\ 0 & 0 \end{bmatrix}.$$

Under the assumption that $\det(H) \neq 0$, it is straightforward to prove that:

$$L^k = \begin{bmatrix} H^k & H^k H^{-1} v \\ 0 & 0 \end{bmatrix}, \tag{4.2}$$

for any positive integer k, and hence that:

$$\exp(L) = \begin{bmatrix} \exp(H) & (\exp(H) - I) H^{-1} v \\ 0 & 1 \end{bmatrix}. \tag{4.3}$$

If $\det(H) = 0$, the expression $(\exp(H) - I)H^{-1}v$ is ill-written. It should be replaced with the series $\sum_{k=0}^{\infty} \frac{H^k v}{(k+1)!}$ which is nevertheless convergent.

The expression $\exp(L)$ may be evaluated on the basis of the expression (4.3), where the submatrix $\exp(H)$, in turn, is evaluated through its third-order polynomial representation (3.10). In the presence of a decentered system, the quantity $(\exp(H) - I)H^{-1}v$ needs to be evaluated as well. It may be computed either on the basis of the third-order polynomial representation (4.3) or by computing explicitly the third-order polynomial representation of the function $(\exp(H) - I)H^{-1}$ by studying the function $f(z) = \frac{\exp(z)-1}{z}$.

Moreover, let $T \in \text{Opt}$. Recall from Subsection 2.3 that:

$$\log(T) = -\sum_{k=1}^{\infty} \frac{(I - T)^k}{k}, \tag{4.4}$$

where:

$$I_5 - T = \begin{bmatrix} I_4 - S & -\delta \\ 0 & 0 \end{bmatrix}.$$

Under the assumption that $\det(S - I) \neq 0$, it is straightforward to prove that:

$$(I_5 - T)^k = \begin{bmatrix} (I_4 - S)^k & -(I_4 - S)^k (I_4 - S)^{-1} \delta \\ 0 & 0 \end{bmatrix},$$

for any positive integer k, and hence that:

$$\log(T) = \begin{bmatrix} \log(S) & \log(S)(S - I)^{-1} \delta \\ 0 & 0 \end{bmatrix}. \tag{4.5}$$

The latter result coincides with the formula already proven in the contribution [15]. The expression $\log(T)$ may be evaluated on the basis of the expression (4.5), where the submatrix $\log(S)$, in turn, is evaluated through its third-order polynomial representation (3.21). Whenever a decentered system is dealt with, the quantity $\log(S)(S - I)^{-1}\delta$ needs to be evaluated as well. It may be computed either on the basis of the third-order polynomial representation (3.21) or by computing explicitly the third-order polynomial representation of the function $\log(S)(S - I)^{-1}$ by studying the function $f(z) = \frac{\log z}{z-1}$.

4.2. Cayley transform on the optical transference group. Let again $L \in \text{LOpt}$. Recall from Subsection 2.3 that:

$$\text{cay}(L) = (I + L) \sum_{k=0}^{\infty} L^k. \tag{4.6}$$

By the partial result (4.2) of the preceding subsection, it is readily inferred that:

$$\sum_{k=0}^{\infty} L^k = \begin{bmatrix} (I-H)^{-1} & ((I-H)^{-1} - I)H^{-1}v \\ 0 & 1 \end{bmatrix}.$$

Therefore, it results that:

$$\text{cay}(L) = \begin{bmatrix} \text{cay}(H) & (\text{cay}(H) - I)H^{-1}v \\ 0 & 1 \end{bmatrix}. \tag{4.7}$$

The expression $\text{cay}(L)$ may be evaluated on the basis of the expression (4.7), where the submatrix $\text{cay}(H)$, in turn, is evaluated through its third-order polynomial representation (3.49). For a decentered system, the quantity $(\text{cay}(H) - I)H^{-1}v$ needs to be computed. It may be evaluated either on the basis of the third-order polynomial representation (3.49) or by computing explicitly the third-order polynomial representation of the function $(\text{cay}(H)-I)H^{-1}$ by studying the function $f(z) = \frac{2}{z(z-1)}$.

