

ON THE CAYLEY TRANSFORM OF POSITIVITY CLASSES OF MATRICES*

SHAUN M. FALLAT[†] AND MICHAEL J. TSATSOMEROS[‡]

Abstract. The Cayley transform of A, $F = (I+A)^{-1} (I-A)$, is studied when A is a P-matrix, an M-matrix, an inverse M-matrix, a positive definite matrix, or a totally nonnegative matrix. Given a matrix A in each of these positivity classes and using the fact that the Cayley transform is an involution, properties of the ensuing factorization $A = (I+F)^{-1} (I-F)$ are examined. Specifically, it is investigated whether these factors belong to the same positivity class as A and, conversely, under what conditions on these factors does A belong to one of the above positivity classes.

Key words. Cayley transform, *P*-matrices, *M*-matrices, Positive definite matrices, Totally nonnegative matrices, Stable matrices, Matrix factorizations.

AMS subject classifications. 15A23, 15A24, 15A48

1. Introduction. An *n*-by-*n* complex matrix A ($A \in M_n(\mathbb{C})$) is called a *P*-matrix if every principal minor of *A* is positive. Among the *P*-matrices are several important and well-studied matrix subclasses (see [5, 6] for more information on these classes of matrices): An (invertible) *M*-matrix is a real *P*-matrix all of whose off-diagonal entries are non-positive; an *inverse M*-matrix is the inverse of an *M*-matrix and hence a *P*-matrix itself; a (Hermitian) *positive definite* matrix is simply a Hermitian *P*-matrix. Finally, a *P*-matrix all of whose non-principal minors are nonnegative is known as an (invertible) *totally nonnegative* matrix (see [1, 3]).

Our interest here lies in considering the Cayley transform of matrices in the positivity classes above. We begin by making precise the term Cayley transform.

DEFINITION 1.1. Let $A \in M_n(\mathbb{C})$ such that I + A is invertible. The Cayley transform of A, denoted by $\mathcal{C}(A)$, is defined to be

$$C(A) = (I + A)^{-1}(I - A).$$

The Cayley transform, not surprisingly, was defined in 1846 by Cayley. He proved that if A is skew-Hermitian, then $\mathcal{C}(A)$ is unitary and conversely, provided of course that $\mathcal{C}(A)$ exists. This feature is useful e.g., in solving matrix equations subject to the solution being unitary by transforming them into equations for skew-Hermitian matrices. One other important feature of the Cayley transform is that it can be viewed as an extension to matrices of the conformal mapping

$$T(z) = \frac{1-z}{1+z}$$

 $^{^{*}}$ Received by the editors on 22 February 2002. Final manuscript accepted for publication on 28 August 2002. Handling Editor: Bryan L. Shader.

[†] Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, S4S 0A2 (sfallat@math.uregina.ca). Research supported in part by an NSERC research grant.

[‡] Mathematics Department, Washington State University, Pullman, WA 99164-3113 (tsat@math.wsu.edu). Research supported in part by an NSERC research grant.

The Electronic Journal of Linear Algebra ISSN 1081-3810 A publication of the International Linear Algebra Society Volume 9, pp. 190-196, September 2002



S.M. Fallat and M.J. Tsatsomeros

from the complex plane into itself. In this regard, Stein [9] and Taussky [10] both considered the Cayley transform, for the most part indirectly, when they provided connections between *stable matrices* (i.e., matrices for which $\operatorname{Re}(\lambda) < 0$ for all eigenvalues λ) and *convergent matrices* (i.e., those matrices A for which $\operatorname{Im}_{n\to\infty} A^n = 0$). In both of these papers the key connection came via Lyapunov's equation, $AG + GA^* = -I$, and the Cayley transform. The use of the Cayley transform for stable matrices was made explicit in the paper by Haynes [4] in 1991. He proved that a matrix B is convergent if and only if there exists a stable matrix A such that B = C(-A).

