A UNIFIED APPROACH TO DISCRETE AND CONTINUOUS HIGH-GAIN ADAPTIVE CONTROLLERS USING TIME SCALES*

IAN A. GRAVAGNE, JOHN M. DAVIS†, AND JEFFREY J. DACUNHA

ABSTRACT. It has been known for some time that proportional output feedback will stabilize certain classes of linear time-invariant systems under an adaptation mechanism that drives the feedback gain sufficiently high. More recently, it was demonstrated that discrete implementations of the high-gain adaptive controller also require adaptation of the sampling rate. In this paper, we use recent advances in the mathematical field of dynamic equations on time scales to unify the discrete and continuous versions of the high-gain adaptive controller. A novel proof method is presented based on time scales, as is a brief tutorial on the subject of time scales.

CONTENTS

1. INTRODUCTION

The idea of high-gain adaptive feedback was born from a desire to stabilize certain classes of linear continuous systems without the need to explicitly identify the unknown system parameters. This type of adaptive feedback, unlike "traditional" adaptive control, does not identify system parameters at all, but rather adapts the feedback gain itself in order to regulate the overall system. Many papers have discussed the details of various kinds of high-gain adaptive controllers, including

Key words and phrases. adaptive controller, high-gain, stability, hybrid systems, time scales.

*This project was funded by a Baylor University Research Council Summer Grant.

†The second author's research was supported by a Baylor University Summer Sabbatical from the College of Arts and Sciences.

[11, 16, 19, 24] among others. Recently, however, there have been several papers discussing one particularly practical angle on the high-gain adaptive controller, namely how to cope with the possibility that the output is only available at discretely sampled instants, a necessary assumption for computer implementation of a high-gain adaptation algorithm. In particular, Owens [18] illustrated that it is not generally possible to stabilize a linear system with adaptive high-gain feedback under the usual assumption of uniform sampling. Thus, in [18] a mechanism was developed to adapt the sampling rate as well as the gain, a notion that was subsequently improved upon by Ilchmann and Townley [12, 13], and also appeared in a more general form for certain types of nonlinear systems in [14].

The primary objective of this paper is to unify the previously disparate studies of sampled and continuous high-gain adaptive controllers. In so doing, the paper will illustrate the power of a burgeoning new field of mathematics called *dynamic equations on time scales*. Time scale methods allow the examination and manipulation of dynamical systems without regard to the particular domain of the system, i.e. continuous, discrete or mixed. Time scale theory is finding application in the study of population dynamics, financial forecasting models, and adaptive methods for the numerical solution of partial differential equations [23]. From a control perspective, since there is no restriction that the discrete case be uniformly sampled, this theoretical framework easily handles the nonuniform sampling conditions of interest in this paper. To our knowledge, this is the first instance of an engineering application employing time scale analysis (using the definition of "time scale" presented in Appendix A); furthermore, in this context the analysis supports a proof of the stability of high-gain adaptive controllers that is novel in and of itself.

The paper is organized as follows: in Section 2 we discuss the system model and the control objectives and outline the background work. In Section 3 we convert the model and objectives into their time scale realizations, and then in Section 4 introduce three prerequisite lemmas to the main stability theorem appearing in Section 5. A short tutorial about dynamic equations on time scales appears in Appendix A.

2. Background

Before continuing, we first state several assumptions that are required in the subsequent text.

(A1) The system model and feedback law are given by

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad y(t) = Cx(t), \qquad x(0) = x_0; \tag{2.1}$$

$$u(t) = -k(t)y(t),\ k(t) > 0. \tag{2.2}$$

System parameters $A \in \mathbb{R}^{n\times n}$ and $B, C^T \in \mathbb{R}^{n\times m}$ define a linear, time-invariant, minimum phase system with positive high-frequency gain, and $\operatorname{spec}(CB) \subset \mathbb{C}_+ := \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 0\}$.

(A2) There is a gain k_1 and small constant $\varepsilon_1 > 0$ such that $\operatorname{spec}(A - kBC) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < -\varepsilon_1\}$ for all $k(t) > k_1$. Note that (A1) plays a necessary role in the validity of (A2), as it is not obvious that increasing the gain arbitrarily beyond some threshold will necessarily *keep* all poles in the left-half plane, even if they start out there. We rely here on extensive investigation into this phenomenon; see [11, 19, 24].

(A3) The feedback gain $k(t)$ increases monotonically as $t \to \infty$ (by which we mean that k is increasing except possibly for points of inflection).

Under these conditions it has been known for some time (e.g. [11]) that there are a wide class of gain adaptation laws $k(t) = f(y(t))$, $f : \mathbb{R}^m \to \mathbb{R}$, where f is designed so that $k(t)$ satisfies (A2), that can asymptotically stabilize system (2.1) in the sense that

$$y(t) \in L^2[t_0, \infty), \quad \lim_{t\to\infty} k(t) = k_\infty < \infty,$$

under a few requisite assumptions. (Notationally, $w(t) \in L^2[t_0, \infty)$ means that each element of w belongs to L^2 if w is a vector or matrix.)

Subsequent to this result, various authors [12, 13, 18, 20] have investigated the possibility that the output is available only at discretely sampled instants, $y_i = y(t_i)$. Under this condition, it is necessary to adapt not only the feedback gain but also the sampling period, $h_i = t_{i+1} - t_i$. Thus, the control objective becomes

$$y_i \in \ell^2[i_0, \infty), \quad \lim_{i\to\infty} k_i = k_\infty < \infty, \quad \lim_{i\to\infty} h_i = h_\infty > 0.$$

Though several variations and adaptations of these results have appeared, these remain the core control results for continuous and discrete high-gain adaptive controllers. Until now, the two cases have been treated quite differently, and the remainder of the paper is dedicated to unifying the two sets of results under the umbrella of time scale theory.

3. Stabilization Via High-Gain Adaptation

Using the concepts introduced in Appendix A, we see that (2.1) is simply a linear ODE defined on the time scale $\mathbb{T} = \mathbb{R}$, with graininess $\mu \equiv 0$. In general, however, (2.1) can be replaced with

$$x^\Delta(t) = \widehat{A}x(t) + \widehat{B}u(t), \qquad y(t) = Cx(t), \qquad x(0) = x_0, \tag{3.1}$$

on $t \in \mathbb{T}$, an arbitrary time scale. The relationship between (2.1) and (3.1) follows from a series expansion similar to that in [11]; namely,

$$\widehat{A} := \xi_\mu^{-1}(A) = \operatorname{expc}(\mu A)A, \qquad \widehat{B} := \frac{1}{\mu}\int_0^\mu e^{A\tau} B\, d\tau = \operatorname{expc}(\mu A)B. \tag{3.2}$$

Implementing control law (2.2) in (3.1) produces

$$x^\Delta(t) = \mathcal{A}x(t) := \operatorname{expc}(\mu A)(A - kBC)x(t) \tag{3.3}$$

which will be the main dynamical system referenced in the remainder of the paper.

Using the notation introduced above, the control objectives are now to design adaptation laws

$$k^\Delta(t) = f(y) \geq 0, \tag{3.4}$$

$$\mu(t) = g(y) \geq 0, \tag{3.5}$$

which exist on some time scale $t \in \mathbb{T}$ (with $t = 0$ the left-most point in $\mathbb{T}$), such that

$$y(t) \in L^2[0, \infty)_\mathbb{T} := \left\{ w(t) : \int_0^\infty w^2(t)\Delta t < \infty \right\}, \quad \lim_{t\to\infty} k(t) = k_\infty < \infty.$$

The case covering discretely sampled systems will also require

$$\lim_{t\to\infty} \mu(t) = \mu_\infty > 0.$$

Expression (3.4) looks similar to previous works, requiring that the positive gain $k(t)$ increase monotonically. The point of equation (3.5) is to dynamically define the time scale itself, a process we term *μ-dynamics*. In the special case that output equation (3.5) has $g \equiv 0$, then we are talking about a purely continuous linear system, i.e. $\mathbb{T} = \mathbb{R}$. Furthermore, by definition of a time scale, (3.5) must be designed to that $\mu \geq 0$ for the system to be defined on a valid time scale. Note that there is no ambiguity in the fact that k^{Δ} is defined in a way that depends on the graininess: one may order the evaluation of (3.5) before the evaluation of (3.4).

4. Stability Preliminaries

The stability properties of systems of the form $x^{\Delta}(t) = \mathcal{A}x(t)$ have quite recently been extensively studied [8, 21]. Since $\mathcal{A}(t)$ is time-varying and not obviously Jordan reducible in this case, the results from [8, 21] for matrix systems are not directly applicable. However, our approach in the next section will be to formulate an exponentially convergent quadratic functional, which will involve proving stability for a scalar system rather than a multi-dimensional matrix system. Thus, we summarize one of the main results of the excellent paper [21].

