Internat. J. Math. & Math. Sci.
VOL. 14 NO. 1 (1991) 77-82

STABLE MATRICES, THE CAYLEY TRANSFORM, AND CONVERGENT MATRICES

TYLER HAYNES

Mathematics Department
Saginaw Valley State University
University Center, Michigan 48710

(Received January 17, 1990)

ABSTRACT. The main result is that a square matrix D is convergent ($\lim_{n\to\infty} D^n = 0$) if and only if it is the Cayley transform $C_A = (I-A)^{-1}(I+A)$ of a stable matrix A, where a stable matrix is one whose characteristic values all have negative real parts. In passing, the concept of Cayley transform is generalized, and the generalized version is shown closely related to the equation AG + GB = D. This gives rise to a characterization of the non-singularity of the mapping $X \to AX + XB$. As consequences are derived several characterizations of stability (closely related to Lyapunov's result) which involve Cayley transforms.

KEY WORDS AND PHRASES. Stable matrix, Cayley transform, convergent matrix.
1980 AMS SUBJECT CLASSIFICATION CODES. 15A04, 15A24

Both Taussky and Stein [Stein, 1965] have written on the connection between stable matrices and convergent matrices. The link joining the two is the Cayley transform: a matrix is convergent $\Leftrightarrow$ it is the Cayley transform of a stable matrix (theorem 8).

Cayley transforms are introduced by considering the matrix equation AX+XB = C. But first a lemma:

Lemma 1: Over field F let matrix A be nxn and let x be either indeterminate over F or in F but not a characteristic value of A. Then

$$(xI-A)^{-1}(xI+A) = (xI+A)(xI-A)^{-1}. \tag{1}$$

If either expression in (1) is denoted by $C_{A,x}$, then $C_{A',x}=C'_{A,x}$. If $x \neq 0$, then

$$A = x(C_{A,x}-I)(C_{A,x}+I)^{-1}. \tag{2}$$

Proof: Since x is not a characteristic value of A, $(xI-A)^{-1}$ exists. (1) follows from

$$(xI+A)(xI-A) = (xI-A)(xI+A). \tag{3}$$

Before (2) can be derived, the non-singularity of $C_{A,x} + I$ must be proven. This equation holds:

$$C_{A,x} + I = (xI+A)(xI-A)^{-1} + (xI-A)(xI-A)^{-1}$$
$$= 2x(xI-A)^{-1}.$$

Therefore, $|C_{A,x} + I| = 2x|xI-A|^{-1} \neq 0$ since $x \neq 0$ and $|xI-A| \neq 0$ (for xI-A is non-singular); hence, $C_{A,x} + I$ is non-singular. (2) then follows directly. QED

$C_{A,x}$ of (1) is the generalized Cayley transform of A. If x = 1 is not a characteristic value of A, then $C_{A,1}$ is the Cayley transform of A; it will be denoted C_A. Note that the mapping $A \to C_A$ is bijective from the set of matrices having no characteristic value = 1 onto those having no characteristic value = -1, the inverse transformation being determined by (2).

Theorem 2*: Let matrix A be mxm, G and D be mxn, and B be nxn, all with entries in field F.

$$AG+GB = D \Leftrightarrow G-C_{A,x}GC_{B,x} = -2x(xI_m-A)^{-1}D(xI_n-B)^{-1}, \tag{4}$$

where x is either indeterminate over F or in F but $\neq 0$ and a characteristic value of neither A nor B.