Moreover, let $T \in \text{Opt}$ again. It is important to note that the quantity $\text{cay}(T)$ is *ill-defined*, because the matrix $T - I$ is singular (the lower row is all zero). This is the reason why the Kolmogoroff-Nagumo mean in the Cayley-transform-based version is defined in terms of the matrix $\text{cay}(-T)$. The aim of the following calculations is to find the inner structure of the matrix $\text{cay}(-T)$.

It is readily seen that $I_5 + T = \left[\begin{smallmatrix} I_4+S & \delta \\ 0 & 2 \end{smallmatrix}\right]$. The inverse matrix $(I + T)^{-1}$ may be calculated directly and it is found to be equal to $\left[\begin{smallmatrix} (I+S)^{-1} & -(I+S)^{-1}\delta/2 \\ 0 & 1/2 \end{smallmatrix}\right]$. Conversely, it holds that $I_5 - T = \left[\begin{smallmatrix} I_4-S & -\delta \\ 0 & 0 \end{smallmatrix}\right]$. As a result, it holds that:

$$\text{cay}(-T) = (I - T)(I + T)^{-1} = \begin{bmatrix} \text{cay}(-S) & -(\text{cay}(-S) + I)\delta/2 \\ 0 & 0 \end{bmatrix}. \tag{4.8}$$

The expression $\text{cay}(-T)$ may be evaluated on the basis of the expression (4.8), where the submatrix $\text{cay}(-S)$, in turn, is evaluated through its third-order polynomial representation (3.52). If a decentered system is to be dealt with, the quantity $(\text{cay}(-S)+I)\delta/2$ needs to be evaluated as well. It may be computed directly on the basis of the third-order polynomial representation (3.52).

5. Conclusion. The present manuscript investigates on the relationships between the exponential-mean-log rule to compute the average out of a set of optical transference matrices (either centered or decentered) and the general notion of Kolmogoroff-Nagumo mean extended to the Lie group of affine symplectic matrices. It was shown that, indeed, the exponential-mean-log rule may be regarded as a special case of the Kolmogoroff-Nagumo φ-mean-φ^{-1} averaging rule where the φ-map needs to coincide with a Lie-group exponential map. A second choice of φ-map, based on the matrix Cayley transform, was also discussed.

The present manuscript also focuses on the efficient, closed-form computation of the φ-maps as well as of the φ^{-1}-maps as third-order matrix polynomials. The computation efforts focused on the closed-form determination of the coefficients of the polynomial associated to each discussed map, which depend on the eigenvalues of the matrix (either symplectic or Hamiltonian) that the mentioned maps are applied to. In particular, the Section 3 discussed all the possible cases corresponding to all kind of multiplicity of such eigenvalues for the exponential and the logarithmic maps and a few cases for the Cayley maps.

6. Acknowledgments. I had the great pleasure of meeting Prof. William Harris during a Workshop on matrix analysis in Manchester in April 2013. I would like to thank Prof. Harris for sharing his papers on optical system transference theory and for his encouragement on pursuing the present research endeavor.

REFERENCES

[1] S. Agoujil, A.H. Bentbib and A. Kanber, *A structure preserving approximation method for Hamiltonian exponential matrices*, Applied Numerical Mathematics, Vol. 62, 1126 – 1138, 2012

[2] M.J. Bastiaans and T. Alieva, *Classification of the linear canonical transformation and its associated real symplectic matrix*, in Proceedings of the 9th International Symposium on Signal Processing and Its Applications (ISSPA 2007, February 12-15, 2007), pp. 1 – 4, 2007

[3] G.S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods*, Birkhäuser, 2009

[4] O. Chisini, *Sul concetto di media*, Periodico di Matematiche, Vol. 4, pp. 106 – 116, 1929

