ABSTRACT

Stability and Control on Stochastic Time Scales

Dylan Richard Poulsen, Ph.D.

Advisor: John M. Davis, Ph.D

We develop the stability theory for two classes of dynamic equations evolving on a time domain that is non-uniform and stochastic. In particular, we examine the mean-square exponential stability and almost sure exponential stability of linear, time invariant systems and of linear, time-varying systems, where the variation in time is only due to the local time step.

With the stability theory in hand, we apply our results to control systems evolving on stochastic, non-uniform time domains. We design stabilizing closed-loop feedback controllers, observers, observer-based closed-loop feedback controllers, and optimal closed-loop feedback controllers.

Stability and Control on Stochastic Time Scales

by

Dylan Richard Poulsen, B.S.

A Dissertation

Approved by the Department of Mathematics

Lance L. Littlejohn, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

John M. Davis, Ph.D, Chairperson

Ian A. Gravagne, Ph.D.

Johnny Henderson, Ph.D.

Lance L. Littlejohn, Ph.D.

Qin Sheng, Ph.D.

Accepted by the Graduate School
May 2015

J. Larry Lyon, Ph.D., Dean

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I owe my love of learning to my parents and the many inspiring teachers in my life. Three teachers in particular deserve special mention. Thank you to Adella Croft for seeing my potential in mathematics. Thank you to Sigrun Bodine for introducing me to the joys of mathematical research. Thank you to Mike Spivey for showing me the beauty and fun of applied mathematics.

To the faculty and graduate students in the Department of Mathematics at Baylor University, thank you for creating a genuine sense of community. Through discussions both inside and outside the classroom, I have learned a great amount over the past five years.

This work would not be possible without the support of the time scales research group at Baylor University. Thank you to Matthew Mosley for taking the ideas in this work to the laboratory and for generating new questions along the way. Thank you to Bob Marks for helping me to hone the idea of stochastic time scales with wonderful questions. Thank you to Geoff Eisenbarth for the many great discussions and for sharing insights about time scales. Thank you to Ian Gravagne for guiding my ideas toward the applicable and for insightful questions each week during seminar.

Finally, I am grateful that John Davis encouraged me to study with him at Baylor University while we sat at a lunch table in Laramie, Wyoming. This was not the last time he gave me good advice. Thank you to John Davis for the immense support and guidance.

To the love of my life,

Kayla

CHAPTER ONE

Introduction

From cruise-control systems to rocket dynamics, control theory forms a foundation for our modern society. Much of control theory relies on updates to the system occurring at uniform, predictable moments in time. As control systems become distributed over large scales or become controlled by low-speed devices, the uniformity and predictability of the underlying time domain cannot be guaranteed. In this dissertation, we develop a stability theory for linear systems evolving on non-uniform and random time domains. In order to study such systems, we utilize and develop the theory of *dynamic equations on time scales*, a recent theory which unifies and extends continuous and discrete analysis. Many of the results will consider the case where the time scale is generated in a stochastic manner, allowing us to study uncertainty in the time domain.

The dissertation is organized as follows. In Chapter Two, we introduce background material necessary to understand the rest of the dissertation. In particular, we will introduce the theory of dynamic equations on time scales, focusing on the existence and uniqueness of solutions to the first-order, linear dynamic equation on time scales, which generalizes the exponential function. Next, we will introduce the concept of stochastic time scales and their relation with standard time scales. Then, we will review the various notions of stability used in both the deterministic and stochastic setting. Finally, we will introduce concepts, models, and results from control theory that we study later in the work.

In Chapter Three, we develop stability theory for both linear time invariant and the linear μ-varying systems evolving on stochastic time scales. Utilizing both direct and indirect methods, we completely classify the notions of exponential stability almost surely and of mean-square exponential stability for both classes of dynamic equations. We show that the mean-square exponential stability of both classes of dynamic equations is equivalent to a positive definite solution of an associated matrix Lyapunov equation, generalizing known results on $\mathbb{R}$ and $\mathbb{Z}$. In the linear μ-varying case, the matrix Lyapunov equation cannot be solved directly. Therefore, we provide sufficient conditions for a solution to the Lyapunov equation using a fixed point theorem of Ran and Reurings [38].

In Chapter Four, we apply the stability theory of Chapter Three to the two main design problems in control engineering: the observer problem and the controller problem. We begin with the design of a novel time scale observer and apply it to the problem of estimating battery state-of-charge. This observer, however, requires knowledge of the duration of a future time step, making it unsuitable for state feedback. We attempt to fix this issue by designing an observer which does not require future knowledge of the time step. Using the corollary to the fixed point theorem of Ran and Reurings, we find a sufficient condition that guarantees the effectiveness of the observer. Next, we switch our focus to the controller problem. We provide a theorem which produces the control law of both the linear time invariant control system as well as the linear μ-varying control system which minimizes a quadratic cost functional involving the state and the controls. For both classes of control systems, the optimal control is obtained by solving an associated Riccati matrix equation, generalizing celebrated results on $\mathbb{R}$ and $\mathbb{Z}$.

In Chapter Five, we use the results concerning stability theory on stochastic time scales to arrive at results for deterministic time scales. Inspired by the geometric relationship between the region of mean-square exponential stability and exponential stability almost surely in the linear time invariant case on stochastic time scales, we give a formula for the largest Hilger circle contained in the region of exponential stability for the linear time invariant case on deterministic time scales. This formula resolves an open question recently posed by Doan *et al.* [17]. Finally, we discuss how the optimal control theory of stochastic time scales can be applied and interpreted on deterministic time scales.

CHAPTER TWO

Preliminaries

2.1 Time Scales Preliminaries

Stefan Hilger developed the theory of *Dynamic Equations on Time Scales* in his 1988 dissertation [25]. He successfully unified the theory of difference equations and differential equations by showing they are special classes of a broader theory. A time scale, which we denote by $\mathbb{T}$, is an arbitrary closed subset of the real numbers. If $\mathbb{T} = \mathbb{R}$, the theory yields calculus and differential equations. If $\mathbb{T} = \mathbb{Z}$, the theory yields difference calculus and difference equations. The power of the theory lies in the arbitrary choice of the time scale; one could choose a mixture of discrete points and closed intervals for the time scale, even a Cantor ternary set, and the theory would describe how to analyze dynamic equations defined on the time scale. Especially since the book by Bohner and Peterson [7], the field has grown substantially and supports an active community of researchers examining many areas of mathematics with a time scales viewpoint. In this section, we introduce the background necessary for this thesis.

For a point $t \in \mathbb{T}$, the *forward jump operator*, $\sigma(t)$, is defined as the point immediately to the right of t, in the sense that

$$\sigma(t) = \inf\{s \in \mathbb{T} \mid s > t\}.$$

The *graininess* is the distance between points defined as

$$\mu(t) := \sigma(t) - t.$$

When $\mathbb{T} = \mathbb{R}$, $\sigma(t) = t$ and $\mu(t) = 0$, whereas when $\mathbb{T} = \mathbb{Z}, \sigma(t) = t+1$ and $\mu(t) = 1$. We can classify whether a given point $t \in \mathbb{T}$ is behaving as a continuous or discrete point.

Definition 2.1. An element $t \in \mathbb{T}$ is called *right-dense* if $\mu(t) = t$ and *right-scattered* if $\mu(t) > t$. A time scale is *purely discrete* if $\mu(t) > 0$ for all $t \in \mathbb{T}$.

The *time scale* or *Hilger derivative* of a function $x(t)$ on $\mathbb{T}$ is defined as

$$x^{\Delta}(t) := \frac{x(\sigma(t)) - x(t)}{\mu(t)},$$

and is interpreted in the limit as $\mu \to 0^+$ when $\mu(t) = 0$. When $\mathbb{T} = \mathbb{R}$, $x^{\Delta}(t) = \frac{d}{dt}x(t)$, while when $\mathbb{T} = \mathbb{Z}, x^{\Delta}(t) = x(t+1) - x(t) = \Delta x(t)$, where Δ is the forward difference operator.

The *Hilger integral* can be viewed as the antiderivative or Cauchy integral in the sense that, if $y(t) = x^{\Delta}(t)$, then for $s, t \in \mathbb{T}$,

$$\int_{\tau=s}^{t} y(\tau)\Delta\tau = x(t) - x(s).$$

For each $\gamma > 0$, define the *Hilger Circle* by

$$\mathcal{H}_{\gamma} := \left\{ z \in \mathbb{C} \,\middle|\, |1 + z\mu(t)| < 1, z \neq -\frac{1}{\gamma} \right\} \tag{2.1}$$

Note that $\mathcal{H}_{\gamma}$ is a disc of radius $1/\mu(t)$ contained in the left half–plane tangent to the imaginary axis. We interpret $\mathcal{H}_0$ as the open left-half complex plane. We note that the misnomer Hilger circle is prominent in the literature, despite the fact that the Hilger circle is a disc.

The Hilger circle is extremely important for defining the generalization of the exponential function to time scales. In order to define the time scale exponential function, we study the first order initial value problems

$$x^{\Delta} = \lambda(t)x; \qquad x(t_0) = x_0, \tag{2.2}$$

where $\lambda : \mathbb{T} \to \mathbb{C}$, and

$$x^{\Delta} = \lambda x; \qquad x(t_0) = x_0, \tag{2.3}$$

where $\lambda \in \mathbb{C}$. In order to obtain existence and uniqueness results for (2.2) and (2.3), we require a generalization of the complex plane and some conditions on λ. First, we examine the generalization of the complex plane. Given $t \in \mathbb{T}$ with $\mu(t) > 0$, we can map the complex plane into the set $\mathbb{C}_{\mu(t)} := \mathbb{C}\backslash\{1/\mu(t)\}$, called the *Hilger complex plane* as follows.

First, define

$$\mathbb{Z}_{\mu(t)} := \left\{ z \in \mathbb{C} \,\middle|\, -\frac{\pi}{\mu(t)} \leq \operatorname{Im}(z) \leq \frac{\pi}{\mu(t)} \right\}.$$

Next, define the *cylinder transformation* $\xi_{\mu(t)} : \mathbb{C}_{\mu(t)} \to \mathbb{Z}_{\mu(t)}$ by

$$\xi_{\mu(t)}(z) := \frac{1}{\mu(t)} \operatorname{Log}(1 + z\mu(t)),$$

where Log is the principal logarithm. The *inverse cylinder transformation* is then given by

$$\xi_{\mu(t)}^{-1}(z) = \frac{e^{z\mu(t)} - 1}{\mu(t)}. \tag{2.4}$$

In the limiting case where $\mu(t) = 0$, we define $\mathbb{C}_0 := \mathbb{C}$, and $\xi_0(z) := z$. The effect of the cylinder transformation is shown in Figure 2.1

The *Hilger real axis* $\mathbb{R}_{\mu(t)}$, the *Hilger alternating axis* $\mathbb{A}_{\mu(t)}$, and the *Hilger imaginary circle* $\mathbb{H}_{\mu(t)}$ are defined as follows:

$$\mathbb{R}_{\mu(t)} := \left\{ z \in \mathbb{C}_{\mu(t)} \,\middle|\, z \in \mathbb{R} \text{ and } z > -\frac{1}{\mu(t)} \right\},$$

$$\mathbb{A}_{\mu(t)} := \left\{ z \in \mathbb{C}_{\mu(t)} \,\middle|\, z \in \mathbb{R} \text{ and } z < -\frac{1}{\mu(t)} \right\},$$

$$\mathbb{H}_{\mu(t)} := \left\{ z \in \mathbb{C}_{\mu(t)} \,\middle|\, \left| z + \frac{1}{\mu(t)} \right| = \frac{1}{\mu(t)} \right\}.$$

When $\mu(t) = 0$, we define $\mathbb{R}_0 = \mathbb{R}$, $\mathbb{A}_0 = \emptyset$ and $\mathbb{H}_0 = \mathbb{I}$, the imaginary axis. Note that the boundary of the Hilger circle $\mathcal{H}_{\mu(t)}$ is the Hilger imaginary circle $\mathbb{H}_{\mu(t)}$. Within Hilger's complex plane, we can now discuss the Hilger real part and Hilger imaginary part of a complex number z. The *Hilger real part* of $z \in \mathbb{C}$, $\operatorname{Re}_{\mu(t)}(z)$, is given by

$$\operatorname{Re}_{\mu(t)}(z) := \lim_{s \downarrow \mu(t)} \frac{|1 + sz| - 1}{s},$$

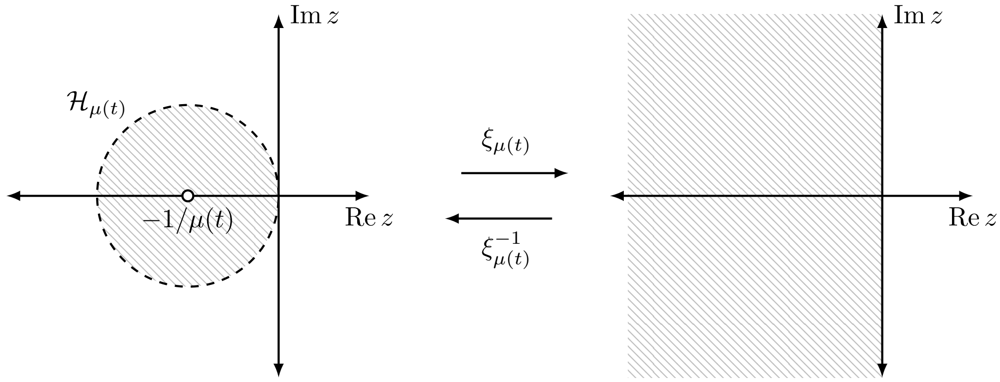

Figure 2.1. The effect of the cylinder transformation $\xi_{\mu(t)}$ and its inverse $\xi^{-1}_{\mu(t)}$.

and the *Hilger imaginary part* of $z \in \mathbb{C}$, $\text{Im}_{\mu(t)}(z)$, is given by

$$\text{Im}_{\mu(t)}(z) := \lim_{s \downarrow \mu(t)} \frac{\text{Arg}(1+sz)}{s},$$

where $\text{Arg}(z)$ is the principal argument of z. Note that $\text{Re}_0(z) = \text{Re}(z)$ and $\text{Im}_0(z) = \text{Im}(z)$. Finally, we define the Hilger purely imaginary number $\mathring{\imath}\omega$, with $-\pi/\mu(t) \leq \omega \leq \pi/\mu(t)$, by

$$\mathring{\imath}\omega := \frac{e^{i\omega\mu(t)} - 1}{\mu(t)}.$$

The Hilger complex plane is shown in Figure 2.2

As we continue towards the uniqueness of solutions to (2.2), we require two additional definitions.

Definition 2.2. Let $\mathbb{T}$ be a time scale. A function $f : \mathbb{T} \to \mathbb{R}$ is *rd-continuous* provided it is continuous at right-dense points in $\mathbb{T}$ and its left-sided limits exist (and are finite) at left-dense points in $\mathbb{T}$. If $A : \mathbb{T} \to \mathbb{R}^{n\times n}$ is a matrix valued function, then we say A is rd-continuous provided A is entry-wise rd-continuous.

Definition 2.3. Let $\mathbb{T}$ be a time scale, then $\lambda(t)$ is said to be regressive if $\lambda(t) \neq -\frac{1}{\mu(t)}$ for all $t \in \mathbb{T}$ for which $\mu(t) \neq 0$ and *positively regressive* if $\lambda(t)$ is regressive and contained on the Hilger real axis $\mathbb{R}_{\mu(t)}$ for all $t \in \mathbb{T}$. Furthermore, $\lambda(t)$ is said to be *uniformly regressive* if there exists a $\gamma > 0$ for which $1/\gamma \geq |1 + \mu(t)\lambda(t)|$ for all $t \in \mathbb{T}$.

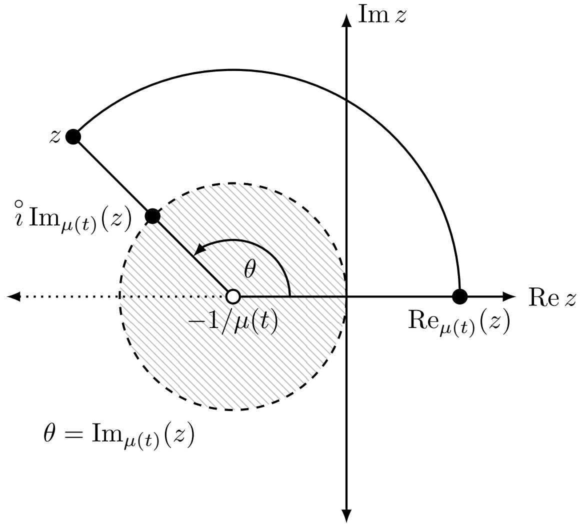

Figure 2.2. The Hilger complex plane.

When $\lambda(t) \equiv \lambda$ is constant, the definition of regressivity requires that λ is not the center of any Hilger circle $\mathcal{H}_{\mu(t)}$ for all $t \in \mathbb{T}$. Uniform regressivity implies λ is bounded away from every possible Hilger circle center.

Theorem 2.1. *Let $\mathbb{T}$ be a time scale and suppose that $\lambda(t)$ is rd-continuous and regressive. Then there exists a unique solution to the initial value problem* (2.2) *with $x_0 = 1$, denoted $\Phi_\lambda(t, t_0)$. When $\lambda(t) \equiv \lambda$, then $\Phi_\lambda(t, t_0) := e_\lambda(t, t_0)$, which we call the time scale exponential function.*

The time scale exponential function generalizes the standard exponential function. When $\mathbb{T} = \mathbb{R}$, λ is rd-continuous and regressive, and $e_\lambda(t, t_0) = e^{\lambda(t-t_0)}$. When $\mathbb{T} = \mathbb{Z}$, as long as $\lambda \neq -1$, $e_\lambda(t, t_0) = (1 + \lambda)^{t-t_0}$. The behavior of the time sale exponential can be very irregular, depending on the time scale. For example, we see that when $\mathbb{T} = \mathbb{Z}$, the exponential function can be oscillatory if λ is real valued with $\lambda < -1$.

In this work, we will also be interested in first order *systems* of dynamic equations of the form

$$x^\Delta = A(t)x; \qquad x(t_0) = x_0, \tag{2.5}$$

where $A : \mathbb{T} \to \mathbb{R}^{n \times n}$ and

$$x^{\Delta} = Ax; \qquad x(t_0) = x_0, \tag{2.6}$$

where $A \in \mathbb{R}^{n \times n}$. Just as in the scalar case, we will require additional conditions to guarantee the existence and uniqueness of (2.5) and (2.6).

Definition 2.4. Let $\mathbb{T}$ be a time scale. A matrix $A : \mathbb{T} \to \mathbb{C}^{n \times n}$ is *(uniformly) regressive* if and only if each eigenvalue $\lambda(t)$ is (uniformly) regressive. Equivalently, a matrix is regressive if and only if $I_n + \mu(t)A(t)$ is invertible for all $t \in \mathbb{T}$.

Theorem 2.2. *Let $\mathbb{T}$ be given and suppose that $A(t)$ is rd-continuous and uniformly regressive. Then there exists a unique solution to the matrix initial value problem* (2.5) *with $x(t_0) = I$, which is given by the time scale transition matrix, denoted $\Phi_A(t, t_0)$. When $A(t) \equiv A$, this is known as the time scale matrix exponential, denoted $e_A(t, t_0)$.*

For a purely discrete time scale, the function $\mu(t)$ completely classifies the time scale. We can therefore generate time scales by generating a sequence of graininesses stochastically. In this work, we consider time scales which are stochastically generated as follows.

Definition 2.5. Let $t_0 \in \mathbb{R}$ and $\{\mu_i\}_{i=0}^{\infty}$ be a sequence of random variables with range $(0, \infty)$. A *stochastic time scale* with initial time t_0 generated by $\{\mu_i\}_{i=0}^{\infty}$ is the set

$$\tilde{\mathbb{T}} := \{t_0\} \cup \left\{ t_0 + \sum_{i=0}^{n} \mu_i \mid n \in \mathbb{N}_0 \right\}.$$

If the set of random variables $\{\mu\} \cup \{\mu_i\}_{i=0}^{\infty}$ consists of independent, identically distributed random variables, we call the resulting stochastic time scale an *i.i.d. stochastic time scale generated by μ.*

Note that the realization of any stochastic time scale is an unbounded above, purely discrete time scale. Now we show how to define the solution of (2.2) on a stochastic time scale.

Definition 2.6. Let $\tilde{\mathbb{T}}$ be a stochastically generated time scale with initial time t_0 generated by $\{\mu_i\}_{i=0}^{\infty}$ and let $t_i = t_0 + \sum_{i=0}^{i-1} \mu_i$. Let $\lambda : \mathbb{R} \to \mathbb{C}$. Assume that $\lambda(t_i) \neq -1/\mu_i$ almost surely. Suppose the sequence $\{x(t_i)\}_{i=1}^{\infty}$ satisfies

$$x(t_i) = (1 + \lambda(t_{i-1})\mu_{i-1})x(t_{i-1}); \qquad x(t_0) = 1. \tag{2.7}$$

Define $\tilde{\Phi}_\lambda(t_i, t_0) := x(t_i)$. If $\lambda(t) \equiv \lambda$ is constant, denote $x(t_i) = \tilde{e}_\lambda(t_i, t_0)$ and call $\tilde{e}_\lambda(t_i, t_0)$ the *stochastic time scale exponential*. We say that $\tilde{e}_\lambda(t_i, t_0)x_0$ is the solution of (2.2) on the stochastic time scale $\tilde{\mathbb{T}}$

Note that $\tilde{\Phi}_\lambda(t_i, t_0)x_0$ solves (2.2) on any realization of the stochastic time scale $\tilde{\mathbb{T}}$. Moreover, the function λ is rd-continuous and regressive almost surely. This definition is therefore a natural extension of the time scales exponential function. We now make the same extension for the matrix case.

Definition 2.7. Let $\tilde{\mathbb{T}}$ be a stochastically generated time scale with initial time t_0 generated by $\{\mu_i\}_{i=0}^{\infty}$ and let $t_i = t_0 + \sum_{i=0}^{i-1} \mu_i$. Let $A : \mathbb{R} \to \mathbb{C}^{n \times n}$. Assume that $I + \mu_i A(t_i)$ is invertible almost surely. Suppose the random sequence of matrices $\{x(t_i)\}_{i=1}^{\infty}$ satisfies

$$x(t_i) = (I + A(t_{i-1})\mu_{i-1})x(t_{i-1}); \qquad x(t_0) = I. \tag{2.8}$$

Then we denote $x(t_i) = \tilde{\Phi}_A(t_i, t_0)$ and call $\tilde{\Phi}_A(t_i, t_0)$ the *stochastic time scale transition matrix*. If $A(t) \equiv A$ is constant, we denote $x(t_i) = \tilde{e}_A(t_i, t_0)$ and call $\tilde{e}_A(t_i, t_0)$ the *stochastic time scale matrix exponential*. We say that $\tilde{\Phi}_A(t_i, t_0)x_0$ is the solution of (2.5) on the stochastic time scale $\tilde{\mathbb{T}}$

2.2 Stability Theory

In systems and control theory, fundamental questions revolve around the stability of the matrix exponential function when $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. In the following section, we introduce various notions of stability in both the deterministic and stochastic sense. We then review some classical approaches to stability theory via both direct and indirect methods.

2.2.1 Deterministic Notions of Stability

The notion of stability is vague and very broad. We now define precisely the notions of stability that will appear in this work.

Definition 2.8. The equilibrium $x(t) \equiv 0$ of a linear dynamic system (2.5) is *Lyapunov stable* if for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that if $\|x(t_0)\| < \delta$, then $\|x(t)\| < \varepsilon$ for every $t > t_0$. The equilibrium is *globally Lyapunov stable* if there exists a finite constant $\gamma > 0$ such that for any initial conditions t_0 and $x(t_0)$ the corresponding solution of (2.5) satisfies $\|x(t)\| \leq \gamma \|x(t_0)\|$.

Definition 2.9. The equilibrium $x(t) \equiv 0$ of the linear dynamic system (2.5) is *asymptotically stable* if it is Lyapunov stable and there exists a $\delta > 0$ such that if $\|x(t)\| < \delta$, then $\lim_{t\to\infty} \|x(t)\| = 0$. Furthermore, the equilibrium is *globally asymptotically stable* if it is globally Lyapunov stable and given any $\delta > 0$ there exists a $T > 0$ such that for any initial conditions t_0 and $x(t_0)$ the corresponding solution satisfies $\|x(t)\| \leq \delta \|x(t_0)\|$, $t \geq t_0 + T$.

Definition 2.10. The equilibrium $x(t) \equiv 0$ of the linear dynamic system (2.5) is *exponentially stable* if for every t_0, there exists $K(t_0) \geq 1$ and $\alpha > 0$ such that

$$\|\Phi_A(t, t_0)\| \leq K e^{-\alpha(t-t_0)}, \qquad t \geq t_0.$$

Definition 2.11. The equilibrium $x(t) \equiv 0$ of the linear dynamic system (2.5) is *uniformly exponentially stable* if K can be chosen independently of t_0 in the definition of exponential stability.

Throughout, we abuse semantics and say that the dynamic equation is stable, rather than saying that the equilibrium is stable.

We now present theorems which help us ascertain the stability properties of (2.3), (2.2) and (2.5).

Pötzsche *et al.* [33] took a direct approach by computing the solution of (2.3) and (2.5) in the case that $A(t) \equiv A$ in order to classify stability properties. Their results relate the stability of first order, linear time invariant systems of dynamic equations on time scales to a region of the complex plane, which we define next.

Definition 2.12. Let $\mathbb{T}$ be a time scale unbounded above. Then for any $t_0 \in \mathbb{T}$, define

$$\mathcal{S}_{\mathbb{C}} := \left\{ \lambda \in \mathbb{C} \mid \limsup_{T\to\infty} \frac{\int_{t_0}^{T} \lim_{s\downarrow\mu(t)} \frac{\ln|1+\lambda s|}{s} \Delta t}{T - t_0} \right\},$$

and

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

Finally, define the *region of exponential stability* for $\mathbb{T}$ by

$$\mathcal{S} := \mathcal{S}_{\mathbb{C}} \cup \mathcal{S}_{\mathbb{R}}.$$

The name "region of exponential stability" is appropriate, as the following theorem shows.

Theorem 2.3 ([33]). *Let $\mathbb{T}$ be a time scale unbounded above and $\lambda \in \mathbb{C}$. The scalar dynamic equation* (2.3) *is exponentially stable if and only if $\lambda \in \mathcal{S}$.*

The above theorem can be generalized for scalar, linear time varying dynamic equations where the time dependence occurs only via the graininess, that is, dynamic equations of the form

$$x^{\Delta} = \lambda(\mu(t))x; \qquad x(t_0) = x_0. \tag{2.9}$$

Lemma 2.1. *Let $\mathbb{T}$ be a time scale which is unbounded above and let $\lambda : [0, \infty) \to \mathbb{C}$. The scalar dynamic equation* (2.9) *is exponentially stable if and only if one of the following conditions is satisfied:*

(1) $\limsup_{T\to\infty} \frac{1}{T-t_0} \int_{t_0}^{T} \lim_{s\to\mu(t)} \frac{\ln|1+\lambda(s)s|}{s} \Delta t < 0.$

(2) for all $T \in \mathbb{T}$, *there exists* $t \in \mathbb{T}$ *with* $t > T$ *such that* $1 + \mu(t)\lambda(\mu(t)) = 0$.

Proof. Follow the proof of Theorem 2.3 in Pötzsche et al. [33]. No step in the proof relies explicitly on $\lambda(\mu(t))$ being a constant. □

Theorem 2.3 extends to (2.6) using eigenvalues, but it requires an additional assumption.

Theorem 2.4 ([33]). *Let* $\mathbb{T}$ *be a time scale unbounded above and let* $A \in \mathbb{R}^{n\times n}$ *be regressive. Then the following hold:*

(1) If (2.6) *is exponentially stable, then* $\text{spec}(A) \subset \mathcal{S}_{\mathbb{C}}$.