Definition 4.1. The *set of exponential stability* for the equation $x^{\Delta} = \lambda x$ on the time scale $\mathbb{T}$ is $\mathcal{S}(\mathbb{T}) = \mathcal{S}_{\mathbb{C}}(\mathbb{T}) \cup \mathcal{S}_{\mathbb{R}}(\mathbb{T})$ where

$$\mathcal{S}_{\mathbb{C}}(\mathbb{T}) = \left\{ \lambda \in \mathbb{C} : \alpha = -\limsup_{t\to\infty} \frac{1}{t-t_0} \int_{t_0}^{t} \frac{\log|1+\mu(\tau)\lambda|}{\mu(\tau)} \Delta\tau > 0 \right\},$$

$$\mathcal{S}_{\mathbb{R}}(\mathbb{T}) = \left\{ \lambda \in \mathbb{R} : \forall t \in \mathbb{T},\ \exists \tau > t \text{ with } \tau \in \mathbb{T} \text{ such that } 1+\mu(\tau)\lambda = 0 \right\}.$$

Theorem 4.1. [21] *Solutions of the the scalar equation $x^{\Delta} = \lambda x$ are exponentially stable on an arbitrary $\mathbb{T}$ if and only if $\lambda \in \mathcal{S}(\mathbb{T})$.*

The set $\mathcal{S}_{\mathbb{R}}(\mathbb{T})$ can immediately be seen to define nonregressive eigenvalues, and a loose interpretation of $\mathcal{S}_{\mathbb{C}}(\mathbb{T})$ suggests that it is necessary for a regressive eigenvalue to reside in the area of the complex plane defined by $|1+\mu(\tau)\lambda| < 1$ "most" of the time (reading the integral as a time average). This area is defined as the *Hilger circle*, illustrated in Figure 3 of Appendix A, and is of primary importance in the stability analysis of linear systems. Since the solution of a scalar system $x^{\Delta} = \lambda x$ is, by definition, $x(t) = x_0 e_{\lambda}(t,t_0)$,, Theorem 4.1 states that, if $\lambda \in \mathcal{S}(\mathbb{T})$ then there is some $K = K(t_0) \geq 1$ so that

$$|x(t)| = |x_0|\,|e_{\lambda}(t,t_0)| \leq |x_0|\,K e^{-\alpha(t-t_0)},$$

where $e^{-\alpha(t-t_0)}$ is the usual exponential function (as opposed to the generalized time scale exponential which we will always denote with a subscript for clarity). Note that, as $\mu \to 0$, the Hilger circle widens to include the entire left-hand complex plane, the expected result for the special case where $\mathbb{T} = \mathbb{R}$, and if $\mu = 1$, the Hilger circle is a unit circle centered at -1, again the desired result when $\mathbb{T} = \mathbb{Z}$. Converting from a difference equation to a recursion equation shifts the circle center to the origin, the familiar result for the special case of classical discrete linear systems.

5. System Stability

In this section, first we develop time scale generalizations of concepts we need in order to state and prove our main stability results. We establish the exponential stability of solutions to (3.3) when BC is full rank and when it is not full rank. The stark contrast of methods in the two proofs is very interesting. These results then put us in a position to prove the main result of the paper, Theorem 5.4.

Consider the *generalized* or *time scale Lyapunov matrix equation*

$$A^T(t)P(t) + P(t)A(t) + \mu(t)A^T(t)P(t)A(t) = -Q. \tag{5.1}$$

It is important to point out that the time scale Lyapunov matrix equation is the unification (with $B(t) \equiv A^T(t)$) of the *Sylvester matrix equation* [2]

$$XA(t) + B(t)X = -Q$$

in the continuous case ($\mathbb{T} = \mathbb{R}$) and the *Stein matrix equation*

$$B(t)XA(t) - X = -Q \tag{5.2}$$

in the discrete case ($\mathbb{T} = \mathbb{Z}$). The Stein matrix equation above is written in recursive form. It can easily be transformed into equivalent difference form

$$XA(t) + B(t)X + B(t)XA(t) = -Q, \tag{5.3}$$

but (5.2) certainly is found more frequently in the literature than (5.3).

The following is a very powerful result recently established by DaCunha [6, 7]. It will play an important role in our main stability theorems.

Theorem 5.1. *If the $n \times n$ matrix $A(t)$ has all eigenvalues in the corresponding Hilger circle for every $t \geq t_0$, then for each $t \in \mathbb{T}$ there exists some time scale $\mathbb{S}$ such that integration over $I := [0, \infty)_{\mathbb{S}}$ yields a unique solution to* (5.1) *given by*

$$P(t) = \int_I e_{A^T(t)}(s, 0)Qe_{A(t)}(s, 0)\Delta s. \tag{5.4}$$

Moreover, if Q is positive definite, then $P(t)$ is positive definite for all $t \geq t_0$.

We will need the following basic lemma for our stability arguments.

Lemma 5.1. *Let $\mathbb{T}$ be a time scale, and consider $y^\Delta \leq \lambda y$, $\lambda(t) \in \mathbb{R}$. If $y(t) > 0$ for all $t \geq t_0 \in \mathbb{T}$, then $\lambda(t) \in \mathcal{R}^+$.*

Proof. The inequality gives rise to the initial value problem

$$y^\Delta = \lambda(t)y + f(t), \qquad y(t_0) = y_0, \tag{5.5}$$

where $f(t) \leq 0$ and $y_0 > 0$.

First, suppose $\lambda(t) \in \mathcal{R}$ but $\lambda(t) < -1/\mu(t)$. Then $\mu(t) > 0$ and (5.5) yields $y^\sigma = \mu\lambda y + y + \mu f = y(1 + \mu\lambda) + \mu f$. However, $1 + \mu\lambda < 0$ which implies $y^\sigma(t_0) < 0$ (since $y_0 > 0$). This is a contradiction.

On the other hand, suppose $\lambda(T)$ is nonregressive for some $T > t_0$, $T \in \mathbb{T}$. If their exists a $t_0 < t < T$ such that $\lambda(t) < -1/\mu(t)$, then invoke the preceding argument. If $\lambda(t) > -1/\mu(t)$ for $t < T$, then solve (5.5) to get

$$y(t) = e_\lambda(t, t_0)y_0 + \int_{t_0}^t e_\lambda(t, \sigma(\tau))f(\tau)\,\Delta\tau.$$

Since $e_\lambda(t, t_0) > 0$ for $t < T$, we see $\int_{t_0}^t e_\lambda(t, \sigma(\tau))f(\tau)\,\Delta\tau < 0$ for $t < T$. However, for $t \geq T$,

$$y(t) = 0 + \int_{t_0}^{\rho(T)} e_\lambda(t, \sigma(\tau))f(\tau)\,\Delta\tau.$$

By the argument above, $y(t) < 0$ for $t \geq T$ which again contradicts the positivity of $y(t)$. □

We now state and prove two central exponential stability results. If BC is not full rank, we require $\operatorname{spec}(A - k(t)BC) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < 0\}$ for all $t > t^*$. However, we relinquish this assumption when BC is full rank.

Theorem 5.2. **(Exponential Stability:** BC **Full Rank Case)** *For* (3.3)*, suppose the following:*

(i) $\mathbb{T}$ *is a time scale which is unbounded above but with* μ*,* μ^Δ *bounded,*
(ii) $k(t) \to \infty$ *as* $t \to \infty$*,* $\mu(t) \downarrow 0$ *as* $t \to \infty$ *whenver* $\mu(t) > 0$*,*
(iii) $\mu(t)k(t) \leq M$ *for all* $t > t^*$*,*
(iv) BC *has full rank.*

Then the system (3.3) *is exponentially stable in the sense that there exists some* $t_0 \in \mathbb{T}$ *and constant* $K \geq 1$ *such that* $\|x(t)\| \leq \|x(t_0)\| Ke^{-\alpha(t-t_0)}$ *for all* $t \geq t_0$*.*

Proof. For notational convenience, let $D = BC$ and for $t \in \mathbb{T}$ define

$$A^*(t) := [\xi_\mu^{-1}(A)A^{-1}]A = \operatorname{expc}(\mu A)A,$$
$$D^*(t) := [\xi_\mu^{-1}(A)A^{-1}]D = \operatorname{expc}(\mu A)D.$$

Let $P(t) = P(t)^T > 0$ for all $t \in \mathbb{T}$ and $Q = Q^T > 0$ and consider the time scale Lyapunov equation

$$D^{*^T}(t)P(t) + P(t)D^*(t) - \mu(t)k(t)D^{*^T}(t)P(t)D^*(t) = Q. \tag{5.6}$$

By Theorem 5.1, for each $t \in \mathbb{T}$, the solution $P(t)$ of (5.6) is given by

$$P(t) = \int_I e_{-D^{*T}(t)}(s,0)Qe_{-D^*(t)}(s,0)\Delta s. \tag{5.7}$$

Lemmas 8.1–8.3 of Appendix B show that for $k(t)$ sufficiently large (and thus $\mu(t)$ sufficiently small), $\operatorname{spec}(-D^*(t))$ is contained in the Hilger circle for all $t > t^*$; hence (5.7) is well defined. Also, the correct interpretation of the improper integral (5.7) is crucial: for each $t \in \mathbb{T}$, the time scale over which the integration is performed is $I := [0,\infty)_{\mathbb{S}} := [0,\infty) \cap \mathbb{S}$ where $\mathbb{S} = \mu(t)k(t)\mathbb{Z}$ and hence has constant graininess for each fixed t.