In this paper, given a matrix A in each of the aforementioned positivity classes, we examine properties of its Cayley transform F = C(A). Specifically, since A can be factored into $A = (I + F)^{-1}(I - F)$, we investigate whether the factors $(I + F)^{-1}$ and (I - F) belong to same positivity class as A and, conversely, under what conditions on these factors does A belong to one of these positivity classes. Our interest in this topic grew out of work on factorizations of the form $A = X^{-1}Y$, where X and Y have certain properties such as diagonal dominance and stability. We obtain results of this type by using the fact that the Cayley transform is an involution and by employing the factorization of A in terms of its Cayley transform. The analysis for P-matrices is in section 3 and the other classes are treated in section 4. Section 2 contains some definitions and auxiliary results that are used frequently throughout the paper.

2. Preliminaries. For $A = [a_{ij}] \in M_n(\mathbb{C})$, we let $\sigma(A)$ and $\rho(A)$ denote the spectrum and spectral radius of A, respectively, and for $\alpha \subseteq \{1, 2, \ldots, n\}$ we let $A[\alpha]$ denote the principal submatrix of A lying in rows and columns indexed by α . A *Z*-matrix is a matrix all of whose off-diagonal entries are non-positive, and an essentially nonnegative matrix is the negative of a *Z*-matrix.

We continue with two basic lemmas regarding the Cayley transform.

LEMMA 2.1. Let $A \in M_n(\mathbb{C})$ such that $-1 \notin \sigma(A)$ and set $F = \mathcal{C}(A)$. Then,

$$A = \mathcal{C}(F) = (I + F)^{-1}(I - F).$$

Proof. As $F = \mathcal{C}(A)$, we have (I + A)F = I - A or, equivalently,

$$A(I+F) = I - F. (2.1)$$

Now notice that if Fx = -x, then x = 0; that is, $-1 \notin \sigma(F)$. Thus by (2.1) and since $(I + F)^{-1}$ and I + F commute (being rational functions of F), it follows that $A = \mathcal{C}(F)$. \square

LEMMA 2.2. Let $A \in M_n(\mathbb{C})$ such that $-1 \notin \sigma(A)$ and set $F = \mathcal{C}(A)$. Then,

$$I + F = 2(I + A)^{-1}.$$
(2.2)

If, in addition, A is invertible, then

$$I - F = 2(I + A^{-1})^{-1}.$$
(2.3)

The Electronic Journal of Linear Algebra ISSN 1081-3810 A publication of the International Linear Algebra Society Volume 9, pp. 190-196, September 2002



Cayley Transform and Positivity Classes

Proof. As $F = \mathcal{C}(A)$, we have

$$I + F = I + (I + A)^{-1}(I - A) = (I + A)^{-1}(I + A + I - A) = 2(I + A)^{-1}.$$

Similarly, $I - F = 2(I + A)^{-1}A$. So if A is invertible,

$$I - F = 2(A^{-1}(I + A))^{-1} = 2(I + A^{-1})^{-1}$$

as claimed. \blacksquare

Finally, notice that if $F = \mathcal{C}(A)$, then

$$\lambda \in \sigma(A) \iff \lambda = \frac{1-\mu}{1+\mu}$$
; for some $\mu \in \sigma(F)$.

That is, we have a conformal and univalent mapping $z \to \frac{1-z}{1+z}$ from the spectrum of A onto the spectrum of F and conversely. In particular, this mapping transforms the right half-plane onto the unit disc. Thus -A is stable if and only if $\rho(F) < 1$, in which case we refer to A as a *positive stable* matrix.

3. Cayley transforms of *P*-matrices. In this section we pursue the relation of *P*-matrices and Cayley transforms.

THEOREM 3.1. Let $A \in M_n(\mathbb{C})$ be a *P*-matrix. Then $F = \mathcal{C}(A)$ is well-defined and both I - F and I + F are *P*-matrices. In particular, $A = (I + F)^{-1}(I - F)$ is a factorization of a *P*-matrix into (commuting) *P*-matrices.

