Strong stability and the Cayley transform

G. Halikias[1]

Abstract: The general notion of "strong" stability for internal autonomous system descriptions has been recently introduced for continuous and discrete-time systems. This is a stronger notion of stability compared to alternative definitions (asymptotic, Lyapunov), which prohibits systems described by natural coordinates to have overshooting responses, for arbitrary initial conditions in state-space. The paper reviews three refined notions of strong stability, along with the necessary and sufficient conditions corresponding to each notion. Using the Cayley transformation it is shown that the notions in the two domains are essentially equivalent and that the strong stability conditions can be transformed from one domain to the other in a straightforward way.

Keywords: Strong stability, Cayley (bilinear) transformation

1. Introduction

Stability is a crucial system property that has been extensively studied from many aspects [1], [2], [5], [7]. The paper reviews a new definition of stability, defined as "strong stability", which has been studied independently for both continuous and discrete systems [3], [4], [8].

Essentially, strong stability prohibits "overshoots" in the autonomous trajectory of the system, defined in state-space, for arbitrary initial conditions. Non-overshooting response is a desirable property in many applications and can be considered as a special case of constrained control. The strong stability property is also related to low degree of eigen-frame skewness (and hence low sensitivity of eigenvalues to data uncertainty in stabilisation problems [3], [8]) and the transient response of a system, e.g. its overshooting behaviour, initial exponential growth or its transient energy [6], [10], [11] and could prove useful for analysing stability properties of systems under switching regimes [9].

The Cayley transform is and extension to matrices of the conformal mapping: $f(z) = (z-1)(z+1)^{-1}$, $z \neq -1$. It has been used extensively in Control Systems as a tool for translating asymptotic/Lyapunov notions of stability for state-space systems between the continuous and discrete domains. This can also be extended to the notion of strong stability introduced earlier.

The paper reviews three refined notions of strong stability in the discrete and continuous domains, along with sets of necessary and sufficient conditions corresponding to each notion in each domain. Using the Cayley transformation it is shown that the two notions of strong stability are essentially equivalent and that the strong stability conditions can be transformed from one domain to the other in a straightforward way. Note that this applies to each of the three refined strong stability notions, so that the correspondence between the two domains is complete. This result is important for control synthesis problems, since intuition and strong stabilisation conditions (e.g. applying to state or output

[1]Department of Mathematics, University of Athens, Panepistimioupolis GR-157 84, Athens Athens, Greece, e-mail: ghalikias@math.uoa.gr

feedback problems) can be transferred from one domain to the other. In this way, numerically ill-conditioned problems/algorithms in one domain may be solved more effectively when transformed to the other domain.

The structure of the paper is as follows. Section 2 reviews the notions of strong stability in the continuous and discrete domains, and the corresponding sets of necessary and sufficient conditions. Section 3 shows that, using the Cayley transform, the strong stability conditions described in section 2 can be translated between the continuous and discrete domains in a straightforward way. In this way, certain aspects of the definitions of strong stability in the two domains are illuminated.

The notation of the paper is standard and is summarized here for convenience. $\mathcal{N}$, $\mathcal{R}$ and $\mathcal{C}$ denote the sets of natural, real and complex numbers, respectively. The set of non-negative integers is $\mathcal{N}_0 = \mathcal{N} \cup \{0\}$. The set of complex numbers with negative (non-positive) real part is denoted by $\mathcal{C}_-$ ($\bar{\mathcal{C}}_-$). $\mathcal{R}^{m\times n}$ denotes the space of all $m \times n$ real matrices. If A is a square matrix, then $\lambda(A)$ denotes the spectrum of A and $\rho(A)$ is the spectral radius of A. $\|\cdot\|$ denotes the Euclidian norm of a vector or the spectral norm of a matrix depending on context. A positive definite matrix A (positive semi-definite, negative definite, negative semi-definite) is denoted as $A > 0$ ($A \geq 0$, $A < 0$, $A \leq 0$, respectively). The (right) null-space of a matrix A is denoted by $\mathcal{N}_r(A)$, while the range (column-span) of A is denoted as Range(A). The (right) nullity of A is null(A) $= \dim(\mathcal{N}_r(A))$. Finally, a matrix $A \in \mathcal{R}^{n\times n}$ is called Hurwitz if $\lambda(A) \subseteq \mathcal{C}_-$ and Schur if $\rho(A) < 1$.