We are now in a position to show that $V(x(t)) = x^T(t)P(t)x(t)$ is a Lyapunov function. To verify this,

$$\begin{aligned}
V^\Delta(x(t)) &= x^{T^\Delta}Px + x^{T^\sigma}P^\Delta x + x^\sigma P^\sigma x^\Delta \\
&= x^T(\mathcal{A}^TP + P\mathcal{A} + \mu\mathcal{A}^TP\mathcal{A} + (I+\mu\mathcal{A}^T)P^\Delta(I+\mu\mathcal{A}))x \\
&= x^T((A^{*^T} - kD^{*^T})P + P(A^* - kD^*) + \mu(A^{*^T} - kD^{*^T})P(A^* - kD^*) \\
&\quad + (I+\mu\mathcal{A}^T)P^\Delta(I+\mu\mathcal{A}))x \\
&= x^T(A^{*^T}P + PA^* + \mu A^{*^T}PA^*)x - kx^T(D^{*^T}P + PD^*)x \\
&\quad + \mu x^T(-kA^{*^T}PD^* - kD^{*^T}PA^* + k^2D^{*^T}PD^* + (I+\mu\mathcal{A}^T)P^\Delta(I+\mu\mathcal{A}))x \\
&= x^T(A^{*^T}P + PA^* + \mu A^{*^T}PA^*)x - kx^T(D^{*^T}P + PD^* - \mu kD^{*^T}PD^*)x \\
&\quad + \mu x^T(-kA^{*^T}PD^* - kD^{*^T}PA^*)x + x^T((I+\mu\mathcal{A}^T)P^\Delta(I+\mu\mathcal{A}))x \\
&= x^T(A^{*^T}P + PA^* + \mu A^{*^T}PA^*)x - kx^TQx \\
&\quad - \mu kx^T(A^{*^T}PD^* + D^{*^T}PA^*)x + x^T((I+\mu\mathcal{A}^T)P^\Delta(I+\mu\mathcal{A}))x.
\end{aligned} \tag{5.8}$$

The first and third terms are bounded; this follows from Proposition 7.1(v). In the last term, notice $\mathcal{A}^T$ and $\mathcal{A}$ contain a factor of $k(t)$ so this term will be bounded as $t \to \infty$ so long as $P^\Delta(t)$ is bounded. However, under our hypotheses DaCunha [6, 7] recently showed that $P^\Delta(t)$ exists and is bounded for each $t \in \mathbb{T}$. As a result,

$$V^\Delta \le -(k(t)+\varepsilon)x^Tx \le \frac{-(k(t)+\varepsilon)}{\lambda_{\min}(P)}V := \eta(t)V, \qquad t \ge t^*,$$

where ε is the collective upper bound on all but the second term in (5.8). By Theorem 7.7 and Lemma 5.1, $V(t) \le V(t^*)e_{-\eta}(t,t^*)$, $t \ge t^*$ and $\eta(t) \in \mathcal{R}^+$, $t \ge t^*$. Then there exists some $K \ge 1$ such that $e_\eta(t,t^*) \le K^2(t^*)e^{-2\alpha(t-t^*)}$, where

$$\alpha = -\frac{1}{2}\limsup_{t\to\infty}\frac{1}{t-t^*}\int_{t^*}^{t}\frac{\operatorname{Log}|1-\eta(\tau)\mu(\tau)|}{\mu(\tau)}\Delta\tau > 0.$$

Consequently,

$$||x(t)|| \le K||x(t^*)||e^{-\alpha(t-t^*)}.$$

□

From [19], we know that there exists some k^* such that for all $k \ge k^*$, our system $S(A,B,C)$ under high gain feedback has a positive real realization $S(A - kBC, B(CB)^{-1}, C)$ under regulation. By the Kalman-Yakubovich Lemma [22], there exists $P, Q > 0$ such that

$$(A - k^*BC)^TP + P(A - k^*BC) = -Q \text{ and } PB = C^TCB. \tag{5.9}$$

Theorem 5.3. (Exponential Stability: BC Not Full Rank Case) *For (3.3), suppose the following:*

(i) $\mathbb{T}$ *is a time scale which is unbounded above but with bounded graininess,*
(ii) $k(t) \to \infty$ *as* $t \to \infty$, $\mu(t) \downarrow 0$ *as* $t \to \infty$ *whenver* $\mu(t) > 0$,
(iii) $(CB)^T + CB - \mu(t)k(t)(CB)^TCB > 0$ *for all* $t > t^*$.

Then the system (3.3) is exponentially stable in the sense that there exists some $t_0 \in \mathbb{T}$ *and constant* $K \ge 1$ *such that* $\|x(t)\| \le \|x(t_0)\|\,Ke^{-\alpha(t-t_0)}$ *for all* $t \ge t_0$.

Proof. Consider the Lyapunov function $V(x(t)) = x^T Px$ in light of (5.9). Taking the time scale derivative with respect to t (again suppressing the time dependence in k, μ, and $\mathcal{A}$), and writing $I + \Sigma = \mathrm{expc}(\mu A)$, we see

$$\begin{aligned}
V^\Delta &= x^T \left[(A - kBC)^T (I + \Sigma)^T P + P(I + \Sigma)(A - kBC) \right.\\
&\qquad \left. +\mu(A - kBC)^T (I + \Sigma)^T P(I + \Sigma)(A - kBC)\right] x\\
&\le x^T \left[(A - kBC)^T P + P(A - kBC) + \mu(A - kBC)^T P(A - kBC)\right] x\\
&\qquad + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right)\\
&= x^T \left[(A - k^*BC)^T P + P(A - k^*BC) - (k - k^*)((BC)^T P + PBC) \right.\\
&\qquad \left. +\mu(A - kBC)^T P(A - kBC)\right] x + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right)\\
&= -x^T Qx - (k - k^*)x^T ((BC)^T P + PBC)x + \mu x^T (A - kBC)^T P(A - kBC)x\\
&\qquad + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right)\\
&= -x^T Qx - (k - k^*)x^T (C^T B^T C^T C + C^T CBC)x\\
&\qquad + \mu x^T [A^T PA - k(C^T B^T PA + A^T PBC) + k^2 C^T B^T C^T CBC]x\\
&\qquad + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right)\\
&= -x^T Qx + k^* x^T (C^T B^T C^T C + C^T CBC)x - ky^T (B^T C^T + CB)y\\
&\qquad + \mu x^T [A^T PA - k(C^T B^T PA + A^T PBC)]x\\
&\qquad + \mu k^2 y^T [(CB)^T CB]y + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right)\\
&= -x^T Qx + k^* x^T (C^T B^T C^T C + C^T CBC)x\\
&\qquad + \mu x^T [A^T PA - k(C^T B^T PA + A^T PBC)]x\\
&\qquad - ky^T [(CB)^T + CB - \mu k(CB)^T CB]y + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right).
\end{aligned} \tag{5.10}$$

Now let

$$||B^T C^T C + C^T CB|| := \eta_1, \quad ||A^T PA|| := \eta_2, \quad ||B^T PA|| := \eta_3,$$
$$Q_2(t) := (CB)^T + CB - \mu(t)k(t)(CB)^T CB > \varepsilon > 0 \text{ for all } t > t^*.$$

Using these, we obtain the following estimates from (5.10):

$$\begin{aligned}
V^\Delta &\le -(\lambda_{\min}(Q) - \mu\eta_2 - \mu k\beta\eta_3)x^T x - (k\lambda_{\min}(Q_2) - k^*\eta_1 - \mu k\frac{1}{\beta}\eta_3)y^T y + k \left(\alpha_1||\Sigma|| + \alpha_2||\Sigma||^2\right)\\
&\le -(\lambda_{\min}(Q) - \mu\eta_2 - \mu k\beta\eta_3 - k\alpha_1||\Sigma|| - k\alpha_2||\Sigma||^2)x^T x - (k\varepsilon - k^*\eta_1 - \mu k\frac{1}{\beta}\eta_3)y^T y.
\end{aligned}$$

Note that

$$k||\Sigma|| = k||\,\mathrm{expc}(\mu A) - I|| \le k(\exp(||\mu A||) - 1) \to 0 \text{ as } \mu \to 0.$$

Hence,

$$V^\Delta \le -\gamma x^T x \le \frac{-\gamma}{\lambda_{\min}(P)} x^T Px = \frac{-\gamma}{\lambda_{\min}(P)} V := -\eta V.$$

By Theorem 7.7 and Lemma 5.1, $V(t) \le V(t^*)e_{-\eta}(t, t^*)$, $t \ge t^*$ and $\eta \in \mathcal{R}^+$. Then there exists some $K \ge 1$ such that $e_\eta(t, t^*) \le K^2(t^*)e^{-2\alpha(t-t^*)}$, where

$$\alpha = -\frac{1}{2} \limsup_{t\to\infty} \frac{1}{t - t^*} \int_{t^*}^{t} \frac{\mathrm{Log}\,|1 - \eta(\tau)\mu(\tau)|}{\mu(\tau)} \Delta\tau > 0.$$

Consequently,

$$||x(t)|| \leq K||x(t^*)||e^{-\alpha(t-t^*)}.$$

□

The previous theorem, and the lemmas leading to it, did not assume that the system under consideration was necessarily continuous or discrete. However, in the special case where it is desired that $\mu(t) > 0$ over all time (i.e. the system is purely discrete, or in time scales jargon, consists only of isolated points), the notation we employ somewhat betrays the fact that we still have a fundamentally sampled-data system architecture, and therefore it cannot be guaranteed that the plant output or state (both functions of truly continuous time) will necessarily converge. There are a number of ways to examine the problem of intersample behavior, with both [13] and [18] giving coverage of the topic. In particular, Ilchmann and Townley in [13] discuss how convergence of the intersample state (and therefore output) can be guaranteed under an assumption of detectability. We refer the interested reader to these works.

