



Mueller matrix cone and its application to filtering

TIM ZANDER^{1,2,*}  AND JUERGEN BEYERER^{1,2}

¹Karlsruhe Institute of Technology, Vision and Fusion Laboratory (IES), Karlsruhe, Germany

²Fraunhofer Institute of Optonics, System Technologies and Image Exploitation (IOSB), Karlsruhe, Germany

*tim.zander@iosb.fraunhofer.de

Abstract: We show that there is an isometry between the real ambient space of all Mueller matrices and the space of all Hermitian matrices that maps the Mueller matrices onto the positive semidefinite matrices. We use this to establish an optimality result for the filtering of Mueller matrices, which roughly says that it is always enough to filter the eigenvalues of the corresponding “coherency matrix.” Then we further explain how the knowledge of the cone of Hermitian positive semidefinite matrices can be transferred to the cone of Mueller matrices with a special emphasis towards optimisation. In particular, we suggest that means of Mueller matrices should be computed within the corresponding Riemannian geometry.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

In polarisation optics Mueller matrices are of great importance, as they describe the change of polarisation of light after interacting with a medium in a linear fashion. In order to be a Mueller matrix the matrix has to satisfy the Stokes criterion, which states that every Stokes vector has to be mapped onto a Stokes vector. Cloude then showed in [1], after establishing an additional criterion for the realisability of Mueller matrices, that Mueller matrices can be associated with Hermitian matrices with non-negative eigenvalues the so called coherency or covariance matrices. This was then used for filtering measured matrices in order to make them physically meaningful, i.e. satisfying the Stokes criterion and Cloude’s criterion. Moreover, it was shown that any coherency matrix of a non-depolarising Mueller matrix also known as a Jones-Mueller matrix has only one non-zero eigenvalue. This then easily suggests that any Mueller matrix can be expressed as the sum of no more than four non-depolarising matrices. In [2] and [3] matrices, which can be decomposed into a non-depolarising part and one perfectly depolarising part, have been analysed and in the latter a filtering method was proposed. Optimality of filtering was analysed in [4] by using a maximum likelihood method originally developed for quantum process tomography and does such as [5] rely on the Cholesky decomposition of the coherency matrix for filtering. More about optimality filtering of Mueller matrices was derived in [6], [7], [8] and [5]. In [9] then the optimality of the Cloude filter was rigorously proved.

The purpose of this work is to connect the methodologies of filtering of measured Mueller matrices to well-established mathematical theories. We will show how this can be used to prove a more general theorem about the optimality of filtering of Mueller matrices. This simplifies and generalises part of the results of [9]. Moreover, we then review the mathematical theory about the Hermitian positive semidefinite cone and explain, along with reviewing existing results, how this gives rise to the differential geometry of the manifold of all Mueller matrices.

2. Isometry of the ambient space

In this section we explain how a well know result about the connection of Mueller matrices and Hermitian positive definite matrices establishes an isometry between them. But first we will

define some elementary concepts. The polarisation state and intensity of a (partially) polarised light beam can be described by a four-dimensional real vector (I, Q, U, V) , where I is the intensity and (Q, U, V) describe the intensity and state of the polarised part. This vector satisfies the Stokes criterion

$$I^2 \geq Q^2 + U^2 + V^2, \quad (1)$$

which implies that the degree of polarisation is ≤ 1 and is called the Stokes vector. We then call a real 4×4 -matrix which maps Stokes vectors onto Stokes vector a pre-Mueller matrix [10]. If a system is mapping pure-states, i.e. fully polarised beams, onto pure-states, i.e. fully polarised beams, then this system is called a Jones system. Note that in a pure-state the degree of polarisation is 1 which means that the Eq. (1) is an equality. Normally such systems are described by 2×2 -complex matrices known as Jones matrices, where the phase and amplitude of the light beam is described by a two-dimensional complex vector. But a Jones matrix J can be also be transferred in to the Mueller calculus by a simple transformation [11]. This transformed Jones matrix is then called a pure Mueller matrix (alternatively a non-depolarising or Jones-Mueller matrix) [12]. Note that a pre-Mueller is not necessarily physically realisable, which was noted in [1]. Consequently, in the same paper a criterion for physical realisability was developed, which is known as Cloude's criterion. The criterion can be phrased as follows; a pre-Mueller matrix M is a Mueller matrix if it can be written as a positive linear combination of Jones-Mueller matrices, i.e. there exists some Jones-Mueller matrices M_i and some positive real parameter c_i with $1 \leq i \leq N$ such that

$$M = \sum_{i=1}^N c_i M_i.$$

That this introduces an additional constraint which cannot be derived from the Stokes criterion, i.e. that not all pre-Mueller matrices are Mueller matrices, can be seen by two Mueller matrix examples in [13]. Now after we having setup a definition of Mueller matrices, we will restate the theorem which establishes the connection between Mueller matrices and Hermitian positive definite matrices. The Theorem implicitly first appeared in [1].