(2) If $\text{spec}(A) \subset \mathcal{S}_{\mathbb{C}}$ *and each eigenvalue of* A *is uniformly regressive, then* (2.6) *is exponentially stable.*

Alternatively, we can drop the regressivity of A and the uniform regressivity assumption on the eigenvalues of A by imposing a condition on the geometric and algebraic multiplicities on the eigenvalues of A.

Theorem 2.5 ([33]). *Let* $\mathbb{T}$ *be a time scale unbounded above and let* $A \in \mathbb{R}^{n\times n}$. *Then the following hold*

(1) If (2.6) *is exponentially stable, then* $\text{spec}(A) \subset \mathcal{S}_{\mathbb{C}}$.

(2) Suppose $\text{spec}(A) \subset \mathcal{S}$, $\mu(t)$ *is bounded above, and for all eigenvalues* λ *with unequal geometric and algebraic multiplicities, the scalar system* (2.3) *is uniformly exponentially stable. Then* (2.6) *is exponentially stable.*

The last theorem shows the importance of uniform exponential stability for showing the exponential stability of (2.5).

The paper by Pötzsche *et al.* raised two issues which are still guiding current research. The first issue is, unlike the cases where $\mathbb{T} = \mathbb{R}$ or $\mathbb{T} = h\mathbb{Z}, h > 0$, on a general time scale the notion of exponential stability is not equivalent to the other varieties of stability, such as uniform exponential stability [16]. The second issue is that the region of exponential stability is in general difficult to compute.

In research regarding the first issue, in a recent paper, Doan *et al.* [17] focused on properties of the region of uniform exponential stability in order to prove results concerning the stability radii of positive systems on time scales. In particular, they focused on the "ball of uniform exponential stability," the largest circle contained in the region of stability which is tangent to the origin. While both the region of uniform exponential stability and the ball of uniform exponential stability had useful theoretical properties, they could only be calculated in some very special cases; no general description or formula for these regions was given. In particular, they posed as an open question the radius of the ball of uniform exponential stability when the time scale consists of repeated Cantor sets. We will offer a solution to this open problem in Chapter Five by providing a formula for the radius of the ball of uniform exponential stability under a mild condition on the time scale.

To get around the computational difficulties of the region of exponential stability, Gard and Hoffacker [20] found that $\mathcal{H}_{\mu_{\max}} \subset \mathcal{S}$, where $\mu_{\max} = \sup_{t \in \mathbb{T}}\{\mu(t)\}$. While this region is a subset of the region of exponential stability, in many cases this disc is a small portion of the entire region of exponential stability.

In contrast to the direct method of Pötzsche *et al.*, we can analyze stability using an indirect method where the explicit solution of the dynamic equation is not necessary. The indirect method we will discuss here is Lyapunov's Second Method, introduced in his 1892 thesis, and its generalization to time scales. Heuristically, Lyapunov's method relies on a generalized energy function. If the energy never increases in time, then the solution must remain bounded. If the energy is always decreasing, then the system must reach equilibrium. We now formally introduce the notion of Lyapunov functions and state Lyapunov's theorem for time scales.

Definition 2.13. A functional $V : \mathbb{R}^n \to \mathbb{R}$ is called a *Lyapunov function* for the system $x^\Delta = A(t)x$, if

(1) $V(x(t)) \geq 0$ with equality if and only if $x(t) = 0$.

(2) $V^\Delta(x(t)) \leq 0$, where the Δ-differentiation is with respect to t.

Theorem 2.6 ([30]). *The equilibrium $x(t) \equiv 0$ of $x^\Delta = A(t)x$ is Lyapunov stable if there exists an associated Lyapunov function. Furthermore, if $V^\Delta(x(t)) < 0$ then the equilibrium is globally asymptotically stable.*

We can analyze the stablity of (2.6) using Lyapunov's method, even though an indirect method is not absolutely necessary. We search for quadratic Lyapunov functions, that is, functions of the form

$$V(x(t)) = x^T(t)Px(t),$$

where P is a symmetric, positive definite matrix. This approach leads to the following condition.

Theorem 2.7 ([18]). *The equilibrium $x(t) \equiv 0$ of $x^\Delta = Ax$ is globally asymptotically stable if there exists $P > 0$ such that*

$$A^TP + PA + \mu(t)A^TPA < 0, \tag{TSALI}$$

where the inequalities are interpreted in the sign-definiteness sense. We call this inequality the time scale algebraic Lyapunov inequality (TSALI).

When $\mathbb{T} = \mathbb{R}$, this condition is equivalent to the existence of a solution $P > 0$ to the *continuous algebraic Lyapunov equation* (CALE)

$$A^TP + PA = -M; \qquad M > 0. \tag{CALE}$$

Similarly, on when $\mathbb{T} = \mathbb{Z}$, the existence of a quadratic Lyapunov function is equivalent to the existence of a positive definite solution to the matrix equation, called the *discrete algebraic Lyapuonov equation* (DALE), given by

$$A^T P + PA + A^T PA = -M; \qquad M > 0. \tag{DALE}$$

Additionally, each equation can be solved if and only if the eigenvalues of A are in the respective regions of exponential stability $\mathcal{S}$. Therefore, on $\mathbb{R}$ and $\mathbb{Z}$, global asymptotic stability is equivalent to exponential stability for $x^\Delta = Ax$. This result is not true for a general time scale, as (TSALI) can only be solved if $\mathrm{spec}(A) \subset \mathcal{H}_{\mu_{\max}}$.

2.2.2 Stochastic Notions of Stability

In this section, we consider notions of stability on a stochastic time scale. We will be especially concerned with linear time-varying systems of dynamic equations of the form

$$x^\Delta = A(\mu)x; \qquad x(t_0) = x_0, \tag{2.10}$$

where $A : \mathbb{R} \to \mathbb{R}^{n\times n}$.

Definition 2.14. Let $\tilde{\mathbb{T}}$ be a stochastically generated time scale with initial time t_0 generated by $\{\mu_i\}_{i=0}^\infty$ and let $A : \mathbb{R} \to \mathbb{R}^{n\times n}$. Let $t_i = t_0 + \sum_{i=0}^{i-1} \mu_i$. We say the equilibrium $x(t_i) \equiv 0$ of (2.10) is *exponentially stable almost surely* if and only if with probability one there exists a constant $\alpha > 0$ such that for every $t_i \in \tilde{\mathbb{T}}$ there exists a $K := K(t_i) \geq 1$ with

$$\left|\tilde{\Phi}_{A\circ\mu}(t_k, t_i)\right| \leq Ke^{-\alpha(T_k - T_i)}, \quad \text{for } k \geq i.$$

We can arrive at an entirely different notion of stability if we regard each point in the sample path as a random variable and define the norm of each point in the sample path to be the L^2 norm in the underlying probability space Ω, that is, we use the norm $\|x(t_n)\|_\Omega := \mathrm{E}\left[x^T(t_n)x(t_n)\right]$, where E denotes expectation.

Definition 2.15. The zero solution of (2.5) on $\tilde{\mathbb{T}}$ is called *mean-square stable*, if for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for each initial condition $x_0 \in \mathbb{R}^n$ with $\|x_0\| < \delta$ and all $m \in \mathbb{N}$, we have $\|\tilde{\Phi}_A(t_m, t_0)x_0\|_\Omega < \varepsilon$. If there is no dependence on δ, we call the zero solution *globally mean-square stable*. If in addition, $\|\tilde{\Phi}_A(t_m, t_0)x_0\|_\Omega \to 0$ as $m \to \infty$ for sufficiently small $\|x_0\|$, then the zero solution is called *mean-square asymptotically stable*. If the convergence is exponential, the zero solution is called *mean-square exponentially stable*.

Definition 2.16. The zero solution of (2.5) on $\tilde{\mathbb{T}}$ is called *second moment (exponentially, asymptotically) stable* if the corresponding property in Definition 2.15 above holds, when $\|\tilde{\Phi}_A(t_n, t_0)x_0\|_\Omega$ is replaced by

$$\|\mathrm{E}[\tilde{\Phi}_A(t_m, t_0)x_0 x_0^T \tilde{\Phi}_A^T(t_m, t_0)]\|_{\mathbb{R}^{n\times n}},$$

where $\|\cdot\|_{\mathbb{R}^{n\times n}}$ is the induced matrix norm from the vector 2-norm.

These two notions of stability can be used interchangeably, as the next lemma shows.

Lemma 2.2 (Damm [13]). *Second moment stability and mean-square stability are equivalent.*

In the later part of the dissertation, we restrict ourselves to i.i.d stochastic time scales. In this case, $\{\tilde{\Phi}_{A\circ\mu}(t_n, t_0)\}_{n=1}^{\infty}$ forms a Markov chain. The following two key theorems due to Kushner establish a Lyapunov-like theorem for this case. These theorems will lead us to a notion similar to Lyapunov stability.

Definition 2.17. Let $\mathcal{M} := \{X_n\}_{n=0}^{\infty}$ be a sequence of random variables with outputs in the state space S. We say $\mathcal{M}$ is a *Markov chain* provided

$$\Pr[X_n = s \mid X_{n-1} = s_{n-1}, \ldots, X_0 = s_0] = \Pr[X_n = s \mid X_{n-1} = s_n],$$

where $\Pr[\cdot \mid \cdot]$ represents conditional probability and $s, s_m \in S$ for $0 \le m \le n-1$.

Less formally, a Markov chain is a sequence of independent random variables whose next value is conditionally independent of the past values, given the present value.

Theorem 2.8 (Kushner [29]). *Let* $\{x_n\}$ *be a Markov Chain on a state space* S. *Suppose there exists nonnegative* V *such that*

$$\mathrm{E}_x[V(x_1)] - V(x) = -k(x),$$

where E_x *is the expectation given that the Markov chain has initial state* x *and where* $k(x) \geq 0$ *on* $Q_\lambda := \{x \mid V(x) < \lambda\}$. *Then*

$$\Pr_x \left[\sup_{0 \leq n < \infty} V(x_n) \geq \lambda \right] \leq \frac{V(x)}{\lambda}.$$

Therefore,

- *Solution paths stay in* Q_λ *with probability at least* $1 - V(x)/\lambda$.
- $k(x_n) \to 0$ *for all paths remaining in* Q_λ.
- *There is some random* $v \geq 0$ *such that* $V(x_n) \to v \geq 0$ *with probability one.*

Theorem 2.9 (Kushner [29]). *Let* $V(x) \geq 0$ *and* $\mathrm{E}_x[V(x_1)] - V(x) \leq -\alpha V(x)$ *for some* $0 < \alpha < 1$. *Then*

- $\mathrm{E}_x[V(x_n)] \leq (1-\alpha)^n V(x)$;
- $V(x_n) \to 0$ *with probability one*;
- $\mathrm{P}_x\left[\sup_{N \leq n < \infty} V(x_n) \geq \lambda\right] \leq \frac{(1-\alpha)^N}{\lambda} V(x)$.

Definition 2.18. If there exists a $V(x_n) \geq 0$ satisfying the conditions of Theorem 2.8 for $x_n = \tilde{\Phi}_{A \circ \mu}(t_n, t_0)x_0$ on an i.i.d. stochastic time scale, we say that the zero solution of (2.10) is *asymptotically stable in the sense of Kushner.* If such a V exists satisfying the conditions of Theorem 2.9, we say that the zero solution of (2.10) is *exponentially stable in the sense of Kushner.*

In Chapter Three, we will show that, for i.i.d stochastic time scales, mean-square exponential stability is equivalent to exponential stability in the sense of Kushner.

2.3 Control Theory

We now introduce the necessary background from control theory.

2.3.1 Linear System Model

Many physical systems of interest to engineers can be approximated by the *continuous time linear time invariant system model*

$$\dot{x} = Ax + Bu \tag{2.11}$$

$$y = Cx + Du, \tag{2.12}$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$, and $D \in \mathbb{R}^{p\times m}$, are constant matrices. $x : \mathbb{R} \to \mathbb{R}^n$ is called the *state vector*, $y : \mathbb{R} \to \mathbb{R}^n$ is called the *output vector* and $u : \mathbb{R} \to \mathbb{R}^m$ is called the *control vector*. Also commonly considered is the *discrete time linear time invariant system model*

$$\Delta x = Ax + Bu \tag{2.13}$$

$$y = Cx + Du, \tag{2.14}$$

where A, B, C, D, x, u are as before. Two primary goals of control theory are:

(1) *Controller Problem:* Can we choose $u(t)$ so that the state $x(t)$ has a desired property (such as stability).

(2) *Observer Problem:* If we cannot directly measure the state x, can we estimate or reconstruct x using only the information contained in the output y.

We will only discuss relevant results for equations (2.13) and (2.11). A vast theory exists for more general control systems, such as linear time-varying and non-linear systems [3].

2.3.1.1 Controllability We now introduce a fundamental concept in control theory essential for determining if we can solve the controller problem.

Definition 2.19. A linear system is said to be *controllable* at t_0 if it is possible to find $u(t)$ defined over $t_0 \leq t \leq t_1 < \infty$ so that $x(t_1) = 0$. If this is true for all initial times t_0 and all initial states $x(t_0)$, the system is said to be *completely controllable.*

For the linear time invariant systems under consideration, the following rank condition determines controllability.

Theorem 2.10. *The systems* (2.11) *and* (2.13) *are completely controllable if and only if the $n \times mn$ controllability matrix*

$$\mathcal{C} := \big(B \mid AB \mid A^2B \mid \cdots \mid A^{n-1}B\big)$$

has rank n.

If the condition in Theorem 2.10 holds, we call the pair (A, B) controllable. Controllability is an essential feature of systems which we want to stabilize. Notice that if we use state-feedback, that is, controls of the form $u(t) = Kx(t), K \in \mathbb{R}^{m \times n}$, then the system dynamics of (2.11) and (2.13) are given by

$$\dot{x} = (A + BK)x$$

and

$$\Delta x = (A + BK)x,$$

respectively. Therefore, if $\text{spec}(A + BK) \subset \mathbb{C}^-$ in the continuous case or $\text{spec}(A + BK) \subset \mathcal{H}_1$ in the discrete case, then the system will be exponentially stable. When the pair (A, B) is controllable, we can position the eigenvalues of $A+BK$ arbitrarily (up to complex conjugates) through judicious choice of K, as the next lemma shows.

Lemma 2.3 ([9, 3]). *Let Γ be a set of n complex numbers such that if $\lambda \in \Gamma$, then $\overline{\lambda} \in \Gamma$. There exists $K \in \mathbb{R}^{m\times n}$ such that* $\operatorname{spec}(A + BK) = \Gamma$ *for any admissible set Γ if and only if the pair (A, B) is controllable.*

In certain cases, we can achieve our control goals even if the pair (A, B) is not controllable. To see this, we introduce the Kalman controller canonical form of (2.11) and (2.13).

Lemma 2.4 ([9, 3]). *The linear system* (2.11) *can be written, via the orthogonal transformation $Tx = \begin{pmatrix} T_1 & T_2 \end{pmatrix} x = \begin{pmatrix} w_1 & w_2 \end{pmatrix}^T = w$, as*

$$\dot{\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}} = \begin{pmatrix} T_1^T A T_1 & T_1^T A T_2 \\ 0 & T_2^T A T_2 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} T_1^T B \\ 0 \end{pmatrix} u$$
$$y = \begin{pmatrix} CT_1 & CT_2 \end{pmatrix} w + Du,$$

where the pair $(T_1^T A T_1, T_1^T B)$ is controllable.

Similarly, the linear system (2.13) *can be written, via the same transformation,*

$$\Delta \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} T_1^T A T_1 & T_1^T A T_2 \\ 0 & T_2^T A T_2 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} T_1^T B \\ 0 \end{pmatrix} u$$
$$y = \begin{pmatrix} CT_1 & CT_2 \end{pmatrix} w + Du,$$

where the pair $(T_1^T A T_1, T_1^T B)$ is controllable.

Notice that the dynamics of w_2 are given by

$$\dot{w}_2 = T_2^T A T_2 w_2.$$

Therefore the trajectory of w_2 is independent of the control u and decoupled from w_1. We see that as long the eigenvalues of $T_2^T A T_2$ are in the correct stability region, stability can be achieved by using Lemma 2.3 for the dynamics of w_1. This leads to the definition of stabilizability.

Definition 2.20. The continuous time linear system (2.11) is said to be *stabilizable* if $\text{spec}(T_2^T A T_2) \in \mathbb{C}^-$. The discrete time linear system (2.13) is said to be *stabilizable* if $\text{spec}(T_2^T A T_2) \in \mathcal{H}_1$.

2.3.1.2 Observability We now introduce a condition similar to controllability helps us determine whether we can solve the observer problem.

Definition 2.21. A linear system is said to be *observable* at t_0 if $x(t_0)$ can be determined from the output function $y(t)$ defined over $t_0 \leq t \leq t_1 < \infty$. If this is true for all t_0 and $x(t_0)$, the system is said to be *completely observable.*

For the linear time invariant systems under consideration, the following rank condition determines observability.

Theorem 2.11. *The systems* (2.11) *and* (2.13) *are completely observable if and only if the $mn \times n$ observability matrix*

$$\mathcal{O} := \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

has rank n.

If the condition in Theorem 2.11 holds, we call the pair (A, C) observable. Observability is a fundamental concept for systems that we want to estimate, as the following construction shows. Suppose we guess the state $\hat{x}$ with the system model

$$\dot{\hat{x}} = A\hat{x} + Bu + H(y - \hat{y});$$

$$\hat{y} = C\hat{x} + Du,$$

where $H \in \mathbb{R}^{n\times p}$ is some constant matrix which we can design. Then the error $\varepsilon := x - \hat{x}$ has the dynamics

$$\dot{\varepsilon} = (A - HC)\varepsilon.$$

A similar design holds in discrete time. Therefore, if $\operatorname{spec}(A - HC) \subset \mathbb{C}^-$ in the continuous case, or $\operatorname{spec}(A - HC) \subset \mathcal{H}_1$ in the discrete case, then the system will be exponentially stable. When the pair (A, C) is observable, we can position the eigenvalues of $A - HC$ arbitrarily (up to complex conjugates) through judicious choice of H, as the next lemma shows.

Lemma 2.5 ([9, 3]). *Let Γ be a set of n complex numbers such that if $\lambda \in \Gamma$, then $\overline{\lambda} \in \Gamma$. There exists $H \in \mathbb{R}^{n\times p}$ such that* $\operatorname{spec}(A - HC) = \Gamma$ *for any admissible set Γ if and only if the pair (A, C) is observable.*

In certain cases, we can achieve our estimation goals even if the pair (A, C) is not observable. To see this, we introduce an analogous Kalman observable canonical form of (2.11) and (2.13) for the observer case.

Lemma 2.6 ([9, 3]). *The continuous linear system* (2.11) *can be written, via the orthogonal transformation $Vx = \begin{pmatrix} V_1 & V_2 \end{pmatrix} x = \begin{pmatrix} v_1 & v_2 \end{pmatrix}^T = v$, as*

$$\dot{\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}} = \begin{pmatrix} V_1^T A V_1 & 0 \\ V_2^T A V_1 & V_2^T A V_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} V_1^T B \\ V_2^T B \end{pmatrix} u$$

$$y = \begin{pmatrix} CV_1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + Du,$$

where the pair $(V_1^T A V_1, CV_1)$ is observable.

Similarly, the discrete time linear system (2.13) *can be written, via the same transformation,*

$$\Delta \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} V_1^T A V_1 & 0 \\ V_2^T A V_1 & V_2^T A V_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} V_1^T B \\ V_2^T B \end{pmatrix} u$$

$$y = \begin{pmatrix} CV_1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + Du,$$

where the pair $(V_1^T A V_1, CV_1)$ *is observable.*

Notice that the states v_2 do not contribute to the output y. Moreover, v_2 does not contribute to y via v_1, as the dynamics of v_1 are independent of v_2. We will not be able to observe estimate v_2, so if the dynamics of v_2 are not stable, we will not be able to estimate the state v. If the dynamics of v_2 are stable, we know the destination of the v_2, and can therefore estimate the state v.

Definition 2.22. The linear system (2.11) is said to be *detectable* if $\text{spec}(V_2^T A V_2) \in \mathbb{C}^-$. The linear system (2.13) is said to be *detectable* if $\text{spec}(V_2^T A V_2) \in \mathcal{H}_1$.

2.3.2 Discretizing onto a Time Scale

We will be interested in generalizing (2.11) and (2.13) to time scales. Here, we introduce a generalization which is useful in practical applications such as real-world engineering problems. We consider the system dynamics

$$\dot{x} = Ax + Bu + Fw_t \tag{2.15}$$

$$y = Cx + Du,$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}, D \in \mathbb{R}^{p\times m}$, and $F \in \mathbb{R}^{n\times l}$ are constant matrices and w_t is an orthogonal increments process [10] with $\text{E}[w_t] = 0$ and $\text{E}[w_s w_t^T] = I \min\{t, s\}$. Similarly, the discrete LTI system model is given by

$$\Delta x = Ax + Bu + Fw, \qquad x(t_0) = x_0,$$

$$y = Cx + Du.$$

When encountering these two models, it is tempting to simply study the time scale LTI control system

$$x^{\Delta} = Ax + Bu + Fw$$
$$y = Cx + Du.$$

While this approach is perfectly acceptable from a mathematical perspective, in this subsection, we argue that the simplistic approach given above does not suffice if one wishes to study practical control systems. We will derive the time scales control system that, while more mathematically difficult to work with, provides a usable model for applications. In this way, both theorists and practitioners benefit: the theorists discover new questions to ask and the engineers receive accurate models.

Let us consider the control system (2.15). Suppose we wish to control the system using a piecewise constant function which can only change values at a discrete set of points denoted $\mathbb{T} := \{t_k\}_{k=0}^{\infty}$. It is well known [3] that the solution of (2.15) is given by

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau + \int_{t_0}^{t} e^{A(t-\tau)}Cdw_{\tau}, \tag{2.16}$$

where the last integral is a Weiner integral [15].

Now, as $u(t) = u(t_k)$ on $[t_k, t_{k+1})$, we can write for any $t \in [t_k, t_{k+1}]$,

$$x(t) = e^{A(t-t_0)}x(t_k) + \int_{t_k}^{t} e^{A(t-\tau)}d\tau Bu(t_k) + \int_{t_k}^{t} e^{A(t-\tau)}Cdw_{\tau}.$$

Therefore,

$$\begin{aligned} x^{\Delta}(t_k) &= \frac{x(t_{k+1}) - x(t_k)}{\mu(t_k)} \\ &= \frac{e^{A(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)}d\tau Bu(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)}Cdw_{\tau} - x(t_k)}{\mu(t_k)} \\ &= \frac{(e^{A\mu_k} - I)}{\mu(t_k)}x(t_k) + \frac{A^{-1}(e^{A\mu_k} - I)B}{\mu(t_k)}u(t_k) + \frac{\int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)}Cdw_{\tau}}{\mu(t_k)} \end{aligned}$$

$$:= \mathcal{A}(\mu(t_k))x(t_k) + \mathcal{B}(\mu(t_k))u(t_k) + \int_{t_k}^{t_{k+1}} \frac{e^{A(t_{k+1}-\tau)}C}{\mu(t_k)} dw_\tau.$$

The last term above [15] is a random variable with

$$\mathrm{E}\left[\int_{t_k}^{t_{k+1}} \frac{e^{A(t_{k+1}-\tau)}C}{\mu(t_k)} dw_\tau\right] = 0,$$

and

$$\begin{aligned} \mathrm{Cov}\left[\int_{t_k}^{t_{k+1}} \frac{e^{A(t_{k+1}-\tau)}C}{\mu(t_k)} dw_\tau\right] &= \frac{1}{\mu(t_k)^2} \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)}CC^T e^{A^T(t_{k+1}-\tau)} d\tau \\ &= \frac{1}{\mu(t_k)^2} \int_0^{\mu(t_k)} e^{A\tau}CC^T e^{A^T\tau} d\tau, \end{aligned}$$

where Cov denotes the variance-covariance matrix of a random vector. Since most of the modeling we do will involve only the first and second moments, a model which agrees with the current model up to the first and second moments is given by

$$x^\Delta = \mathcal{A}(\mu)x + \mathcal{B}(\mu)u + \mathcal{C}(\mu)w,$$

where w is a random variable with $\mathrm{E}[w] = 0, \mathrm{E}[ww^T] = I$ and

$$\mathcal{C}(\mu) = \frac{1}{\mu}\sqrt{\int_0^\mu e^{A\tau}CC^T e^{A^T\tau} d\tau}.$$

Note that

$$\mathrm{E}[\mathcal{C}(\mu)w] = 0,$$

and

$$\begin{aligned} \mathrm{Cov}[\mathcal{C}(\mu)w] &= \mathrm{E}[\mathcal{C}(\mu)ww^T\mathcal{C}^T(\mu)] \\ &= \mathcal{C}(\mu)\mathrm{E}[ww^T]\mathcal{C}^T(\mu) = \mathcal{C}(\mu)^2 \\ &= \frac{1}{\mu(t_k)^2} \int_0^{\mu(t_k)} e^{A\tau}CC^T e^{A^T\tau} d\tau. \end{aligned}$$

In the case where A is not invertible, we can still define $\mathcal{B}(\mu)$ via the convergent power series

$$\mathrm{expc}(X) := \sum_{n=1}^{\infty} \frac{X^{n-1}}{n!}. \tag{2.17}$$

Therefore, in general, we have

$$\mathcal{A}(\mu(t)) = \text{expc}(A\mu(t))A$$

and

$$\mathcal{B}(\mu) = \text{expc}(A\mu(t))B.$$

In conclusion, this derivation shows that in order to understand the behavior of the continuous time linear system (2.15) which updates only on points in the time scale $\mathbb{T}$, we should study the time scale control system

$$x^{\Delta} = \mathcal{A}(\mu)x + \mathcal{B}(\mu)u + \mathcal{C}(\mu)w, \qquad t \in \mathbb{T}$$

2.3.3 Optimal Control Theory

When studying the controller problem, especially time varying or nonlinear systems, one of the largest difficulties is finding valid control laws. The contributions to optimal control theory by Kalman [27], perhaps one of the most important developments in control theory during the twentieth century, helped to alliviate this problem. The usefulness of the theory was not due to the fact that the control was optimal with respect to some criterion, as the criterion was somewhat arbitrary, but that the engineer could now obtain a stabilizing control relatively easily. The most well-known optimal control theory is the classic infinite horizon linear quadratic regulator (LQR), a theory that develops a control which minimizes the quadratic cost function over either a finite or an infinite time frame. In continuous time, the finite horizon cost functional is given by

$$J_c^f(x,u,T) := \int_{t_0}^{T} x^T(t)Qx(t) + u^T(t)Ru(t)dt + x^T(T)Q^f x(T), \tag{2.18}$$

where $Q, Q^f \geq 0$ and $R > 0$. The continuous time infinite horizon cost functional is given by

$$J_c(x,u) := \int_{t_0}^{\infty} x^T(t)Qx(t) + u^T(t)Ru(t)dt. \tag{2.19}$$

Similarly, in discrete time, the finite horizon cost functional is given by

$$J_d^f(x,u,N) := \sum_{n=0}^{N-1} x^T(n)Qx(n) + u^T(n)Ru(n) + x^T(N)Q^f x(N), \tag{2.20}$$

and the infinite horizon cost functional is given by

$$J_d(x,u) := \sum_{n=0}^{\infty} x^T(n)Qx(n) + u^T(n)Ru(n). \tag{2.21}$$

Just as the stability of the LTI system (2.6) on $\mathbb{R}$ and $\mathbb{Z}$ is connected to the existence of a solution to an associated Lyapunov matrix equation, the design of optimal controllers in the finite horizon and infinite horizon LQR case for (2.6) on $\mathbb{R}$ and $\mathbb{Z}$ is connected to the solution of an associated Riccati matrix equation.