2. Strong Stability of Discrete and Continuous Systems

Consider the autonomous linear time-invariant (LTI) discrete-time system:

$$\Sigma_d(A) : x_{k+1} = Ax_k,\ k \in \mathcal{N}_0,\ x_0 \in \mathcal{R}^n$$

$\Sigma_d(A)$ is said to be Lyapunov-stable if for every $\epsilon > 0$ there exists $\delta = \delta(\epsilon) > 0$ such that $\|x_k\| < \epsilon$ for all $k \in \mathcal{N}_o$ whenever $\|x_0\| < \delta$. $\Sigma_d(A)$ is asymptotically stable, if it is Lyapunov-stable and there exists $\eta > 0$ such that, if $\|x_0\| < \eta$ then $\lim_{k\to\infty} \|x_k\| = 0$. For discrete LTI systems simple necessary and sufficient conditions can be derived for these two fundamental notions of stability: $\Sigma_d(A)$ is asymptotically stable if and only if $\rho(A) < 1$ [1]. $\Sigma_d(A)$ is Lyapunov-stable if and only if $\rho(A) \leq 1$ and every eigenvalue that lies on the unit circle has equal algebraic and geometric multiplicity. The equivalent conditions for continuous-time systems $\Sigma_c(A) : \dot{x} = Ax(t),\ x(0) = x_0 \in \mathcal{R}^n$, are: (i) $\lambda(A) \subseteq \mathcal{C}_-$ (asymptotic stability) and, (ii) $\lambda(A) \subseteq \bar{\mathcal{C}}_-$ and any eigenvalue which lies on the imaginary axis has equal algebraic and geometric multiplicity (Lyapunov stability). In discrete-time strong stability is defined as follows:

Definition 2.1: *The system $\Sigma_d(A)$ is:*

(i) Strong Lyapunov stable (SLS) if and only if $\|x_{k+1}\| \leq \|x_k\|$ for all $k \in \mathcal{N}_o$.

(ii) Strong asymptotically stable in the wide sense (SAS w.s.) if and only if it is asymptotically stable and $\|x_{k+1}\| \leq \|x_k\|$ for all $k \in \mathcal{N}_o$.

(iii) Strong asymptotically stable in the strict sense (SAS s.s.) if and only if $\|x_{k+1}\| < \|x_k\|$ for all $k \in \mathcal{N}_o : x_k \neq 0\}$.

The following Theorem gives simple necessary and sufficient conditions for the three notions of strong stability in discrete-time:

Theorem 2.1: *The system $\Sigma_d(A)$ is:*

(i) SLS if and only if $\|A\| \leq 1$.

(ii) SAS w.s. if and only if either one of the following two equivalent contitions hold: (a) $\|A\| \leq 1$ and $\rho(A) < 1$; (b) $\|A\| \leq 1$ and the pair $(A, I_n - A^t A)$ is observable.

(iii) SAS s.s. if and only if $\|A\| < 1$.

The corresponding notions and conditions of strong stability in continuous time are as follows:

Definition 2.2***:*** *The system $\Sigma_c(A) : \dot{x}(t) = Ax(t),\ x(0) = x_0 \in \mathcal{R}^n$ is:*

1. *Strong Lyapunov stable (SLS) if $\|x(t)\| \leq \|x(t_0)\|$, $\forall t > 0$ and $\forall x_0 \in \mathcal{R}^n$.*
2. *Strong asymptotically stable in the wide sense (SAS w.s.) if $\|x(t)\| < \|x_0\|$ for all $t > 0$ and $x_0) \neq 0$.*
3. *Strong asymptotically stable in the strict sense (SAS s.s.) if $\frac{d\|x(t)\|}{dt} < 0$ for all $t \geq 0$ and $x_0 \neq 0$.* $\square$

Strong Lyapunov stability does not allow state trajectories to exit (at any time $t > 0$) the (closed) hyper-sphere with centre the origin and radius $r_0 = \|x_0\|$ (although motion on the boundary of the sphere $\|x(t)\| = r_0$ is allowed, e.g. an oscillator's trajectory). Strong asymptotic stability (s.s.) requires that trajectories enter each hyper-sphere $\|x(t)\| = r \leq r_0$ from a non-tangential direction, whereas for systems which are strong asymptotically stable (w.s.), tangential entry is allowed. For examples of each type of strong stability see [8].

Theorem 4.2: *The system $\Sigma_c(A)$ is:*

(i) SLS if and only if $A + A^t \leq 0$.

(ii) SAS w.s. if and only if one of the following two equivalant conditions hold: (a) $A + A^t \leq 0$ and A is Hurwitz; (b) $A + A^t \leq 0$ and the pair $(A, A + A^t)$ is observable.