Theorem 1 (Theorem A.1 of [14]) *Every matrix $M \in \mathbb{R}^{4 \times 4}$ with $M = (m_{ij})$ is a Mueller matrix if and only if the Hermitian matrix $H = (h_{ij})$ defined by the following linear equations has non-negative eigenvalues. Moreover, if the matrix H has only one non-zero eigenvalue, then it is a pure Mueller matrix.*

$$\begin{aligned} h_{00} &= \frac{1}{2}(m_{11} + m_{22} + m_{33} + m_{44}), \\ h_{11} &= \frac{1}{2}(m_{11} + m_{22} - m_{33} - m_{44}), \\ h_{22} &= \frac{1}{2}(m_{11} - m_{22} + m_{33} - m_{44}), \\ h_{33} &= \frac{1}{2}(m_{11} - m_{22} - m_{33} + m_{44}) \\ h_{03} &= \frac{1}{2}(m_{14} + m_{41} - im_{23} + im_{32}), \\ h_{30} &= \frac{1}{2}(m_{14} + m_{41} + im_{23} - im_{32}), \\ h_{12} &= \frac{1}{2}(m_{14} - im_{41} + m_{23} + m_{32}), \\ h_{21} &= \frac{1}{2}(-m_{14} + im_{41} + m_{23} + m_{32}) \end{aligned} \quad (2)$$

$$\quad (3)$$

$$\begin{aligned}
 h_{01} &= \frac{1}{2}(m_{12} + m_{21} - im_{34} + im_{43}), \\
 h_{10} &= \frac{1}{2}(m_{12} + m_{21} + im_{34} - im_{43}), \\
 h_{23} &= \frac{1}{2}(m_{12} - im_{21} + m_{34} + m_{43}), \\
 h_{32} &= \frac{1}{2}(-m_{12} + im_{21} + m_{34} + m_{43})
 \end{aligned} \tag{4}$$

$$\begin{aligned}
 h_{02} &= \frac{1}{2}(m_{13} + m_{31} - im_{24} + im_{42}), \\
 h_{20} &= \frac{1}{2}(m_{13} + m_{31} + im_{24} - im_{42}), \\
 h_{13} &= \frac{1}{2}(m_{13} - im_{31} + m_{24} + m_{42}), \\
 h_{31} &= \frac{1}{2}(-m_{13} + im_{31} + m_{24} + m_{42})
 \end{aligned} \tag{5}$$

Note that we altered the result of these linear equations by a factor of 2 in order to simplify oncoming observations, but this does of course not change the validity of the theorem.

What to our knowledge has not yet discussed explicitly about the above result and the above equations is the following simple observation. The whole trick is to realise that the Mueller matrices and Hermitian matrices are vectors and then conclude that the Frobenius inner product coincide with the Hermitian/Euclidean inner product.

Lemma 2 *Let T be a linear map, which is defined by the Eqs. (2), (3), (4) and (5). The linear automorphism T of the Hilbert space $\mathbb{C}^{4 \times 4}$ with the usual Hermitian inner product is unitary. Moreover, the eigenvalues are $\{1, -1\}$ with multiplicity 12 and 4 respectively and has therefore determinant 1.*

Proof: We may assume that we are working in \mathbb{C}^{16} by taking the canonical bijection from $\mathbb{C}^{4 \times 4}$ to \mathbb{C}^{16} . It will be enough now to write down the complex 16×16 -matrix T corresponding to the Eqs. (2), 3, 4 and 5. Compute the eigenvalues of T with your favourite solver and then conclude that $T^\dagger T = TT^\dagger = 1$ where T^\dagger is the conjugate transpose follows as we show in Code 1 “symbolic.ipynb” (Ref. [15]). \square

The nice thing about unitary operators is that they preserve the Hermitian inner product, i.e. we have that $\langle x, y \rangle = \langle T(x), T(y) \rangle$ for any $x, y \in \mathbb{C}^{4 \times 4}$. Hence, the Hermitian norm (which coincides with the euclidean norm, in case there are only real entries) is preserved under these map.