One more lemma is required before the main theorem is presented.

Lemma 5.2. *If $\mathbb{T}$ is time scale with bounded graininess (i.e. $\mu_\infty := \sup_{t\in\mathbb{T}} \mu(t) < \infty$), then*

$$c_1 \int_{t_0}^{\infty} e^{\alpha t}\, dt \leq \int_{t_0}^{\infty} e^{\alpha t}\, \Delta t \leq c_2 \int_{t_0}^{\infty} e^{\alpha t}\, dt$$

where $c_1, c_2, \alpha \in \mathbb{R}$ with $c_1, c_2 > 0$.

Proof. Consider the case when $\alpha > 0$. The process of time scale integration is akin to the approximation of a continuous integral via a left-endpoint sum of (variable width) rectangles. If the function to be "summed" is increasing (as in this case), the sum of retangular areas will be less than the continuous integral, meaning $c_2 = 1$ and $c_1 < 1$. One estimate of the lower bound, then, would follow by simply integrating a multiple of the original function, $c_1 e^{\alpha t}$ and gradually increasing c_1 until $c_1 e^{\alpha t}$ just passes through one of the rectangle right endpoints, which are given by $e^{\alpha\rho(t)}$. Thus $c_1 e^{\alpha t} \leq e^{\alpha\rho(t)}$, or equivalently, $c_1 e^{\alpha\sigma(t)} \leq e^{\alpha t}$, which in turn yields

$$\log c_1 + \alpha(\sigma(t) - t) \leq 0 \implies \log c_1 + \alpha\mu(t) \leq 0 \implies c_1 \leq e^{-\alpha\mu(t)}.$$

Therefore, the most conservative bound is given by $c_1 = e^{-\alpha\mu_\infty}$. The case $\alpha < 0$ can be argued similarly, leading to the following bounds:

$$c_1 = \begin{cases} e^{-\alpha\mu_\infty}, & \alpha > 0, \\ 1, & \alpha \leq 0, \end{cases} \qquad c_2 = \begin{cases} 1, & \alpha \geq 0, \\ e^{-\alpha\mu_\infty}, & \alpha < 0. \end{cases}$$

□

We are now in a position to state the main theorem of the paper.

Theorem 5.4. *Suppose the system* (3.3) *satisfies* (A1)–(A3)*, update laws* (3.4)*,* (3.5)*, and*

- (i) *the k-dynamics in* (3.4) *exhibit radial unboundedness, that is, $f(y) \leq \|y\|^2$ for all $y \in \Omega := \{w(t) \in \mathbb{R}^n : |w_i| < \bar{w}_i \quad \forall t \in \mathbb{T}\}$,*
- (ii) *the μ-dynamics defined by* (3.5) *generate $\mu(t)$ adhering to the relationship* (8.1) *for $t > t_0 \in \mathbb{T}$, where μ decreases monotonically if $\mu > 0$.*

Then $\lim_{t\to\infty} k(t) < \infty$ *and* $y(t) \in L^2[0,\infty)_{\mathbb{T}}$.

Proof. For the sake of contradiction, suppose $k(t) \to \infty$ as $t \to \infty$. Theorems 5.2, 5.3 yield $x(t) \in L^{\infty}[t_0, \infty)_{\mathbb{T}}$ and consequently $y(t) \in L^{\infty}[t_0, \infty)_{\mathbb{T}}$. Since the system $x^{\Delta}(t) = \mathcal{A}(t)x(t)$ has a unique solution over the finite time interval $t \in [0, t_0)$, the state must remain bounded in that interval and therefore $y(t) \in L^{\infty}[0, \infty)_{\mathbb{T}}$. Consequently, y is bounded, ensuring that $\Omega \neq \emptyset$. By Theorem 7.6 of Appendix A, the solution for $k(t)$ is

$$k(t) = k_0 + \int_0^t f(y(\tau))\Delta\tau \le k_0 + \int_0^t \|y\|^2 \Delta\tau \le k_0 + \|C\| \int_0^t \|x\|^2 \Delta\tau,$$

and it therefore follows that

$$\begin{aligned}
\lim_{t\to\infty} k(t) &\le k_0 + \|C\| \int_0^{\infty} \|x\|^2 \Delta\tau \\
&= k_0 + \|C\| \left[\int_0^{t_0} \|x\|^2 \Delta\tau + \int_{t_0}^{\infty} \|x\|^2 \Delta\tau \right] \\
&\le k_0 + \|C\| \left[\int_0^{t_0} \|x\|^2 \Delta\tau + \|x_0\|^2 K^2 \int_{t_0}^{\infty} e^{-\alpha(\tau - t_0)} \Delta\tau \right] \\
&= k_0 + \|C\| \left[\int_0^{t_0} \|x\|^2 \Delta\tau + \|x_0\|^2 K^2 \int_0^{\infty} e^{-\alpha\tau} \Delta\tau \right] \\
&< \infty.
\end{aligned}$$

Note that Lemma 5.2 was used above, because μ_{∞} exists due to condition (ii). The original assumption that $k(t) \to \infty$ as $t \to \infty$ is contradicted, so it must be that $k(t) < \infty$ for $t > 0$. By the arguments above, it also immediately follows that $\int_0^{\infty} \|y\|^2 \Delta\tau < \infty$. □

The arguments above reveal what a wide variety of gain update laws will work for the high-gain adaptive controller. In fact, the possibilities are even more diverse than what we have specified, c.f. [11]. However, the simple choice $k^{\Delta} = \|y\|^2$ pervades the literature (probably for its simplicity), and that is the law we will adopt in this paper. Perhaps more interesting is the inherent graininess update scheme because the constraints are a little tighter: $\mu(t)$ must not only decrease monotonically, but decrease "fast enough" to stay underneath some function of order $\frac{1}{k}$. The relationship between k and μ is more generally given in (8.1), but it is not obvious how to construct a graininess update law that adheres to (8.1) directly. In light of this general time scale expression, we now briefly examine the methods proposed in [13] and [18].

Owens [18] originally proposed that

$$0 < \lim_{\mu\to 0} \mu k(\mu) CB < 2$$

for a single-output, single-input system. Since this condition is equivalent to

$$0 < \lim_{k\to\infty} \mu(k) k CB < 2,$$

we define

$$\lim_{k\to\infty} \mu(k) k := \widehat{\mu k},$$

and note the following manipulation of (8.1):

$$\begin{aligned}\mu\frac{|\lambda_i[\mathcal{A}]|^2+\varepsilon_3^2}{-(\mathrm{Re}(\lambda_i[\mathcal{A}])+\varepsilon_3)} &\geq \mu\frac{|\lambda_i[\mathcal{A}]|^2+\varepsilon_3^2}{|\lambda_i[\mathcal{A}]|-\varepsilon_3}\\ &= \mu(|\lambda_i[\mathcal{A}]|+\varepsilon_3)\\ &\geq \mu\,|\lambda_i[(A-kBC)+\Sigma(\mu)(A-kBC)]|\\ &= |\lambda_i[(\mu A-\mu kBC)+\Sigma(\mu)(\mu A-\mu kBC)]|.\end{aligned}$$

Allowing $\mu k \to \widehat{\mu k}$ (implying that $k \to \infty$ and $\mu \to 0$), we see

$$|\lambda_i[(\mu A-\mu kBC)+\Sigma(\mu)(\mu A-\mu kBC)]| \to \left|\lambda_i[\widehat{\mu k}BC]\right| = \begin{cases}0, & i \neq n,\\ \widehat{\mu k}CB, & i = n.\end{cases}$$

The last step follows from the fact the CB is nonzero (which we know via the minimum phase assumption). So, upon enforcing the strict inequality $\widehat{\mu k}CB < 2$, we can rewrite (8.1) as

$$0 < \mu\frac{|\lambda_i|^2+\varepsilon_3^2}{-(\mathrm{Re}(\lambda_i)+\varepsilon_3)} = \widehat{\mu k}CB + v(\mu) \leq 2,$$

where the function $v(\mu) \to 0$ as $\mu \to 0$, and therefore Lemma 8.3 still holds: there is some $\mu > 0$ that will give $\mathrm{Re}_\mu(\lambda_i[\mathcal{A}]) < 0$. The Owens condition is the least restrictive possible; in other words, it allows for the slowest possible decline of $\mu(t)$, with perhaps the simplest update law being

$$\mu(t) = \frac{CB-\varepsilon}{k(t)}.$$

This will converge to $\mu(t) > 0$ simply because $k(t) \to k_\infty < \infty$. Clearly, condition (iii) of Theorem 5.3 reduces to $0 < \mu kCB < 2$ when CB is a scalar (i.e. the system is in SISO).