(iii) SAS s.s. if and only if $A + A^t < 0$.

Note also that, both for the discrete and continuous systems, SAS s.s. implies SAS w.s. which implies SLS. Table 2.1 below summarizes the necessary and sufficient conditions for each strong stability notion in the two domains, along with the standard conditions for Lyapunov and asymptotic stability.

	Continuous-time: $\dot{x} = Ax$	Discrete-time: $x_{k+1} = Ax_k$
Lyapunov stability	$\mathrm{Re}(\lambda_i(A)) \leq 0$ for all i, simple Jordan structure for any $\lambda_i(A)$ on $j\omega$-axis	$\rho(A) \leq 1$, simple Jordan structure for any $\lambda_i(A)$ with $\lvert\lambda_i(A)\rvert = 1$
Asymptotic stability	$\mathrm{Re}(\lambda_i(A)) < 0$ for all i	$\rho(A) < 1$
Strong Lyapunov stability	$A + A^t \leq 0$	$\|A\| \leq 1$
Strong asymptotic stability (w.s.)	$A + A^t \leq 0$ and $\mathrm{Re}(\lambda_i(A)) < 0$, or $A + A^t \leq 0$ and $(A, A + A^t)$ obs.	$\|A\| \leq 1$ and $\rho(A) < 1$, or $\|A\| \leq 1$ and $(A, I - A^t A)$ obs.
Strong asymptotic stability (s.s.)	$A + A^t < 0$	$\|A\| < 1$

Table 1: Summary of stability conditions

3. Strong Stability and Cayley transform

In this section the Cayley (bilinear) transformation is introduced. It is shown that, using the transformation, the strong stability conditions described in section 2 can be can be translated from the discrete to the continuous domain (and vive versa) in a straightforward way. If $A \in \mathcal{R}^{n\times n}$ with $-1 \notin \lambda(A)$ the Cayley transformation is defined by $\hat{A} = (A - I_n)(A + I_n)^{-1}$. The properties of the transformation are summarized next:

Theorem 3.1: *If (λ, ξ) is an eigenvalue/(right) eigenvector pair of A, then $\left(\frac{\lambda-1}{\lambda+1}, \xi\right)$ is the corresponding eigenvalue/(right) eigenvector pair of $\hat{A}$. Conversely, if (λ, ξ) is an eigenvalue/(right) eigenvector pair of $\hat{A}$, then $\left(\frac{1+\lambda}{1-\lambda}, (A + I_n)^{-1}\xi\right)$ is an eigenvalue/(right) eigenvector pair of A.*

Proof: Follows by direct calculations. □

Theorem 3.2***:*** *Consider the systems discrete and continuous systems $\Sigma_d(A)$ and $\Sigma_c(\hat{A})$, respectively, where $-1 \notin \lambda(A)$ and $\hat{A} = (A - I)(A + I)^{-1}$. Then,*
(i) $\Sigma_d(A)$ is SAS (s.s.) if and only if $\Sigma_c(\hat{A})$ is SAS (s.s.).
(ii) $\Sigma_d(A)$ is SAS (w.s.) if and only if $\Sigma_c(\hat{A})$ is SAS (w.s.); and
(iii) $\Sigma_d(A)$ is SLS if and only if $\Sigma_c(\hat{A})$ is SLS.

Proof: Part (i) follows from Theorems 2.1(iii) and 2.2(iii) and the following sequence of equivalent statements:

$$\begin{aligned}
\Sigma_c(\hat{A}) \text{ is SAS (s.s.)} \Leftrightarrow & (I - A)(I + A)^{-1} + (I + A^t)^{-1}(I - A^t) > 0 \\
\Leftrightarrow & \quad (I + A^t)^{-1}\{(I - A^t)(I + A) + (I + A^t)(I - A)\}(I + A)^{-1} > 0 \\
\Leftrightarrow & \quad (I + A^t)^{-1}\{2I - 2A^tA\}(I + A)^{-1} > 0 \\
\Leftrightarrow & \quad A^tA < I \Leftrightarrow \quad \|A\| < 1 \Leftrightarrow \quad \Sigma_d(A) \ \text{ is SAS (s.s.)}
\end{aligned}$$

An almost identical sequence of arguments shows that $\hat{A} + \hat{A}^t \leq 0 \Leftrightarrow \|A\| \leq 1$ proving part (iii), using Theorems 2.1(i) and 2.2(i). Finally, part (ii) follows from part (iii), the first set of (equivalent) conditions from Theorems 2.1(ii) and 2.2(ii) and the fact that under the Cayley transformations the eigenvalues of A and $\hat{A}$ are related as:

$$\lambda_i(\hat{A}) = \frac{\lambda_i(A) - 1}{\lambda_i(A) + 1}, \quad i = 1, 2, \ldots, n$$