Moreover, by Theorem 1 we know that T maps the set of all Mueller matrices to the set of all positive semidefinite Hermitian matrices. We further investigate some properties of the map T . We denote as \mathcal{C} the \mathbb{R} -vector space of all Hermitian 4×4 -matrices and denote as \mathcal{R} the \mathbb{R} -vector space of all 4×4 - \mathbb{R} -matrices both with the usual trace scalar product. As a side note we believe that it could be of interest to some, that we can define a Lie group structure on 4×4 -Hermitian positive semidefinite matrices corresponding to the non-singular Mueller matrices by defining $A \cdot B = T(T^{-1}(A)T^{-1}(B))$.

Lemma 3 *The restriction $T \upharpoonright \mathcal{R}$ of the linear map T (as defined in Lemma 2) is some non-singular orthogonal linear transformation from \mathcal{R} to \mathcal{C} .*

Proof: We first consider T to be a map from \mathcal{R} to 4×8 - \mathbb{R} -matrices (map the complex numbers to \mathbb{R}^2). Moreover, the space of all Hermitian matrices can be considered a 16-dimensional subspace of $\mathbb{R}^{4 \times 8}$. The orthogonality and non-singularity follows as the eigenvalues of the T are 1, -1 by Lemma 2. \square

Now the next result follows by Theorem 1.

Corollary 4 *The map T is an isometry on $\mathbb{C}^{4 \times 4}$ (with the Hermitian norm) and an isometry between $\mathbb{R}^{4 \times 4}$ and \mathcal{C} (with the Euclidean norm) which maps the set of all Mueller matrices onto the set of all semidefinite matrices.*

3. Optimal filtering revisited

Now we are able to translate the following problem into a question about Hermitian matrices: Given a real 4×4 -matrix (a Mueller matrix one got from a measurement), we ask what the nearest (in terms of the euclidean distance) physically feasible Mueller matrix is. The same holds for the question, which asks for the nearest non-depolarising matrix to a given measurement. Which now by Lemma 2, Lemma 3 and Corollary 4 can be translated to the question; What is the nearest positive semidefinite matrix (with rank 1 in the non-depolarising case) to a given Hermitian matrix. The answer to the first question by implicitly answering the second was already given in [9]. But we can now rely on well-established mathematical theory to show this. We will further derive a more general result and apply it in way such that we derive two additional filters.

Notation; By $[a]$ we denote the diagonal matrix with entries $a_n \leq \dots \leq a_1$ and by \mathcal{U}_n the set of all unitary $n \times n$ -matrices.

The following is true in fact for any unitarily-invariant matrix norm $\|\cdot\|$ such as the Hermitian norm. It can be considered a Hermitian version of Theorem 4.5 of [16].

Theorem 5 *Let c be some fixed real number. Let Y be some non-empty closed subset of*

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \geq \dots \geq x_n \geq c\}.$$

Further, let S_Y be the set

$$\{V^\dagger[d]V : V \in \mathcal{U}_n, d \in Y\}.$$

Given some Hermitian $A = U^\dagger[a]U$ with $a_1 \geq \dots \geq a_n \geq c$ and $U \in \mathcal{U}_n$, we then have that for some $b \in Y$ the following holds

$$\|A - U^\dagger[b]U\| \leq \|A - X\| \quad \text{for all } X \in S_Y.$$

Proof: The proof of Theorem 4.5 in [16] can be easily modified. Conclude in the same way that there exists some $b \in Y$ such that $\|A - U^\dagger[b]U\| \leq \|[a - x]\|$ for any $x \in Y$. Since our matrices are Hermitian, we know then by Theorem 2 of [17] that $\|[a] - [x]\| \leq \|A - X\|$ for any $X \in S_Y$ which has x as its eigenvalues. \square

This Theorem together with Corollary 4 now lets us translate any nearness problems of Mueller matrices into a problem of nearness of the eigenvalues. The following corollary roughly states that any system which limitations can be described as an eigenvalue criterion of the corresponding coherency matrix can be optimally filtered by filtering the eigenvalues of the corresponding coherency matrix.