Proof: x satisfies the requirements for $C_{A,x}$ and $C_{B,x}$ to exist, according to the lemma, and the dimensions of $C_{A,x}$, $C_{B,x}$, $(xI_m-A)^{-1}$, and $(xI_n-B)^{-1}$ are such that the expression on the right of (4) is well-defined.

$$AG+GB = D$$
$$\Leftrightarrow \quad (xG-AG)(xI_n-B) - (xG+AG)(xI_n+B) = -2xD$$
$$\Leftrightarrow \quad (xI_m-A)G(xI_n-B) - (xI_m+A)G(xI_n+B) = -2xD$$
$$\Leftrightarrow \quad G-(xI_m-A)^{-1}(xI_m+A)G(xI_n+B)(xI_n-B)^{-1} = -2x(xI_m-A)^{-1}D(xI_n-B)^{-1}$$
$$\Leftrightarrow \quad G-C_{A,x}GC_{B,x} = -2x(xI_m-A)^{-1}D(xI_n-B)^{-1}$$
QED

One consequence of the preceding theorem is the celebrated result that every properly orthogonal** matrix P can be expressed as $P = (I+K)^{-1}(I-K)$, where K is a real skew matrix. To derive it, in the theorem let F = real number field, G = I, D = O, x = -1, and B = A'. Then it follows that $A+A' = 0 \Leftrightarrow PP' = I$, where $P = (-I-A)^{-1}(-I+A) = (I+A)^{-1}(I-A)$, the relationship between P and A being determined by (1) and (2) of the lemma (cf. the remark on the bijective character of $A \to C_A$). Likewise the Cayley parametrization of unitary matrices follows [Gantmacher, Vol. I; p. 279 (95)].

Over a field F let A be an mxm matrix, X an mxn matrix and B an nxn matrix. Let $\mathcal{L}_{A,B} = AX + XB$. Clearly the mapping $\mathcal{L}_{A,B}$: $X \to AX + XB$ is a linear transformation on the

---

*This theorem generalizes a lemma of Weyl's [Weyl: p. 57, lemma (2.10.A)].

**An orthogonal matrix is proper $\Leftrightarrow$ none of its characteristic values = -1.

linear space of mxn matrices. Denote $\mathcal{L}_{A,A^*}$ by $\mathcal{L}_A$: $\mathcal{L}_A(X) = AX + XA^*$, where all matrices are of the same dimension.

Corollary 3: Let A, B, G, x, and F be as in theorem 2. Then the mapping $G \to G - C_{A,x}GC_{B,x}$ is linear from the set of all mxn matrices into itself. This mapping is non-singular $\Leftrightarrow$ $\mathcal{L}_{A,B}$ is non-singular.

Proof: The linearity of the mapping is obvious. $\mathcal{L}_{A,B}$ is non-singular $\Leftrightarrow$ for every D there exists a solution of $AX+XB = D$ $\Leftrightarrow$ for every E there exists a solution of $X - C_{A,x}XC_{B,x} = E$ (theorem 2 and the non-singularity of $xI_m - A$ and $xI_n - B$) $\Leftrightarrow$ the mapping $G \to G - C_{A,x}GC_{B,x}$ is non-singular. QED

In the rest of this article, let F be the field of complex numbers and let all matrices be square.

The inertia of an nxn matrix X is the ordered triple of integers $(\pi(X), \nu(X)\ \delta(X)) = \mathrm{In}(X)$, where $\pi(X)$ is the number of characteristic values of X whose real parts are positive, $\nu(X)$, the number whose real parts are negative, and $\delta(X)$ the number whose real parts are 0.

Corollary 4: If A has no characteristic value =1, then $\mathrm{In}(I - C_A C_A^*) = \mathrm{In}(-(A+A^*))$.

Proof: $C_{A^*} = C_A^*$ by a slight modification of lemma 1. In theorem 2, let $B = A^*$, $G = I$, and $x = 1$; then $D = A+A^*$. Therefore, $I - C_AC_A^* = I - C_AIC_{A^*} = -2(I-A)^{-1}(A+A^*)(I-A^*)^{-1} = (I-A)^{-1}[-2(A+A^*)][(I-A)^{-1}]^*$. Since the last expression is congruent to $-2(A+A^*)$, their inertias are the same, and $\mathrm{In}(-2(A+A^*)) = \mathrm{In}(-(A+A^*))$. QED

A square matrix is stable $\Leftrightarrow$ all its characteristic values have negative real parts. $\mathcal{S}$ denotes the set of all stable nxn matrices, Π denotes the set of all positive-definite hermitian matrices and $\mathcal{N}$ denotes the set of all negative-definite hermitian matrices.