On the other hand, the Owens condition requires knowledge of CB; furthermore, it is not obvious how the condition extends to multi-input, multi-output systems. For this reason, Ilchmann and Townley [13] proposed the update law

$$\mu(t) = \frac{1}{k(t)\log k(t)},$$

which obviously results in $\lim_{k\to\infty}\mu(k)k = 0$ and thus guarantees the fulfillment of (8.1) at some point in time, without the need to know or identify any system parameters.

Lastly, we comment on a motivating example found early in [18]. The example points out that, given feedback law $u(t) = -k(t)y(t)$, the scalar system $\dot{y}(t) = cbu(t)$ with $cb > 0$ may diverge when discretized, regardless of the chosen sample period h, merely by choosing initial conditions such that

$$cbh(k_0 + hy_0^2(1-cbhk_0)^2) > 2.$$

In the context of time scale methods, this is again an illustrative example. In canonical form, the system is given by

$$\dot{x}(t) = bu(t), \quad y(t) = cx(t), \quad t \in \mathbb{R},$$

and its time scale realization is simply

$$x^\Delta(t) = bu(t), \quad y(t) = cx(t), \quad t \in \mathbb{T}.$$

Clearly $\mathcal{A} = -kcb$, so of course $\lambda[\mathcal{A}] = -kcb$. One easily obtains $\mathrm{Re}_\mu(\lambda[\mathcal{A}]) < 0$ by choosing μ and k in obedience to (8.1), which simplifies to $0 < \mu kcb < 2$. Intuitively, this says we only have to adjust $\mu(t)$ to place the system pole inside of the Hilger circle. Conversely, if the pole is not placed in the circle, e.g. $|\mu kcb| > 2$, the system may be unstable. Thus, it is clear that instability is indeed a function of both the graininess and the initial condition k_0.

6. Conclusions

In concluding, we again highlight the fact that several variations of the basic high-gain controller presented here exist in the literature, including extensions to infinite-dimensional systems [17], systems under integral control [20] and certain classes of nonlinear systems [14]. The main contributions of this paper are the unification and generalization of the previously disparate discrete and continuous time analyses and doing so with a novel proof that employs the power of time scale methods to handle the problem of nonuniform sampling in a straightforward manner.

Taking a broader view, however, there is no apparent reason why many other types of systems could not be adapted for nonuniform sampling using a similar type of time scale analysis. Nonuniform sampling is actually a fairly common phenomenon, existing in the form of "jitter" any time a computer system controls the sampling. Synchonization jitter can be significantly amplified in systems with networked distributed and embedded controllers [5], a problem for which relatively few analysis and design methods exist. The utility of time scale analysis in this context, however, becomes truly clear when one considers the possibility of dynamically generating the time scale itself *on the fly*. In fact, the adaptive controllers in this paper are an example of this idea, in the special case where the sample rate is forced to continually increase. However, in general, systems of the form

$$x^\Delta = f(x), \quad \mu = g(x) \geq 0,$$

have not been extensively studied in a control context, to our knowledge. The ability to dynamically adapt the system time scale (which we term *μ-dynamics*) has the subtle but powerful consequence that the system software (which calculates $\mu = g(x)$) and the system hardware (which evolves as $x^\Delta = f(x)$) can be designed in a holistic way. For example, the exponential stability set of Definition 1 suggests that it may not always be necessary to sample at some maximum uniform rate; sometimes, perhaps if network traffic is especially heavy in a distributed controller, it might be desirable to decrease the sample rate even to the point of temporary instability, if done for "short enough" periods of time. When and how to do this are subjects of future study, but subjects that now seem significantly more tractable in view of modern mathematical advances like dynamical systems on time scales.

We thank our colleague Robert Marks for fruitful conversations throughout the writing of this paper.

7. Appendix A: Time Scales Primer

7.1. **What Are Time Scales?** A thorough introduction to dynamic equations on time scales is beyond the scope of this appendix. In short, the theory springs from the 1988 doctoral dissertation of Stefan Hilger [10] that resulted in his seminal paper

[9] in 1990. These works aimed to unify and generalize various mathematical concepts from the theories of discrete and continuous dynamical systems. Afterwards, the body of knowledge concerning time scales advanced fairly quickly, culminating in the excellent introductory text by Bohner and Peterson [4] and their more recent advanced monograph [3]. A succint survey on time scales can be found in [1].

A *time scale* $\mathbb{T}$ is any nonempty closed subset of the real numbers $\mathbb{R}$. Thus time scales can be any of the usual integer subsets (e.g. $\mathbb{Z}$ or $\mathbb{N}$), the entire real line $\mathbb{R}$, or any combination of discrete points unioned with continuous intervals. The bulk of engineering systems theory to date rests on two time scales, $\mathbb{R}$ and $\mathbb{Z}$ (or more generally $h\mathbb{Z}$, meaning discrete points separated by distance h). However, as this paper illustrates, there are occasions when necessity or convenience dictates the use of an alternate time scale. The question of how to approach the study of dynamical systems on time scales then becomes relevant, and in fact the majority of research on time scales so far has focused on expanding and generalizing the vast suite of tools available to the differential and difference equation theorist. We now briefly outline the portions of the time scales theory that are needed for this paper to be as self-contained as is practically possible.

The *forward jump operator of* $\mathbb{T}$, $\sigma(t) : \mathbb{T} \to \mathbb{T}$, is given by $\sigma(t) = \inf_{s\in\mathbb{T}}\{s > t\}$. The *backward jump operator of* $\mathbb{T}$, $\rho(t) : \mathbb{T} \to \mathbb{T}$, is given by $\rho(t) = \sup_{s\in\mathbb{T}}\{s < t\}$. The *graininess function* $\mu(t) : \mathbb{T} \to [0, \infty)$ is given by $\mu(t) = \sigma(t) - t$. Here we adopt the conventions $\inf \emptyset = \sup \mathbb{T}$ (i.e. $\sigma(t) = t$ if $\mathbb{T}$ has a maximum element t), and $\sup \emptyset = \inf \mathbb{T}$ (i.e. $\rho(t) = t$ if $\mathbb{T}$ has a minimum element t).

A point $t \in \mathbb{T}$ is *right-scattered* if $\sigma(t) > t$ and *right dense* if $\sigma(t) = t$. A point $t \in \mathbb{T}$ is *left-scattered* if $\rho(t) < t$ and *left dense* if $\rho(t) = t$. If t is both left-scattered and right-scattered, we say t is *isolated*. If t is both left-dense and right-dense, we say t is *dense*. The set $\mathbb{T}^\kappa$ is defined as follows: if $\mathbb{T}$ has a left-scattered maximum m, then $\mathbb{T}^\kappa = \mathbb{T} - \{m\}$; otherwise, $\mathbb{T}^\kappa = \mathbb{T}$. If $f : \mathbb{T} \to \mathbb{R}$ is a function, then the composition $f(\sigma(t))$ is often denoted by $f^\sigma(t)$.

For $f : \mathbb{T} \to \mathbb{R}$ and $t \in \mathbb{T}^\kappa$, define $f^\Delta(t)$ as the number (when it exists), with the property that, for any $\varepsilon > 0$, there exists a neighborhood U of t such that

$$\big|[f(\sigma(t)) - f(s)] - f^\Delta(t)[\sigma(t) - s]\big| \le \epsilon|\sigma(t) - s|, \quad \forall s \in U.$$

The function $f^\Delta : \mathbb{T}^\kappa \to \mathbb{R}$ is called the *delta derivative* or the *Hilger derivative* of f on $\mathbb{T}^\kappa$. We say f is *delta differentiable* on $\mathbb{T}^\kappa$ provided $f^\Delta(t)$ exists for all $t \in \mathbb{T}^\kappa$.

The following theorem establishes several important observations regarding delta derivatives.

Theorem 7.1. [4, p. 5] *Suppose $f : \mathbb{T} \to \mathbb{R}$ and $t \in \mathbb{T}^\kappa$.*

(i) *If f is delta differentiable at t, then f is continuous at t.*
(ii) *If f is continuous at t and t is right-scattered, then f is delta differentiable at t and $f^\Delta(t) = \frac{f(\sigma(t))-f(t)}{\mu(t)}$.*
(iii) *If t is right-dense, then f is delta differentiable at t if and only if $\lim_{s\to t} \frac{f(t)-f(s)}{t-s}$ exists. In this case, $f^\Delta(t) = \lim_{s\to t} \frac{f(t)-f(s)}{t-s}$.*
(iv) *If f is delta differentiable at t, then $f(\sigma(t)) = f(t) + \mu(t)f^\Delta(t)$.*

Note that f^Δ is precisely f' from the usual calculus when $\mathbb{T} = \mathbb{R}$. On the other hand, $f^\Delta = \Delta f = f(t+1) - f(t)$ (i.e. the forward difference operator) on the time scale $\mathbb{T} = \mathbb{Z}$. These are but two very special (and rather simple) examples of time scales. Moreover, the realms of differential equations and difference equations can now be viewed as but special, particular cases of more general *dynamic equations on time scales*, i.e. equations involving the delta derivative(s) of some unknown function.