On $\mathbb{R}$, the finite horizon problem is approached by solving the *continuous differential Riccati equation*

$$\dot{V} = Q + A^T V + VA - VBR^{-1}B^T V; \qquad V(T) = Q^f. \tag{CDRE}$$

The infinite horizon problem involves the solution of the *continuous algebraic Riccati equation* (CARE), given by

$$V = Q + V + A^T V + VA - VBR^{-1}B^T V. \tag{CARE}$$

On $\mathbb{Z}$, the finite horizon problem is approached by solving the *discrete difference Riccati equation*

$$\begin{aligned} -\Delta V_j =& Q + A^T V_{j+1} + V_{j+1}A + A^T V_{j+1} A \\ &- (I+A)^T V_{j+1} B \left(R + B^T V_{j+1} B\right)^{-1} B^T V_{j+1} (I+A); \qquad V_N = Q^f. \end{aligned} \tag{DDRE}$$

The infinite horizon problem involves the solution of the *discrete algebraic Riccati equation* (DARE), given by

$$V = Q + (I+A)^T V(I+A) - (I+A)^T VB(R + B^T VB)^{-1} B^T V(I+A). \tag{DARE}$$

Both (CDRE) and (DDRE) are backward dynamic equations. In both continuous and discrete time, the solution of the infinite horizon problem is achieved by solving for the initial value of V in (CDRE) and (DDRE), respectively, and taking the limit as $T \to \infty$.

In the continuous time case, the state-feedback control

$$u^*(t) = R^{-1}B^T V x(t),$$

where V is the unique positive definite solution of (CARE), minimizes the infinite horizon cost functional (2.19) and the total cost is given by $J_c(x, u^*) = x_0^T V x_0$. Similarly, in the discrete time case, the state-feedback control

$$u^*(t) = (R + B^T V B)^{-1} B^T V (I + A) x(t),$$

where V is the unique positive definite solution of (DARE), minimizes the infinite horizon cost functional (2.21) and the total cost is given by $J_d(x, u^*) = x_0^T V x_0$.

On a general μ-varying systems, finding a suitable stabilizing control has been a difficult and open problem. In Chapter Four, we will develop an optimal control theory for μ-varying and LTI systems on stochastic time scales and generalize (CDRE), (CARE), (DDRE), and (DARE) in order to find stabilizing control laws.

CHAPTER THREE

Stability Theory on Stochastic Time Scales

3.1 Direct Method: Exponential Stability Almost Surely

In this section, we generalize the direct method of Pötzsche et al. [33] to stochastic time scales. This yields an easily verifiable condition for almost sure exponential stability of scalar dynamic equations on stochastic time scales. We present only the scalar case here for ease of exposition. We begin with a lemma, which is a modest generalization of Theorem 2.3.

Theorem 3.1. *Let $\tilde{\mathbb{T}}$ be an i.i.d stochastic time scale generated by μ and let $\lambda : (0,\infty) \to \mathbb{C}$. Assume $\lambda(\mu) \neq -1/\mu$ almost surely. Let $t_i = t_0 + \sum_{i=0}^{i-1} \mu_i$. Then the scalar dynamic equation* (2.9) *is exponentially stable almost surely on $\tilde{\mathbb{T}}$ if and only if*

$$\mathrm{E}[\ln|1 + \lambda(\mu)\mu|] < 0.$$

Proof. If $\mathrm{E}[\ln|1+\lambda(\mu)\mu|] < 0$, then by the Strong Law of Large Numbers [28],

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n} \ln|1+\lambda(\mu_i)\mu_i|}{n} = \lim_{n\to\infty} \frac{\sum_{i=0}^{n} \ln|1+\lambda(\mu_i)\mu_i|}{n} = \mathrm{E}[\ln|1+\lambda(\mu)\mu|] < 0$$

almost surely. This implies

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \ln|1+\lambda(\mu_i)\mu_i| < 0$$

almost surely, or equivalently,

$$\begin{aligned} 0 &> \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln|1+\lambda(\mu_i)\mu_i| \\ &= \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1}\mu_i}{n} \frac{1}{t_n - t_0} \sum_{i=0}^{n-1} \ln|1+\lambda(\mu_i)\mu_i| \\ &= \mathrm{E}[\mu] \limsup_{n\to\infty} \frac{1}{t_n - t_0} \int_{t_0}^{t_n} \frac{\ln|1+\lambda(\mu_i)\mu_i|}{\mu_i} \Delta t \end{aligned}$$

almost surely. Hence,

$$\limsup_{n\to\infty} \frac{1}{t_n - t_0} \int_{t_0}^{t_n} \frac{\ln|1+\lambda(\mu_i)\mu_i|}{\mu_i} \Delta t$$

almost surely. Thus by Lemma 2.1, the dynamic equation (2.9) is exponentially stable almost surely.

On the other hand, first note for all $t > t_0$, $1 + \mu(t)\lambda(\mu(t)) \neq 0$ almost surely since $P[1 + \lambda(\mu_i)\mu_i = 0] = 0$, $i \in \mathbb{N}^0$, thus the second condition of Lemma 2.1 does not hold.

If $\mathrm{E}[\ln|1+\lambda(\mu)\mu|] \geq 0$, then by the Strong Law of Large Numbers,

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n} \ln|1+\lambda(\mu_i)\mu_i|}{n} = \lim_{n\to\infty} \frac{\sum_{i=0}^{n} \ln|1+\lambda(\mu_i)\mu_i|}{n} = \mathrm{E}[\ln|1+\lambda(\mu)\mu|] \geq 0$$

almost surely. The above implies

$$\begin{aligned} 0 &\leq \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln|1+\lambda(\mu_i)\mu_i| \\ &= \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mu_i}{n} \frac{1}{t_n - t_0} \sum_{i=0}^{n-1} \ln|1+\lambda(\mu_i)\mu_i| \\ &= \mathrm{E}[\mu] \limsup_{n\to\infty} \frac{1}{t_n - t_0} \int_{t_0}^{t_n} \frac{\ln|1+\lambda(\mu_i)\mu_i|}{\mu_i} \Delta t \end{aligned}$$

almost surely. Therefore,

$$\limsup_{n\to\infty} \frac{1}{t_n - t_0} \int_{t_0}^{t_n} \frac{\ln|1+\lambda(\mu_i)\mu_i|}{\mu_i} \Delta t$$

almost surely. Thus by Lemma 2.1, the dynamic equation (2.9) is not exponentially stable almost surely. □

Remark 3.1. If M is an a continuous random variable which admits a probability density function $f : D \to [0, \infty)$ with support D, the condition $\mathrm{E}[\ln|1+\lambda(\mu)\mu|] < 0$ becomes

$$\int_D f(\mu) \ln|1+\lambda(\mu)\mu| d\mu < 0.$$

We note that the above gives a straightfowardly checkable test for whether a given function λ makes (2.9) exponentially stable. The function space of all such functions λ is quite complicated. We can, however, study certain classes of functions within the space. Letting $\lambda(\mu(t)) = \frac{e^{z\mu(t)}-1}{\mu(t)}$, i.e., the inverse cylinder transformation defined in (2.4), we find

$$\mathrm{E}[\ln|1+\lambda(\mu)\mu|] = \mathrm{E}\left[\ln\left|1+\frac{e^{z\mu}-1}{\mu}\mu\right|\right] = \mathrm{Re}(z)\mathrm{E}[\mu] < 0$$

if and only if $\mathrm{Re}(z) < 0$ and μ has finite mean. This agrees with our intuition, as the region of exponential stability for the equation $\dot{x} = zx$ on $\mathbb{R}$ is $\{z \in \mathbb{C} \mid \mathrm{Re}(z) < 0\}$.

Remark 3.2. If μ is a discrete random variable with finitely many possible values $\mu_1, \mu_2, \ldots, \mu_n$ with a probability mass function $g : D \to [0, \infty)$, the condition $\mathrm{E}[\ln|1+\lambda(\mu)\mu|] < 0$ becomes

$$\sum_{i=1}^{n} f(\mu_i)\ln|1+\lambda(\mu_i)\mu_i| < 0,$$

or, equivalently,

$$\prod_{i=1}^{n} |1+\lambda(\mu_i)\mu_i|^{f(\mu_i)} < 1. \tag{3.1}$$

In the special case where $\lambda(\mu)$ is constant, then (3.1) agrees with the result of Davis *et al.* [14] where the asymptotic weights are given by $d_k = f(\mu_k)$. This work gives a broader interpretation of their concept of asymptotic equivalence class as the set of all time scales which are distributed the same in the tail.

Remark 3.3. In the proof of Proposition 6 of Pötzsche *et al.* [33], a formula for a suitable α in the bounding exponential function $Ke^{-\alpha t}$ is given by, in our case,

$$\begin{aligned}
\alpha &= -\limsup_{n\to\infty} \frac{\sum_{i=0}^{n}\ln|1+\lambda(\mu_i)\mu_i|}{t_n - t_0} \\
&= -\limsup_{n\to\infty} \frac{n}{\sum_{i=0}^{n-1}\mu_i}\frac{\sum_{i=0}^{n}\ln|1+\lambda(\mu_i)\mu_i|}{n} \\
&= -\frac{\mathrm{E}[\ln|1+\lambda(\mu)\mu|]}{\mathrm{E}[\mu]} \\
&> 0.
\end{aligned}$$

Remark 3.4. We can view the solution of the deterministic equation (2.9) on a stochastic time scale as the solution of the *stochastic* equation $x_{n+1} = (1+\lambda(\mu_n)\mu_n)x_n$ on the deterministic time scale $\mathbb{Z}$. The problem of stability of stochastic systems has been studied in [6]. It is known that the stochastic difference equation $x_{n+1} = a_n x_n$, where $\{a_n\}$ is a sequence of ergodic scalar random variables is exponentially stable almost surely if and only if $\mathrm{E}[\ln |a_n|] < 0$. This result matches our result, as the sequence of random variables $\{1 + \lambda(\mu_n)\mu_n\}$ is a sequence of independent random variables, and hence is a sequence of ergodic random variables.

Corollary 3.1. *Let $\{\mu_i\}_{i=0}^{\infty}$ be states at step i of an ergodic Markov chain with finitely many states $\mu_1, \mu_2, \ldots \mu_n$, all of which are nonzero. Let $\lambda \in \mathbb{C}$ such that $|1+\lambda\mu_k| \neq 0$ for $1 \leq k \leq n$. Let $\tilde{\mathbb{T}}$ be a stochastically generated time scale with initial time t_0 generated by $\{\mu_i\}_{i=0}^{\infty}$. Define π to be the unique stationary discrete distribution associated with the Markov chain. Then the scalar dynamic equation* (2.3) *is exponentially stable almost surely on $\tilde{\mathbb{T}}$ if and only if*

$$\sum_{i=1}^{n} \pi(\mu_i) \ln |1 + \lambda\mu_i| < 0.$$

Corollary 3.2. *Let $\{\mu_i\}_{i=0}^{\infty}$ be a sequence of nonnegative independent random variables and let $\tilde{\mathbb{T}}$ be a stochastically generated time scale with initial time t_0 generated by $\{\mu_i\}_{i=0}^{\infty}$. Assume, for $\lambda : [0, \infty) \to \mathbb{C}$,*

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)^2} \mathrm{var}[\ln |1 + \lambda(\mu_k)\mu_k|] < \infty.$$

Assume further that $\Pr[\mu_i = 0] = 0$ and $\Pr[1 + \lambda(\mu_i)\mu_i = 0] = 0$ for $i \in \mathbb{N}^0$. Then the scalar dynamic equation (2.3) *is exponentially stable almost surely on $\tilde{\mathbb{T}}$ if and only if*

$$\lim_{n\to\infty} \mathrm{E}\left[\frac{\sum_{i=0}^{n} \ln |1 + \lambda(\mu_i)\mu_i|}{n+1}\right] < 0.$$

Proof. Use Kolmogorov's Strong Law of Large Numbers [28] and follow the proof of Theorem 1. □

The next proof requires a corollary to the Borel-Cantelli Lemma, which we state here

Lemma 3.1 ([28]). *Let* $\{X_n\}_{n=0}^{\infty}$ *be a sequence of random variables and let* $a \in \mathbb{R}$ *be such that*

$$\sum_{n=0}^{\infty} P[X_n \geq a] < \infty.$$

Then

$$\limsup_{n\to\infty} X_n < a.$$

We can now prove the following corollary.

Corollary 3.3. *Let* $\{\mu_i\}_{i=0}^{\infty}$ *be a sequence of nonnegative random variables and let* $\lambda : [0,\infty) \to \mathbb{C}$ *such that*

$$\sum_{n=0}^{\infty} \Pr\left[\sum_{k=0}^{n} \ln|1+\lambda(\mu_i)\mu_i| \geq 0\right] = \sum_{n=0}^{\infty} \Pr\left[\frac{\sum_{k=0}^{n} \ln|1+\lambda(\mu_i)\mu_i|}{n+1} \geq 0\right] < \infty. \quad (3.2)$$

Assume further that $\Pr[\mu_i = 0] = 0$ *and* $\Pr[1+\lambda(\mu_i)\mu_i = 0] = 0$ *for* $i \in \mathbb{N}^0$. *Let* $\tilde{\mathbb{T}}$ *be a stochastically generated time scale with initial time* t_0 *generated by* $\{\mu_i\}_{i=0}^{\infty}$. *Then the scalar dynamic equation* (2.3) *is exponentially stable almost surely on* $\tilde{\mathbb{T}}$.

Proof. Condition (3.2) yields, by Lemma 3.1, that

$$\limsup_{n\to\infty} \frac{\sum_{k=0}^{n} \ln|1+\lambda(\mu_k)\mu_k|}{n+1} < 0$$

almost surely. With this, we proceed as in the second part of the proof of Theorem 3.1. □

In the case where $\lambda(\mu) \equiv \lambda$, we can consider the region of the complex plane where the condition in Theorem 3.1 holds.

Definition 3.1. The *region* $\tilde{\mathcal{S}}$ *of almost sure exponential stability* for the scalar dynamic equation $x^{\Delta} = \lambda x$ on a stochastically generated time scale generated by independent, identically distributed random variables is defined to be

$$\tilde{\mathcal{S}} := \{\lambda \in \mathbb{C} \mid \mathrm{E}[\ln|1+\lambda\mu|] < 0\}.$$

3.2 Examples

We now examine the behavior of (2.3) on a stochastically generated stochastic time scale denoted $\tilde{\mathbb{T}}_\Gamma$ which is generated by independently identically distributed random variables taken from a Gamma Distribution with shape parameter 2 and rate parameter 2, whose probability density function we call f. Notice that such a stochastically generated time scale falls under the scope of Theorem 3.1, and by Remark 3.2, given $\lambda \in \mathbb{C}$, (2.3) is exponentially stable on $\mathbb{T}_\Gamma$ if and only if

$$\int_0^\infty f(\mu) \ln |1 + \lambda\mu| d\mu < 0.$$

The stability region $\tilde{\mathcal{S}}$ is shown in Figure 3.1.

We choose two values of λ, $\lambda_1 = 1 + .25i$ and $\lambda_2 = -2 + .67i$ and generate six realizations of the time scale using each λ_i, $i = 1, 2$. The results are shown in Figure 3.2 and Figure 3.3 along with the theoretical decay rate as in a Remark 3.3.

Note that the solution of (2.3) with $\lambda = \lambda_1$ decays fairly regularly and does not require an extremely large multiplier on the bounding exponential. The solution of (2.3) with $\lambda = \lambda_2$, on the other hand, is very irregular in its behavior, having swings on the order of magnitude of x as large as 10^{20} in Figure 3.3. Amazingly (2.3) is exponentially stable by Theorem 3.1, but it has a very slow decay rate and does not decay at each time step.

We note that this analysis informs use about deterministic time scales. If we know the frequency with which different graininesses appear in the tail of the time scale, similar results hold. To see this, we consider

$$\mathbb{T}_{1,2} = \{0, 1, 3, 4, 6, \ldots, k, k+1, k+3, k+4, \ldots\},$$

which is a time scale where the graininess alternates between 1 and 2. Thus we can think of this as a particular instance of a time scale generated by a random variable

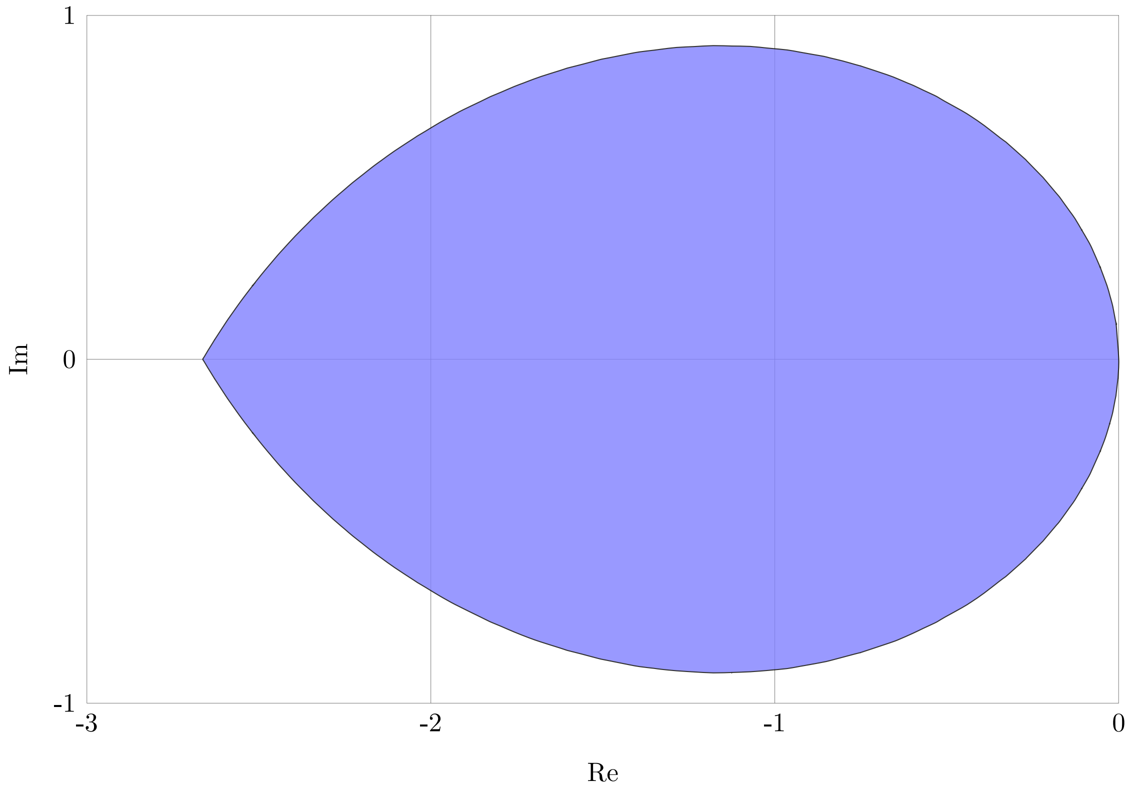

Figure 3.1. The region of stability for the stochastically generated time scale $\mathbb{T}_\Gamma$.

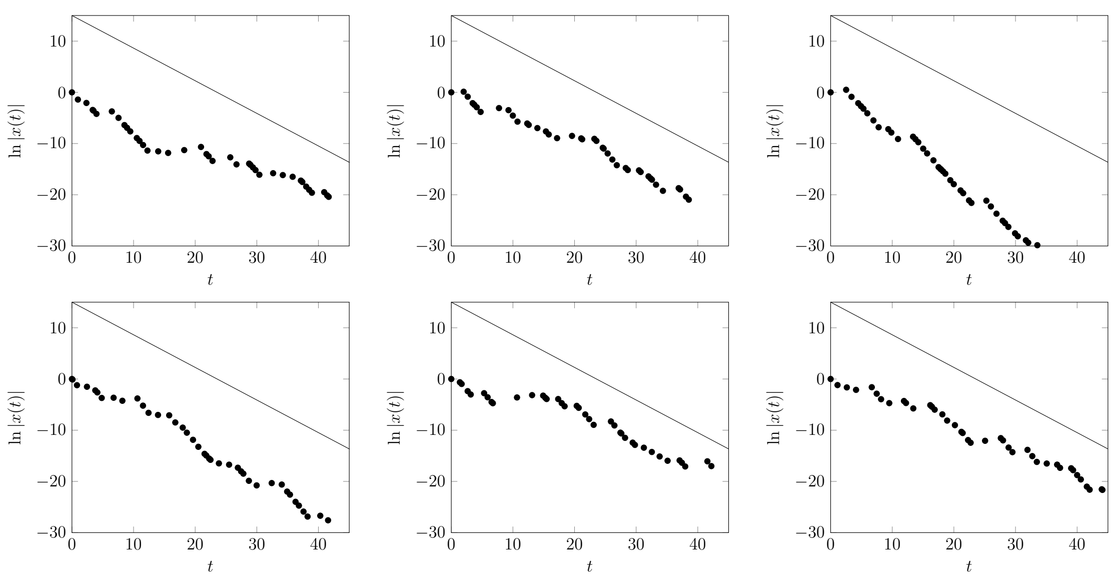

Figure 3.2: $\ln|x(t)|$ (dots) and the theoretical decay rate (line) with $\lambda_1 = -1 + .25i$ on six different time scales generated from the gamma distribution with shape parameter 2 and rate parameter 2.

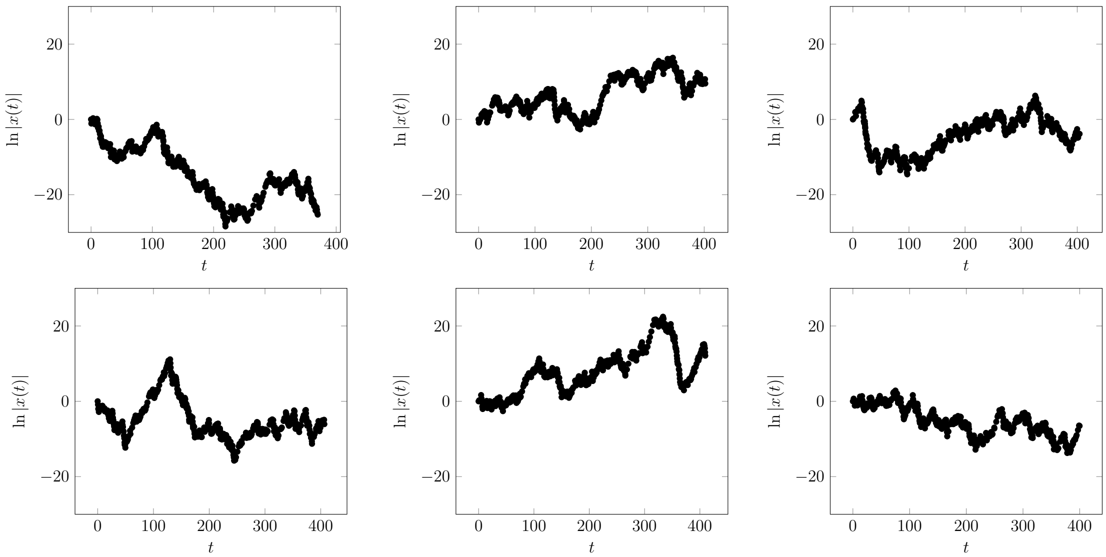

Figure 3.3: $\ln(|x(t)|)$ for $\lambda_2 = -2 + .67i$ on six different time scales generated from the gamma distribution with shape parameter 2 and rate parameter 2.

with probability mass function

$$f(t) = \begin{cases} 1/2 \text{ if } t = 1, \\ \\ 1/2 \text{ if } t = 2. \end{cases} \quad .$$

The condition on λ for stability of (2.3) on $\mathbb{T}_{1,2}$ is

$$\sum_{i=1}^{2} f(i) \ln|1 + \lambda i| < 0.$$

Solutions of (2.3) for λ satisfying the above condition is shown in Figure 3.4 with along with the theoretical decay rate which we mentioned in Remark 3.3. We will expand upon this observation in Chapter Five, where we discuss how the theory of stochastic time scales informs the theory of general time scales.

3.2.1 Decay Analysis

The example that showed the exponential stability of (2.3) with $\lambda = \lambda_2$ on $\mathbb{T}_\Gamma$ should give us some concern with this framework. After all, in applications we would not call a system with a state variable whose magnitude reached 10^{20} "stable"! We now consider how to analyze the probability that the state variable will have a

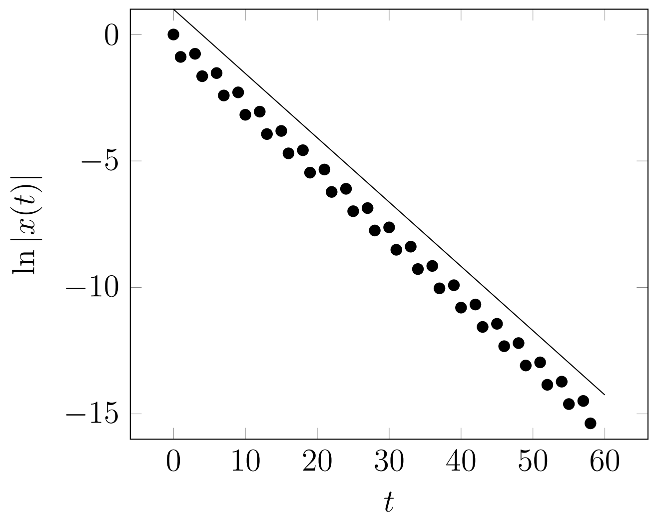

Figure 3.4: The solution of (2.3) on $\mathbb{T}_{1,2}$ for $\lambda = -.9 + .4i$ along with the predicted decay rate.

magnitude below a certain tolerance $\tau > 0$. We will see in Section 3.3 that we can also tackle this problem by considering a different kind of stability, mainly mean square exponential stability. Throughout this section we will denote the conditional probability of an event A given another event B by $\Pr[A; B]$.

Let $\{x(t_k)\}_{k=0}^{\infty}$ be the solution of (2.3) with initial condition $x(t_0) = k$ where $|k| = 1$ on an i.i.d. time scale $\tilde{\mathbb{T}}$ generated by μ. Note

$$\Pr[|x(t_0)| < \tau] = \begin{cases} 0 \text{ if } \tau \leq 1, \\ 1 \text{ if } \tau > 1. \end{cases}$$

For the sake of the simplicity, assume further that the μ is a continuous random variable which admits a probability distribution function f with support $(0, \infty)$. To find the probability that the magnitude of the state variable is beneath the tolerance after one step, write $\lambda = x + iy$ and calculate

$$\begin{aligned} \Pr[|x(t_1)| < \tau] &= \Pr[|x(t_0)(\lambda\mu_0 + 1)| < \tau] \\ &= \Pr[|x(t_0)||(x+iy)\mu_0 + 1| < \tau] \\ &= \Pr[(\mu_0 x + 1)^2 + \mu_0^2 y^2 < \tau^2] \end{aligned}$$

$$= \Pr[\mu_0^2(x^2+y^2) + 2\mu_0 x + (1-\tau^2) < 0]$$
$$= \Pr[\mu_0^2|\lambda|^2 + 2\mu_0 \operatorname{Re}(\lambda) + (1-\tau^2) < 0]$$
$$= \Pr[c_1(\tau) < \mu_0 < c_2(\tau)]$$
$$= \int_{c_1(\tau)}^{c_2(\tau)} f(\mu)d\mu,$$

where

$$c_1(\tau) = \frac{-\operatorname{Re}(\lambda) - \sqrt{(\operatorname{Re}(\lambda))^2 - |\lambda|^2(1-\tau^2)}}{|\lambda|^2}$$

and

$$c_2(\tau) = \frac{-\operatorname{Re}(\lambda) + \sqrt{(\operatorname{Re}(\lambda))^2 - |\lambda|^2(1-\tau^2)}}{|\lambda|^2}$$

are obtained via the quadratic formula with the assumption $c_i(\tau) = 0$ if the equations above yield imaginary or negative numbers, $i = 1, 2$. Note that if $\tau \geq 1$ then $c_1(\tau)$ and $c_2(\tau)$ are real–valued. This is not necessarily the case if $\tau < 1$, since the solution cannot decay arbitrarily fast. The smallest factor the solution can decay by is $\hat{\tau}$ such that $\operatorname{Re}(\lambda)^2 - |\lambda|^2(1-\hat{\tau}^2) = 0$.

Since $\tilde{\mathbb{T}}$ is an i.i.d. time scale, the method above shows the probability that the solution grows by a factor bounded by τ on any single time step. By letting $\tau = 1$, we obtain the probability that the solution will not grow the next time step,

$$p := \Pr[|x(t_1)| < 1] = P[|x(t_k)| < c; |x(t_{k-1})| = c] \tag{3.3}$$

for any $c > 0$. This can be a very useful design parameter, as we would like to choose λ so that the probability of decay in the state variable is sufficiently large (or one), ensuring "local stability." The design parameter p may be more convenient than calculating the probability that the magnitude of the state variable is beneath a tolerance at t_k, as it involves k integrations, as we now show.