Corollary 6 *Let c be some fixed real number. Let Y be a non-empty closed set in*

$$\{(x_1, \dots, x_4) \in \mathbb{R}^n : x_1 \geq \dots \geq x_4 \geq c\}$$

and let M be a real 4×4 -matrix such that the eigendecomposition of $T(M)$ is $U^\dagger[a]U$. Let

$$\mathcal{M}_Y = \{T^{-1}(V^\dagger[d]V) : V \in \mathcal{U}_n, d \in Y\}.$$

Then the nearest Mueller matrix in \mathcal{M}_Y in terms of the euclidean norm to M is the matrix $T^{-1}(U^\dagger[d]U)$ where $d \in Y$ is chosen such that $\|d - a\|$ is minimal among all elements of Y .

Now we can now easily conclude that the filtering proposed by Cloude [1] is optimal. For this let M be the measured matrix and let $T(M) = U^\dagger[a]U$ has a minimal eigenvalue of c . We further

assume that $c < 0$ as otherwise we do not have to apply any filter. Let $[a']$ be the tuple where we set all negative eigenvalues of $[a]$ to 0. As the set

$$\{(x_1, \dots, x_4) \in \mathbb{R}^4 : x_1 \geq \dots \geq x_4 \geq 0\}$$

is closed in

$$\{(x_1, \dots, x_4) \in \mathbb{R}^4 : x_1 \geq \dots \geq x_4 \geq c\},$$

Corollary 6 lets us conclude that $T(U^\dagger[a']U)$ is the nearest Mueller matrix estimate of M . Note that \mathcal{M}_Y in this case is the space of a Mueller matrices satisfying Cloude's criterion. In similar fashion we can conclude that setting all but the biggest eigenvalue of $T(M)$ to 0 will give us the best estimate for non-depolarising Mueller matrices.

Let us now consider all Mueller matrices M which can be decomposed as a sum of a non-depolarising part P and a perfectly depolarising matrix D , i.e. D 's only non-zero element is the upper left entry. Now if we map D via T onto the Hermitian matrices we will see that $T(D)$ is a diagonal matrix $[(d \dots d)]$. We continue by stating some simplified version of Weyl's inequality.

Theorem 7 Let A, B, C be Hermitian matrices. If $A + B = C$ and $a_n \leq \dots \leq a_1, b_n \leq \dots \leq b_1$ and $c_n \leq \dots \leq c_1$ be their eigenvalues, then we have that $a_i + b_n \leq c_i \leq a_i + b_1$ for all $1 \leq i \leq n$.

Now in our case, if we take $T(P) = A, T(D) = B$ and $T(M) = C$ and let p_i be the eigenvalues of $T(P)$ and x_i be the eigenvalues of $T(M)$, we get that $p_i + d = x_i$ for all i where $1 \leq i \leq 4$ holds. Of course, it is not logically necessary to apply Weyl's inequality here as B is diagonal under any basis. But we do for the sake of introducing longstanding mathematical results. Hence the matrix $T(M)$ has eigenvalues of the form $x_1 \geq x_2 = x_3 = x_4 \geq 0$. On the other hand, take any matrix H which has eigenvalues of this form. Then subtracting $[x_4, \dots, x_4]$ will give us by applying Weyl's inequality again, that $C = H - [x_4, \dots, x_4]$ has eigenvalues $x_1 - x_4, 0, 0, 0$ and therefore $T^{-1}(C)$ is non-depolarising. Hence we know that $T^{-1}(S_E)$ with $E = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 \geq x_2 = x_3 = x_4 \geq 0\}$ is the set of all Mueller matrices which can be decomposed in a non-depolarising part and perfectly depolarising part. This was already mentioned in [18]. Now asking what the best estimate for a measured matrix, which has this type of composition, can be answered by applying Corollary 6 with E (which is closed). Now if $a_1 \geq a_2 \geq a_3 \geq a_4$ are the eigenvalues of $T(M)$ then $b = (a_1, c, c, c)$ with $c = \frac{1}{3} \sum_{i=2}^4 a_i$ is the best estimate in E . And hence, we have that for a measurement M the best estimate in $T^{-1}(S_E)$ is $T^{-1}(U^\dagger[b]U)$. This also shows us that Equation (17) of [3] is in fact the best *a priori* estimate for the perfectly depolarising part, contrary to what was stated in that paper.