Proof. First, since A is a P-matrix, A is nonsingular and has no negative real eigenvalues. Hence F = C(A) is well-defined. By Lemmas 2.1 and 2.2 and as addition of positive diagonal matrices and inversion are operations that preserve P-matrices, it follows that I - F and I + F are (commuting) P-matrices. \Box

One consequence of the above result is that if A is a P-matrix, then the main diagonal entries of the matrix F = C(A) all have absolute value less than one.

The converse to Theorem 3.1 is not true in general, as the following example demonstrates.

EXAMPLE 3.2. Let

$$F = \left[\begin{array}{rrrr} 0 & 1 & 1.1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{array} \right].$$

Then

$$I - F = \begin{bmatrix} 1 & -1 & -1.1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } (I + F)^{-1} = \begin{bmatrix} .4762 & -.5 & .0238 \\ 0 & .5 & -.5 \\ .4762 & 0 & .4762 \end{bmatrix}$$

are both *P*-matrices. However,

$$(I+F)^{-1}(I-F) = \begin{bmatrix} -.0476 & -1 & -.0476 \\ 0 & 0 & -1 \\ .9524 & 0 & -.0476 \end{bmatrix}$$



S.M. Fallat and M.J. Tsatsomeros

is not a *P*-matrix.

To obtain a characterization of a *P*-matrix in terms of its Cayley transform, we need the following lemma, which is known for real matrices (see e.g., [6, 2.5.6.5, p. 120]). The extension below to complex matrices with self-conjugate spectra is straightforward. The spectrum of a complex matrix A is self-conjugate if $\sigma(A) = \overline{\sigma(A)}$.

LEMMA 3.3. Let $B \in M_n(\mathbb{C})$ so that $\sigma(B[\alpha]) = \sigma(B[\alpha])$ for all $\alpha \subseteq \{1, 2, ..., n\}$. Then B is a P-matrix if and only if every real eigenvalue of every principal submatrix of B is positive.

REMARK 3.4. Based on Lemma 3.3, Theorem 3.1 can be interpreted as follows: If A is a P-matrix, then $(I + F)^{-1}$ is a P-matrix and thus every real eigenvalue of every principal submatrix of $(I + F)^{-1}$ is positive. The following result states that a stronger condition on the real eigenvalues of the principal submatrices of $(I + F)^{-1}$ is necessary and sufficient for A to be a P-matrix.

THEOREM 3.5. Let $A \in M_n(\mathbb{C})$ such that $\sigma(A[\alpha]) = \overline{\sigma(A[\alpha])}$ for all $\alpha \subseteq \{1, 2, ..., n\}$ and $-1 \notin \sigma(A)$. Let $F = \mathcal{C}(A)$. Then A is a P-matrix if and only if every real eigenvalue of every principal submatrix of $(I + F)^{-1}$ is greater than 1/2.

Proof. In view of Lemma 2.1 and by reversing the roles of A and F in Lemma 2.2, we obtain $A = 2(I + F)^{-1} - I$. The result now follows by applying Lemma 3.3 to $2(I + F)^{-1} - I$. \square

Notice that in Example 3.2 all 1-by-1 and 2-by-2 principal submatrices of $(I+F)^{-1}$ fail to satisfy the condition in Theorem 3.5.

Recall now the following characterization of real *P*-matrices. THEOREM 3.6. [7] Let $B, G \in M_n(\mathbb{R})$. The set of matrices

$$\{BT + G(I - T) : T = \text{diag}(t_1, \dots, t_n), t_i \in [0, 1] \ (1 \le i \le n)\}$$

contains only nonsingular matrices if and only if $G^{-1}B$ is a P-matrix.

Based on the above result, we can prove the following necessary and sufficient condition for real P-matrices; see [8, Theorem 4.1] for a related result.

THEOREM 3.7. Let $A \in M_n(\mathbb{R})$. Then A is a P-matrix if and only if F = C(A) is well-defined and I - FD is nonsingular for all diagonal matrices $D = (d_{ij})$ with $-1 \leq d_{ii} \leq 1$ $(1 \leq i \leq n)$.