Note that

$$\Pr[|x(t_1)| < \tau] = \int_{c_1(\tau)}^{c_2(\tau)} f(\mu)d\mu := F(\tau)$$

where $F(\tau)$ is increasing as $c_2(\tau)$ is increasing and $c_1(\tau)$ is decreasing and $f(\mu) \geq 0$ for $\mu \geq 0$. Thus $F(\tau)$ is a CDF for the random variable $|x(T_1)|$. Note

$$\begin{aligned} F'(\tau) &= f(c_2(\tau))c_2'(\tau) - f(c_1(\tau))c_1'(\tau) \\ &= \frac{\tau(f(c_2(\tau)) + f(c_1(\tau)))}{\sqrt{(\mathrm{Re}(\lambda))^2 - |\lambda|^2(1-\tau^2)}} \\ &:= h(\tau) \end{aligned}$$

is therefore a PDF for $|x(t_1)|$. Now, by the Law of Total Probability [32],

$$\begin{aligned} \Pr[|x(t_2)| < \tau] &= \int_0^\infty h(l) \Pr[|x(T_2)| < \tau; |x(T_1)| = l] dl \\ &= \int_0^\infty h(l) \Pr[|1 + \lambda M_1| < \tau/l] dl \\ &= \int_0^\infty h(l) \Pr[|1 + \lambda M_0| < \tau/l] dl \\ &= \int_0^\infty h(l) \int_0^{\tau/l} h(\mu) d\mu dl \\ &= \int_0^\infty \int_0^\tau h(l) h\left(\frac{\mu}{l}\right) \frac{1}{l} d\mu dl \\ &= \int_0^\tau \int_0^\infty h(l) h\left(\frac{\mu}{l}\right) \frac{1}{l} dl d\mu, \end{aligned}$$

so $k(\tau) := \int_0^\infty h(l) h\left(\frac{\tau}{l}\right) \frac{1}{l} dl$ is a probability distribution function for the random variable $|x(t_2)|$.

By induction it is easy to show

$$\begin{aligned} &\Pr[|x(t_k)| < \tau] \\ &= \int_0^\tau \underbrace{\int_0^\infty \cdots \int_0^\infty}_{k-1 \text{ times}} h(s_1) h\left(\frac{s_2}{s_1}\right) \cdots h\left(\frac{\mu}{s_{k-1}}\right) \frac{1}{s_1 s_2 \cdots s_{k-1}} ds_1 \ldots ds_{k-1} d\mu. \end{aligned}$$

Rather than calculate the above integral, we may use the parameter p in (3.3), an easy–to–calculate number that yields important information about tendency of the system to decay. On one hand, choosing the pole λ such that p is near one helps ensure the magnitude of the state variable will not become extremely large. On the other hand, this choice of λ may yield a slow decay rate. The "best" performance is obtained by balancing the p and the decay rate α according to some metric. For example, we may wish to maximize the decay rate subject to $p > c$ with $0 \le c < 1$. Instead of building an optimization algorithm for this problem, we simply plot both p and α as a function of λ. Such a plot is shown in Figure 3.5 for $\mathbb{T}_\Gamma$.

The balance of two opposing criteria is a theme within control theory. We will do a more in-depth analysis of the pole-placement problem in Chapter Four via optimal control theory.

We note that the value of p is constant along any Hilger circle since

$$p = \int_0^{-2\operatorname{Re}(\lambda)/|\lambda|^2} f(\mu)d\mu = F\left(-\frac{2\operatorname{Re}(\lambda)}{|\lambda|^2}\right),$$

where F is the cumulative probability distribution of the random variable μ. For every λ on the boundary of $\mathcal{H}_\gamma$, $-2\operatorname{Re}(\lambda)/|\lambda|^2 = \gamma$. Therefore the contour plot of p as a function of λ consists of Hilger circles! If the support of the distribution of the graininess is bounded by $\mu_{\max} \in (0, \infty)$ and $\lambda \in \mathcal{H}_{\mu_{\max}}$, then $-2\operatorname{Re}(\lambda)/|\lambda|^2 > \mu_{\max}$. Thus,

$$p \ge F(\mu_{\max}) = 1,$$

so $p = 1$. This shows that if λ is in the smallest possible Hilger circle $\mathcal{H}_{\mu_{\max}}$, then the solution will decay at each step with probability one. This behavior helps explain why we can only find deterministic time scale Lyapunov functions by solving (TSALI) if and only if the eigenvalues of the system matrix lie in $\mathcal{H}_{\mu_{\max}}$. We will see in the next section that stochastic Lyapunov functions remove this constraint.

Remark 3.5. Recall from Section 2.2 that the smallest Hilger circle is contained in the region of exponential stability. In general, the Hilger circle corresponding to a probability of decay $\beta < 1$ is not contained in the region of stability. To see this, Consider a time scale generated by the probability mass function

$$f(\mu) = \begin{cases} \beta \text{ if } \mu = 1, \\ 1 - \beta \text{ if } \mu = 2. \end{cases}$$

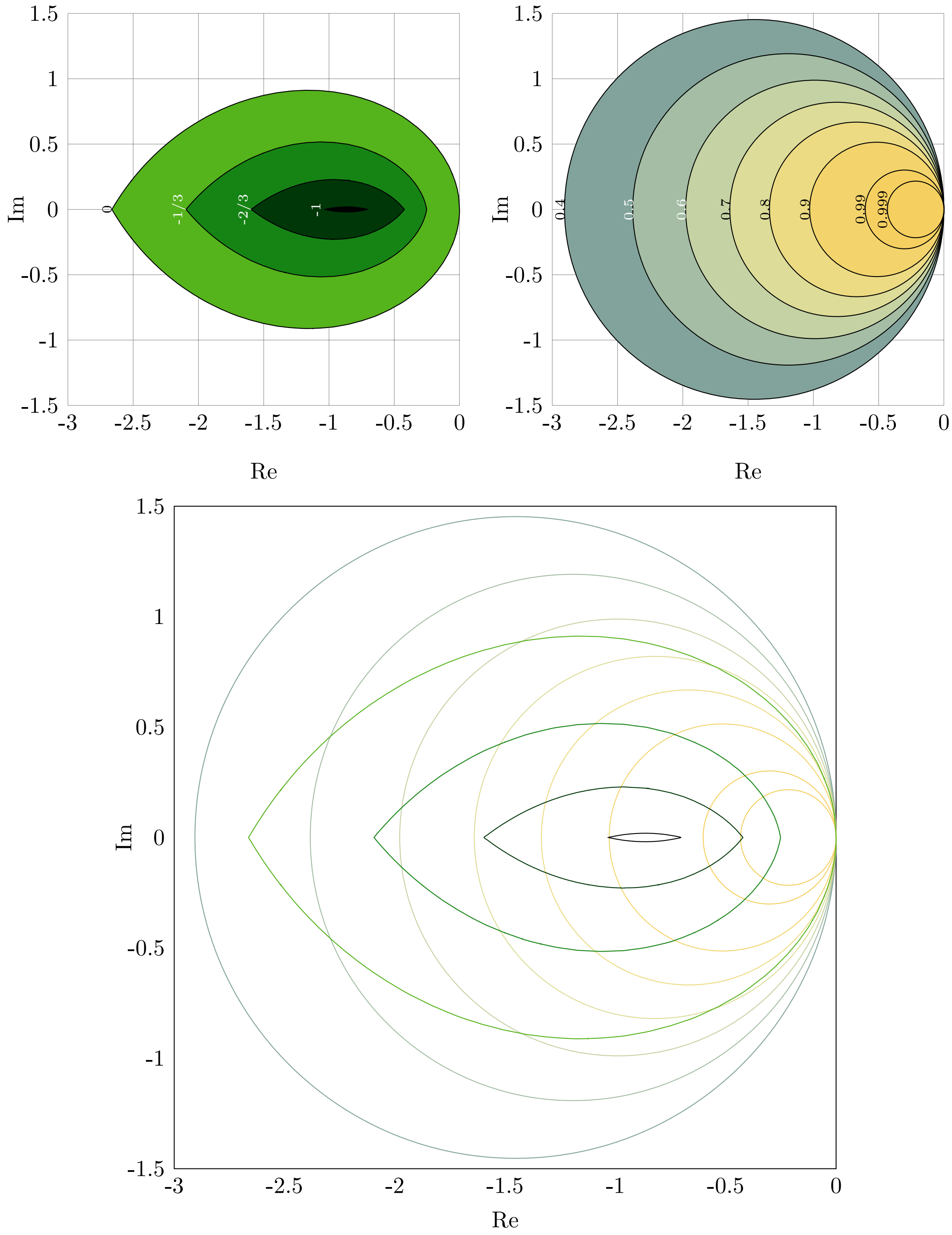

Figure 3.5: All figures are for the stochastically generated time scale $\mathbb{T}_\Gamma$. Top Left: Contour plot of the decay rate α in the region of stability. Top Right: Contour plot of p in the left–half complex plane. Bottom: Contour plot of p in the left–half complex plane with the decay rate α in the region of stability overlaid.

Then the probability β contour is a Hilger circle of radius one, but the region of exponential stability is strictly contained in the Hilger circle of radius one.

3.3 Indirect Method: Mean Square Exponential Stability

The direct method just described classifies exponential stability almost surely. As we saw, this brand of stability does not always lead to desirable qualitative behavior. In this section, rather than appeal to a decay analysis, as in the previous section, we will instead explore other types of stability. We now move from scalar dynamic equations to systems of dynamic equations.

Since solutions of the vector dynamic equation

$$x^{\Delta} = A(\mu(t))x$$

on an i.i.d stochastic time scale generated by μ form a Markov chain, the study of dynamic equations on stochastic time scales can be viewed as the study of stability theory of Markov chains. For the remainder of this dissertation, we refer to an i.i.d. stochastic time scale generated by μ simply as a *stochastic time scale* and denote this by $\tilde{\mathbb{T}}$.

3.3.1 Quadratic Stochastic Lyapunov Functions

We choose to search for quadratic Lyapunov functions, i.e. where V is of the form

$$V(x) = x^T P x,$$

and $P = P^T > 0$. We immediately have $V(0) = 0$ and $V(x) > 0$ for $x \neq 0$. Let x be a solution of the stochastic dynamic equation (2.10) on the stochastic time scale generated by $\{\mu_n\}$ with initial point t_0. Then,

$$x(t_{n+1}) = (I + A(\mu_n)\mu_n)x(t_n).$$

In particular, for $x(t_0) = x_0$, $x(t_1) = x_1$, and $\mu_0 = \mu$,

$$x_1 = (I + A(\mu)\mu)x_0.$$

Thus,

$$\begin{aligned}
\mathrm{E}_{x_0}[V(x(t_1))] - V(x_0) =& \mathrm{E}_{x_0}[x(t_1)^T P x(t_1)] - x_0^T P x_0 \\
=& \mathrm{E}_{x_0}[((I + A(\mu)\mu)x_0))^T P(I + A(\mu)\mu)x_0)] - x_0^T P x_0 \\
=& x_0^T \mathrm{E}[\mu A^T(\mu)P + \mu P A(\mu) + \mu^2 A^T(\mu) P A(\mu)] x_0.
\end{aligned}$$

Hence, the hypotheses in Theorems 2.8 and 2.9 are satisfied provided the *stochastic time scale algebraic Lyapunov equation*

$$\mathrm{E}[\mu A^T(\mu)P + \mu P A(\mu) + \mu^2 A^T(\mu) P A(\mu)] = -M, \qquad M > 0 \qquad (\mu\text{-STSALE})$$

admits a positive definite solution P ($M = \alpha P$ in Theorem 2.9).

When $A(\mu) \equiv A$, (μ-STSALE) becomes

$$\mathrm{E}[\mu] A^T P + \mathrm{E}[\mu] P A + \mathrm{E}[\mu^2] A^T P A = -M. \qquad \text{(STSALE)}$$

The left-hand side of (STSALE) in this case is reminiscent of (TSALI). An important difference between (STSALE) and (TSALI) is that (STSALE) does not depend on μ, since μ integrates out in the expected value calculation. Due to the time varying nature of μ in the TSALI, theorists only know there is a solution to TSALI if $\mathrm{spec}(A) \subset \mathcal{H}_{\mu_{\max}}$, where $\mu_{\max} = \limsup_{t\to\infty} \mu(t)$. We will see that solutions of STSALE exist for $\mathrm{spec}(A)$ in a larger region of the complex plane than $\mathcal{H}_{\mu_{\max}}$.

3.3.2 Mean-Square Stability

We now introduce some technical lemmas before arriving at the main result in this section. We will show that the existence of a quadratic stochastic time scale Lyapunov function turns out to be equivalent to mean-square exponential stability.

Lemma 3.2. *Let u, v be a random vectors with $\mathrm{E}[u], \mathrm{E}[v], \mathrm{Cov}(u, v) < \infty$, where $\mathrm{Cov}(u, v)$ is the cross-covariance matrix of u and v with entries $[\mathrm{Cov}(u, v)]_{ij} = \mathrm{cov}(u_i, v_j)$, where* cov *denotes scalar covarience. Let X be a random matrix such* *that the entries of u and v are independent of the entries of X. Then*

$$\mathrm{E}[u^T X v] = \mathrm{tr}[\mathrm{E}[X]\mathrm{Cov}(u,v)] + \mathrm{E}[u]^T \mathrm{E}[X]\mathrm{E}[v].$$

Proof. By direct calculation, we see

$$\begin{aligned}
\mathrm{E}[u^T X v] &= \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} u_i X_{ij} v_j\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X_{ij} u_i v_j] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X_{ij}]\mathrm{E}[u_i v_j] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X]_{ij}(\mathrm{cov}(u_i, v_j) + \mathrm{E}[u_i]\mathrm{E}[v_j] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X]_{ij}\mathrm{Cov}(u,v)_{ij} + \mathrm{E}[u_i]\mathrm{E}[X]_{ij}\mathrm{E}[v_j] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X]_{ij}\mathrm{Cov}(u,v)_{ji} + \mathrm{E}[u]^T\mathrm{E}[X]\mathrm{E}[v] \\
&= \sum_{i=1}^{n} (\mathrm{E}[X]\mathrm{Cov}(u,v))_{ii} + \mathrm{E}[u]^T\mathrm{E}[X]\mathrm{E}[v] \\
&= \mathrm{tr}(\mathrm{E}[X]\mathrm{Cov}(u,v)) + \mathrm{E}[u]^T\mathrm{E}[X]\mathrm{E}[v].
\end{aligned}$$

□

Corollary 3.4. *Let X, Y be random matrices such that the entries of X and the entries of Y are independent of each other. Then*

$$\mathrm{E}[X^T Y X] = \mathrm{E}[X^T \mathrm{E}[Y] X].$$

Proof. Let X_i be the i^{th} column of X and consider

$$\begin{aligned}
[\mathrm{E}[X^T Y X] - \mathrm{E}[X^T \mathrm{E}[Y] X]]_{ij} =& [\mathrm{E}[X^T Y X - X^T \mathrm{E}[Y] X]]_{ij} \\
=& [\mathrm{E}[X^T (Y - \mathrm{E}[Y]) X]]_{ij} \\
=& \mathrm{E}[X_i^T (Y - \mathrm{E}[Y]) X_j]
\end{aligned}$$

$$
\begin{aligned}
&= \operatorname{tr}(\mathrm{E}[Y - \mathrm{E}[Y]]\operatorname{Cov}(X_i, X_j)) \\
&\quad + \mathrm{E}[X_i]^T \mathrm{E}[Y - \mathrm{E}[Y]]\mathrm{E}[X_j] \\
&= \operatorname{tr}(0 \operatorname{Cov}(X_i, X_j)) + \mathrm{E}[X_i]^T 0 \mathrm{E}[X_j] \\
&= 0.
\end{aligned}
$$

The result follows. □

The following theorem shows that certain essential Lyapunov stability results for dynamic equations on $\mathbb{R}$ and $\mathbb{Z}$ carry over to this more general stochastic time scales setting. Here, we link the concept of mean-square stability with the existence of a solution of (μ-STSALE).

Theorem 3.2. *Let $\tilde{\mathbb{T}}$ be a stochastic time scale. Define the operator*

$$S_{A(\mu)}(P) := \mathrm{E}[(\mu A(\mu))^T P + P(\mu A(\mu)) + (\mu A(\mu))^T P(\mu A(\mu))].$$

Then the adjoint of $S_{A(\mu)}(P)$ is given by

$$S^*_{A(\mu)}(P) = \mathrm{E}[(\mu A(\mu))P + P(\mu A(\mu))^T + (\mu A(\mu))P(\mu A(\mu))^T].$$

The following are equivalent:

(i) $x^\Delta = A(\mu)x$ *is globally asymptotically mean-square stable.*

(ii) $x^\Delta = A(\mu)x$ *is globally exponentially mean-square stable.*

(iii) $\operatorname{spec}(S_{A(\mu)}) \subset \mathcal{H}_1$.

(iv) There exists $P > 0$ such that $S_{A(\mu)}(P) < 0$.

(v) For all $M > 0$, there exists $P > 0$ such that $S_{A(\mu)}(P) = -M$.

Proof. (i) $\iff$ (ii) $\iff$ (iii):

Define $X(n) = \mathrm{E}[x(t_n)x^T(t_n)]$. Then

$$\begin{aligned}\Delta X(n) &= X(n+1) - X(n)\\ &= \mathrm{E}[(I+\mu A(\mu))x(t_n)x^T(t_n)(I+\mu A(\mu))^T] - X(n)\\ &= \mathrm{E}[(I+\mu A(\mu))X(n)(I+\mu A(\mu))^T] - X(n)\\ &= \mathrm{E}[\mu A(\mu)X(n) + \mu X(n)A^T(\mu) + \mu^2 A(\mu)X(n)A^T(\mu)]\\ &= S^*_{A(\mu)}(X(n)).\end{aligned}$$

Therefore, the linear difference equation above is globally (asymptotically, exponentially) mean-square stable iff

$$\mathrm{spec}(S^*_{A(\mu)}) = \mathrm{spec}(S_{A(\mu)}) \subset \mathcal{H}_1.$$

(iv) $\Longrightarrow$ (ii):

Let $x_0 \in \mathbb{R}^n$ be given. Let $c_1, c_2, c_3 > 0$ such that $c_1 I \leq P \leq c_2 I$ and $\mathrm{E}[(\mu A(\mu))^T P + P(\mu A(\mu)) + (\mu A(\mu))^T P(\mu A(\mu))] < -c_3 I$ with $c_3/c_2 < 1$. Then

$$\begin{aligned}\Delta\mathrm{E}[x^T(t_n)Px(t_n)] =& \mathrm{E}[x^T(t_{n+1})Px(t_{n+1})] - \mathrm{E}[x^T(t_n)Px(t_n)]\\ =& \mathrm{E}[x^T(t_n)(I+\mu A(\mu))^T P(I+\mu A(\mu))x(t_n) - x^T(t_n)Px(t_n)]\\ =& \mathrm{E}[x^T(t_n)[(\mu A(\mu))^T P + P(\mu A(\mu)) + (\mu A(\mu))^T P(\mu A(\mu))]x(t_n)]\\ \leq& \mathrm{E}[x^T(t_n)(-c_3 I)x(t_n)]\\ \leq& \frac{c_3}{c_2}\mathrm{E}[x^T(t_n)(-c_2 I)x(t_n)]\\ \leq& -\frac{c_3}{c_2}\mathrm{E}[x^T(t_n)Px(t_n)].\end{aligned}$$

Therefore, by the Gronwall inequality,

$$\mathrm{E}[x^T(t_n)Px(t_n)] \leq \left(1 - \frac{c_3}{c_2}\right)^n x^T(t_0)Px(t_0).$$

So,

$$\mathrm{E}[\|x(t_n)\|^2] = \frac{1}{c_1}\mathrm{E}[x^T(t_n)(c_1 I)x(t_n)]$$

$$\leq \frac{1}{c_1}\mathrm{E}[x^T(t_n)Px(t_n)]$$
$$\leq \frac{1}{c_1}\left(1-\frac{c_3}{c_2}\right)^n x^T(t_0)Px(t_0)$$
$$\leq \frac{1}{c_1}\left(1-\frac{c_3}{c_2}\right)^n x^T(t_0)(c_2I)x(t_0)$$
$$= \frac{c_2}{c_1}\left(1-\frac{c_3}{c_2}\right)^n \|x(t_0)\|^2.$$

(ii) $\Longrightarrow$ (v):

Let $M > 0$ be given. Denote $\tilde{\Phi}_{A^T \circ \mu}(t, t_0)$, the transition matrix of $x^\Delta = A^T(\mu)x$, by $\Phi(t)$. Define

$$P := \sum_{n=0}^{\infty} \mathrm{E}[\Phi(t_n)M\Phi^T(t_n)] = M + \sum_{n=1}^{\infty} \mathrm{E}[\Phi(t_n)M\Phi^T(t_n)] > 0.$$

Note that the sum converges by the assumption of mean-square exponential stability. Setting $X(n) := \mathrm{E}[\Phi(t_n)M\Phi^T(t_n)]$,

$$\begin{aligned}
\Delta X(n) &= \mathrm{E}[(I+\mu A^T(\mu)\Phi(t_n)M\Phi^T(t_n)(I+\mu A(\mu))] - \mathrm{E}[\Phi(t_n)M\Phi^T(t_n)] \\
&= \mathrm{E}[(I+\mu A^T(\mu)X(n)(I+\mu A(\mu))] - X(n) \\
&= \mathrm{E}[\mu A^T(\mu)X(n) + \mu X(n)A(\mu) + \mu^2 A^T(\mu)X(n)A(\mu)] \\
&= S_{A(\mu)}(X(n)).
\end{aligned}$$

Thus,

$$\begin{aligned}
S_{A(\mu)}(P) &= \mathrm{E}\left[\mu A^T(\mu)\sum_{n=0}^{\infty} X(n) + \mu \sum_{n=0}^{\infty} X(n)A(\mu) + \mu^2 A^T(\mu)\sum_{n=0}^{\infty} X(n)A(\mu)\right] \\
&= \sum_{n=0}^{\infty} \mathrm{E}[\mu A^T(\mu)X(n) + \mu X(n)A(\mu) + \mu^2 A^T(\mu)X(n)A(\mu)] \\
&= \sum_{n=0}^{\infty} S_{A(\mu)}(X(n)) \\
&= \sum_{n=0}^{\infty} \Delta X(n) \\
&= -X(0)
\end{aligned}$$

$= -M.$

□

In the theory of stochastic dynamic systems, mean-square asymptotic stability implies almost sure (or stochastic) asymptotic stability, but the converse is not true [13]. In the LTI case, we can explore the relationship between these two stability concepts from a geometric viewpoint. To do this, we require the following corollary to Theorem 3.2.

Corollary 3.5. *Let $\tilde{\mathbb{T}}$ be a stochastic time scale generated by μ. Let $A \in \mathbb{R}^{n\times n}$ be given. Consider the dynamic equation on $\tilde{\mathbb{T}}$*

$$x^{\Delta} = Ax, \qquad x(t_0) = x_0. \tag{3.4}$$

Define the operator

$$S_A(P) := \mathrm{E}[\mu]A^T P + \mathrm{E}[\mu]PA + \mathrm{E}[\mu^2]A^T PA.$$

The following are equivalent:

(i) The system (3.4) *on $\tilde{\mathbb{T}}$ is globally asymptotically mean-square stable.*

(ii) The system (3.4) *on $\tilde{\mathbb{T}}$ is globally exponentially mean-square stable.*

(iii) $\mathrm{spec}(S_A) \subset \mathcal{H}_1$

(iv) There exists $P > 0$ such that $S_A(P) < 0$.

(v) $\forall\ M > 0$, $\exists\ P > 0$ *such that* $S_A(P) = -M$.

(vi) $\mathrm{spec}(A) \subset \mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$.

Proof. It remains to show $(v) \iff (vi)$. Multiplying the equation in (v) by $\mathrm{E}[\mu^2]/(\mathrm{E}[\mu])^2$, we get

$$\frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A^TP + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}PA + \left(\frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}\right)^2 A^TPA = -\frac{\mathrm{E}[\mu^2]}{(\mathrm{E}[\mu])^2}M := -L,$$

where $L > 0$. Factoring yields

$$-P + \left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A\right)^T P \left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A\right) = -L,$$

which we recognize as a discrete Lyapunov or Stein equation [37]. This has a solution if and only if

$$\operatorname{spec}\left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A\right) \subset \mathcal{B}_1,$$

the unit ball, which is equivalent to $\operatorname{spec}(A) \subset \mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$. □

It is worth noting that the preceding proof relied on the fact that (STSALE) can be written in the form of a Stein equation. This fact makes solving (STSALE) easy via modern computer algebra systems. Unfortunately, solving (μ-STSALE) is not as simple, as it cannot be written as a Stein equation. We can, however, write (μ-STSALE) as a *linearly perturbed Stein equation* [2] of the form

$$A^TP + PA + A^TPA + \Pi_1(P) = -M,$$

where Π_1 is a positive linear operator. We will explore this structure more in the next section.

Next, we show that the region of mean-square exponential stability $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$ is the osculating circle at the origin of the region of almost sure exponential stability $\tilde{\mathcal{S}}$.

3.3.3 Geometry of Solutions to (STSALE) *in the LTI Case*

Since mean square exponential stability implies stochastic stability, we expect $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]} \subset \tilde{\mathcal{S}}$. Next we show this is indeed the case.

Lemma 3.3. *Let $\tilde{\mathbb{T}}$ be a stochastic time scale generated by μ. Then $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]} \subset \tilde{\mathcal{S}}$.*

Proof. Let $\lambda \in \mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$. Then $2\,\mathrm{Re}(\lambda) + \mathrm{E}[\mu^2]/\mathrm{E}[\mu]|\lambda|^2 < 0$. Therefore,

$$\begin{aligned}\mathrm{E}[\ln|1+\lambda\mu|] &= \frac{1}{2}\mathrm{E}[\ln(1+2\,\mathrm{Re}(\lambda)\mu + |\lambda|^2\mu^2)] \\ &\leq \frac{1}{2}\mathrm{E}[2\,\mathrm{Re}(\lambda)\mu + |\lambda|^2\mu^2] \\ &< 0,\end{aligned}$$

where we used the inequality $\ln(1+x) \leq x$ for all $x > -1$. □

Now we explore the relationship between $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$ and $\tilde{\mathcal{S}}$ in more depth. Note that the boundary of $\tilde{\mathcal{S}}$ is given by

$$\{\lambda \in \mathbb{C} \mid \mathrm{E}[\ln|1+\lambda\mu|] = 0\} = \left\{x + iy \in \mathbb{C} \mid \mathrm{E}\left[\frac{1}{2}\ln[(1+\mu x)^2 + (\mu y)^2]\right] = 0\right\}.$$

Therefore, the boundary of $\tilde{\mathcal{S}}$ is described by the implicit equation

$$g(x,y) := \mathrm{E}\left[\frac{1}{2}\ln[(1+\mu x)^2 + (\mu y)^2]\right] = 0.$$

Its curvature can be computed with the aid of the following theorem.

Theorem 3.3 (Grey [22]). *The curvature for a two-dimensional curve given implicitly by $g(x,y) = 0$ is given by*

$$\kappa = \frac{g_{xx}g_y^2 - g_{xy}g_xg_y - g_{yx}g_xg_y + g_{yy}g_x^2}{(g_x^2 + g_y^2)^{3/2}}.$$

Lemma 3.4. *Consider the implicitly defined curve*

$$g(x,y) = \mathrm{E}\left[\frac{1}{2}\ln[(1+\mu x)^2 + (\mu y)^2]\right] = 0.$$

Suppose we can write

$$\mathrm{E}[g(\mu)] = \int_0^M f_\mu(m) g(m)\, dm,$$

where $g : (0,\infty) \to \mathbb{R}$, $M > 0$, f_μ is the probability distribution function of the random variable μ, and the above integral is with respect to some appropriate measure (counting measure for discrete distributions, Lebesgue measure for continuous).