Theorem 5: $A \in \mathcal{S} \Leftrightarrow$ for any $G_1 \in \Pi$ there exists $G \in \Pi : G - C_AGC_A^* = G_1$

$\Leftrightarrow$ there exists $G_1 \in \Pi : G - C_AGC_A^* = G_1$ for some $G \in \Pi$.

Proof: In theorem 2, let $B = A^*$, $x = 1$ (for 1 is not characteristic of a stable matrix and C_A presupposes that $x \neq 1$), and $D = -\frac{1}{2}(I-A)G_1(I-A^*)$. Then the last term of (4) is G_1, and (4) becomes

$$AG+GA^* = D \Leftrightarrow G - C_AGC_A^* = G_1.$$

D is hermitely congruent to $-\frac{1}{2}G_1$, and so $\mathrm{In}(D) = \mathrm{In}(-\frac{1}{2}G_1)$. Therefore, $G_1 \in \Pi \Leftrightarrow D \in \mathcal{N}$.

First equivalence: Assume $A \in \mathcal{S}$. For any $G_1 \in \Pi$, $D \in \mathcal{N}$. Therefore, $\exists G \in \Pi$: $AG+GA^* = D$ [Taussky], so $G - C_AGC_A^* = G_1$. Conversely, if for any $G_1 \in \Pi$ there exists $G \in \Pi$: $G - C_AGC_A^* = G_1$, then $AG+GA^* = D$; since G_1 is arbitrary, so is D, for I-A and I-A* are non-singular, otherwise C_A and $C_A^* = C_{A^*}$ would not be defined. Since $D \in N$, $A \in S$ [Taussky].

Second equivalence: Assume $A \in S$. Then $\exists G \in \Pi$: $AG+GA^* = D$ for some $D \in N$, and so $G\text{-}C_AGC_A^* = G_1$; $G_1 \in \Pi$ as above. Conversely, if, for some $G_1 \in \Pi$, $G\text{-}C_AGC_A^* = G_1$ for some $G \in \Pi$, then $AG+GA^* = D$ and $D \in N$. Hence, $A \in S$. QED

<u>Corollary 6</u>: $A \in S \iff \exists G \in \Pi$: $I\text{-diag}(g_1,...,g_n) \in \Pi$, where $\{g_i\}_1^n$ are the roots of $|\lambda G\text{-}C_AGC_A^*| = 0$; furthermore, g_i is real $(i=1,...,n)$.

<u>Proof</u>: $\Rightarrow$ Assume $A \in S$. By the first equivalence of the preceding theorem $\exists G \in \Pi$: $G\text{-}C_AGC_A^* = I$. Since both G and $C_AGC_A^*$ are hermitian and $G \in \Pi$, $\exists R$: R is non-singular and $R'GR = I$, $R'(C_AGC_A^*)R = \text{diag}(g_1,...,g_n)$ where $\{g_i\}$ are the roots of $|\lambda G\text{-}C_AGC_A^*| = 0$. Then $R'R = R'IR = R'(G\text{-}C_AGC_A^*)R = R'GR\text{-}R'(C_AGC_A^*)R = I\text{-diag}(g_1,...,g_n)$. $R'R \in \Pi$ because $R'GR = I \Rightarrow R'^{-1}R^{-1} = G \in \Pi \Rightarrow RR' \in \Pi \Rightarrow R'R \in \Pi$. Therefore, $I\text{-diag}(g_1,...,g_n) \in \Pi$.