Proof. By our discussion so far, since A is a P-matrix, we have A = C(F) and F = C(A). By Theorem 3.6 applied to the factorization $A = (I + F)^{-1}(I - F)$, we have that (I - F)T + (I + F)(I - T) is nonsingular for all diagonal matrices $T = \text{diag}(t_1, \ldots, t_n), t_i \in [0, 1]$ $(1 \le i \le n)$. Hence I - F(2T - I) is nonsingular for all such diagonal matrices T, from which the conclusion follows. \Box

We conclude the section with a comment on a stabilization of *P*-matrices. Let *A* be a *P*-matrix and F = C(A). By Theorem 3.1, I + F and I - F are also *P*-matrices. Consequently, by a result of Ballantine [2], there exist positive diagonal matrices D, E such that (I + F)D and (I - F)E are positive stable. It follows that

$$D^{-1}AE = [(I+F)D]^{-1} [(I-F)E)]$$



Cayley Transform and Positivity Classes

is a factorization of the *P*-matrix $D^{-1}AE$ into a product of two positive stable *P*-matrices.

4. Cayley transforms of other positivity classes. In this section, we look at Cayley transforms of *M*-matrices, inverse *M*-matrices, positive definite matrices and totally nonnegative matrices.

THEOREM 4.1. Let $A \in M_n(\mathbb{R})$. Then A is an M-matrix if and only if $F = \mathcal{C}(A)$ is well-defined, and both $(I + F)^{-1}$ and I - F are M-matrices.

Proof. If A is an M-matrix, then A is a P-matrix and thus F = C(A) is welldefined. By Lemma 2.2 and the fact that M-matrices and inverse M-matrices are preserved by the addition of positive diagonal matrices, it follows that $(I + F)^{-1} = \frac{1}{2}(I + A)$ and $I - F = 2(I + A^{-1})^{-1}$ are both M-matrices.

For the converse, suppose that $(I + F)^{-1}$ and I - F are *M*-matrices and that $F = \mathcal{C}(A)$ is well-defined. As in the proof of Theorem 3.5, $A = 2(I + F)^{-1} - I$ and thus *A* is a *Z*-matrix and nonsingular. For the sake of a contradiction assume that *A* is not an *M*-matrix. Then we have A = tI - B, where *B* is a nonnegative matrix and $0 \le t < \rho(B)$. Hence *A* has a negative real eigenvalue, namely $\lambda = t - \rho(B)$. Since $F = \mathcal{C}(A)$ we know that any eigenvalue of *A* is equal to $(1 - \mu)/(1 + \mu)$, where μ is some eigenvalue of *F*. Now there are two cases to consider: 1) $-1 < \lambda < 0$; or 2) $\lambda < -1$. It is straightforward to verify that if 1) holds, then $\mu > 1$ and real, otherwise if 2) holds, then $\mu < -1$ and real. However, since both $(I + F)^{-1}$ and I - F are *M*-matrices, any eigenvalue μ of *F* satisfies $|\text{Re}(\mu)| < 1$. Thus we have reached a contradiction. Hence *A* is an *M*-matrix. \square

We note here in passing that if A is an M-matrix and F = C(A), then $\rho(F) < 1$ follows from properties of the conformal mapping between the eigenvalues of A and F and the fact that M-matrices are positive stable.

The proof of the next result is similar to the previous one and thus is omitted.

THEOREM 4.2. Let $A \in M_n(\mathbb{R})$. Then A is an inverse M-matrix if and only if $F = \mathcal{C}(A)$ is well-defined, and both $(I + F)^{-1}$ and I - F are inverse M-matrices.

It is worth noting that if A is an M-matrix, then [-I + F, I + F] is a matrix interval from minus an M-matrix to an inverse M-matrix. In particular, the interval [-I+F, I+F] is an interval of nonsingular essentially nonnegative matrices. Similarly, if A is an inverse M-matrix, then [-I+F, I+F] is an interval from minus an inverse M-matrix to an M-matrix. In particular, the interval [-I+F, I+F] is an interval of nonsingular z-matrices.