(see Theorem 3.1). Thus, for each $i = 1, 2, \ldots, n$, $\mathrm{Re}(\lambda_i(A)) < 0 \Leftrightarrow |\lambda_i(\hat{A})| < 1$ and hence A is Hurwitz if and only if $\hat{A}$ is Schur. □

Next we investigate in more detail the properties of the transformation when $\|A\| = 1$ and clarifies the conditions under which the system is SAS (w.s.):

Theorem 3.3: *Let $A \in \mathcal{R}^{n\times n}$ with $-1 \notin \lambda(A)$. Define $\hat{A} = (A - I)(A + I)^{-1}$. Then:*

(i) $\|A\| = 1$ if and only if $\hat{A} + \hat{A}^t \leq 0$ and $\hat{A} + \hat{A}^t$ is singular.

(ii) Suppose that $\|A\| = 1$. Then null$(I - A^tA)$ = null$(\hat{A} + \hat{A}^t)$. Let A have a singular value decomposition $A = U_1V_1^t + U_2\Sigma_2V_2^t$ with $[U_1\ U_2]$ and $[V_1\ V_2]$ orthogonal and $\Sigma_2 = diag(\Sigma_2)$ such that $\|\Sigma_2\| < 1$. Then $\mathcal{N}_r(\hat{A} + \hat{A}^t) = Range(V_1 + U_1)$.

(iii) Suppose that $\|A\| = 1$. Then $(A, I - A^tA)$ is observable if and only if $(\hat{A}, \hat{A} + \hat{A}^t)$ is observable. Further, any unobservable mode of $(A, I - A^tA)$ has modulus one and corresponds to an unobservable mode of $(\hat{A}, \hat{A} + \hat{A}^t)$ which is imaginary.

Proof: Part (i) follows immediately from the proof of Theorem 3.1. (ii) Introduce the singular value decomposition $A = U_1V_1^t + U_2\Sigma_2V_2^t$ with $[U_1\ U_2]$ and $[V_1\ V_2]$ orthogonal and $\Sigma_2 = \mathrm{diag}(\Sigma_2)$ such that $\|\Sigma_2\| < 1$. Then $AV_1 = U_1$, $A^tU_1 = V_1$ and $I - A^tA = V_2(I - \Sigma_2^2)V_2^t$ which implies that $\mathcal{N}_r(I - A^tA) = \mathrm{Range}(V_1)$. A straightforward calculation also shows that:

$$\hat{A} + \hat{A}^t = 2(I + A^t)^{-1}(I - A^tA)(I + A)^{-1} = 2(I + A^t)^{-1}V_2(I - \Sigma_2^2)V_2^t(I + A)^{-1}$$

Thus $\mathrm{null}(I - A^tA) = \mathrm{null}(\hat{A} + \hat{A}^t)$ and

$$\mathcal{N}_r(\hat{A} + \hat{A}^t) = \mathrm{Range}((A + I)V_1) = \mathrm{Range}((A^t + I)U_1) = \mathrm{Range}(V_1 + U_1)$$

as required. (iii) If $(A, I - A^tA)$ is unobservable there exists $\lambda \in \mathcal{C}$ and $\xi \neq 0$ such that

$$(\lambda I - A)\xi = 0 \tag{1}$$

and

$$(I - A^tA)\xi = 0 \tag{2}$$

Equation (2) implies that $\xi \in \mathcal{N}_r(I - A^tA)$ and hence from the proof of part (ii) $\xi = V_1\theta$, $\theta \neq 0$. Then equation (1) gives:

$$A\xi = \lambda\xi \Rightarrow (U_1V_1^t + U_2\Sigma_2V_2^t)V_1\theta = \lambda V_1\theta \Rightarrow U_1\theta = \lambda V_1\theta$$