TABLE 1. Basic notions from time scales calculus.

$\mathbb{T}=\mathbb{R}$	$\mathbb{T}=\mathbb{Z}$	Any $\mathbb{T}$
$(k\,f)'=k\cdot f'$	$\Delta(k\,f)=k\Delta f$	$(k\,f)^\Delta=k\cdot f^\Delta$
$(f+g)'=f'+g'$	$\Delta(f+g)=\Delta f+\Delta g$	$(f+g)^\Delta=f^\Delta+g^\Delta$
$(fg)'=fg'+f'g$	$\Delta\,[fg]=f\Delta g+\Delta f\cdot g(t+1)$	$(fg)^\Delta=f\cdot g^\Delta+f^\Delta\cdot g^\sigma$
$(f/g)'=\frac{f'g-fg'}{g^2}$	$\Delta(f/g)=\frac{\Delta f\cdot g-f\Delta g}{g\cdot g(t+1)}$	$(f/g)^\Delta=\frac{f^\Delta g-f\cdot g^\Delta}{g\cdot g^\sigma}$

A function $f:\mathbb{T}\to\mathbb{R}$ is called *regulated* provided its right- and left-sided limits exist at all right- and left-dense points in $\mathbb{T}$, respectively. A function $f:\mathbb{T}\to\mathbb{R}$ is *rd-continuous* if f is continuous at every right dense point $t\in\mathbb{T}$, and its left hand limit exists at each left dense point $t\in\mathbb{T}$. The set of rd-continuous functions $f:\mathbb{T}\to\mathbb{R}$ will be denoted by $C_{\mathrm{rd}}=C_{\mathrm{rd}}(\mathbb{T})=C_{\mathrm{rd}}(\mathbb{T},\mathbb{R})$.

A function $F:\mathbb{T}\to\mathbb{R}$ is called an *antiderivative* of $f:\mathbb{T}\to\mathbb{R}$ provided $F^\Delta(t)=f(t)$ holds for all $t\in\mathbb{T}^\kappa$. The *indefinite integral* of a regulated function f is defined as $\int f(t)\Delta t=F(t)+C$, where C is an arbitrary constant and F is any antiderivative of f. On the other hand, the *Cauchy integral* or *definite integral* is then

$$\int_a^b f(t)\Delta t=F(b)-F(a)\quad\text{ for all } a,b\in\mathbb{T},$$

where F is any antiderivative of f.

Theorem 7.2. [4, p. 27] *Every rd-continuous function has an antiderivative. If $t_0\in\mathbb{T}$, then $F(t)=\int_{t_0}^t f(\tau)\,\Delta\tau$, $t\in\mathbb{T}$, is an antiderivative of f.*

The following two theorems are useful in computing time scales integrals.

Theorem 7.3. [4, p. 28] *If $f\in C_{\mathrm{rd}}$ and $t\in\mathbb{T}^\kappa$, then $\int_t^{\sigma(t)} f(\tau)\,\Delta\tau=f(t)\mu(t)$.*

Theorem 7.4. [4, p. 29] *Suppose $a,b\in\mathbb{T}$ and $f\in C_{\mathrm{rd}}$.*

(i) *If $\mathbb{T}=\mathbb{R}$, then $\int_a^b f(t)\,\Delta t=\int_a^b f(t)\,dt$, where the integral on the right is the usual Riemann integral.*
(ii) *If $[a,b]\cap\mathbb{T}:=[a,b]_{\mathbb{T}}$ consists only of isolated points, then*

$$\int_a^b f(t)\,\Delta t=\begin{cases}\sum_{t\in[a,b]_{\mathbb{T}}} f(t)\mu(t), & a<b,\\ 0, & a=b,\\ -\sum_{t\in[a,b]_{\mathbb{T}}} f(t)\mu(t), & a>b.\end{cases}$$

The theorem above reveals that in the continuous case, $\mathbb{T}=\mathbb{R}$, definite integrals are the usual definite integrals from calculus. When $\mathbb{T}=\mathbb{Z}$, definite integrals correspond to definite sums from the difference calculus; see [15].

7.2. **The Hilger Complex Plane.** For $h > 0$, define the *Hilger complex numbers*, the *Hilger real axis*, the *Hilger alternating axis*, and the *Hilger imaginary circle* by

$$\mathbb{C}_h := \left\{ z \in \mathbb{C} : z \neq -\frac{1}{h} \right\}, \quad \mathbb{R}_h := \left\{ z \in \mathbb{R} : z > -\frac{1}{h} \right\},$$
$$\mathbb{A}_h := \left\{ z \in \mathbb{R} : z < -\frac{1}{h} \right\}, \quad \mathbb{I}_h := \left\{ z \in \mathbb{C} : \left| z + \frac{1}{h} \right| = \frac{1}{h} \right\},$$

respectively. For $h = 0$, let $\mathbb{C}_0 := \mathbb{C}$, $\mathbb{R}_0 := \mathbb{R}$, $\mathbb{A}_0 := \emptyset$, and $\mathbb{I}_0 := i\mathbb{R}$. See Figure 1.

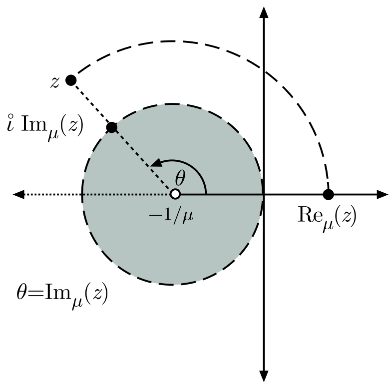

FIGURE 1. The Hilger complex plane.

Let $h > 0$ and $z \in \mathbb{C}_h$. The *Hilger real part of* z is defined by

$$\mathrm{Re}_h(z) := \frac{|zh+1| - 1}{h},$$

and the *Hilger imaginary part of* z is defined by

$$\mathrm{Im}_h(z) := \frac{\mathrm{Arg}(zh+1)}{h},$$

where Arg(z) denotes the principal argument of z (i.e., $-\pi < \mathrm{Arg}(z) \leq \pi$). See Figure 1.

For $h > 0$, define the strip $\mathbb{Z}_h := \{z \in \mathbb{C} : -\frac{\pi}{h} < \mathrm{Im}(z) \leq \frac{\pi}{h}\}$, and for $h = 0$, set $\mathbb{Z}_0 := \mathbb{C}$. Then we can define the *cylinder transformation* $\xi_h : \mathbb{C}_h \to \mathbb{Z}_h$ by

$$\xi_h(z) = \frac{1}{h} \mathrm{Log}(1 + zh), \ h > 0, \tag{7.1}$$

where Log is the principal logarithm function. When $h = 0$, we define $\xi_0(z) = z$, for all $z \in \mathbb{C}$. It then follows that the *inverse cylinder transformation* $\xi_h^{-1} : \mathbb{Z}_h \to \mathbb{C}_h$ is given by

$$\xi_h^{-1}(z) = \frac{e^{zh} - 1}{h}. \tag{7.2}$$

See Figure 2.

Since the graininess may not be constant for a given time scale, we will interchangeably subscript various quantities (such as ξ and ξ^{-1}) with $\mu = \mu(t)$ instead of h to reflect this.

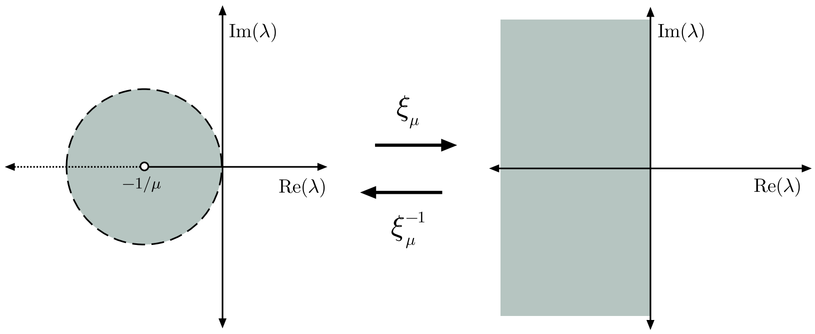

FIGURE 2. The cylinder (7.1) and inverse cylinder (7.2) transformations map the familiar stability region in the continuous case to the interior of the Hilger circle in the general time scale case.

7.3. **Generalized Exponential Functions.** Before using the cylinder transformation to define the generalized exponential function on a time scale, we need the concept of regressivity.

The function $p : \mathbb{T} \to \mathbb{R}$ is *regressive* if $1 + \mu(t)p(t) \neq 0$ for all $t \in \mathbb{T}^\kappa$, and this concept motivates the definition of the following sets:

$$\mathcal{R} = \{p : \mathbb{T} \to \mathbb{R} : p \in C_{\mathrm{rd}}(\mathbb{T}) \text{ and } 1 + \mu(t)p(t) \neq 0 \ \forall t \in \mathbb{T}^\kappa\},$$
$$\mathcal{R}^+ = \{p \in \mathcal{R} : 1 + \mu(t)p(t) > 0 \text{ for all } t \in \mathbb{T}^\kappa\}.$$

A matrix is regressive if and only if all of its eigenvalues are in $\mathcal{R}$.