Now we want to preserve the upper left entry of the measured Mueller matrix M , e.g. the change in intensity of an unpolarised beam is believed to be measured correctly. Note that the upper left element m_{11} is equal to the trace of the coherency matrix. And hence as the trace of matrix is equal to the sum of all eigenvalues, we need to find the closest element in

$$E = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1, x_2, x_3, x_4 \geq 0 \wedge x_1 + x_2 + x_3 + x_4 = m_{11}\}.$$

So let a_1, a_2, a_3, a_4 be the sorted eigenvalues of the coherency matrix corresponding to the measured Mueller matrix. We assumed that $a_4 < 0, |a_4| < |a_3|$ and $a_1, a_2, a_3 > 0$. We define the filtered eigenvalues b by first setting $b_4 = 0$. It will be not hard to see, that we find the closest element in E by further setting the filtered eigenvalue to $b_i = a_i + \frac{a_4}{3}$ for all i with $1 \leq i \leq 3$. Now Corollary 6 lets us again conclude $T^{-1}(U^\dagger[b]U)$ is the nearest Mueller matrix estimate. An implementation of this filter and the other two above are made in Code 1 "numeric.ipynb" (Ref. [15]). We also test out this filter on some example measurement and see an 8 to 14 percent improvement against naive scaling of the Mueller matrix and a 7 percent improvement against the maximum-likelihood estimate example of [4].

4. Geometry of the semidefinite cone

The reader may wonder why we could so easily compute the nearest Mueller matrix to a given real matrix or respectively solve the corresponding problem namely the nearest semidefinite matrix to a given Hermitian matrix. This ultimately has to do with the nature of the object consisting of all complex semidefinite matrices. It turns out that this is a deeply studied object which is known under the name *complex semidefinite cone* or more generally symmetric cones and is used among other things for *complex semidefinite programming*. Many very nice properties such as convexity are known about it. In fact, it is a cone. So it is closed under positive linear combinations, i.e. $\alpha H_1 + \beta H_2$ is also positive definite with H_1, H_2 positive definite and α, β positive numbers. Of course, this implies that the set of all Mueller matrices is a cone by linearity of T . It is also known what the interior of this cone is, namely all positive definite matrices and also the boundary of this cones is known, namely all singular positive semidefinite matrices. Moreover, there is a Riemannian metric tensor on its interior (see [19] and Chapter 6 of [20]).

Now the practitioner can use this knowledge and grab ready available tools and mathematical theory. For example, take a subset of Mueller matrices S and a function $f : S \rightarrow \mathbb{R}$ one wants to optimise. We have just seen such functions namely the distance of the Mueller matrices (or certain subsets of them) to a given measurement M . As done before, we can translate the problem by optimising the map $f \circ T^{-1}$ from $T(S)$ to \mathbb{R} instead. This can be either done by finding suitable theory about the semidefinite cone such as Theorem 5 and then solve the problem directly. Or a more general approach would be to use available tools for solving optimisation problems. As a start one would transfer the complex optimisation problem into a real one (with tools as YALMIP [21]). Although voices have been raced to consider optimisation in the complex numbers directly [22]. In any way, there are many available software tools for computing the optimum of a function on the complex or real semidefinite cone such as Manopt [23], Pymanopt [24] and SeDuMi [25].

We highlight one approach of characterising the space of semidefinite matrices of some fixed rank taken from [26] and [27] which is also described in the code of [23] and [24]. Now if the rank is 1 then this space is in correspondence via T with the non-depolarising Mueller matrices. The differential geometry of the non-depolarising Mueller matrices was already studied in [28]. We going to outline now the differential geometry of the Hermitian positive semidefinite cone.

A semidefinite matrix H from $\mathbb{C}^{4 \times 4}$ of rank k can be written as an outer product YY^\dagger of a matrix Y of $\mathbb{C}^{4 \times k}$ of full rank. On the other hand any such outer product YY^\dagger is positive semidefinite and of rank k . Note that the same factorisation was already used in [29], although they did not consider the subtleties of the rank and the oncoming uniqueness properties. As in [26] we define an equivalence relation on $\mathbb{C}^{4 \times k}$ by identifying YU with Y for all unitary matrices U (as the outer product does not change, i.e. $YY^\dagger = YU(YU)^\dagger$). We denote the manifold of all $\mathbb{C}^{4 \times k}$ matrices of full rank as C_{4k} . Now by the quotient manifold theorem (see for example Theorem 9.16 of [30]) the manifold $C_{4k}/U(k)$ is a Riemann quotient manifold, if $U(k)$ is the Lie group of all unitary matrices.