Recall that $X = [x_{ij}] \in M_n(\mathbb{C})$ is called *strictly column diagonally dominant* if $|x_{jj}| > \sum_{i \neq j} |x_{ij}| \quad (j = 1, 2, ..., n)$. The proof of the following result is based on the fact that the rows of *M*-matrices can be scaled via left multiplication by a positive diagonal matrix to become strictly column diagonally dominant.

THEOREM 4.3. Let $A \in M_n(\mathbb{R})$ be an *M*-matrix. There exist a > 0 and positive diagonal matrix *D* such that $F_a = \mathcal{C}(aA)$ is well-defined and

$$A = [a D(I + F_a)]^{-1} [D(I - F_a))]$$

is a factorization of A into the product of the inverse of a strictly column diagonally dominant M-matrix and a strictly column diagonally dominant M-matrix.

The Electronic Journal of Linear Algebra ISSN 1081-3810 A publication of the International Linear Algebra Society Volume 9, pp. 190-196, September 2002



Proof. The Cayley transform $F_a = C(aA) = [f_{ij}]$ is well-defined as aA is a P-matrix for all a > 0. Choose a > 0 large enough so that the diagonal entries of $(I + aA)^{-1}$ are less then 1/2. This is possible as aA and I + aA are *M*-matrices for all a > 0 and thus their inverses exist and are entrywise nonnegative. It follows from Lemma 2.2 that the diagonal entries of

$$F_a = 2(I + aA)^{-1} - I$$

are negative. Next, as $(I + F_a)$ is by Theorem 4.1 an *M*-matrix, we can choose $D = \text{diag}(d_1, d_2, \ldots, d_n), d_i > 0$, such that $D(I + F_a)$ is strictly column diagonally dominant. Notice then that as $f_{jj} < 0$ and as $|f_{jj}| < 1$ (see comment following Theorem 3.1), we have

$$d_j(1-f_{jj}) > d_j(1+f_{jj}) > \sum_{i \neq j} |d_j f_{ij}| \quad (j = 1, 2, ..., n);$$

that is, both $D(I-F_a)$ and $D(I+F_a)$ are strictly column diagonally dominant. Thus,

$$aA = (I + F_a)^{-1} (I - F_a) = [D(I + F_a)]^{-1} [D(I - F_a)]$$

provides the claimed factorization of A.

We continue with positive definite and totally nonnegative matrices.

THEOREM 4.4. Let $A \in M_n(\mathbb{C})$. Then A is a positive definite matrix if and only if $F = \mathcal{C}(A)$ is well-defined, and both I + F and I - F are positive definite matrices.

Proof. If A is positive definite, then by Lemma 2.2 and properties of positive definite matrices, $(I + F) = 2(I + A)^{-1}$ and $I - F = 2(I + A^{-1})^{-1}$ are both positive definite matrices.

For the converse, observe that if I + F and I - F are positive definite, then $I + A^{-1}$ and I + A are positive definite. Hence $A = A^*$ and $\sigma(A) \subset \mathbb{R}$. Also, if $\lambda \in \sigma(A)$, then $1 + \lambda > 0$ and $1 + 1/\lambda > 0$, which implies $\lambda > 0$. Thus A is positive definite. \square

For totally nonnegative matrices, a complete picture of the properties of Cayley transforms is currently unknown. However, we do have the following. Recall first that if A is totally nonnegative and invertible, then $SA^{-1}S$ is totally nonnegative for $S = \text{diag}(1, -1, \ldots, (-1)^{n-1})$ (see [3, p. 109]). In particular, if A is a tridiagonal M-matrix, then A^{-1} is a totally nonnegative matrix. (In the symmetric tridiagonal case A^{-1} is sometimes referred to as a Green's matrix.) Recall also that a square matrix A is called *irreducible* if there does not exist a permutation matrix P such that

$$PAP^T = \left[\begin{array}{cc} A_{11} & A_{12} \\ 0 & A_{22} \end{array} \right],$$

where A_{11} and A_{22} are square non-vacuous matrices.