If $p \in \mathcal{R}$, then we define the *generalized time scale exponential function* by

$$e_p(t,s) = \exp\left(\int_s^t \xi_{\mu(\tau)}(p(\tau))\Delta\tau\right), \qquad \text{for all } s,t \in \mathbb{T}.$$

The following theorem is a compilation of properties of $e_p(t,t_0)$ (some of which are counterintuitive) that we need in the main body of the paper.

Theorem 7.5. [4, Chapter 2] *The function $e_p(t,t_0)$ has the following properties:*

(i) *If $p \in \mathcal{R}$, then $e_p(t,r)e_p(r,s) = e_p(t,s)$ for all $r,s,t \in \mathbb{T}$.*
(ii) $e_p(\sigma(t),s) = (1 + \mu(t)p(t))e_p(t,s)$.
(iii) *If $p \in \mathcal{R}^+$, then $e_p(t,t_0) > 0$ for all $t \in \mathbb{T}$.*
(iv) *If $1 + \mu(t)p(t) < 0$ for some $t \in \mathbb{T}^\kappa$, then $e_p(t,t_0)e_p(\sigma(t),t_0) < 0$.*
(v) *If $\mathbb{T} = \mathbb{R}$, then $e_p(t,s) = e^{\int_s^t p(\tau)d\tau}$. Moreover, if p is constant, then $e_p(t,s) = e^{p(t-s)}$.*
(vi) *If $\mathbb{T} = \mathbb{Z}$, then $e_p(t,s) = \prod_{\tau=s}^{t-1}(1 + p(\tau))$. Moreover, if $\mathbb{T} = h\mathbb{Z}$, with $h > 0$ and p is constant, then $e_p(t,s) = (1 + hp)^{\frac{t-s}{h}}$*

If $p \in \mathcal{R}$ and $f \in C_{\mathrm{rd}}$, then the dynamic equation $y^\Delta(t) = p(t)y(t) + f(t)$ is called *regressive*.

Theorem 7.6. [4, p. 77] *Consider the first order dynamic initial value problem*

$$y^\Delta(t) = p(t)y(t) + f(t), \qquad y(t_0) = y_0. \tag{7.3}$$

If (7.3) *is regressive and* $t_0 \in \mathbb{T}$ *is fixed, then the unique solution is*

$$y(t) = y_0 e_p(t, t_0) + \int_{t_0}^{t} e_p(t, \sigma(\tau)) f(\tau) \Delta\tau.$$

Theorem 7.7. [4, p. 255] *Let* $y, f \in C_{\mathrm{rd}}$ *and* $\in \mathcal{R}^+$. *Then*

$$y^{\Delta}(t) \leq p(t)y(t) + f(t), \quad \text{for all } t \in \mathbb{T},$$

implies

$$y(t) \leq y(t_0)e_p(t, t_0) + \int_{t_0}^{t} e_p(t, \sigma(\tau)) f(\tau)\, \Delta\tau, \quad \text{for all } t \in \mathbb{T}.$$

7.4. **The** expc **Function.**

Definition 7.1. For any scalar or matrix argument X, define

$$\operatorname{expc}(X) = I + \frac{1}{2}X + \frac{1}{6}X^2 + \cdots + \frac{1}{n!}X^{n-1} + \cdots$$

Proposition 7.1. (Properties of $\operatorname{expc}(X)$**)**

(i) $\operatorname{expc}(X)X = X\operatorname{expc}(X)$
(ii) $\operatorname{expc}(X) = (e^X - I)X^{-1}$ *when* X^{-1} *exists.*
(iii) $\xi_\mu^{-1}(X) = \operatorname{expc}(\mu X)X$ *where* ξ_μ^{-1} *is the Hilger inverse cylinder transformation*
(iv) *For real, scalar arguments* x,

$$\operatorname{expc}(ix) = e^{ix/2}\frac{e^{ix/2} - e^{-ix/2}}{ix} = e^{ix/2}\operatorname{sinc}(x/2),$$

where sinc *here denotes the sine cardinal function.* (*This is the motivation for the* expc *notation.*)
(v) $||\operatorname{expc}(X)|| \leq \exp(||X||)$
(vi) $||\operatorname{expc}(X) - I|| \leq \exp(||X||) - 1$

Proof. Parts (i)–(iv) follow immediately from the definition. To verify (v), note

$$\left|\left|\sum_{n=1}^{\infty}\frac{X^{n-1}}{n!}\right|\right| \leq \sum_{n=1}^{\infty}\frac{\|X\|^{n-1}}{n!} \leq \sum_{n=1}^{\infty}\frac{\|X\|^{n-1}}{(n-1)!} = \exp(||X||).$$

Finally, (vi) follows from an argument similar to the one above. □

A final observation that we need is the following. By (vi), we have the decomposition $\operatorname{expc}(\mu X) = I + \Sigma(\mu)$ and $||\Sigma(\mu)|| \leq \exp(\mu||A||) - 1$ so that $||\Sigma(\mu)|| \to 0$ as $\mu \downarrow 0$.

8. Appendix B: More On the Hilger Circle

Knowing the importance of the Hilger circle, it is perhaps not surprising that the ultimate objective of a high-gain adaptive controller will be to move the system eigenvalues into the Hilger circle in such a way that they remain there forever thereafter. Referring again to Figure 3, the roles of $k(t)$ and $\mu(t)$ now become a little clearer: k must increase enough to place the poles of $\mathcal{A}$ strictly in the left-hand complex plane, and $\mu(t)$ must decrease enough to completely enclose all poles inside of the Hilger circle.

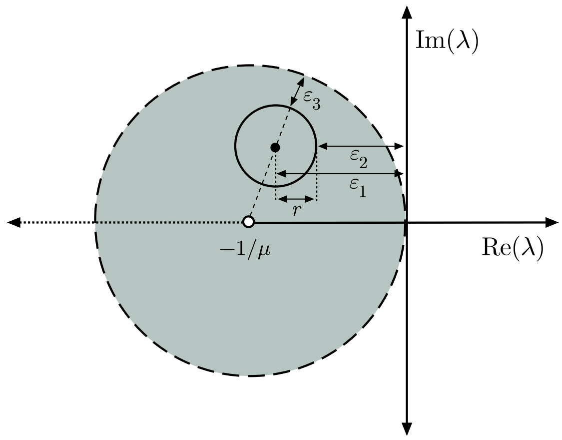

FIGURE 3. A complex eigenvalue $\lambda_i(A - kBC)$ is located inside the Hilger circle. An eigenvalue $\lambda_i(\mathcal{A})$ must be located inside the small circle of radius r.

Lemma 8.1. *Given a matrix* $X \in \mathbb{R}^{n\times n}$*, and a perturbation* $Y \in \mathbb{R}^{n\times n}$*, there exists a permutation* $\rho(i)$ *of the integers* $i = 1, 2, \ldots, n$ *such that*

$$\left|\lambda_i(X+Y) - \lambda_{\rho(i)}(X)\right| \le c\,\|Y\|, \qquad \text{for some } c \ge 1.$$

Proof. Although the proof may arise as a special case of a more general principle, we present one that is straightforward. The proof relies on a consequence of Frobenius' Theorem, that for some positive constant $k > 0$ and matrix X, $\lambda(kX) = k\lambda(X)$. Therefore, we can write the Jordan decomposition of kX as

$$kX = TJ(k)T^{-1}, \quad J(k) = \begin{bmatrix} k\lambda_1 & \delta_1 & & \\ & k\lambda_2 & \delta_1 & \\ & & k\lambda_3 & \ddots \\ & & & \ddots \end{bmatrix},$$

where $\delta_i = 1$ if λ_i and λ_{i+1} are non-simple repeated eigenvalues, and $\delta_i = 0$ otherwise. Given a perturbation matrix Y and defining $\tilde{Y} := T^{-1}YT$, we obtain the spectral equality

$$\lambda_i(kX + kY) = \lambda_i(J(k) + k\tilde{Y}).$$

By Gerschgorin's Theorem, it then follows that some integer permutation $\kappa(i)$ exists for $i = 1, 2, \ldots, n$, such that

$$\left|\lambda_{\kappa(i)}(kX + kY) - J_{ii}(k) - k\tilde{Y}_{ii}\right| \le \sum_{j=1, j\neq i}^{n} \left|k\tilde{Y}_{ij}\right| + \delta_i$$

$$k\left|\lambda_{\kappa(i)}(X+Y) - \lambda_i(X) - \tilde{Y}_{ii}\right| + k\left|\tilde{Y}_{ii}\right| \le k\sum_{j=1}^{n}\left|\tilde{Y}_{ij}\right| + \delta_i$$

$$\left|\lambda_{\kappa(i)}(X+Y) - \lambda_i(X)\right| \le \sum_{j=1}^{n}\left|\tilde{Y}_{ij}\right| + \frac{\delta_i}{k}.$$