One can note a striking similarity to Cholesky decomposition, which is used in [4] and [5]. In particular, in case of positive definite matrices the Cholesky decomposition is unique and C_{44} can be replaced with all triangular matrices with real diagonal entries. In the case $k < 4$ one can find a unique decomposition after a twisting with permutation matrices [31] and hence one would end up with a finite-to-one map (bounded by 24, the number of 4×4 -permutation matrices). For all k the metric of the manifold is given by the real-trace inner product, if identifying the complex numbers with \mathbb{R}^2 . Moreover, when $k = 1$ we can find a representative of the equivalence classes by requiring that the first non-zero element of the tuple $c \in \mathbb{C}^4$ is a real number. This lets us conclude that its dimension is 7.

Further, if we identify the \mathbb{C} with \mathbb{R}^2 then the quotient manifold theorem tells us also that the dimension of the Riemannian manifold of all complex positive semi-definite matrices of rank k has dimension $4 \cdot k - k^2$. Of course, all this analysis extends to the Mueller matrices by extending

the mapping via T^{-1} . This then implies that the manifold of Mueller matrices is a decomposition of this quotient manifolds $C_{4k}/U(k)$ with $1 \leq k \leq 4$ and the zero element. Furthermore, in case where the Mueller matrices M are assumed to be the sum of a non-depolarising matrix and an ideal depolariser and hence the corresponding coherency matrices $T(M)$ are a sum of a rank-1 positive semidefinite matrix and a diagonal matrix with positive entries, it is not hard to see that this manifold is the product manifold of the positive real numbers \mathbb{R}_+ and the manifold of all complex rank-1 positive semidefinite matrices.

What also can be interfered from the above analysis is the following. We set the above together to receive a map F which is defined as follows

$$\mathbb{R}^{2 \times 4 \times k} \rightarrow_{\mathbb{R}^2 = \mathbb{C}} \mathbb{C}^{4 \times k} \rightarrow_{YY^\dagger} \text{HPSD} \rightarrow_{T^{-1}} \mathcal{M} \quad (6)$$

where *HPSD* is the space of all Hermitian positive semidefinite matrices and \mathcal{M} the space of the Mueller matrices. Now a short calculation gives us then that F is a quadratic homogeneous polynomial and hence any $F(\lambda x) = \lambda^2 F(x)$. In order to do that calculation follow the arrows of Eq. (6). Hence we start with an matrix of symbols of size $\mathbb{R}^{2 \times 4 \times k}$, transform it into a matrix in $\mathbb{C}^{4 \times k}$. Then compute the outer product of this matrix with its conjugate transpose, here the quadratic part of the polynomial comes in. Then applying T^{-1} to this result give us that F is the quadratic homogeneous polynomial. The exact symbolic computation can be made with a computer algebra system as we show in [Code 1](#) “symbolic.ipynb” (Ref. [15]). Moreover, we can see that $\|x\|_{\text{Euclidean}}^2 = F(x)_{11}$, where $F(x)_{11}$ is the upper left element of the Mueller matrix. This means that it is almost always enough to study the reduced case of Mueller matrices which have upper left element 1 or treat it as scalar if one is interested in the polarise-independent loss of light (see [32]). But this of course an practise which is already well established in the field.

Another question which now arises is that of the mean of two or more matrices. In the euclidean space this of course just the standard Arithmetic mean. But in manifolds the geodesic might look very different from a straight line and hence the average of two matrices, i.e. the middle point on the geodesic between these two, might be significantly different from the arithmetic mean. This case of the geometric average of two Mueller matrices was already covered in [33]. The generalisation of this concept namely the Riemannian barycenter of matrices $A_1 \dots A_n$, i.e. the matrix which is the minimum of the function $\sum_{i=1}^n d(X, A_i)$ where d is the distance measure on the manifold. Again we can rely on a well studied area of means of semidefinite linear operators. Studying of the mean of two linear operators began through a study of connections of electrical networks [34]. This was then followed by more axiomatic studies on general Hermitian operators [35], [36]. Means between more than two matrices have been studied in [37]. In [38] means have been studied in case of real semidefinite matrices of fixed rank. An exposition of the geometric nature of means can be found in Chapter 6 of [20]. All together this suggests that computing the mean of multiple Mueller matrices should be done using the Riemannian geometric mean. In practice this would be done by transferring them via T to the semidefinite cone and then using available implementation of the Riemannian mean such as the tool Yalmip [21].