$\Rightarrow$ Since G and $C_AGC_A^*$ are hermitian and $G \in \Pi$, $\exists R$: R is non-singular and $R'GR = I$, $R'(C_AGC_A^*)R = \text{diag}(g_1,...,g_n)$ where $\{g_i\}$ are the (real) roots of $|\lambda G\text{-}C_AGC_A^*| = 0$. Then $R'^{-1}[I\text{-diag}(g_1,...,g_n)]R^{-1} = R'^{-1}R^{-1}\text{-}R'^{-1}\text{diag}(g_1,...,g_n)R^{-1} = G\text{-}C_AGC_A^* \in \Pi$. By the second equivalence of the preceding theorem, $A \in S$.

g_i is real $(i=1,...,n)$ [Gantmacher, Vol. I; p. 338, thm. 22]. QED

<u>Corollary 7</u>: $A \in S \iff \exists G \in \Pi$: $g_i < 1$ $(i=1,...,n)$ where $\{g_i\}_1^n$ are the characteristic values of $G^{-1}C_AGC_A^*$.

<u>Proof</u>: In the preceding corollary, G is non-singular since $G \in \Pi$. Hence, $\{g_i\}_1^n$, the roots of $|\lambda G\text{-}C_AGC_A^*| = 0$, are the characteristic values of $G^{-1}C_AGC_A^*$, for $|\lambda G\text{-}C_AGC_A^*| = 0 \iff |G| \cdot |\lambda I\text{-}G^{-1}C_AGC_A^*| = 0$. $I\text{-diag}(g_1,...,g_n) \in \Pi$ is equivalent to $1\text{-}g_i > 0$ $(i\text{-}1,...,n)$. QED

The algebraic properties of the Cayley transform previously developed will be applied to prove theorems about convergent matrices.

The nxn matrix A is <u>convergent</u> $\iff \lim_{m\to\infty} A^m = 0$.

<u>Theorem 8</u>: D is convergent $\iff \exists A \in S : D = C_A$.

<u>Proof</u>: D is convergent $\iff D^*$ is convergent.

$\Rightarrow$ Assume that D is convergent. Then D^* is convergent. By Stein's theorem [Stein, 1952; p. 82, thm. 1] $(\exists G \in \Pi)(\exists G_1 \in \Pi)$: $G\text{-}DGD^* = G_1$. Define A by $A = (D\text{-}I)(D+I)^{-1}$; then $D = C_A$. By theorem 2, $AG+GA^* = -\frac{1}{2}(I\text{-}A)G_1(I\text{-}A^*)$. Since $-\frac{1}{2}(I\text{-}A)G_1(I\text{-}A^*)$ is hermitely congruent to $-G_1$, $AG+GA^* \in N$ and by [Taussky] $A \in S$.

$\Leftarrow$ Assume that $A \in S$. Then by theorem 5, $(\exists G \in \Pi)(\exists G_1 \in \Pi)$: $G\text{-}C_AGC_A^* = G_1$. By

Stein's theorem, C_A^* is convergent, and so C_A is convergent. QED

Corollary 9: D is convergent $\Leftrightarrow (\forall G_1 \epsilon \Pi)(\exists G \epsilon \Pi)$: $G\text{-}DGD^* = G_1$

$\Leftrightarrow (\exists G_1 \epsilon \Pi)(\exists G \epsilon \Pi)$: $G\text{-}DGD^* = G_1$.

Proof: By the preceding theorem, D is convergent $\Leftrightarrow$ $D = C_A$, where $A \in \mathcal{S}$. The two equivalences follow from this fact and theorem 5. QED

The preceding corollary is a theorem of Taussky's [Taussky; p. 7, thm. 5], which is itself a strengthening of Stein's theorem.

REFERENCES

1. Gantmacher, F. R. *The Theory of Matrices*. 2 vols. Translated by K. A. Hirsch. New York: Chelsea Publishing Company, 1960.

2. Stein, P. "Some General Theorems on Iterants." *Journal of Research of the National Bureau of Standards*, Vol. 48, No. 1 (1952 January), 82-83.

3. Stein, P. "On the Ranges of Two Functions of Positive Definite Matrices." *Journal of Algebra*, Vol. 2, No. 3 (1965 September) 350-353.

4. Taussky, Olga. "Matrices C with $C^n \to 0$." *Journal of Algebra*, Vol. 1, No. 1 (1964 April), 5-10.

5. Weyl, Hermann. *The Classical Groups: Their Invariants and Representations*. 2nd ed. Princeton, N.J.: Princeton University Press, 1946.