By equivalence of norms, the largest absolute column sum of $\tilde{Y}$ is upper bounded by the usual matrix 2-norm, and $||\tilde{Y}|| \leq ||T||||T^{-1}||||Y||$, so that

$$\left|\lambda_{\kappa(i)}(X+Y) - \lambda_i(X)\right| \leq c\,\|Y\| + \frac{\delta_i}{k}, \quad c \geq 1.$$

Since this holds for arbitrary $k > 0$, the tightest bound is given when $k \to \infty$, and the lemma is proved. □

Lemma 8.2. *Suppose that $k(t) \to \infty$ as $t \to \infty$, and $\mu(t) \downarrow 0$ as $t \to \infty$ whenver $\mu(t) > 0$. For the system* (3.3) *which satisfies* (A1)–(A3)*, there is a time $t^* \in \mathbb{T}$ and constant $\varepsilon_2 > 0$ such that* $\operatorname{spec}(\mathcal{A}) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < -\varepsilon_2\}$ *for $t > t^*$.*

Proof. The first step is to observe the decomposition

$$\operatorname{expc}(\mu A) = I + \Sigma(\mu), \qquad \lim_{\mu \to 0} \|\Sigma(\mu)\| = 0.$$

Applying Lemma 8.1, we can now state that

$$\left|\lambda_i\left[(A - kBC) + \Sigma(A - kBC)\right] - \lambda_{\rho(i)}\left[A - kBC\right]\right| \leq c\,\|\Sigma(A - kBC)\|,$$

or

$$\left|\lambda_i\left[\mathcal{A}\right] - \lambda_{\rho(i)}\left[A - kBC\right]\right| \leq c\,\|\Sigma(\mu)(A - kBC)\| := r(\mu).$$

When $k(t) > k_1$, (A2) guarantees that $\operatorname{Re}\lambda_{\rho(i)}\left[A - kBC\right] < -\varepsilon_1$, and therefore it is possible to guarantee that $\operatorname{Re}\lambda_i\left[\mathcal{A}\right] < -\varepsilon_2$ if μ is chosen sufficiently small (say, $\mu(t) = \mu_3$) such that $\varepsilon_2 = \varepsilon_1 - r(\mu) > 0$. Figure 3 illustrates this situation. Let k_1 be achieved at time t_1, μ_2 at time t_2, and μ_3 at time t_3. Under the conditions of monotonicity imposed on $k(t)$ and $\mu(t)$, the result of the lemma occurs when $t^* \geq \max\{t_1, t_2, t_3\}$. □

The preceding lemma does not deal directly with the exact magnitude of $\mu(t^*)$ (or, equivalently, the time t^* itself) necessary to produce the result. It merely guarantees that, if $k \to \infty$ as $t \to \infty$, there must come a time after which eigenvalues of $\mathcal{A}$ will be placed in the left-hand complex plane and they will remain there. For a continuous system, this is usually sufficient to begin a Lyapunov analysis of the system stability. However, in the case of a system on an arbitrary time scale, it is still desired to enclose the eigenvalues in the Hilger circle. Intuitively, once an eigenvalue is known to be restricted to the the left-hand plane, it is evident that some sufficiently small $\mu > 0$ will exist to enclose the eigenvalues in the Hilger circle. The *Hilger real* operator defined in Appendix A captures this notion precisely, and brings us to the following lemma.

Lemma 8.3. *Given $k > k^* := k(t^*)$, there is some $\mu < \mu^* := \mu(t^*)$ and positive constant $\varepsilon_3 < \varepsilon_2$ such that the eigenvalues of $\mathcal{A}$ have $\operatorname{Re}_\mu(\lambda_i\left[\mathcal{A}\right]) \leq -\varepsilon_3$ for $i = 1, 2, \ldots, n$, with t^* and ε_2 as defined in Lemma 8.2.*

Proof. In the case that $\mu = 0$, the lemma is proved by observing that $\operatorname{Re}_\mu(\lambda_i\left[\mathcal{A}\right]) = \operatorname{Re}\left(\lambda_i\left[\mathcal{A}\right]\right)$ and applying Lemma 8.2. In the case that $\mu > 0$, we fix k and simply note that $\operatorname{Re}_\mu(\lambda_i) \leq -\varepsilon_3$ is by definition

$$\frac{|\lambda_i \mu + 1| - 1}{\mu} \leq -\varepsilon_3,$$

which leads to

$$0 < \mu \frac{|\lambda_i|^2 + \varepsilon_3^2}{-(\operatorname{Re}(\lambda_i) + \varepsilon_3)} \leq 2. \tag{8.1}$$

Since the middle term tends to zero with decreasing μ, there must be some sufficiently small $\mu > 0$ for which (8.1) holds. □

From this result, it becomes apparent that $\mu(t)$ and $k(t)$ must share an inverse relationship of at least the first power (meaning $\mu \propto k^{-\beta}$ with $\beta \geq 1$) because as k increases, the ratio $\frac{|\lambda_i|^2}{|\mathrm{Re}\,(\lambda_i)|}$ dominates the expression above and all λ_i are linearly dependent on k. Several such relationships have been proposed [13, 18], but part of the fascination with this type of adaptive controller lies in the incredibly broad freedom to choose the exact relationship.

References

[1] R. Agarwal, M. Bohner, D. O'Regan, A. Peterson, Dynamic equations on time scales: a survey, *Journal of Computational and Applied Mathematics* **141** (2002), 1–26.
[2] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1970.
[3] M. Bohner and A. Peterson, *Advances in Dynamic Equations on Time Scales*, Birkhauser, Boston, 2003.
[4] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales*, Birkhauser, Boston, 2001.
[5] A. Cervin, D. Henriksson, B. Lincoln, J. Eker, K.-E. Årzén, How does control timing affect performance? Analysis and simulation of timing using Jitterbug and TrueTime, *IEEE Control Systems Magazine*, June 2003, 16–30.
[6] J.J. DaCunha, Stability for time varying linear dynamic systems on time scales, submitted (Available: `http://www.timescales.org/`).
[7] J.J. DaCunha, Stabilization Criteria for Nonautonomous Linear Dynamic Systems on Time Scales, Ph.D. Dissertation, Baylor University, 2004.
[8] T. Gard and J. Hoffacker, Asymptotic behavior of natural growth on time scales, *Dynamic Systems and Applications* **12** (2002), 131–147.
[9] S. Hilger, Analysis on measure chains—a unified approach to continuous and discrete calculus, *Results in Mathematics* **18** (1990), 18–56.
[10] S. Hilger, *Ein Masskettenkalkül mit Anwendung auf Zentrumsmannigfaltigkeiten*. Ph.D. thesis, Universität Würzburg, 1988.
[11] A. Ilchmann, D.H. Owens, and D. Prätzel-Wolters, High-gain robust adaptive controllers for multivariable systems, *Systems & Control Letters* **8** (1987), 397–404.
[12] A. Ilchmann and S. Townley, Adaptive high-gain λ-tracking with variable sample rate, *Systems & Control Letters* **36** (1999), 285–293.
[13] A. Ilchmann and S. Townley, Adaptive sampling control of high-gain stabilizable systems, *IEEE Trans. Automatic Control* **44** (1999), 1961–1966.
[14] A. Ilchmann and E.P. Ryan, On gain adaptation in adaptive control, *IEEE Trans. Automatic Control* **48** (2003), 895–899.
[15] W. Kelly and A. Peterson, *Difference Equations: An Introduction with Applications*, Harcourt/Academic Press, Burlington, MA, 2001.
[16] H.K. Khalil and A. Saberi, Adaptive stabilization of a class of nonlinear systems using high-gain feedback, *IEEE Trans. Automatic Control* **43** (1987), 1031–1035.
[17] H. Logemann and B. Martensson, Adaptive stabilization of infinite-dimesional systems, *IEEE Trans. Automatic Control* **37** (1992), 1869–1883.
[18] D.H. Owens, Adaptive stabilization using a variable sampling rate, *Int. J. Control* **63** (1996), 107–119.
[19] D.H. Owens, D. Prätzel-Wolters, and A. Ilchmann, Positive-real structure and high-gain adaptive stabilization, *IMA J. Mathematical Control & Information* **4** (1987), 167–181.
[20] N. Ozdemir and S. Townley, Integral control by variable sampling based on steady-state data, *Automatica* **39** (2003), 135–140.
[21] C. Pötzsche, S. Siegmund, and F. Wirth, A spectral charaterization of exponential stability for linear time-invariant systems on time scales, *Discrete and Continuous Dynamical Systems* **9** (2003), 1223–1241.
[22] J.E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, Upper Saddle River, 1991.
[23] V. Spedding, Taming nature's numbers, *New Scientist*, July 2003, 28–31.

[24] J.C. Willems and C.I. Byrnes, Global adaptive stabilization in the absence of information on the sign of the high frequency gain, *Lecture Notes in Control and Information Sciences* **62** (1984), Springer, Berlin, 49–57.

Department of Engineering, Baylor University, Waco, TX 76798 USA
E-mail address: Ian_Gravagne@baylor.edu

Department of Mathematics, Baylor University, Waco, TX 76798 USA
E-mail address: John_M_Davis@baylor.edu
E-mail address: Jeffrey_DaCunha@baylor.edu