5. Conclusion

We have established a connection between the area of Mueller matrices and the areas of general matrix analysis, Riemannian geometry and optimisation. All basically by interpreting existing results and making the simple observation that the real ambient space of the Hermitian positive semidefinite matrices and the space of Mueller matrices isometrically map onto each other. With this new knowledge, we showed how matrix analysis can be directly used to prove an optimality result (see Corollary 6) for the filtering of measured Mueller matrices.

We further reviewed mathematical results about the complex semidefinite cone and noted how this can be used with our previous results and how this suggests a new mean for Mueller matrices. Of course, such connection have been partly discovered in the past or general results

about semidefinite matrices have been reproved in the special case of Mueller matrices and 4×4 -Hermitian semidefinite matrices. But our connection makes this precise and provides a way to bring well-established mathematical theories and tools into the polarimetric world. One can also speculate that the analysis which we have established here, might bring new insight to quantum optics and quantum information as they share some mathematical objects [4].

What is still missing in our analysis is to bring together this analysis with the study of the Lie group structure of invertible Mueller matrices, or more generally the Lie semigroup structure of all Mueller matrices. Of course, by our analysis of the geometry it is easy now to compute the tangent space at the identity and therefore the Lie algebra. But this is nothing new, the study of the Lie group and Lie algebra was already done in [39]. What is still missing is a study how the geometry of the additive structure of the Mueller matrix, which corresponds to parallel optical elements [40], and the geometry of the multiplicative structure, which corresponds to successive optical elements [40], interact.

Funding

Bundesministerium für Bildung und Forschung; Karlsruher Institut für Technologie; Fraunhofer-Gesellschaft.

Disclosures

The authors declare no conflicts of interest.

References

1. S. R. Cloude, "Conditions For The Physical Realisability Of Matrix Operators In Polarimetry," in *Polarization Considerations for Optical Systems II*, vol. 1166 R. A. Chipman, ed., International Society for Optics and Photonics (SPIE, 1990), pp. 177–187.
2. A. B. Kostinski, "Depolarization criterion for incoherent scattering," *Appl. Opt.* **31**(18), 3506–3508 (1992).
3. F. Le Roy-Bréhonnet, B. Le Jeune, P. Elies, J. Cariou, and J. Lotrian, "Optical media and target characterization by Mueller matrix decomposition," *J. Phys. D: Appl. Phys.* **29**(1), 34–38 (1996).
4. A. Aiello, G. Puentes, D. Voigt, and J. Woerdman, "Maximum-likelihood estimation of Mueller matrices," *Opt. Lett.* **31**(6), 817–819 (2006).
5. S. Faisan, C. Heinrich, G. Sfikas, and J. Zallat, "Estimation of Mueller matrices using non-local means filtering," *Opt. Express* **21**(4), 4424–4438 (2013).
6. F. Boulvert, G. Le Brun, B. Le Jeune, J. Cariou, and L. Martin, "Decomposition algorithm of an experimental Mueller matrix," *Opt. Commun.* **282**(5), 692–704 (2009).
7. G. Anna, F. Goudail, and D. Dolfi, "Optimal discrimination of multiple regions with an active polarimetric imager," *Opt. Express* **19**(25), 25367–25378 (2011).
8. F. Goudail and J. S. Tyo, "When is polarimetric imaging preferable to intensity imaging for target detection?" *J. Opt. Soc. Am. A* **28**(1), 46–53 (2011).
9. J. J. Gil, "On optimal filtering of measured Mueller matrices," *Appl. Opt.* **55**(20), 5449–5455 (2016).
10. B. N. Simon, S. Simon, N. Mukunda, F. Gori, M. Santarsiero, R. Borghi, and R. Simon, "A complete characterization of pre-Mueller and Mueller matrices in polarization optics," *J. Opt. Soc. Am. A* **27**(2), 188–199 (2010).
11. R. Simon, "The connection between Mueller and Jones matrices of polarization optics," *Opt. Commun.* **42**(5), 293–297 (1982).
12. J. J. Gil and R. Ossikovski, *Polarized Light and the Mueller Matrix Approach* (CRC Press, 2017).
13. R. Simon, "Mueller matrices and depolarization criteria," *J. Mod. Opt.* **34**(4), 569–575 (1987).
14. C. Van Der Mee, "An eigenvalue criterion for matrices transforming stokes parameters," *J. Math. Phys.* **34**(11), 5072–5088 (1993).
15. T. Zander, "Logikerkit/muellerconefilter (Jupyter/iPython notebooks)," *Zenodo* (2020) [retrieved 7 May 2020], <https://doi.org/10.5281/zenodo.3813681>.
16. C.-K. Li and N.-K. Tsing, "On the unitarily invariant norms and some related results," *Linear Multilinear Algebra* **20**(2), 107–119 (1987).
17. H. Wielandt, "An extremum property of sums of eigenvalues," *Proc. Am. Math. Soc.* **6**(1), 106 (1955).
18. R. Ossikovski, M. Anastasiadou, S. Ben Hatit, E. Garcia-Caurel, and A. De Martino, "Depolarizing Mueller matrices: how to decompose them?" *Phys. Status Solidi A* **205**(4), 720–727 (2008).
19. R. D. Hill and S. R. Waters, "On the cone of positive semidefinite matrices," *Linear Algebra. its Appl.* **90**, 81–88 (1987).