Suppose further that $\mathrm{E}[\mu], \mathrm{E}[\mu^2] < \infty$. *Then the curvature* κ *of* g *at the origin is given by*

$$\kappa = \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}.$$

Proof. Let $h(x, y, m) := \frac{1}{2}\ln[(1 + mx)^2 + (my)^2]$. Note that $g(0, 0) = 0$, so the origin is on the implicit curve and that $g(x, y)$ is defined on the open neighborhood $B_{-1/(2M)}(0, 0)$.

A standard application of Lebesgue's Dominated Convergence Theorem yields

$$\begin{aligned}
g_x(0,0) &= \int_0^M f_\mu(m) h_x(0,0,m)\, dm = \int_0^M 2 f_\mu(m) m\, dm = 2\mathrm{E}[\mu],\\
g_{xx}(0,0) &= \int_0^M f_\mu(m) h_{xx}(0,0,m) dm = \int_0^M -2 f_\mu(m) m^2 dm = -2\mathrm{E}[\mu^2],\\
g_y(0,0) &= \int_0^M f_\mu(m) h_y(0,0,m) dm = 0,\\
g_{yy}(0,0) &= \int_0^M f_\mu(m) h_{yy}(0,0,m) dm = \int_0^M 2 f_\mu(m) m^2 dm = 2\mathrm{E}[\mu^2],\\
g_{xy}(0,0) &= \int_0^M f_\mu(m) h_{xy}(0,0,m) dm = 0 = g_{yx}(0,0).
\end{aligned}$$

Hence, by Theorem 3.3,

$$\kappa = \frac{8\mathrm{E}[\mu^2]\,(\mathrm{E}[\mu])^2}{\left([2\mathrm{E}[\mu]]^2\right)^{3/2}} = \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}.$$

□

Since the circle $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$ is tangent to the boundary of $\tilde{\mathcal{S}}$ at the origin and the two share curvature at the origin, we will refer to $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$ as *the osculating circle of the boundary* $\tilde{\mathcal{S}}$ *at the origin* and write

$$\mathcal{H}_{\mathrm{osc}} := \mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}. \tag{3.5}$$

Geometrically, $\mathcal{H}_{\mathrm{osc}}$ is the "best" circular approximation of the region of exponential stability $\tilde{\mathcal{S}}$ at the origin in the sense that $\mathcal{H}_{\mathrm{osc}} \subset \tilde{\mathcal{S}}$ and that $\mathcal{H}_{\mathrm{osc}}$ and $\tilde{\mathcal{S}}$ share tangents and have the same curvature at the origin.

It is easy to see $\mathcal{H}_{\min} \subset \mathcal{H}_{\mathrm{osc}}$, so in the case of stochastic time scales, we are able to work with a larger region of existence of solutions to Lyapunov functions.

3.3.4 Examples

Next, we look at two useful examples which illustrate these concepts.

3.3.4.1 Probability Distribution with Bounded Support We consider the stochastic dynamic initial value problem

$$x^{\Delta} = Ax = \begin{pmatrix} -1.2 & 1 & 0 \\ -1 & -3.2 & 0 \\ 0.3 & 0.3 & -1.9 \end{pmatrix} x, \quad x(0) = x_0,$$

on a stochastic time scale $\mathbb{T}_\beta$ with initial value 0 and generated by μ with a beta distribution with shape parameters 1 and 1/3. That is,

$$f_\mu(m) = \frac{1}{3(1-m)^{2/3}}, \quad 0 < m < 1.$$

Then $\mathrm{E}[\mu] = 3/4$ and $\mathrm{E}[\mu^2] = 9/14$. Note $\mathcal{H}_{\min} = \mathcal{H}_1$, $\mathcal{H}_{\mathrm{osc}} = \mathcal{H}_{6/7}$. The relationship between $\tilde{\mathcal{S}}, \mathcal{H}_{\min}$, and $\mathcal{H}_{\mathrm{osc}}$ is shown in Figure 3.6. Then $\mathrm{spec}(A) = \{-2.2, -2.2, -1.9\} \subset \mathcal{H}_{\mathrm{osc}}$, but $\mathrm{spec}(A) \not\subset \mathcal{H}_{\min}$. We found a $P > 0$ such that

$$\mathrm{E}[\mu]A^TP + \mathrm{E}[\mu]PA + \mathrm{E}[\mu^2]A^TPA < -0.16P.$$

Therefore, Theorem 2.9 applies, so

- $\mathrm{E}_x[V(x_n)] \leq (0.84)^n V(x_0)$;
- $V(x_n) \to 0$ with probability one;
- $\Pr_x[\sup_{N \leq n < \infty} V(x_n) \geq \lambda] \leq V(x_0)(.84)^N/\lambda$.

The first of these results is illustrated by examining sample paths, plotting $V(x(t))$ and $V(x_0)(.84)^N$, and seeing that $V(x_0)(.84)^N$ eventually is an upper bound for $V(x(t))$. It is therefore empirically an upper bound for the average sample path. The result of four such realizations is shown in Figure 3.7.

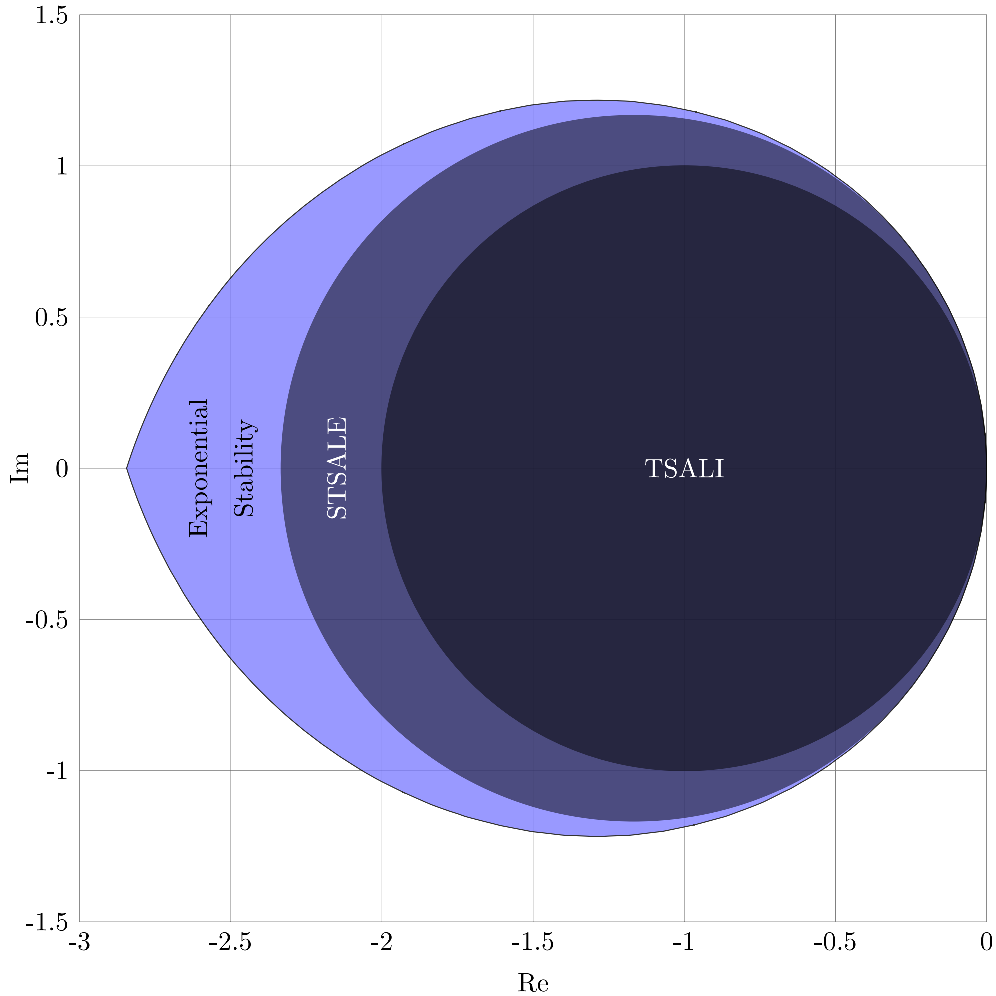

Figure 3.6: The largest region is $\tilde{\mathcal{S}}$, the region of exponential stability. The smallest region is $\mathcal{H}_{\min}$, which is the Hilger circle corresponding to the largest possible graininess. The region $\mathcal{H}_{\text{osc}}$ satisfies $\mathcal{H}_{\min} \subset \mathcal{H}_{\text{osc}} \subset \tilde{\mathcal{S}}$ and hence expands the previously known region for Lyapunov-based stability arguments in the LTI case. We see $\mathcal{H}_{\text{osc}}$ is the best circular approximation to $\tilde{\mathcal{S}}$ at the origin.

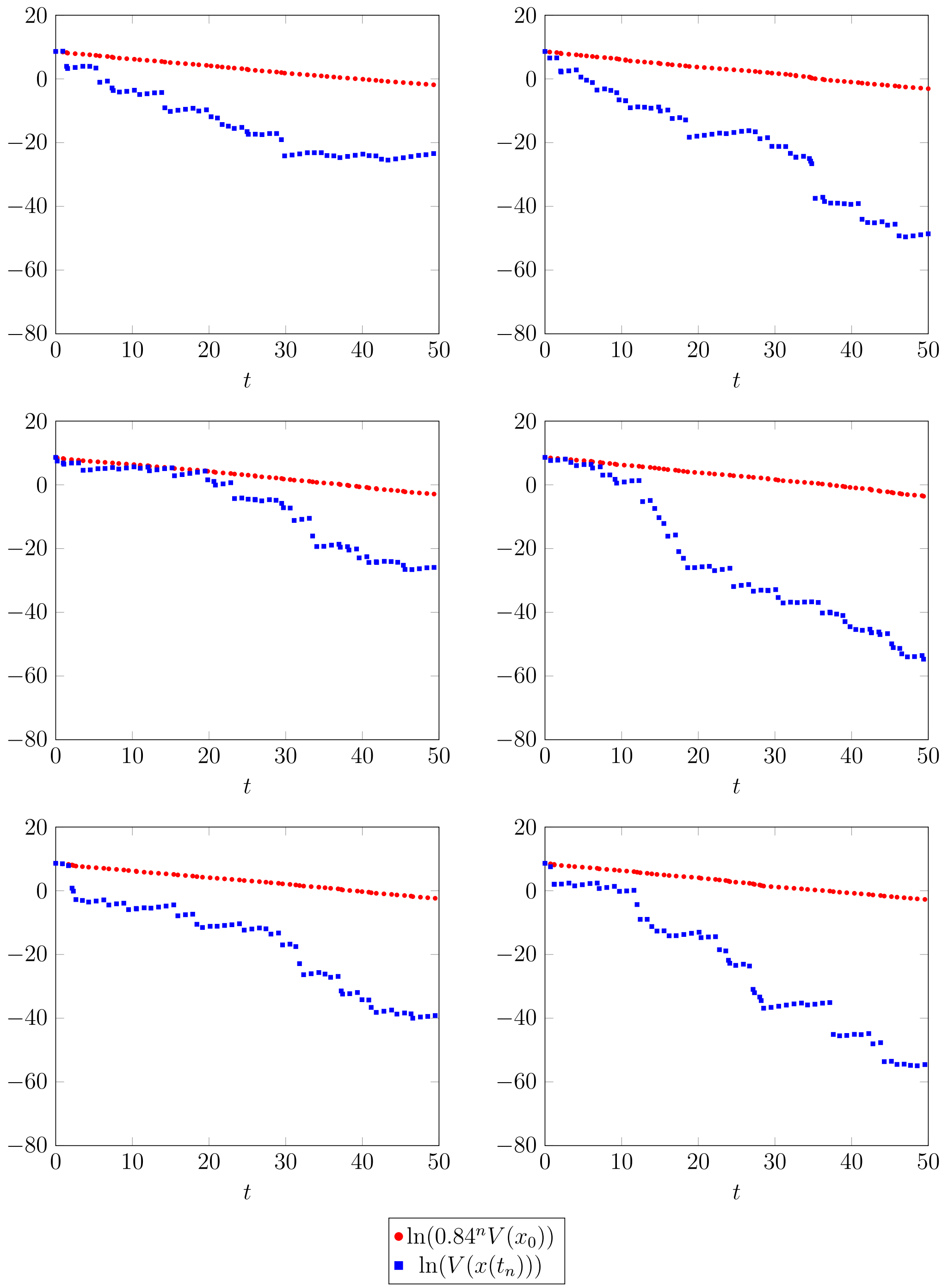

Figure 3.7: $\ln(V(x(t))$ versus t along with the upper bound of the average sample path for six realizations of the stochastic time scale $\mathbb{T}_\beta$.

3.3.4.2 Probability Distribution with Infinite Support We now consider the stochastic dynamic initial value problem

$$x^{\Delta} = Ax = \begin{pmatrix} -0.5 & -0.1 & -0.3 \\ 0.3 & -0.1 & 0.3 \\ 0.1 & 0.1 & -0.1 \end{pmatrix} x, \quad x(0) = x_0,$$

on a stochastic time scale with initial value 0 and generated by μ which is exponentially distributed with rate parameter 1/2. That is,

$$f_{\mu}(m) = \frac{e^{-m/2}}{2}, \quad m > 0.$$

Then $\mathrm{E}[\mu] = 2$ and $\mathrm{E}[\mu^2] = 8$. Hence $\mathcal{H}_{\text{osc}} = \mathcal{H}_4$ and $\mathcal{H}_{\min} = \emptyset$ since μ is an unbounded random variable. See Figure 3.8. Then $\operatorname{spec}(A) = \{-0.4, -0.2, -0.1\} \subset \mathcal{H}_{\text{osc}}$, but $\operatorname{spec}(A) \not\subset \mathcal{H}_{\min} = \emptyset$. There exists a $P > 0$ such that

$$\mathrm{E}[\mu]A^T P + \mathrm{E}[\mu]PA + \mathrm{E}[\mu^2]A^T PA < -0.31P.$$

Therefore, Theorem 2.9 applies, so

- $\mathrm{E}_x[V(x_n)] \leq (.69)^n V(x_0)$;
- $V(x_n) \to 0$ with probability one;
- $\mathrm{P}_x[\sup_{N \leq n < \infty} V(x_n) \geq \lambda] \leq V(x_0)(.69)^N/\lambda$.

Again, by plotting $V(x(t))$ and $V(x_0)(.69)^N$ and seeing that $V(x_0)(.69)^N$ is an eventual upper bound for $V(x(t))$, we illustrate the first of these results empirically. The result of four realizations is shown in Figure 3.9.

3.3.5 μ-Varying Case

We now shift our attention to dynamic equations of the form

$$x^{\Delta} = A(\mu(t)), \qquad x(t_0) = x_0, \tag{3.6}$$

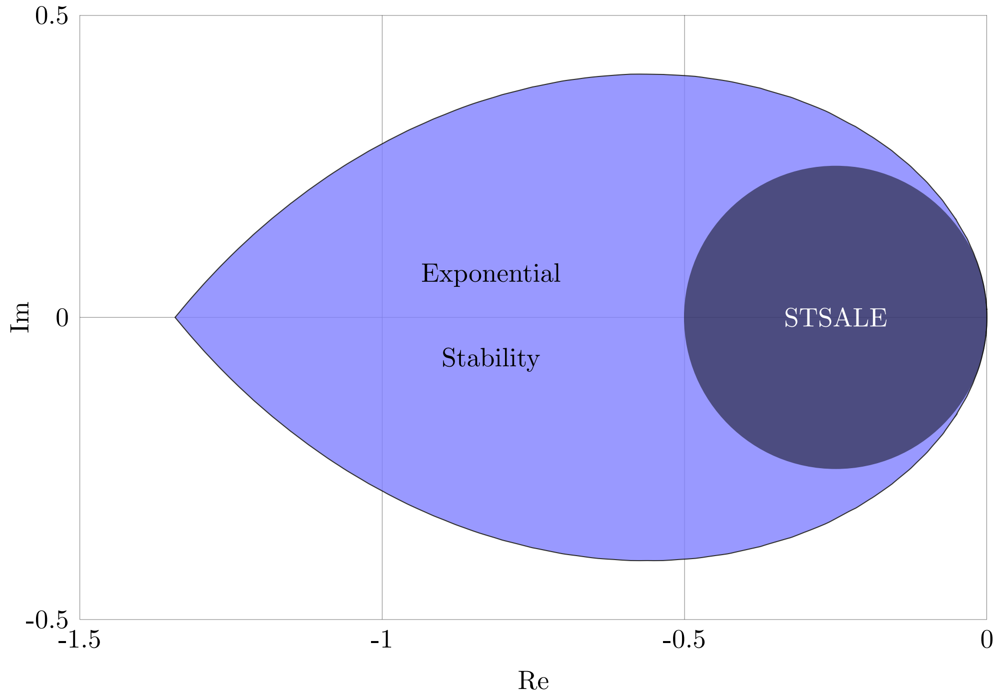

Figure 3.8: The largest region is $\tilde{\mathcal{S}}$, the region of exponential stability. Contained within $\tilde{\mathcal{S}}$ is $\mathcal{H}_{\text{osc}}$. In this case, $\mathcal{H}_{\min} = \emptyset$, and as such, previous Lyapunov theory on time scales would be unable to analyze this example.

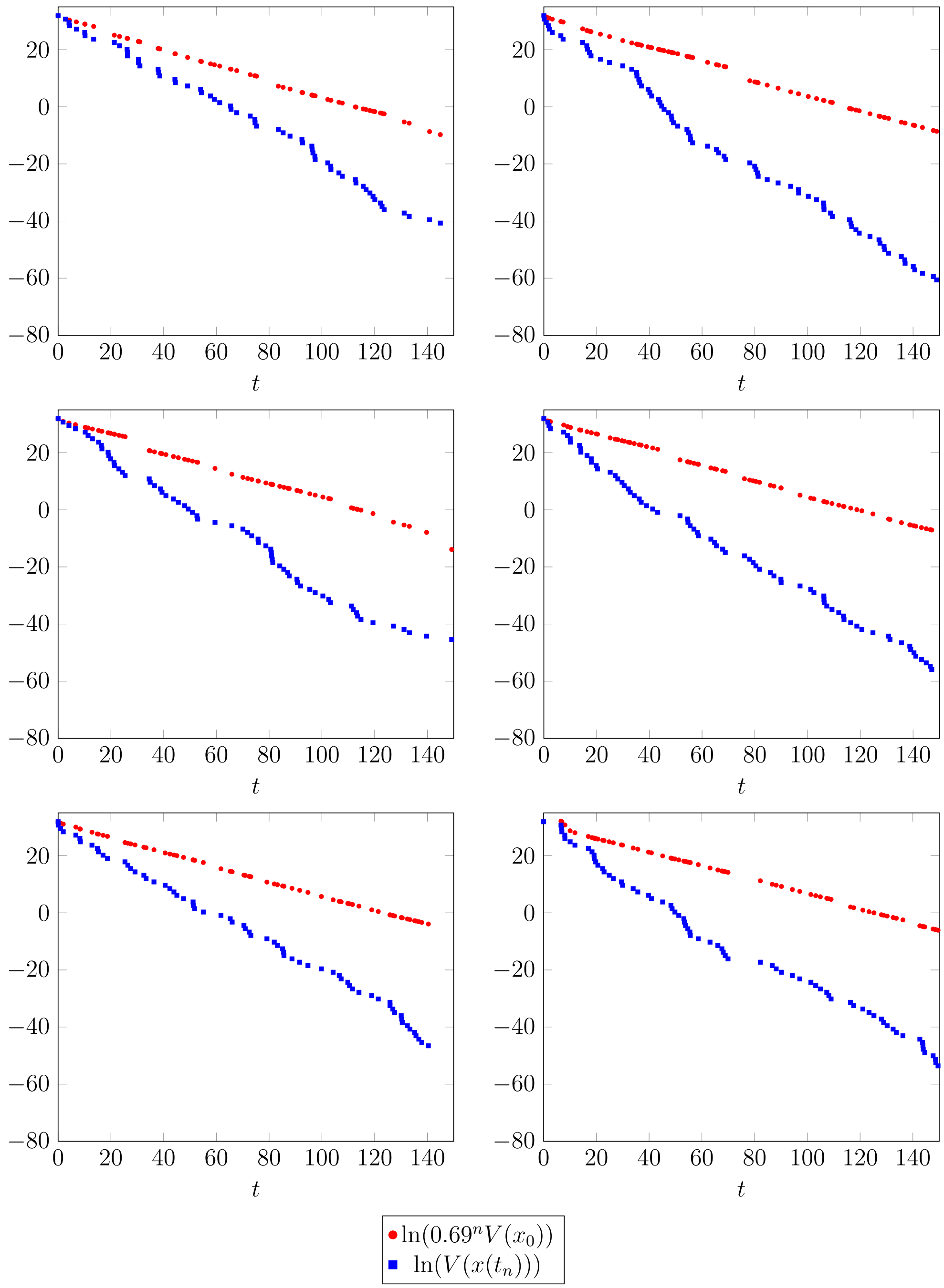

Figure 3.9: $\ln(V(x(t))$ versus t along with the upper bound of the average sample path for six realizations of the stochastic time scale $\mathbb{T}_\Gamma$.

on a stochastically generated time scale, where $A : [0, \infty) \to \mathbb{R}^{n \times n}$. Letting $\hat{A} := \mathrm{E}[\mu A(\mu)]$, (STSALE) can be rewritten as

$$\hat{A}^T P + P\hat{A} + \mathrm{E}[(\mu A(\mu))^T P(\mu A(\mu))] = -M. \tag{3.7}$$

The next lemma provides a more useful formulation of (3.7).

Lemma 3.5. *Let $u, v \in \mathbb{R}^n$ be random variables such that $\mathrm{E}[u], \mathrm{E}[v], \mathrm{Cov}(u, v) < \infty$, where $\mathrm{Cov}(u, v)$ is the cross-covariance matrix of u and v, and let $P \in \mathbb{R}^{n \times n}$ be a nonrandom matrix. Then*

$$\mathrm{E}[u^T P v] = \mathrm{E}[u^T] P \mathrm{E}[v] + \mathrm{E}[(u - \mathrm{E}[u])^T P(v - \mathrm{E}[v])].$$

Proof. Note that

$$\begin{aligned}
\mathrm{E}[u^T P v] =& \mathrm{E}\left[\begin{pmatrix} u^T & v^T \end{pmatrix} \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}\right] \\
=& \mathrm{E}\left[\begin{pmatrix} (u - \mathrm{E}[u]) \\ (v - \mathrm{E}[v]) \end{pmatrix}^T \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u - \mathrm{E}[u] \\ v - \mathrm{E}[v] \end{pmatrix} + \begin{pmatrix} \mathrm{E}[u] \\ \mathrm{E}[v] \end{pmatrix}^T \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}\right. \\
&+ \left.\begin{pmatrix} (u - \mathrm{E}[u]) \\ (v - \mathrm{E}[v]) \end{pmatrix}^T \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathrm{E}[u] \\ \mathrm{E}[v] \end{pmatrix}\right] \\
=& E\left[\begin{pmatrix} (u - \mathrm{E}[u]) \\ (v - \mathrm{E}[v]) \end{pmatrix}^T \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u - \mathrm{E}[u] \\ v - \mathrm{E}[v] \end{pmatrix}\right] \\
&+ \begin{pmatrix} \mathrm{E}[u] \\ \mathrm{E}[v] \end{pmatrix}^T \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathrm{E}[u] \\ \mathrm{E}[v] \end{pmatrix} \\
=& \mathrm{E}[(u - \mathrm{E}[u])^T P(v - \mathrm{E}[v])] + \mathrm{E}[u^T] P \mathrm{E}[v].
\end{aligned}$$

□

By Lemma 3.5, we can rewrite (3.7) as

$$\hat{A}^T P + P\hat{A} + \hat{A}^T P \hat{A} = -(\Pi_1(P) + M), \quad M > 0, \tag{3.8}$$

where

$$\Pi_1(P) := \mathrm{E}[(\mu A(\mu) - \hat{A})^T P(\mu A(\mu) - \hat{A})].$$

Note that Π_1 is a positive, linear operator. Equation (3.8) has the form of a Stein equation, but with a term that depends on P on the right hand side. We find the implicit solution of (3.8) given by

$$P = \sum_{j=0}^{\infty}((\hat{A}+I)^T)^j(\Pi_1(P)+M)(\hat{A}+I)^j.$$

Let $H_{n\times n}$ denote the space of Hermitian $n \times n$ matrices equipped with the trace norm and consider the operator $T : H_{n\times n} \to H_{n\times n}$ given by

$$\begin{aligned} T(X) &:= \sum_{j=0}^{\infty}((\hat{A}+I)^T)^j(\Pi_1(X)+M)(\hat{A}+I)^j \\ &= \sum_{j=0}^{\infty}((\hat{A}+I)^T)^j\Pi_1(X)(\hat{A}+I)^j + R, \end{aligned} \tag{3.9}$$

where

$$R := \sum_{j=0}^{\infty}((\hat{A}+I)^T)^j M(\hat{A}+I)^j > 0.$$

A solution of (3.8) is therefore a fixed point of T. We will use the following fixed point theorem of Ran and Reurings for operators on partially ordered complete metric spaces applied to $H_{n\times n}$ with the trace norm.

Theorem 3.4 (Ran and Reurings [38]). *Let L be a partially ordered set such that every pair $x, y \in L$ has a lower bound and an upper bound. Furthermore, let d be a metric on L such that (L, d) is a complete metric space. If T is a continuous, order preserving map from L into L such that*

(1) there exists $0 < c < 1$ such that $d(T(x), T(y)) \leq d(x, y)$, for all $x \geq y$,

(2) there exists $x_0 \in L$ such that $x_0 \leq T(x_0)$,

then T has a unique fixed point $\hat{x}$. Moreover, for every $x \in L$, $\lim_{n\to\infty} T^n(x) = \hat{x}$.

The calculations necessary to prove that the operator T in (3.9) has a fixed point require some technical lemmas, which we now present.

Lemma 3.6. *Let* $A = \sum_{k=0}^{\infty} A_k$*, where* $A_k \in \mathbb{R}^{n\times n}$*. Then*

$$\operatorname{tr}\left(\sum_{k=0}^{\infty} A_k\right) = \operatorname{tr}(A) = \sum_{k=0}^{\infty} \operatorname{tr}(A_k).$$

Proof. Let $\varepsilon > 0$. Since $A = \sum_{k=0}^{\infty} A_k$, there exists $N \in \mathbb{N}$ such that

$$\left\|\sum_{k=0}^{m} A_k - A\right\|_1 < \varepsilon/n \quad \text{for all } m \geq N.$$

In particular, for each $1 \leq i, j \leq n$,

$$\left|\sum_{k=0}^{m} [A_k]_{ij} - [A]_{ij}\right| < \varepsilon/n \quad \text{for all } m \geq N.$$

Then

$$\begin{aligned}
\left|\sum_{k=0}^{m} \operatorname{tr}(A_k) - \operatorname{tr}(A)\right| &= \left|\sum_{k=0}^{m}\sum_{i=1}^{n} [A_k]_{ii} - \sum_{i=1}^{n} [A]_{ii}\right| \\
&= \left|\sum_{k=0}^{m}\sum_{i=1}^{n} ([A_k]_{ii} - [A]_{ii})\right| \\
&= \left|\sum_{i=1}^{n}\sum_{k=0}^{m} ([A_k]_{ii} - [A]_{ii})\right| \\
&\leq \sum_{i=1}^{n} \left|\sum_{k=0}^{m} [A_k]_{ii} - [A]_{ii}\right| \\
&< \sum_{i=1}^{n} \frac{\varepsilon}{n} \\
&= \varepsilon.
\end{aligned}$$

□

Lemma 3.7. *Let* $A \in \mathbb{R}^{n\times n}$ *with* $\rho(A) < 1$*. Then for* $X, Y \in \mathbb{R}^{n\times n}$*,*

$$\sum_{j=0}^{\infty} (A^T)^j X A^j - \sum_{j=0}^{\infty} (A^T)^j Y A^j = \sum_{j=0}^{\infty} (A^T)^j (X - Y) A^j.$$

Proof. Let $\varepsilon > 0$. As $\sum_{j=0}^{\infty}(A^T)^j X A^j$, $\sum_{j=0}^{\infty}(A^T)^j Y A^j$, and $\sum_{j=0}^{\infty}(A^T)^j (X-Y) A^j$ all converge, there exists $N \in \mathbb{N}$ such that for all $m \geq N$,

$$\left\| \sum_{j=0}^{\infty}(A^T)^j X A^j - \sum_{j=0}^{m}(A^T)^j X A^j \right\| < \varepsilon/3,$$
$$\left\| \sum_{j=0}^{\infty}(A^T)^j Y A^j - \sum_{j=0}^{m}(A^T)^j Y A^j \right\| < \varepsilon/3,$$
$$\left\| \sum_{j=0}^{\infty}(A^T)^j (X-Y) A^j - \sum_{j=0}^{m}(A^T)^j (X-Y) A^j \right\| < \varepsilon/3.$$

Thus,

$$\begin{aligned}
&\left\| \sum_{j=0}^{\infty}(A^T)^j X A^j - \sum_{j=0}^{\infty}(A^T)^j X A^j - \sum_{j=0}^{\infty}(A^T)^j (X-Y) A^j \right\| \\
=&\left\| \sum_{j=m+1}^{\infty}(A^T)^j X A^j - \sum_{j=m+1}^{\infty}(A^T)^j Y A^j - \sum_{j=m+1}^{\infty}(A^T)^j (X-Y) A^j \right\| \\
\leq&\left\| \sum_{j=m+1}^{\infty}(A^T)^j X A^j \right\| + \left\| \sum_{j=m+1}^{\infty}(A^T)^j Y A^j \right\| + \left\| \sum_{j=m+1}^{\infty}(A^T)^j (X-Y) A^j \right\| \\
=&\varepsilon.
\end{aligned}$$

□

Lemma 3.8. *Suppose $\rho(A) < 1$, where $\rho(A)$ is the spectral radius of A. Then the operator*

$$S(X) = \sum_{j=0}^{\infty}(A^T)^j X A^j$$

is continuous.