THEOREM 4.5. Let $A \in M_n(\mathbb{R})$ be an irreducible matrix. Then A is an essentially nonnegative tridiagonal matrix with $\rho(A) < 1$ if and only if I + A and $(I - A)^{-1}$ are totally nonnegative matrices. In particular, $C(-A) = (I - A)^{-1}(I + A)$ is a factorization into totally nonnegative matrices.



Cayley Transform and Positivity Classes

Proof. To verify necessity, observe that if I + A is totally nonnegative, then A is certainly essentially nonnegative. Also, since $(I - A)^{-1}$ is totally nonnegative, it follows that I - A is invertible and has the checkerboard sign pattern (i.e., the sign of its (i, j)-th entry is $(-1)^{i+j}$). Hence $a_{i,i+2} = 0$ and $a_{i+2,i} = 0$ for all $i \in \{1, 2, \ldots, n-2\}$, and since A is irreducible and I + A is totally nonnegative, $a_{i,j} = 0$ for |i - j| > 1. That is, A is tridiagonal. The remaining conclusion now readily follows.

For the converse, if A is an essentially nonnegative irreducible tridiagonal matrix with $\rho(A) < 1$, then I + A is a nonnegative tridiagonal P-matrix and thus totally nonnegative (see [3, p. 117]). Similarly, I - A is a tridiagonal M-matrix since $\rho(A) < 1$, and hence $(I - A)^{-1}$ is totally nonnegative (see [3, p. 109]). Since totally nonnegative matrices are closed under multiplication, $\mathcal{C}(-A) = (I - A)^{-1}(I + A)$ is totally nonnegative. \Box

A natural question arising from Theorem 4.5 is whether in every factorization $\hat{F} = (I - A)^{-1}(I + A)$ of a totally nonnegative matrix \hat{F} the factors $(I - A)^{-1}$ and (I + A) are totally nonnegative or not. We conclude with an example showing that neither of these factors need be totally nonnegative.

EXAMPLE 4.6. Consider the totally nonnegative matrix

$$\hat{F} = \left[\begin{array}{rrr} 1 & .9 & .8 \\ .9 & 1 & .9 \\ 0 & .9 & 1 \end{array} \right]$$

and consider $A = -\mathcal{C}(\hat{F})$. Then $\hat{F} = \mathcal{C}(-A) = (I - A)^{-1}(I + A)$, where neither

	1	.45	.4			.8203	.3994	.2922
$(I-A)^{-1} =$.45	1	.45	nor	I + A =	.6657	.5207	.3994
	0	.45	1			2996	.6657	.8203

is totally nonnegative.

REFERENCES

- [1] T. Ando. Totally positive matrices, Linear Algebra Appl., 90:165–219, 1987.
- [2] C.S. Ballantine. Stabilization by a diagonal matrix. Proc. Amer. Math. Soc., 25:729-734, 1970.
- [3] F.R. Gantmacher and M.G. Krein. Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen Mechanischer Systeme. Akademie-Verlag, Berlin, 1960.
- [4] T. Haynes. Stable matrices, the Cayley transform, and convergent matrices. Internat. J. Math. Math. Sci., 14:77–81, 1991.
- [5] R.A. Horn and C.R. Johnson. *Matrix Analysis*, Cambridge University Press, New York, 1985.
 [6] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, New York, 1991.
- [7] C.R. Johnson and M.J. Tsatsomeros. Convex sets of nonsingular and P-Matrices. Linear and Multilinear Algebra, 38:233–239, 1995.
- [8] S. Rump. On P-matrices. To appear in Linear Algebra Appl.
- P. Stein. On the ranges of two functions of positive definite matrices. J. Algebra, 2:350–353, 1965.
- [10] O. Taussky. Matrices C with $C^n \to 0$. J. Algebra, 1:5–10, 1964.