20. R. Bhatia, *Positive definite matrices*, Princeton series in applied mathematics (Princeton University Press, 2007), illustrated edition ed.
21. J. Löfberg, “Yalmip : A toolbox for modeling and optimization in matlab,” in *In Proceedings of the CACSD Conference* (Taipei, Taiwan, 2004).
22. J. C. Gilbert and C. Jozs, “Plea for a semidefinite optimization solver in complex numbers,” Research report, Inria Paris (2017).
23. N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, “Manopt, a Matlab toolbox for optimization on manifolds,” *Journal of Machine Learning Research* **15**, 1455–1459 (2014).
24. J. Townsend, N. Koep, and S. Weichwald, “Pymanopt: A python toolbox for optimization on manifolds using automatic differentiation,” *Journal of Machine Learning Research* **17**, 1–5 (2016).
25. J. F. Sturm, “Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones,” *Optim. methods software* **11**(1-4), 625–653 (1999).
26. S. Yatawatta, “Radio interferometric calibration using a riemannian manifold,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, (2013), pp. 3866–3870.
27. B. Vandereycken, P.-A. Absil, and S. Vandewalle, “Embedded geometry of the set of symmetric positive semidefinite matrices of fixed rank,” *2009 IEEE/SP 15th Workshop on Statistical Signal Processing* (IEEE, 2009), pp. 389–392.
28. V. Devlaminck and P. Terrier, “Geodesic distance on non-singular coherency matrix space in polarization optics,” *J. Opt. Soc. Am. A* **27**(8), 1756–1763 (2010).
29. C. J. Sheppard, A. Le Gratiet, and A. Diaspro, “Factorization of the coherency matrix of polarization optics,” *J. Opt. Soc. Am. A* **35**(4), 586–590 (2018).
30. J. M. Lee, *Introduction to Smooth Manifolds*, Vol. 218 of Graduate Texts in Mathematics (Springer, 2013).
31. N. J. Higham, “Analysis of the Cholesky decomposition of a semi-definite matrix,” in *Reliable numerical computation*, (Oxford Univ. Press, New York, 1990), Oxford Sci. Publ., pp. 161–185.
32. T. Eftimov, “Müller matrix analysis of pdl components,” *Fiber Integr. Opt.* **23**(6), 453–466 (2004).
33. V. Devlaminck, “Mueller matrix interpolation in polarization optics,” *J. Opt. Soc. Am. A* **27**(7), 1529–1534 (2010).
34. W. N. Anderson Jr and R. J. Duffin, “Series and parallel addition of matrices,” *J. Math. Analysis Appl.* **26**(3), 576–594 (1969).
35. F. Kubo and T. Ando, “Means of positive linear operators,” *Math. Ann.* **246**(3), 205–224 (1980).
36. W. Pusz and S. L. Woronowicz, “Functional calculus for sesquilinear forms and the purification map,” *Rep. Math. Phys.* **8**(2), 159–170 (1975).
37. T. Ando, C.-K. Li, and R. Mathias, “Geometric means,” *Linear Multilinear Algebr.* **385**, 305–334 (2004).
38. S. Bonnabel and R. Sepulchre, “Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank,” *SIAM J. on Matrix Analysis Appl.* **31**(3), 1055–1070 (2010).
39. V. Devlaminck and P. Terrier, “Definition of a parametric form of nonsingular Mueller matrices,” *J. Opt. Soc. Am. A* **25**(11), 2636–2643 (2008).
40. N. G. Parke, “Optical algebra,” *J. Math. Phys.* **28**(1-4), 131–139 (1949).