Lemma 3.9. *Let $A \in \mathbb{R}^{n\times n}$ with $\rho(A) < 1$. Then the operator $S : H_{n\times n} \to H_{n\times n}$ given by*

$$S(X) := \sum_{j=0}^{\infty}(A^T)^j X A^j$$

is order preserving

Proof. This follows since $X - Y \geq 0$ implies $(A^T)^j(X-Y)A^j \geq 0$ for all $j \in \mathbb{N}_0$. □

Lemma 3.10 (Ran and Reurings [38]). *Let* $A \geq 0$ *and* $B \geq 0$ *be* $n \times n$ *matrices. Then* $0 \leq \mathrm{tr}(AB) \leq \|A\|\mathrm{tr}(B)$ *where* $\|\cdot\|$ *is the spectral norm.*

With the technical lemmas in hand, we are now in a position to prove our main result, which gives sufficient conditions for solutions of (2.10) to be mean-square exponentially stable. A crucial step of the result relies on the application of the fixed point theorem of Ran and Reurings.

Theorem 3.5. *The zero solution of* (2.10) *is mean-square exponentially stable provided* $\mathrm{spec}(\hat{A}) \subset \mathcal{H}_1$ *and*

$$\left\|\sum_{j=0}^{\infty}(\hat{A}+I)^j((\hat{A}+I)^T)^j\right\|\left\|\sum_{k=1}^{n}\mathrm{var}([\mu A(\mu)]_k)\right\| = \alpha < 1,$$

where $[\mu A(\mu)]_k$ *denotes the* k^{th} *column of* $A(\mu)$ *and* $\|\cdot\|$ *is the spectral norm.*

Proof. Consider the operator $T : H_{n\times n} \to H_{n\times n}$ defined by

$$\begin{aligned} T(X) &:= \sum_{j=0}^{\infty}((\hat{A}+I)^T)^j(Pi_1(X)+M)(\hat{A}+I)^j \\ &= \sum_{j=0}^{\infty}((\hat{A}+I)^T)^j\mathrm{E}[(\mu A(\mu)-\hat{A})^T X(\mu A(\mu)-\hat{A})](\hat{A}+I)^j + R. \end{aligned}$$

It suffices to show that T has a fixed point $\hat{X}$ such that $\hat{X} > 0$. Once the existence of a fixed point $\hat{X}$ is established, the fact that $\hat{X} > 0$ follows from the conclusion of Theorem 3.4 since T maps positive definite matrices into the set $\{Z \in H_{n\times n} \mid Z \geq R > 0\}$. We now show such a fixed point exists.

The complete metric space $(H_{n\times n}, \|\cdot\|_{\mathrm{tr}})$ is partially ordered by the relation $X \geq Y$ if and only if $X - Y$ is positive semidefinite. Every pair $X, Y \in H_{n\times n}$ has a lower and upper bound. Condition 2 from Theorem 3.4 is satisfied since $0 \leq T(0) = R$. The operator T is well defined because $\mathrm{spec}(\hat{A}) \subset \mathcal{H}_1$.

Since T is continuous and order preserving, it remains to show Condition 1 from Theorem 3.4 holds. To this end, let $Y \geq X$. Then

$$\|T(Y) - T(X)\|_{\mathrm{tr}}$$

$$
\begin{aligned}
&= \operatorname{tr}(T(Y) - T(X)) \\
&= \operatorname{tr}\left(\sum_{j=0}^{\infty}((\hat{A}+I)^T)^j \mathrm{E}[(\mu A(\mu) - \hat{A})^T (Y - X)(\mu A(\mu) - \hat{A})](\hat{A}+I)^j\right) \\
&= \sum_{j=0}^{\infty} \operatorname{tr}\left(((\hat{A}+I)^T)^j \mathrm{E}[(\mu A(\mu) - \hat{A})^T (Y - X)(\mu A(\mu) - \hat{A})](\hat{A}+I)^j\right) \\
&= \sum_{j=0}^{\infty} \operatorname{tr}((\hat{A}+I)^j((\hat{A}+I)^T)^j \mathrm{E}[(\mu A(\mu) - \hat{A})^T (Y - X)(\mu A(\mu) - \hat{A})]) \\
&= \operatorname{tr}\left(\sum_{j=0}^{\infty}(\hat{A}+I)^j((\hat{A}+I)^T)^j \mathrm{E}[(\mu A(\mu) - \hat{A})^T (Y - X)(\mu A(\mu) - \hat{A})]\right) \\
&\leq \left\|\sum_{j=0}^{\infty}(\hat{A}+I)^j((\hat{A}+I)^T)^j\right\| \operatorname{tr}(\mathrm{E}[(\mu A(\mu) - \hat{A})^T (Y - X)(\mu A(\mu) - \hat{A})]) \\
&= \left\|\sum_{j=0}^{\infty}(\hat{A}+I)^j((\hat{A}+I)^T)^j\right\| \sum_{i=1}^{n} \operatorname{tr}(\operatorname{cov}([\mu A(\mu)]_i, [\mu A(\mu)]_i)(Y - X)) \\
&= \left\|\sum_{j=0}^{\infty}(\hat{A}+I)^j((\hat{A}+I)^T)^j\right\| \operatorname{tr}\left(\sum_{i=1}^{n} \operatorname{var}([\mu A(\mu)]_i)(Y - X)\right) \\
&\leq \left\|\sum_{j=0}^{\infty}(\hat{A}+I)^j((\hat{A}+I)^T)^j\right\| \left\|\sum_{i=1}^{n} \operatorname{var}([\mu A(\mu)]_i)\right\| \operatorname{tr}(Y - X) \\
&= \alpha \operatorname{tr}(Y - X) \\
&= \alpha \|Y - X\|_{\mathrm{tr}},
\end{aligned}
$$

where we have used Lemma 3.2 as well as the invariance of trace under cyclic permutations. □

With the sufficient condition for stability that Theorem 3.5 provides, we can answer questions concerning the control theory of μ-varying dynamic equations. We will do this in Section 4.2

CHAPTER FOUR

Control Theory Applications of Stochastic Time Scales Stability Theory

In this chapter, we will apply the results from Chapter Three to the observer and controller problems. We begin with a novel time scales-based observer design which relies on the results of Section 3.1. Later, we will develop an observer design more which is more suitable for use in state-feedback applications by appealing to the results in Section 3.3.5. Finally, we will focus on the controller problem by developing optimal control theories for the control models of interest to us in this work.

4.1 Observer Design for Battery State-of-Charge Estimation

The apparently straightforward question of how to accurately estimate the amount of charge remaining in a battery has long presented an engineering challenge. Unlike a fuel tank with a level gauge, it is impractical to directly measure charge in typical battery. Indirect methods are under investigation, though many will be difficult to practically and cost-effectively implement. The problem of State-of-Charge (*SOC*) estimation is gaining importance, however, as batteries play an increasingly prominent role in the automotive and energy industries, for example [4, 10, 23, 39, 41, 42]. With small consumer applications such as cell phones and laptops, there is a wide margin for error in *SOC* estimation. But with applications in aerospace, transportation, and energy, which require much larger and heavier batteries, oversizing a battery bank comes with a sizable economic penalty, and erroneous estimates of battery charge can be extremely problematic. The two most commonly used methods for *SOC* estimation are Coulomb counting (the direct integration of battery current to obtain charge) and *open circuit voltage* (V_{OC}) correlation (using a known correlation between V_{OC} and *SOC*). Coulomb counting suffers from the problem of unbounded estimation error, stemming from current measurement error that is always present to one degree or another.

$$SOC = Q_{\max} \int_0^t (I(\tau) + b) d\tau = Q_{\max} \int_0^t I(\tau) d\tau - bt$$

The equation above, where $Q_{\max}$ is the maximum battery charge capacity, I is the battery current, and b is the average current measurement bias, shows that Coulomb counting estimates will eventually be overtaken by unbounded error bt.

Open circuit voltage correlation requires either that all loads be removed from the battery periodically, or that V_{OC} be estimated. Also, many batteries have very "flat" SOC vs. V_{OC} curves, so that small errors in V_{OC} estimation lead to large errors in SOC estimation. One promising approach combines both methods, using a battery model observer to reconstruct an estimate of the SOC. The problem introduced by the observer is that it requires battery power to operate itself. The microprocessor updating the observer states requires power to perform computation, A/D conversion, and communication with the outside world. To ensure that the observer remains stable and operates as expected, updates are performed at a sufficiently high, uniform sampling rate. In this section, we show that an observer designed around the theory of dynamic equations on time scales can drastically reduce the update rate relative to traditional observer design, and therefore reduce the parasitic power requirements of the observer circuitry.

This section investigates systems on different time scales, i.e., we have a dynamical system in continuous time (the battery) and another on a discrete time scale (the observer). Context will usually clarify when a variable belongs to one class or the other, but for further clarification we henceforth denote the independent time variables $\tau \in \mathbb{R}$ and $t \in \mathbb{T}$.

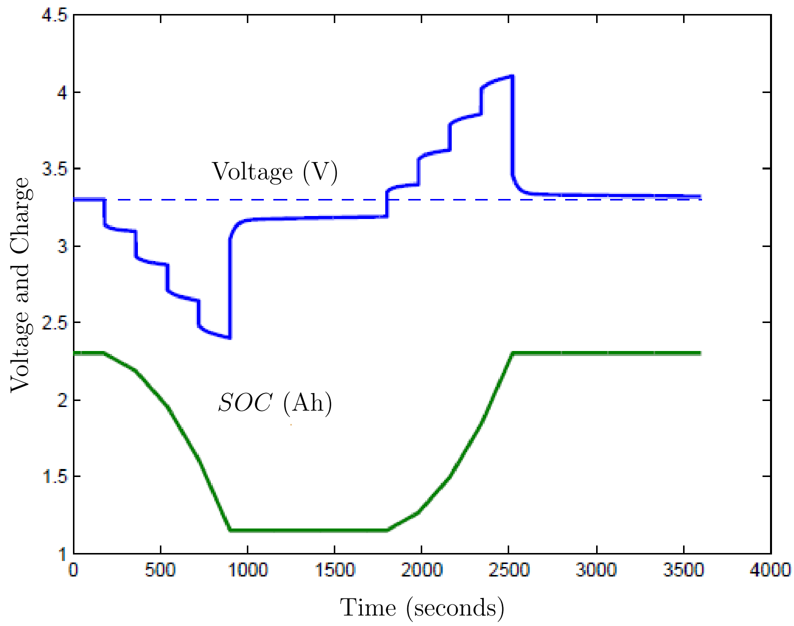

Figure 4.1: An illustration of the battery voltage transient response, excited by a discharge current stair-step of $\{0, 2.3, 3.6, 6.9, 9.2\}$ Amps, followed by a charging sequence of the same magnitude. The *SOC* is also shown.

4.1.1 Battery Model

For this work, we study the 2.3 Ah Lithium-ion battery studied by Codeca, *et al.* [11], with a nominal terminal voltage of 3.3V. The model consists of an *SOC*-dependent voltage source in series with two RC tank circuits that closely approximate the transient response of the batterys terminal voltage to step changes in current. Figure 4.1 illustrates the voltage transient response. An LTI state space model of the battery can be easily derived, after linearly approximating the batterys V_{OC} vs *SOC* curve (which is a very good approximation for *SOC* values between 20% and 80%), in the standard form

$$\dot{x}(\tau) = Ax(\tau) + Bu(\tau), \qquad A \in \mathbb{R}^{3\times 3}, \qquad B \in \mathbb{R}^{3\times 1},$$
$$y(\tau) = Cx(\tau) + Du(\tau), \qquad C \in \mathbb{R}^{1\times 3}, \qquad D \in \mathbb{R}, \tag{4.1}$$

where $x : \mathbb{R} \to \mathbb{R}^3$, and $y, u : \mathbb{R} \to \mathbb{R}$. State vector x represents the cumulative discharge, the voltage across the first RC tank and the voltage across the 2nd RC tank. Input u is the battery current, and output y is the terminal voltage offset. Henceforth, positive current $u > 0$ represents a discharging battery. The corresponding battery variables of interest are: terminal voltage $V = y + 3.3$; battery current $I = u$; and, given as a percentage, state of charge $SOC = (2.3 - x_1)/2.3$. It is noteworthy that the presence of a charge integrator in the model means that A has one zero eigenvalue.

4.1.2 Observer Design

Henceforth, we will assume that $\mathbb{T}$ is a discrete time scale. Next, we note that no discrete-time model will mirror its continuous-time cousin unless the continuous-time input $u(\tau)$ remains constant in between sample points (which will rarely occur). Thus, we define the sample-and-hold error $\varepsilon_t(\tau)$ such that

$$u(\tau) = u(t) + \varepsilon_t(\tau); \qquad t \in \mathbb{T}; \qquad \tau \in [t, \sigma(t)); \qquad \varepsilon_t(t) = 0.$$

Using sample-and-hold discretization as in Section 2.3.2, the continuous model of (4.1) can be written

$$\begin{aligned} x^{\Delta}(t) &= \mathcal{A}(\mu(t))x(t) + \mathcal{B}(\mu(t))u(t) + \frac{1}{\mu(t)} \int_t^{\sigma(t)} \varepsilon_t(\tau) e^{A(t-\tau)} B d\tau, \\ y(t) &= Cx(t) + Du(t), \end{aligned}$$

with $x, y, u : \mathbb{T} \to \mathbb{R}^n$.

If a time scale $\mathbb{T}$ with constant graininess is chosen, then A and B become constant matrices and the problem boils down to standard discrete-time observer design. However, the following design leverages the idea of allowing the time scale to have widely varying graininess.

The proposed observer dynamics are

$$\begin{aligned} \hat{x}^{\Delta}(t) &= \mathcal{A}(\mu(t))\hat{x}(t) + \mathcal{B}(\mu(t))u(t) + \mathcal{H}(\mu(t))[y(t) - \hat{y}(t)], \\ \hat{y}(t) &= C\hat{x}(t) + Du(t), \end{aligned}$$

where $\hat{x}$ is the estimated state, $\hat{y}$ is the estimated output, and

$$\mathcal{H}(\mu(t)) := \text{expc}(A\mu(t))H$$

for some feedback gain matrix H. Then the error dynamics of the system are

$$e^{\Delta}(t) = \text{expc}(A\mu(t))(A - HC)e(t) + \frac{1}{\mu(t)} \int_{t}^{\sigma(t)} \varepsilon(\tau) e^{A(t-\tau)} B d\tau, \qquad (4.2)$$

where $e = x - \hat{x}$.

At this point the analysis begs the question, what is the appropriate time scale $\mathbb{T}$ on which to discretize? One possibility is suggested by the hardware design: A common component in the design of *SOC* estimator circuitry is the Coulomb counter, a low-power device that integrates current and issues an interrupt pulse to a microprocessor every time the integral equals a predetermined quantum of charge, say, q Coulombs. This device, by its very nature, creates a time scale in which the graininess is inversely proportional to the magnitude of the current. However, such a time scale has two drawbacks: On one hand, a high constant current draw would produce a rapid observer update rate (small graininess), even though voltage transients may have settled long ago, sacrificing observer efficiency. On the other hand, large but rapid current fluctuations would not transfer much charge, but would invoke significant voltage transients that could be missed by the observer, sacrificing accuracy. Intuitively, for the observer to exhibit both stability and accuracy, it ought to be updated when the battery current changes "too much." We propose the system illustrated in Figure 4.2. In this design, a new point $\sigma(t) \in \mathbb{T}$ is generated whenever the battery current changes by more than amount β relative to the last point $t \in \mathbb{T}$. Thus, the time scale consists of points chosen to bound the sample-and-hold error, i.e.

$$\mathbb{T} = \left\{ \{t_0 < t_1 < t_2 ...\} \mid \sup_{\tau \in [t_k, t_{k+1})} |u(\tau) - u(t_k)| = \beta \text{ for all } k \in \mathbb{N}^0 \right\}. \qquad (4.3)$$

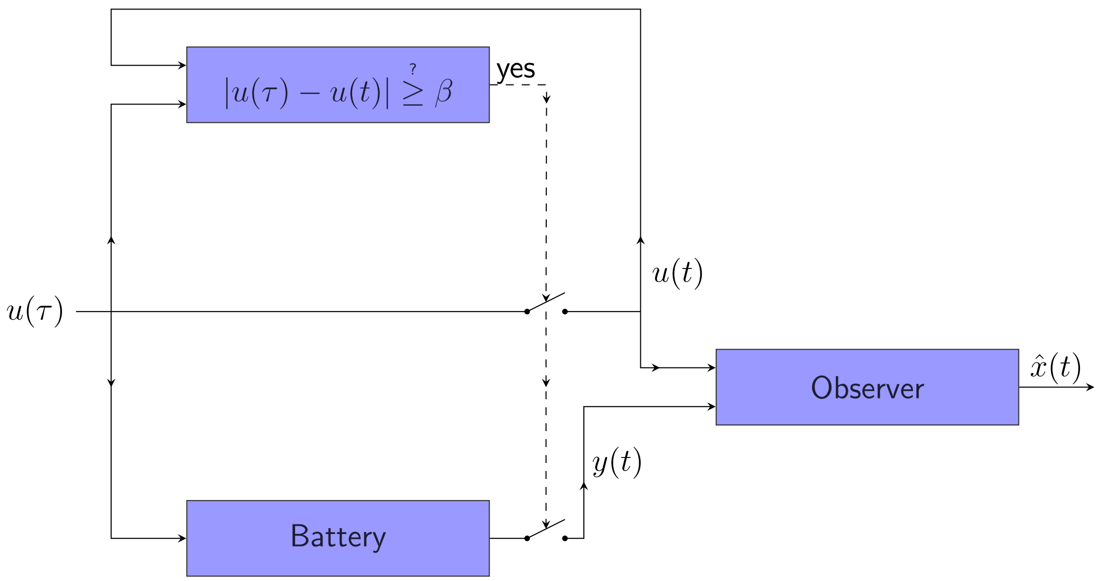

Figure 4.2: The proposed battery observer system, illustrating how the time scale is created. The switches represent sample-and-hold converters.

A graphical illustration of the time scale appears in Figure 4.3. While a complete error analysis is beyond the present scope, it is clear that small β values reduce the norm of the second term of (4.2), yielding error dynamics that are closely approximated by the first order time scale dynamic system

$$e^{\Delta}(t) = \mathrm{expc}(A\mu(t))(A - HC)e(t). \tag{4.4}$$

Using the tools of the previous chapter, we are able to analyze the stability of (4.4) using either Theorem 3.1 or Theorem 3.5. We will apply Theorem 3.1 to this problem and apply Theorem 3.5 to the more challenging problem of observer-based feedback control in the next section.

Proceeding using Theorem 3.1, we can determine the stability of (4.2). Assuming that $\mu(t)$ for $t \in \mathbb{T}$ forms an independent, identically distributed sequence with probability density function $f(\mu)$, the error dynamics will be exponentially stable about the origin if and only if

$$\mathrm{E}[\ln|1 + \mu\lambda(\mu)|] = \int_0^\infty f(\mu)\ln|1 + \mu\lambda(\mu)|d\mu < 0, \tag{4.5}$$

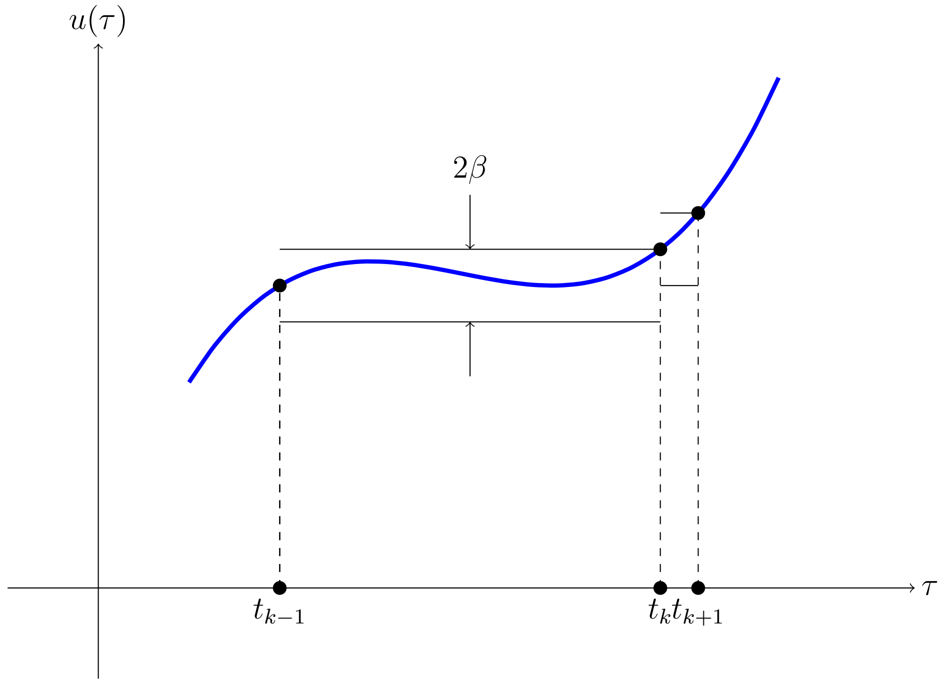

Figure 4.3. An illustration of how points in $\mathbb{T}$ are generated.

where $\lambda(\mu)$ is any eigenvalue of $\text{expc}(A\mu)(A - HC)$. We close this section with some comments.

First, from (4.5), it is evident that we need $\text{Re}(\lambda) < 0$, implying that $A - HC$ must be stable in the continuous-time sense (i.e., real-negative eigenvalues), which means that (A, C) must be observable or at least detectable. If the continuous time observer cannot be stabilized, the time scales observer cannot either.

Next, even given that (A, C) is observable, we can obtain from (4.2) an upper limit, which represents the maximum step size (or sampling interval) beyond which (4.2) would be unstable if the observer sampled at a constant rate. In this case, $\mu(t) \equiv h$ is a constant, and the question is what value of h is "too big". The answer is, the $h = h_{\max}$ corresponding to the smallest Hilger circle that encloses the eigenvalues of $\text{expc}(Ah_{\max})(A - HC)$. For the observer design of this chapter, that value is approximately $h_{\max} = 162$, meaning that the observer will destabilize with updates further apart than 162 seconds. (This may seem large, but consider the time constants involved.) The time scale $\mathbb{T} = h_{\max}\mathbb{Z}$ would produce extremely inaccurate results in general, because $\varepsilon_t(\tau)$ could grow arbitrarily large over $\tau \in [t, t + h_{\max})$. However, a time scale generated according to (4.3) and satisfying criterion (4.5) may actually admit graininess larger than $h_{\max}$, as long as such occurrences are relatively rare. This is a noteworthy result.

4.1.3 Examples

A series of examples are illustrated in Figures 4.4 through 4.9. For all of the examples, $\beta = 200$ mA.

1) Figure 4.4 illustrates that observer, in the absence of measurement error, tracks the continuous model precisely. Note that the time scale model only updates when the battery current changes.

2) Figure 4.5 shows that the time scale model will stabilize even when $\mu(t) > h_{\max}$ for some t as long as (4.5) holds. In this case, the distribution $f(\mu)$ is evenly split between $\mu = 91$ and $\mu = 182$, and all eigenvalues λ have

$$\mathrm{E}[\ln|1 + \mu\lambda(\mu)|] = \frac{1}{2}\ln|1 + 91\lambda(91)| + \frac{1}{2}\ln|1 + 182\lambda(182)| < 0.$$

3) Figure 4.6 verifies that $\mu(t) > h_{\max}$ for all t yields instability.

4) There are two kinds of errors that frequently plague SOC estimates: unknown initial conditions, and measurement bias. Figure 4.7 shows that the observer converges to the correct SOC given incorrect initial conditions.

5) Figure 4.8 illustrates that, in the presence of measurement bias so that $u(t) = I(t) + b$, the observer is more accurate than Coulomb counting alone. This figure also shows the system response to a non-periodic random current draw.

6) Figure 4.9 shows a graininess histogram, illustrating again that $\mu(t) > h_{\max}$ will not disrupt the stability or accuracy of the observer, as predicted by the criterion in (4.5), if such an occurrence is statistically limited.

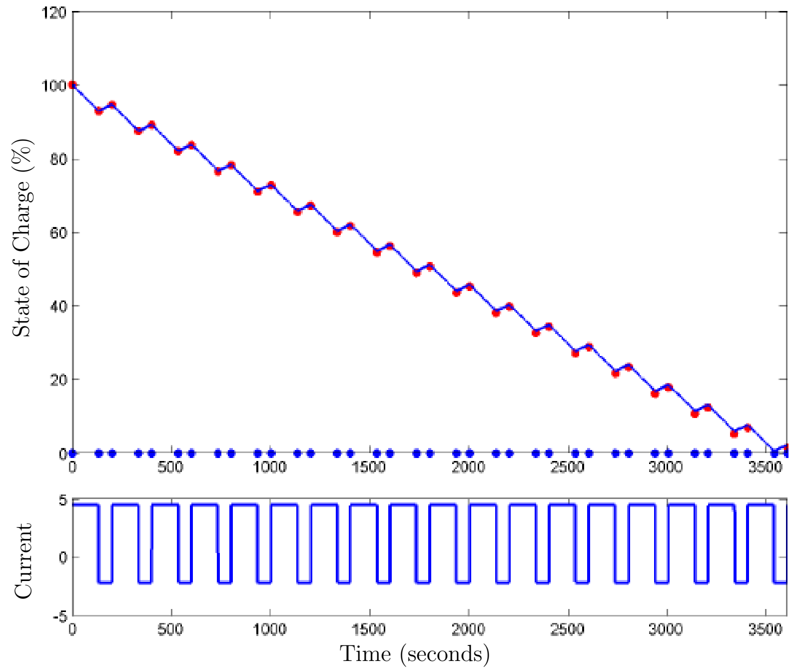

Figure 4.4: A discharge test. Top: the true SOC is the continuous blue line. The observed, discretized SOC is the red dots. The time scale $\mathbb{T}$ is shown along the time axis. Bottom: a plot of the current draw from the battery. (Note the periodic charge and discharge cycles.)