Note that:

$$\|\theta\| = \|U_1\theta\| = \|V_1\theta\| \Rightarrow |\lambda| = 1$$

Since from part (ii) $\mathcal{N}_r(\hat{A} + \hat{A}^t) = \mathrm{Range}(U_1 + V_1)$,

$$(\hat{A} + \hat{A}^t)(V_1 + U_1)\theta = 0 \Rightarrow (1 + \lambda)(\hat{A} + \hat{A}^t)V_1\theta = 0 \Rightarrow (\hat{A} + \hat{A}^t)\xi = 0 \tag{3}$$

since $\lambda \neq -1$. From Theorem 3.1 it also follows that

$$\hat{A}\xi = \sigma\xi \ \text{ where } \ \sigma = \frac{\lambda - 1}{\lambda + 1} \tag{4}$$

in which $\mathrm{Re}(\sigma) = 0$. Equations (3) and (4) imply that the pair $(\hat{A}, \hat{A} + \hat{A}^t)$ is unobservable. Conversely suppose that $(\hat{A}, \hat{A} + \hat{A}^t)$ is unobservable and there exists a pair (λ, ξ), $\xi \neq 0$ such that $\hat{A}\xi = \lambda\xi$ and $(\hat{A} + \hat{A}^t)\xi = 0$. Thus $\xi \in \mathcal{N}_r(\hat{A} + \hat{A}^t)$ and hence from part (ii) ξ can be written as $\xi = (A + I)V_1\psi$, $\psi \neq 0$. Thus, from part (ii):

$$(I - A^tA)(A + I)^{-1}\xi = V_2(I - \Sigma_2^2)V_2^t(A + I)^{-1}\xi = V_2(I - \Sigma_2^2)V_2^tV_1\psi = 0$$

Further, from Theorem 3.1:

$$A(A+I_n)^{-1}\xi = \left(\frac{1+\lambda}{1-\lambda}\right)(A+I_n)^{-1}\xi$$

and hence $(A, I - A^tA)$ is unobservable. $\square$

Remark: If $\|A\| = 1$ then A is necessarily a Lyapunov matrix (i.e. all eigenvalues have modulus less than or equal to one, and any eigenvalue with modulus equal to one has equal algebraic and geometric multiplicity) and hence $\Sigma_d(A)$ is (at least) LSS. Since $\rho(A) \leq \|A\|$, in this case we have either $\rho(A) < 1$ (in which case $\Sigma_d(A)$ is SAS (w.s.)) or $\rho(A) = 1$ (in which case $\Sigma_d(A)$ is just LSS and *not* SAS (w.s.)); When $\|A\| = 1$ any eigenvalue of A with modulus one must be unobservable through $I - A^tA$ [4]; thus when $(A, I - A^tA)$ is observable no eigenvalues with modulus one can exist. Theorem 3.3 shows that the corresponding conclusions can be drawn for continuous-time systems.

4. Conclusions

It has been shown that the Cayley transformation can be applied to translate strong stability conditions between the discrete and continuous domains, for all three refined notions of strong stability defined in the literature. This can help to unify the presentation of the theory, simplify the results related to the solution of strong stabilization problems in the two domains and improve the numerical properties of an ill-conditioned problem/algorithm defined in one domain by transforming it to the other.

References

[1] P. Antsaklis and A.N. Michel, *Linear Systems*, Birkhauser, Boston, USA, 2006.

[2] N.B. Bhatia and G.P. Szego, *Stability theory of Dynamic Systems*, Classics in Mathematics Series, Springer Verlag, New York, 1970.

[3] G. Halikias, A. Papageorgiou, and N. Karcanias, Non-overshooting stabilisation via state and output feedback, *International Journal of Control*, **83**:6, 1232–1247, 2010.

[4] G.D Halikias, L. Dritsas, A. Pantelous and V. Tsoulkas, Strong stability of discrete-time systems, *Linear Algebra and its Applications*, **436**, 1890–1908, 2012.

[5] D. Hinrichsen and A.J. Pritchard, *Mathematical Systems Theory I*, Texts in Applied Mathematics Vol. 41, Springer-Verlag, New York, 2005.

[6] D. Hinrichsen and A.J. Pritchard, On the transient behaviour of stable linear systems. *Proc. 14th Int. Symp. Math. Theory Networks and Systems*, Perpignan, France, 2000.

[7] Horn and Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1980.

[8] N. Karcanias, G.D. Halikias and A. Papageorgiou, Strong stability of Internal System Descriptions, *International Journal of Control*, **83**(1), 182–205, 2010.

[9] D. Liberzon, *Switching in Systems and Control*, Birkhauser, Boston, 2003

[10] von Elmar Plischke, *Transient Effects of Linear Dynamical Systems*, PhD thesis, University of Bremen, 2005.

[11] J.F. Whidborne, J. McKernan and G. Papadakis, Minimal transient energy growth for plane poiseuille flow, *Proc. Inst. Mech. Eng., Part I: Journal of Systems and Control Engineering* **222**(5), 323–331, 2008.