[5] R.D. van Gool and W.F. Harris, *The concept of the average eye*, The South African Optometrist, Vol. 64, No. 2, pp. 38 – 43, 2005

[6] M.A. de Gosson, *Symplectic Methods in Harmonic Analysis and in Mathematical Physics*, Pseudo-Differential Operators 7, Springer Basel AG 2011

[7] S. Fiori, *On vector averaging over the unit hyphersphere*, Digital Signal Processing, Vol. 19, No. 4, pp. 715 – 725, July 2009

[8] S. Fiori, *Learning the Fréchet mean over the manifold of symmetric positive-definite matrices*, Cognitive Computation, Vol. 1, No. 4, pp. 279 – 291, December 2009

[9] S. Fiori, *Averaging over the Lie group of optical systems transference matrices*, Frontiers of Electrical and Electronic Engineering, Special issue of the Sino foreign-interchange Workshop on Intelligence Science and Intelligent Data Engineering 2010 - Part A, Vol. 6, No. 1, pp. 137 – 145, March 2011

[10] S. Fiori, *Solving minimal-distance problems over the manifold of real symplectic matrices*, SIAM Journal on Matrix Analysis and Applications, Vol. 32, No. 3, pp. 938 – 968, 2011

[11] S. Fiori and T. Tanaka, *An algorithm to compute averages on matrix Lie groups*, IEEE Transactions on Signal Processing, Vol. 57, No. 12, pp. 4734 – 4743, December 2009

[12] M. Fréchet, *Les élements aléatoires de nature quelconque dans un espace distancié*, Annales de l'Institut Henri Poincaré, Vol. 10, pp 215 – 310, 1948

[13] W.F. Harris, *Paraxial ray tracing through noncoaxial astigmatic optical systems, and a* 5×5 *augmented system matrix*, Optometry and Vision Science, Vol. 71, No. 4, pp. 282 – 285, 1994

[14] W.F. Harris, *The average eye*, Ophthalmic and Physiological Optics, Vol. 24, pp. 580 – 585, 2004

[15] W.F. Harris and J.R. Cardoso, *The exponential-mean-log-transference as a possible representation of the optical character of an average eye*, Ophthalmic and Physiological Optics, Vol. 26, No. 4, pp. 380 – 383, 2006

[16] G. Hori and T. Tanaka, *Pivoting in Cayley transform-based optimization on orthogonal groups*, in Proceedings of the Second APSIPA Annual Summit and Conference (Biopolis, Singapore, 14-17 December 2010), pp. 181 – 184, 2010

[17] T. Kaneko, S. Fiori and T. Tanaka, *Empirical arithmetic averaging over the compact Stiefel manifold*, IEEE Transactions on Signal Processing, Vol. 61, No. 4, pp. 883 – 894, February 2013

[18] H. Karcher, *Riemannian center of mass and mollifier smoothing*, Communications on Pure and Applied Mathematics, Vol. 30, pp. 509 – 541, 1977

[19] M.P. Keating, *A matrix formulation of spectacle magnification*, Ophthalmic and Physiological Optics, Vol. 2, No. 2, 145 – 158, 1982

[20] J.-L. Marichal, *On an axiomatization of the quasi-arithmetic mean values without the symmetry axiom*, Aequationes Mathematicae, Vol. 59, No. 1-2, pp. 74 – 83, 2000

[21] T. Politi, *A formula for the exponential of a real skew-symmetric matrix of order 4*, BIT Numerical Mathematics, Vol. 41, No. 4, pp. 842 – 845, 2001

[22] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Volume 1, 2nd Edition, Berkeley, CA: Publish or Perish Press, 1979

[23] J. Stillwell, *Naive Lie Theory*, Chapter 7: "The matrix logarithm", pp. 139 – 159, Undergraduate Texts in Mathematics, Springer New York, 2008