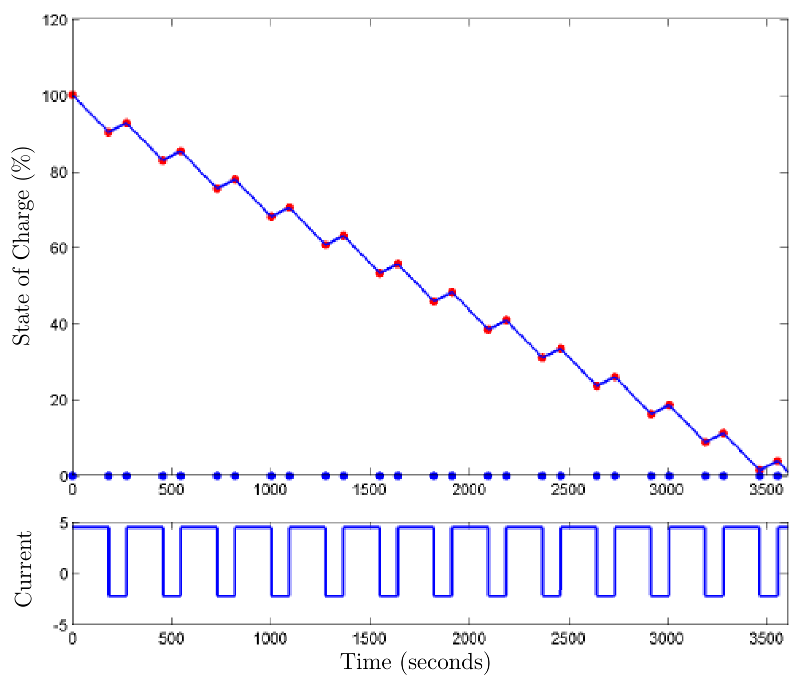

Figure 4.5: In this example, one half of the time steps are greater than $h_{\max}$, yet the time scale observer remains stable and tracks the SOC accurately.

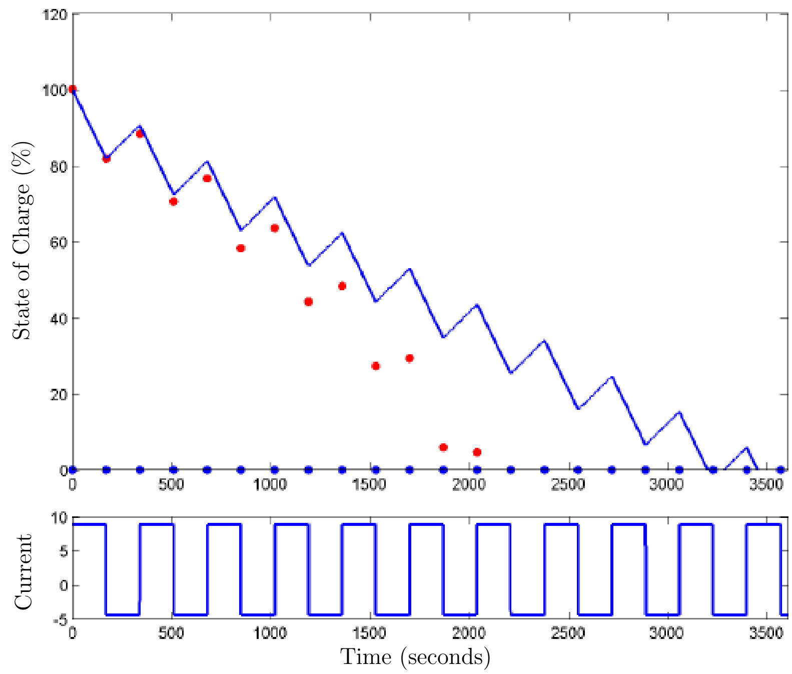

Figure 4.6: A time scale with $\mu(t) > h_{\max}$ cannot be stabilized. In this figure, $\mu(t) = 170$.

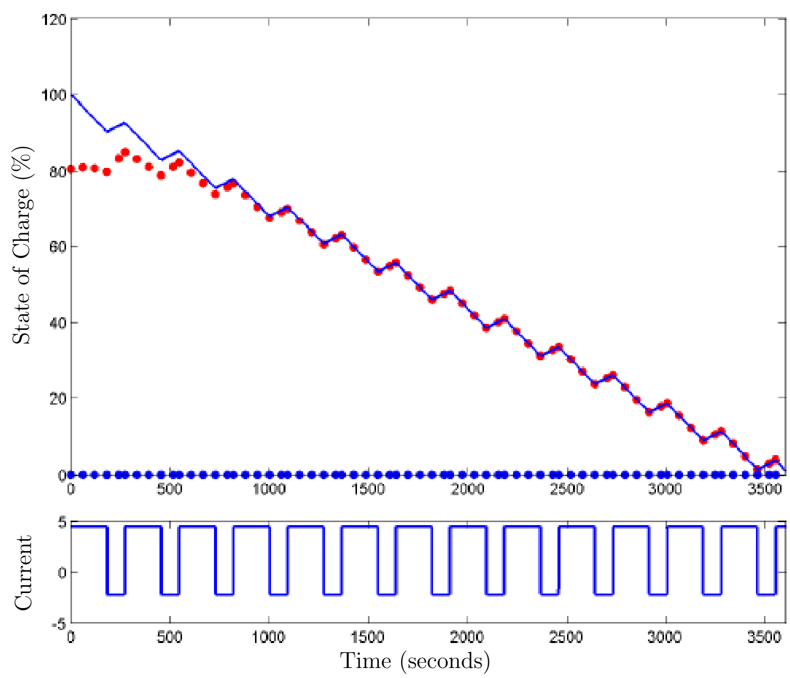

Figure 4.7. The time scale observer corrects for erroneous initial conditions.

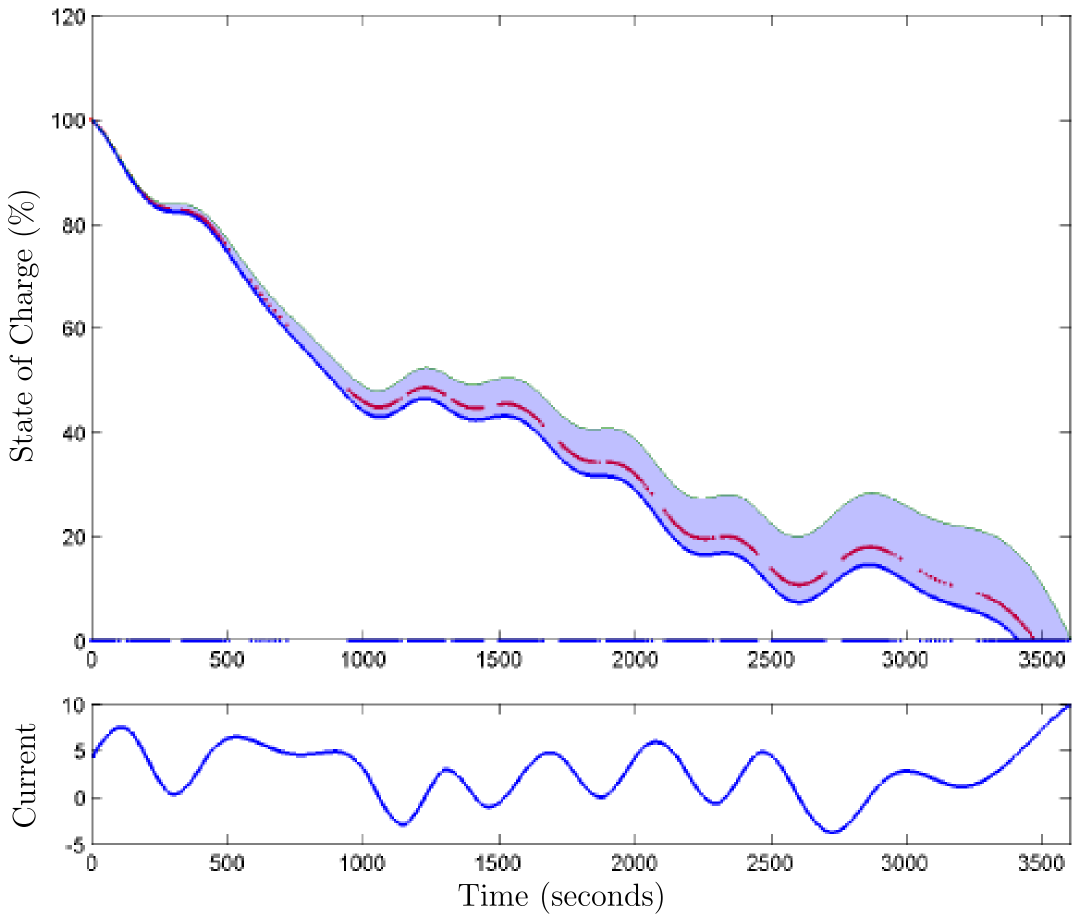

Figure 4.8: Top: The true SOC, in blue, from a random current draw. The top of the shaded area is the estimated SOC via Coulumb counting in the presence of measurement bias. The red dots show that the time scale observer can mitigate the effects of bias. Bottom: A random battery current waveform, normally distributed with mean 2.3 and variance 3.

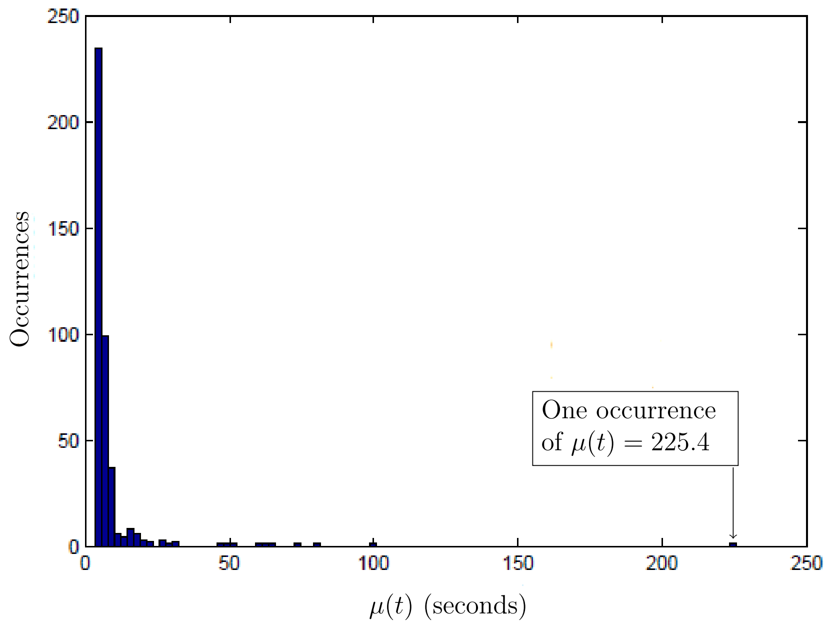

Figure 4.9: The histogram for the example in Figure 4.8 shows one occurrence of graininess significantly larger than $h_{\max}$, which had no adverse effect on observer stability or accuracy since it is statistically rare.

4.1.4 Future Work on the Model

This section proposes a novel design for a battery state-of-charge estimator based on the theory of dynamic equations on time scales. We show that a time scale-based state observer can accurately estimate the state of charge in the presence of initial condition uncertainty and measurement bias. Furthermore, by creating a time scale with graininess that ranges over two orders of magnitude, the observer can function with drastically fewer updates than would be required by traditional observer designs, thereby helping to lower the parasitic power costs of the observer itself.

Future work on this problem involves improving the model to include the actual (nonlinear) V_{OC} vs. SOC curve. Also, while the simulated examples in this section suggest that the design has promise, an embedded instantiation would pose additional challenges including the design of a low power circuit to implement (4.3), and the calculation of the expc$(\cdot)$ function in (2.17). Further investigation will be required to know the limits of this design, e.g. how long can it maintain a reasonably accurate SOC estimate before requiring a reset. Lastly, utilizing the stability criterion of (4.5) requires knowledge of the graininess statistical distribution, which depends on the battery current statistics. Battery current demand statistics would have to be determined experimentally in any given application.

One of the biggest problems of this observer design is that we cannot recover the state instantly for use in feedback control, for the observer depends on the graininess. The next section solves this problem.

4.2 Observer-Based Feedback Control

Beyond the efficient estimate of battery state of charge in batteries with flat voltage versus state of charge curves discussed in the previous section, a promising potential application of time scales theory discussed in the literature is in the arena of networked or distributed control systems where the sampling times are determined via network traffic [26].

In both of these applications, the sampling times form a time scale, and the sampling times are often random and nonuniformly spaced. These both create obstacles in the analysis of the design objective. We can work around the nonuniformity using time scales theory, but the randomness has been difficult to overcome. In fact, both of these applications have been studied with the assumption that the time scale, or at least some portion of it, was known *a priori*. Jackson *et al.* [26] assumed that the time scale was known in some future window and used this information to construct a time-varying stabilizing controller. In the previous section on the battery observer, we designed the observer gain matrix to vary with the graininess. This was only possible, however, with the assumption that the state estimate was not required in real time.

The theory developed in this dissertation allow us to leverage stochastic time scales to address some of these issues. We relax the condition that we know the time scale in some future window, instead assuming that we know the first and second order statistics of the distance between sampling points. This will allow us to build observers whose output can be used in state-feedback without explicit *a priori* knowledge of the step sizes.

We consider the state space model

$$\begin{aligned} \dot{x} &= Ax + Bu, \quad A \in \mathbb{R}^{n\times n}, \quad B \in \mathbb{R}^{n\times m}, \\ y &= Cx + Du, \quad C \in \mathbb{R}^{p\times n}, \quad D \in \mathbb{R}^{p\times m}. \end{aligned}$$

We discretize onto a time scale as in Section 2.3.2 to obtain

$$\begin{aligned} x^{\Delta}(t) &= \mathcal{A}(\mu(t))x(t) + \mathcal{B}(\mu(t))K\hat{x}(t), \\ y(t) &= Cx(t) + DK\hat{x}(t), \quad t \in \tilde{\mathbb{T}}. \end{aligned}$$

The question now is how to design the observer to acquire the state estimate $\hat{x}(t)$. Unlike the observer in the previous section, we cannot design the observer as a function of μ, because doing so would require explicit knowledge of the next time step, which is impossible since μ is a random variable. To overcome this, we propose the observer design

$$\begin{aligned} \hat{x}(t_{k+1}) &= \hat{x}(t_k) + \hat{A}\hat{x}(t_k) + \hat{B}u(t_k) + H(y(t_k) - \hat{y}(t_k)), \\ \hat{y}(t_k) &= C\hat{x}(t_k) + Du(t_k). \end{aligned} \tag{4.6}$$

Since we are using state feedback of the form $u(t_k) = K\hat{x}(t_k)$, (4.6) is equivalent to the stochastic time scale μ-varying dynamic equation

$$\begin{aligned} \hat{x}^{\Delta} &= \frac{1}{\mu}\hat{A}\hat{x} + \frac{1}{\mu}\hat{B}K\hat{x} + \frac{1}{\mu}HC(y - \hat{y}), \\ \hat{y} &= C\hat{x} + DK\hat{x}. \end{aligned}$$

To compute the error dynamics, let $\varepsilon := x - \hat{x}$. Then,

$$\begin{aligned}
\varepsilon^{\Delta} =& \mathcal{A}(\mu)x + \mathcal{B}(\mu)K\hat{x} - \frac{1}{\mu}\hat{A}\hat{x} - \frac{1}{\mu}\hat{B}K\hat{x} - \frac{1}{\mu}H(Cx + DK\hat{x} - (C\hat{x} + DK\hat{x})) \\
=& \mathcal{A}(\mu)x + \mathcal{B}(\mu)K(x-\varepsilon) - \frac{1}{\mu}\hat{A}(x-\varepsilon) - \frac{1}{\mu}\hat{B}K(x-\varepsilon) - \frac{1}{\mu}HC\varepsilon \\
=& \left[(\mathcal{A}(\mu) - \frac{1}{\mu}\hat{A}) + (\mathcal{B}(\mu) - \frac{1}{\mu}\hat{B})K\right]x + \left[\left(\frac{1}{\mu}\hat{B} - \mathcal{B}(\mu)\right)K + \frac{1}{\mu}\hat{A} - \frac{1}{\mu}HC\right]\varepsilon.
\end{aligned}$$

The last line shows how the error dynamics depend on both x and ε. Since we also have, from our state dynamics,

$$x^{\Delta} = \mathcal{A}(\mu)x + \mathcal{B}(\mu)K\hat{x} = [\mathcal{A}(\mu) + \mathcal{B}(\mu)K]x - \mathcal{B}(\mu)K\varepsilon,$$

together these yield the coupled system

$$\begin{pmatrix} x \\ \varepsilon \end{pmatrix}^{\Delta} = \mathcal{L}(\mu)\begin{pmatrix} x \\ \varepsilon \end{pmatrix}, \tag{4.7}$$

where

$$\mathcal{L}(\mu) = \begin{pmatrix} \mathcal{A}(\mu) + \mathcal{B}(\mu)K & -\mathcal{B}(\mu)K \\ (\mathcal{A}(\mu) - \frac{1}{\mu}\hat{A}) + (\mathcal{B}(\mu) - \frac{1}{\mu}\hat{B})K & \left(\frac{1}{\mu}\hat{B} - \mathcal{B}(\mu)\right)K + \frac{1}{\mu}(\hat{A} - HC) \end{pmatrix}.$$

In elementary control books [3, 24], we find similar dynamics, with the exception that the system matrix is block upper triangular, which implies that we can design the observer and controller separately. Although we cannot transform (4.7) into such a form, we can choose $\hat{A}$ and $\hat{B}$ such that the system is block upper triangular "on average," then estimate how close the system behaves to a block upper triangular system by examining covariances.

To illustrate this, if we take $\hat{A} := \mathrm{E}[\mu\mathcal{A}(\mu)]$ and $\hat{B} := \mathrm{E}[\mu\mathcal{B}(\mu)]$, then

$$\mathrm{E}[\mu\mathcal{L}(\mu)] = \begin{pmatrix} \hat{A} + \hat{B}K & -\hat{B}K \\ 0 & \hat{A} - HC \end{pmatrix} := L \tag{4.8}$$

is block upper triangular. A sufficient condition for the stability of (4.7) for the aforementioned choices of $\hat{A}$ and $\hat{B}$ is that $\mathcal{L}(\mu)$ satisfies the associated STSALE:

$$L^T P + PL + L^T PL + \mathrm{E}[(\mu\mathcal{L}(\mu) - L)^T P(\mu\mathcal{L}(\mu) - L)] = -M. \tag{4.9}$$

This leads to the following corollary to Theorem 3.5.

Corollary 4.1. *The coupled error and state dynamics* (4.7) *are mean-square exponentially stable provided there exist* H, K *such that* $\text{spec}(\hat{A}+\hat{B}K) \subset \mathcal{H}_1$*,* $\text{spec}(\hat{A}-HC) \subset \mathcal{H}_1$*, and*

$$\left\| \sum_{j=0}^{\infty} (L+I)^j((L+I)^T)^j \right\| \left\| \sum_{k=1}^{2n} \text{var}([\mu \mathcal{L}(\mu)]_k) \right\| = \alpha < 1. \tag{4.10}$$

Proof. By Theorem 3.5, there is a solution to (4.9) provided (4.10) holds and

$$\text{spec}(L) = \text{spec}(\hat{A} + \hat{B}K) \cup \text{spec}(\hat{A} - HC) \subset \mathcal{H}_1.$$

□

The conditions $\text{spec}(\hat{A} + \hat{B}K) \subset \mathcal{H}_1$ and $\text{spec}(\hat{A} - HC) \subset \mathcal{H}_1$ hold for appropriate choices of K and H provided $(\hat{A}, \hat{B})$ is controllable (stabilizable) and $(\hat{A}, C)$ is observable (detectable) [3, 24].

4.3 Optimal Control Theory

The primary drawback of Corollary 4.1 is that proper gain matrices H and K are difficult to find. In this section, we develop an optimal control theory which can help greatly in the search for suitable gain matrices.

We consider the stochastic time scale Linear μ-Varying (LμV) control system on a stochastic time scale $\tilde{\mathbb{T}}$ generated by μ

$$x^{\Delta} = \mathcal{A}(\mu)x + \mathcal{B}(\mu)u.$$

Our goal is to minimize

$$\text{E}[J(x,u)] := \text{E}\left[x_N^T Q^f x_N + \sum_{j=0}^{N-1} \mu_j x^T(t_j)Qx(t_j) + \mu_j u^T(t_j)Ru(t_j) \right]$$

where $R, Q > 0$ and t_j is the j^{th} point in the time scale. Note that we do not specify an ending time, but we do specify how many times the system updates.

We can view this problem as a discrete stochastic optimal control problem [5] with the following characteristics:

$$\text{Dynamics: } x_{j+1} = (I + \mu\mathcal{A}(\mu))x_j + \mu\mathcal{B}(\mu)u_j,$$

$$\text{Cost Rate: } c(x,u) = \mathrm{E}[\mu]x^TQx + \mathrm{E}[\mu]u^TRu,$$

We can approach this stochastic optimal control problem via Bellman's equation [5]. Bellman's equation is a necessary, but not sufficient, condition for optimality. The equation arises from a stochastic generalization of Bellman's optimality principle, which states that any optimal trajectory must be also be optimal over its sub-trajectories. The stochastic version of Bellman's equation seeks a scalar-valued function $v(x)$ called the "value function" or the "cost–to–go function." This function measures the expected cost of the optimal policy starting from a state x. The value function is unique if it exists and satisfies the equation

$$v(x) = \min_{u\in\mathcal{U}(x)}\{c(x,u) + \mathrm{E}[v(((I + \mu\mathcal{A}(\mu))x + \mu\mathcal{B}(\mu)u)]\}.$$

The optimal policy $u^*(x)$ then satisfies

$$u^*(x) = \arg\min_{u\in\mathcal{U}(x)}\{c(x,u) + \mathrm{E}[v(((I + \mu\mathcal{A}(\mu))x + \mu\mathcal{B}(\mu)u)]\}.$$

Although we do not know the form of $v(x)$, we use the ansatz

$$v(x,j) = x^TV_jx, \quad V_j = V_j^T > 0$$

with the boundary condition

$$V_N = Q^f.$$

The Bellman equation now reads

$$\begin{aligned} x^TV_jx = &\min_u\{\mathrm{E}[\mu]x^TQx + \mathrm{E}[\mu]u^TRu \\ &+ \mathrm{E}[((I + \mu\mathcal{A}(\mu))x + \mu\mathcal{B}(\mu)u)^TV_{j+1}((I + \mu\mathcal{A}(\mu))x + \mu\mathcal{B}(\mu)u)\}. \end{aligned}$$

The right-hand side of the above equation is quadratic in u, and hence we can use basic matrix analysis to minimize the quantity in the bracket. Taking the derivative and setting it equal to zero yields

$$2\mathrm{E}[\mu]Ru + 2\mathrm{E}[(\mu\mathcal{B}(\mu))^T V_{j+1}(I+\mu\mathcal{A}(\mu))]x + 2\mathrm{E}[(\mu\mathcal{B}(\mu))^T V_{j+1}(\mu\mathcal{B}(\mu))]u = 0.$$

Thus, the quantity u which minimizes the RHS is given by

$$u = -(\mathrm{E}[\mu]R + \mathrm{E}[(\mu\mathcal{B}(\mu))^T V_{j+1}(\mu\mathcal{B}(\mu))])^{-1}\mathrm{E}[(\mu\mathcal{B}(\mu))^T V_{j+1}(I+\mu\mathcal{A}(\mu))]x$$

since the Hessian of the quantity in the minimization is a positive definite matrix.

Using this value of u as the achieved minimum of the RHS of the Bellman equation, we arrive at the matrix equation which we call the *μ-varying stochastic time scale dynamic Riccati equation*:

$$\begin{aligned} V_j =& \mathrm{E}[\mu]Q + \mathrm{E}[(I+\mu\mathcal{A}(\mu))^T V_{j+1}(I+\mu\mathcal{A}(\mu))] \\ & -\mathrm{E}[(I+\mu\mathcal{A}(\mu))^T V_{j+1}(\mu\mathcal{B}(\mu))] \\ & \times(\mathrm{E}[\mu]R + \mathrm{E}[(\mu\mathcal{B}(\mu))^T V_{j+1}(\mu\mathcal{B}(\mu))])^{-1}\mathrm{E}[(\mu\mathcal{B}(\mu))^T V_{j+1}(I+\mu\mathcal{A}(\mu))]; \quad V_N = Q^f \end{aligned} \tag{μ-STSDRE}$$

In the LTI case, the same derivation applies, and we arrive at the matrix equation which we call the *stochastic time scale dynamic Riccati equation*:

$$\begin{aligned} -\frac{V_{j+1}-V_j}{\mathrm{E}[\mu]} =& Q + A^T V_{j+1} + V_{j+1}A + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A^T V_{j+1} A \\ & - \left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A\right)^T V_{j+1}B \\ & \times \left(R + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}B^T V_{j+1} B\right)^{-1} B^T V_{j+1}\left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]}A\right); \quad V_N = Q^f. \end{aligned} \tag{STSDRE}$$

The forms of (μ-STSDRE) and (STSDRE) generalize and extend (CDRE) and (DDRE). In a similar approach to the the continuous and discrete cases, we will use the solution of (μ-STSDRE) to solve the quadratic cost, infinite horizon control problem on stochastic time scales.

4.4 Infinite Horizon

Bellman's equation is a necessary but not sufficient condition for optimality. Therefore, the previous work does not necessarily yield an optimal control law. Moreover, our primary interest in this dissertation is control problems with infinite time horizons. The following theorem fixes these two issues, using the solution of the finite-horizon problem to guarantee an optimal control for the infinite horizon case.

Theorem 4.1. *Suppose that there exists $K \in \mathbb{R}^{m\times n}$ such that $x^\Delta = (\mathcal{A}(\mu) + \mathcal{B}(\mu)K)$ is mean-square asymptotically stable. Also suppose $Q > 0$ and $R \geq 0$. Define $V_k(N)$ to be the solution of* (μ-STSDRE) *with $Q^f = 0$ with a horizon of N at $j = k$. Then*

(i) $\lim_{N\to\infty} V_0(N) = V$ *exists.*

(ii) V is a positive definite solution of the μ-varying stochastic time scale algebraic Riccati equation:

$$\begin{aligned} V =& \mathrm{E}[\mu]Q + \mathrm{E}[(I+\mu\mathcal{A}(\mu))^T V (I+\mu\mathcal{A}(\mu))] \\ &- \mathrm{E}[(I+\mu\mathcal{A}(\mu))^T V(\mu\mathcal{B}(\mu))] \\ &\times (\mathrm{E}[\mu]R + \mathrm{E}[(\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))])^{-1}\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I+\mu\mathcal{A}(\mu))]. \end{aligned} \tag{μ-STSARE}$$

(iii) $x^\Delta = [\mathcal{A}(\mu) + \mathcal{B}(\mu)K]x$, *with*

$$K = (\mathrm{E}[\mu]R + \mathrm{E}[(\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))])^{-1}\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I+\mu\mathcal{A}(\mu))],$$

is mean-square exponentially stable.

(iv) The optimal control law for $x^\Delta = \mathcal{A}(\mu)x + \mathcal{B}(\mu)u$ with cost functional

$$\mathrm{E}[J_\infty(x_0,u)] = \mathrm{E}\left[\sum_{j=0}^{\infty} \mu_j x^T(t_j)Qx(t_j) + \mu_j u^T(t_j)Ru(t_j)\right]$$

is given by $u^*(t) = Kx(t)$ with expected costs $\mathrm{E}[J_\infty(x_0,u^*)] = x_0^T V x_0$.

Proof. We first show $\inf_u \mathrm{E}[J_\infty(x_0,u)] < \infty$. Since there exists $K \in \mathbb{R}^{m\times n}$ such that $x^\Delta = (\mathcal{A}(\mu)+\mathcal{B}(\mu)K)x$ is mean-square asymptotically stable, it is mean-square exponentially stable by Theorem 3.2. Therefore, for $u = Kx$,

$$\begin{aligned}
\mathrm{E}[J_\infty(x_0,u)] &= \mathrm{E}\left[\sum_{j=0}^{\infty} \mu_j x^T(t_j)(Q+K^TRK)x(t_j)\right] \\
&\le \lambda_{\max}(Q+K^TRK)\sum_{j=0}^{\infty}\mathrm{E}[\mu_j x^T(t_j)x(t_j)] \\
&= \mathrm{E}[\mu]\lambda_{\max}(Q+K^TRK)\sum_{j=0}^{\infty}\mathrm{E}[x^T(t_j)x(t_j)] \\
&< \infty,
\end{aligned}$$

where the last inequality is by the definition of mean square exponential stability. Note we interchange expectation and summation via Tonelli's Theorem and used the independence of μ_j and $x^T(t_j)x(t_j)$.

We now show $\lim_{N\to\infty} V_0(N)$ exists. Let $\varepsilon > 0$ and choose u^* such that $\mathrm{E}[J_\infty(x_0,u^*)] = \inf_u \mathrm{E}[J_\infty(x_0,u)] + \varepsilon$. Let $J_N(x_0,u)$ be the cost functional for the finite horizon problem with $Q^f = 0$ and horizon N. Then

$$\begin{aligned}
x_0^T V_0(N) x_0 &= \min_u \mathrm{E}[J_N(x_0,u)] \\
&\le \sum_{j=0}^{N-1}\mathrm{E}[\mu_j x^T(t_j)Qx(t_j) + \mu_j u^{*T}(t_j)Ru^*(t_j)] \\
&\le \sum_{j=0}^{\infty}\mathrm{E}[\mu_j x^T(t_j)Qx(t_j) + \mu_j u^{*T}(t_j)Ru^*(t_j)] \\
&= \inf_u \mathrm{E}[J_\infty(x_0,u)] + \varepsilon.
\end{aligned}$$

Therefore $\{V_0(N)\}$ has an upper bound. Now choose u^*_{N+1} such that

$$\mathrm{E}[J_{N+1}(x_0,u^*_{N+1})] = \min_u \mathrm{E}[J_{N+1}(x_0,u)].$$

Then

$$x_0^T V_0(N) x_0 = \min_u \mathrm{E}[J_N(x_0,u)]$$

$$\leq \sum_{j=0}^{N-1} \mathrm{E}[\mu_j x^T(t_j) Q x(t_j) + \mu_j u_{N+1}^{*T}(t_j) R u_{N+1}^*(t_j)]$$

$$\leq \sum_{j=0}^{N} \mathrm{E}[\mu_j x^T(t_j) Q x(t_j) + \mu_j u_{N+1}^{*T}(t_j) R u_{N+1}^*(t_j)]$$

$$= \min_u \mathrm{E}[J_{N+1}(x_0, u)]$$

$$= x_0^T V_0(N+1) x_0.$$

Since this holds for any choice of x_0, we see $\{V_0(N)\}$ is an increasing sequence with an upper bound. Therefore, $\lim_{N\to\infty} V_0(N) := V$ exists. This proves (i).

Note that since $\{V_0(N)\}$ is an increasing sequence of positive definite matrices, $V > 0$. To see V satisfies (μ-STSARE), note that $V_0(N) = V_1(N+1)$. This is because the minimal expected cost over N updates starting from $x(t_0) = \chi$ is given by $\chi^T V_0(N)\chi$, and the minimal expected cost from the second state of a system $x(t_1) = \chi$ over the next N updates is given by $\chi^T V_1(N+1)\chi$. But because $V_0(N)$ satisfies the μSTSDRE in the horizon N case, taking the limit as $N \to \infty$ of the μSTSDRE shows V satisfies μSTSARE. This proves (ii).

Consider $K = (\mathrm{E}[\mu]R + \mathrm{E}[(\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))])^{-1}\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I + \mu\mathcal{A}(\mu))]]$. Define $I + \mu\mathcal{A}(\mu) + \mu\mathcal{B}(\mu)K := \mathcal{C}(\mu)$. Then

$$\begin{aligned}
V =& \mathrm{E}[\mu Q + (I + \mu\mathcal{A}(\mu))^T V(I + \mu\mathcal{A}(\mu))] + \mathrm{E}[(I + \mu\mathcal{A}(\mu))^T V(\mu\mathcal{B}(\mu))]K \\
=& \mathrm{E}[\mu Q + (I + \mu\mathcal{A}(\mu) + \mu\mathcal{B}(\mu)K)^T V(I + \mu\mathcal{A}(\mu) + \mu\mathcal{B}(\mu)K)] \\
&+ \mathrm{E}[(I + \mu\mathcal{A}(\mu))^T V(\mu\mathcal{B}(\mu))]K - K^T\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I + \mu\mathcal{A}(\mu))] \\
&- \mathrm{E}[(I + \mu\mathcal{A}(\mu))^T V(\mu\mathcal{B}(\mu))]K - K^T\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))]K \\
=& \mathrm{E}[\mu Q + \mathcal{C}^T(\mu)V\mathcal{C}(\mu)] - K^T[\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I + \mu\mathcal{A}(\mu)) + (\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))K]] \\
=& \mathrm{E}[\mu Q + \mathcal{C}^T(\mu)V\mathcal{C}(\mu)] \\
&- K^T[\mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I + \mu\mathcal{A}(\mu))] + \mathrm{E}[\mu R + (\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))]K - \mathrm{E}[\mu]RK] \\
=& \mathrm{E}[\mu Q + \mathcal{C}^T(\mu)V\mathcal{C}(\mu)] + \mathrm{E}[\mu]K^T RK
\end{aligned}$$

$$=\mathrm{E}[(I+\mu\mathcal{A}(\mu)+\mu\mathcal{B}(\mu)K)^TV(I+\mu\mathcal{A}(\mu)+\mu\mathcal{B}(\mu)K)]+\mathrm{E}[\mu](Q+K^TRK).$$

Hence

$$\begin{aligned}&\mathrm{E}[(\mu\mathcal{A}(\mu)+\mu\mathcal{B}(\mu)K)^TV+V(\mu\mathcal{A}(\mu)+\mu\mathcal{B}(\mu)K)\\&+(\mu\mathcal{A}(\mu)+\mu\mathcal{B}(\mu)K)^TV(\mu\mathcal{A}(\mu)+\mu\mathcal{B}(\mu)K)]\\=&-\mathrm{E}[\mu](Q+K^TRK).\end{aligned}$$

Now, $V>0$ and $\mathrm{E}[\mu](Q+K^TRK)>0$, so we see that $x^\Delta=(\mathcal{A}(\mu)+\mathcal{B}(\mu)K)x$ is mean-square exponentially stable. This proves (iii).

To show (iv), first note that because

$$0\leq\inf_u J_N(x_0,u)=x_0^TV_0(N)x_0\leq\inf_u J_\infty(x_0,u)+\varepsilon$$

for all $\varepsilon>0$, we have

$$x_0^TVx_0\leq J_\infty(x_0,u)+\varepsilon \text{ for all } \varepsilon>0.$$

We will now show $\inf_u J_\infty(x_0,u)\leq x_0^TVx_0$. To see this, first note for choices of u which are mean-square exponentially stable (of which there is at least one),

$$\mathrm{E}[J_\infty(x_0,u)]+\mathrm{E}\left[\sum_{j=0}^\infty x^T(t_{j+1})Vx(t_{j+1})-x^T(t_j)Vx(t_j)\right]=\mathrm{E}[J_\infty(x_0,u)]-x_0^TVx_0.$$

On the other hand, using completion of squares and the definition of K,

$$\begin{aligned}&\mathrm{E}[J_\infty(x_0,u)]+\mathrm{E}\left[\sum_{j=0}^\infty x^T(t_{j+1})Vx(t_{j+1})-x^T(t_j)Vx(t_j)\right]\\=&\sum_{j=0}^\infty\mathrm{E}[\mu x^TQx+\mu u^TRu+(x^T(I+\mu\mathcal{A}(\mu))^T\\&+u^T(\mu\mathcal{B}(\mu))^T)V((I+\mu\mathcal{A}(\mu))x+(\mu\mathcal{B}(\mu))u)-x^TVx]\\=&\sum_{j=0}^\infty\mathrm{E}[\mu x^TQx+\mu u^TRu\\&+x^T(I+\mu\mathcal{A}(\mu))^TV(I+\mu\mathcal{A}(\mu))x+u^T(\mu\mathcal{B}(\mu))^TV(I+\mu\mathcal{A}(\mu))x\end{aligned}$$

$$
\begin{aligned}
&+ x^T(I+\mu\mathcal{A}(\mu))^T V(\mu\mathcal{B}(\mu))u + u^T(\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))u] \\
=& \sum_{j=0}^{\infty} \mathrm{E}[(u-Kx)^T \mathrm{E}[\mu R + ((\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))](u-Kx)] \\
&+ \mathrm{E}[x^T(\mathrm{E}[\mu Q - V + (I+\mu\mathcal{A}(\mu))^T V(I+\mu\mathcal{A}(\mu))] \\
&- \mathrm{E}[(I+\mu\mathcal{A}(\mu))^T V(\mu\mathcal{B}(\mu))]\mathrm{E}[\mu R + (\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))]^{-1} \\
&\times \mathrm{E}[(\mu\mathcal{B}(\mu))^T V(I+\mu\mathcal{A}(\mu))]x] \\
=& \sum_{j=0}^{\infty} \mathrm{E}[(u-Kx)^T \mathrm{E}[\mu R + ((\mu\mathcal{B}(\mu))^T V(\mu\mathcal{B}(\mu))](u-Kx)].
\end{aligned}
$$

Hence,

$$\mathrm{E}[J_\infty(x_0, Kx)] = x_0^T V x_0,$$

so

$$\inf_u \mathrm{E}[J_\infty(x_0, u)] \le x_0^T V x_0.$$

Therefore,

$$\inf_u \mathrm{E}[J_\infty(x_0, u)] \le x_0^T V x_0 \le \inf_u \mathrm{E}[J_\infty(x_0, u)] + \varepsilon \text{ for all } \varepsilon > 0,$$

and thus

$$\inf_u \mathrm{E}[J_\infty(x_0, u)] = x_0^T V x_0.$$

Since the minimal cost is achieved for the choice $u = Kx$, we see $u = Kx$ is the optimal control. This proves (iv). □

In the LTI case, we arrive at the following corollary.

Corollary 4.2. *Suppose that the pair (A, B) is controllable. Also suppose $Q > 0$ and $R \ge 0$. Define $V_k(N)$ to be the solution of* (STSDRE) *with $Q^f = 0$ at a horizon of N at $j = k$. Then*

(i) $\lim_{N\to\infty} V_0(N) = V$ *exists.*

(ii) V is a positive definite solution of the stochastic time scale algebraic Riccati equation:

$$\begin{aligned} 0 =& Q + A^T V + VA + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} A^T V A \\ &- \left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} A\right)^T VB \left(R + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} B^T V B\right)^{-1} B^T V \left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} A\right). \end{aligned} \tag{STSARE}$$

(iii) The system

$$x^\Delta = \left[A - B\left(R + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} B^T V B\right)^{-1} B^T V \left(I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} A\right)\right] x := [A + BK]x$$

is mean-square exponentially stable.

(iv) The optimal control law for $x^\Delta = Ax + Bu$ with cost functional

$$\mathrm{E}[J_\infty(x_0, u)] = E\left[\sum_{j=0}^{\infty} \mu_j x^T(t_j) Q x(t_j) + \mu_j u^T(t_j) R u(t_j)\right]$$

is given by $u^*(t) = Kx(t)$ with expected costs $\mathrm{E}[J_\infty(x_0, u^*)] = x_0^T V x_0$.

We note that (STSARE) can be transformed into (DARE) by multiplying (STSARE) by $\mathrm{E}[\mu^2]/\mathrm{E}[\mu]$ and making the transformations

$$\tilde{Q} = \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} Q, \quad \tilde{A} = (I + \frac{\mathrm{E}[\mu^2]}{\mathrm{E}[\mu]} A) \quad \tilde{R} = \frac{\mathrm{E}[\mu]}{\mathrm{E}[\mu^2]} R \quad \tilde{B} = B.$$

This can be useful when using a solver in a computer algebra system. Just as we cannot transform (μ-STSALE) into the form of (DALE), we cannot transform (μ-STSARE) into the form of (DARE). We can, however, write (μ-STSARE) in the form of a *linearly perturbed Riccati equation* [19, 40, 2] of the form

$$\begin{aligned} V =& Q + (I + A)^T V (I + A) + \Pi_1(V) \\ &+ [(I + A)^T VB + \Pi_{12}(V)][R + B^T V B + \Pi_2(V)]^{-1}[(I + A)^T VB + \Pi_{12}(V)]^T, \end{aligned}$$

where

$$\Pi(V) := \begin{pmatrix} \Pi_1(V) & \Pi_{12}(V) \\ \Pi_{12}^T(V) & \Pi_2(V) \end{pmatrix}$$

is a positive linear operator. Linearly perturbed Riccati equations first appeared in the stochastic control literature in 1968 [40], and they continue to be an area of research within the field [19, 2]. Our analysis gives a novel and natural example of how an equation of this form emerges. In future work, we will explore leveraging the theory of linearly perturbed Riccati equations to study (μ-STSARE) further.

Finally, we remark that if we use the Bellman equation approach in the LTI case when the time scale is *not* stochastic, we arrive at the dynamic equation

$$\begin{aligned} -V^{\Delta} =& Q + A^T V^{\sigma} + V^{\sigma} A + \mu(t) A^T V^{\sigma} A \\ & - (I + \mu(t)A)^T V^{\sigma} B (R + \mu(t) B^T V^{\sigma} B)^{-1} B^T V^{\sigma} (I + \mu(t)A), \end{aligned}$$

which matches the result of Wintz [8]. He did not, however, demonstrate stability with Lyapunov techniques or obtain a steady-state solution in the infinite horizon case. A major contribution of this work and the stochastic time scale approach is that we can accomplish both of these goals through Theorem 4.1 and Corollary 4.2.

CHAPTER FIVE

Insights From Stochastic Time Scales

The study of stochastic time scales informs us about theorems about deterministic time scales. In this chapter, we explore these connections.

5.1 Uniform Exponential Stability

The theory of stochastic time scales not only explores new classes of probabilistically generated time scales, it also sheds light on general, deterministic time scales. In particular, in this section, we show that the relation between $\mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}$ and $\tilde{\mathcal{S}}$ holds for general time scales. Define

$$\delta(\mathbb{T}) := \limsup_{T\to\infty} \frac{\int_{t_0}^{T} \mu(t)\Delta t}{T - t_0}. \tag{5.1}$$

This turns out to be an important constant for understanding how stochastic time scales inform us about general time scales. We will show that $\mathcal{H}_{\delta(\mathbb{T})}$ is always a subset of the region of exponential stability. Moreover, we will show the geometric importance of this region by showing $\mathcal{H}_{\delta(\mathbb{T})}$ is the best circular approximation of $\mathcal{S}$ at the origin in the sense that osculating circle at the origin of the region of exponential stability. Next, we will demonstrate the power and ease of computing $\mathcal{H}_{\delta(\mathbb{T})}$ versus computing the region of exponential stability. Finally, we will focus on uniform exponential stability and show that $\mathcal{H}_{\delta(\mathbb{T})}$ is the osculating circle to the region of uniform exponential stability under a mild condition, which we call *mean-stationary*, on the time scale.

5.1.1 Properties of $\mathcal{H}_{\delta(\mathbb{T})}$

We will begin by showing $\mathcal{H}_{\delta(\mathbb{T})}$ is always contained in $\mathcal{S}(\mathbb{T})$. We begin with a lemma which helps us determine when a complex number is in the disc $\mathcal{H}_{\delta(\mathbb{T})}$.

Lemma 5.1. *Let $\gamma > 0$. $\lambda \in \mathcal{H}_\gamma$ if and only if*

$$2\,\mathrm{Re}(\lambda) + |\lambda|^2\gamma < 0.$$

Proof. An equivalent form of (2.1) is $\mathcal{H}_\gamma = \{\lambda \in \mathbb{C} : |\lambda + 1/\gamma|^2 < (1/\gamma)^2\}$. Expanding the modulus and letting $\lambda = x + iy$, we have $(x + 1/\gamma)^2 + y^2 < (1/\gamma)^2$. This is equivalent to $x^2 + y^2 + 2x/\gamma = |\lambda|^2 + 2\,\mathrm{Re}(\lambda)/\gamma < 0$, which holds if and only if $2\,\mathrm{Re}(\lambda) + |\lambda|^2\gamma < 0$. □

From this point onward, we will assume the time scale is known and write $\delta := \delta(\mathbb{T})$ and $\mathcal{S}(\mathbb{T}) := \mathcal{S}$.

Theorem 5.1. *$\mathcal{H}_\delta \subset \mathcal{S}$.*

Proof. Let $\lambda \in \mathcal{H}_\delta$. Then $2\,\mathrm{Re}(\lambda) + |\lambda|^2\delta < 0$. Let $\lambda = x + iy$ and recall $\ln(1 + x) \leq x$ for all $x > -1$. Then

$$\begin{aligned}
&2 \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \lim_{s \downarrow \mu(t)} \frac{\ln|1 + \lambda s|}{s} \Delta t \\
=& \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \lim_{s \downarrow \mu(t)} \frac{\ln|1 + \lambda s|^2}{s} \Delta t \\
=& \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \lim_{s \downarrow \mu(t)} \frac{\ln[(1 + sx)^2 + (sy)^2]}{s} \Delta t \\
=& \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \lim_{s \downarrow \mu(t)} \frac{\ln[1 + 2sx + (sx)^2 + (sy)^2]}{s} \Delta t \\
\leq& \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \lim_{s \downarrow \mu(t)} \frac{2sx + (sx)^2 + (sy)^2}{s} \Delta t \\
=& \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \left[2x + \mu(t)(x^2 + y^2)\right] \Delta t \\
=& 2\,\mathrm{Re}(\lambda) + |\lambda|^2 \limsup_{T\to\infty} \frac{1}{T - t_0} \int_{t_0}^{T} \mu(t) \Delta t \\
=& 2\,\mathrm{Re}(\lambda) + |\lambda|^2\delta \\
<& 0.
\end{aligned}$$

□

Since $\mathcal{H}_\delta \subset \mathcal{S}$, we know $\mathcal{H}_\delta$ is a circular approximation to the region at the origin. We claim $\mathcal{H}_\delta$ is the best circular approximation to $\mathcal{S}$ at the origin, i.e. $\mathcal{H}_\delta$ is the osculating circle to $\mathcal{S}$ at the origin. The proof of that fact requires several inequalities which are themselves interesting. We present these inequalities next in a sequence of three lemmas.

Lemma 5.2. *Let* $0 \le \mu_{\min} \le \mu \le \mu_{\max}$. *Let* $\mu_{\min} < \gamma < \mu_{\max}$. *Then for* $\lambda \in \partial\mathcal{H}_\gamma$,

$$1 - \left(\frac{\mathrm{Re}(\lambda)}{|\lambda|}\right)^2 \le |1 + \lambda\mu|^2 \le |1 + \lambda\mu_{\max}|^2.$$

Proof. The quadratic function in μ given by

$$|1 + \lambda\mu|^2 = 1 + 2\,\mathrm{Re}(\lambda)\mu + |\lambda|^2\mu^2 \tag{5.2}$$

is minimized at $\mu = -\mathrm{Re}(\lambda)/|\lambda|^2$ and attains the minimum value

$$m(\lambda) := 1 - (\mathrm{Re}(\lambda)/|\lambda|)^2. \tag{5.3}$$

The function (5.2) is maximized at either $\mu = \mu_{\max}$ or $\mu = \mu_{\min}$. Since $\lambda \in \mathcal{H}_{\mu_{\min}}$ and $\lambda \notin \mathcal{H}_{\mu_{\max}}$, $|1 + \lambda\mu_{\min}|^2 < 1$ while $|1 + \lambda\mu_{\max}|^2 > 1$. Therefore, (5.2) attains the maximum value

$$M(\lambda) := |1 + \lambda\mu_{\max}|^2. \tag{5.4}$$

□

The proof of our main result relies on the behavior of a linear function in the variable $\mathrm{Re}(\lambda)$. The following lemma helps us write all expressions involving λ in terms of $\mathrm{Re}(\lambda)$.

Lemma 5.3. *If* $\lambda \in \partial\mathcal{H}_\gamma$, *then*

$$|\lambda|^2 = -\frac{2}{\gamma}\,\mathrm{Re}(\lambda). \tag{5.5}$$

Proof. The result follows from noticing the similarity of the two triangles in Figure 5.1. The largest triangle is a right triangle by Thales' theorem. □

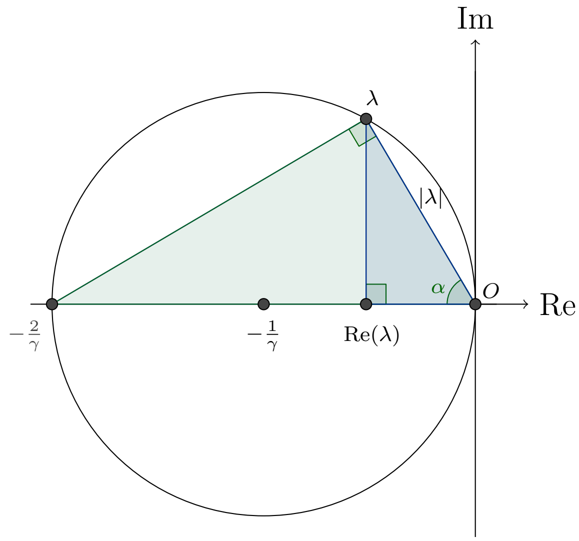

Figure 5.1. The relation between $|\lambda|^2$ and $\mathrm{Re}(\lambda)$ when $\lambda \in \mathcal{H}_\gamma$.

Finally, we require the following equality.

Lemma 5.4. *Let* $0 \leq \mu_{\min} < \gamma < \mu_{\max}$ *and let* $\lambda \in \partial\mathcal{H}_\gamma$. *Then*

$$m(\lambda) - M(\lambda) = \frac{(2\mu_{\max} - \gamma)^2}{2\gamma}\,\mathrm{Re}(\lambda).$$

Proof.

$$\begin{aligned}
m(\lambda) - M(\lambda) &= \left[1 - \left(\frac{\mathrm{Re}(\lambda)}{|\lambda|}\right)^2\right] - \left(1 + 2\,\mathrm{Re}(\lambda)\mu_{\max} + |\lambda|^2\mu_{\max}^2\right) \\
&= -\frac{\mathrm{Re}(\lambda)^2}{|\lambda|^2} - 2\,\mathrm{Re}(\lambda)\mu_{\max} - |\lambda|^2\mu_{\max} \\
&= \frac{\mathrm{Re}(\lambda)^2}{\frac{2}{\gamma}\,\mathrm{Re}(\lambda)} - 2\,\mathrm{Re}(\lambda)\mu_{\max} + \frac{2}{\gamma}\,\mathrm{Re}(\lambda)\mu_{\max} \\
&= \mathrm{Re}(\lambda)\left(\frac{\gamma}{2} - 2\mu_{\max} + \frac{2}{\gamma}\mu_{\max}^2\right) \\
&= \mathrm{Re}(\lambda)\left(\frac{\gamma^2 - 4\gamma\mu_{\max} + 4\mu_{\max}^2}{2\gamma}\right) \\
&= \mathrm{Re}(\lambda)\frac{(2\mu_{\max} - \gamma)^2}{2\gamma}
\end{aligned}$$

□

With these lemmas in hand, we are prepared to classify the osculating circle to the region of stability $\mathcal{S}$.

Theorem 5.2. *$\mathcal{H}_\delta$ is the osculating circle of $\mathcal{S}$ at the origin.*

Proof. Let $\mu_{\min} < \gamma < \delta \le \mu_{\max}$ and $\lambda \in (\partial\mathcal{H}_\gamma - \{(0,0)\}) \cap \mathcal{R}_\varepsilon$, where $\mathcal{R}_\varepsilon$ is the open square with side length 2ε centered at the origin, and

$$0 < \varepsilon < \frac{\left(1-\frac{\delta}{\gamma}\right)}{\left(1-\frac{\mu_{\max}}{\gamma}\right)\left[2\mu_{\max}\left(1-\frac{\delta}{\gamma}\right)+\frac{(2\mu_{\max}-\gamma)^2}{2\gamma}\right]}. \tag{5.6}$$

This implies $\operatorname{Re}(\lambda) \in (-\varepsilon, 0)$. We will argue that the the integral which defines $\mathcal{S}_{\mathbb{C}}$ when evaluated at λ is positive. This will show that $\mathcal{H}_\gamma$ cannot be the osculating circle to $\mathcal{S}$ at the origin.

In the following inequalities, we make use of Lemma 2. Note that

$$\begin{aligned}
&\int_{t_0}^{T} \frac{\ln|1+\lambda\mu(t)|}{\mu(t)}\Delta t \\
\ge&\int_{t_0}^{T} \frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{|1+\lambda\mu(t)|^2}\Delta t \\
=&\int\limits_{\substack{\mu(t)<\gamma\\ t_0\le t<T}} \overbrace{\frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{|1+\lambda\mu(t)|^2}}^{-}\Delta t + \int\limits_{\substack{\mu(t)>\gamma\\ t_0\le t<T}} \overbrace{\frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{|1+\lambda\mu(t)|^2}}^{+}\Delta t \\
\ge&\int\limits_{\substack{\mu(t)<\gamma\\ t_0\le t<T}} \frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{m(\lambda)}\Delta t + \int\limits_{\substack{\mu(t)>\gamma\\ t_0\le t<T}} \frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{M(\lambda)}\Delta t \\
=&\int_{t_0}^{T} \frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{m(\lambda)}\Delta t + \int\limits_{\substack{\mu(t)>\gamma\\ t_0\le t<T}} \frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{M(\lambda)} - \frac{2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)}{m(\lambda)}\Delta t \\
=&\frac{M(\lambda)}{m(\lambda)M(\lambda)}\int_{t_0}^{T} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)\Delta t + \overbrace{\frac{m(\lambda)-M(\lambda)}{m(\lambda)M(\lambda)}\int\limits_{\substack{\mu(t)>\gamma\\ t_0\le t<T}} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)\Delta t}^{-} \\
\ge&\frac{M(\lambda)}{m(\lambda)M(\lambda)}\int_{t_0}^{T} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)\Delta t + \frac{m(\lambda)-M(\lambda)}{m(\lambda)M(\lambda)}\int\limits_{\substack{\mu(t)>\gamma\\ t_0\le t<T}} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max}\Delta t \\
\ge&\frac{M(\lambda)}{m(\lambda)M(\lambda)}\int_{t_0}^{T} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)\Delta t + \frac{m(\lambda)-M(\lambda)}{m(\lambda)M(\lambda)}\int_{t_0}^{T} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max}\Delta t \\
=&\frac{M(\lambda)\int_{t_0}^{T} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu(t)\Delta t + (m(\lambda)-M(\lambda))\int_{t_0}^{T} 2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max}\Delta t}{m(\lambda)M(\lambda)}.
\end{aligned}$$

Dividing by $T - t_0$ and applying the limsup to both sides, we obtain,

$$\begin{aligned}&\limsup_{T\to\infty} \frac{\int_{t_0}^{T} \frac{\ln|1+\lambda\mu(t)|^2}{\mu(t)}\Delta t}{T-t_0}\\ \geq&\frac{M(\lambda)(2\operatorname{Re}(\lambda)+|\lambda|^2\delta)+(m(\lambda)-M(\lambda))(2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max})}{m(\lambda)M(\lambda)},\end{aligned} \tag{5.7}$$

which we will show is positive for all prescribed λ. Since $m(\lambda)M(\lambda) > 0$, (5.7) is postive if and only if

$$M(\lambda)(2\operatorname{Re}(\lambda)+|\lambda|^2\delta)+(m(\lambda)-M(\lambda))(2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max}) > 0.$$

Using Lemmas 5.3 and 5.4 as well as (5.3) and (5.4),

$$\begin{aligned}&M(\lambda)(2\operatorname{Re}(\lambda)+|\lambda|^2\delta)+(m(\lambda)-M(\lambda))(2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max})\\ =&(1+2\operatorname{Re}(\lambda)\mu_{\max}+|\lambda|^2\mu_{\max}^2)(2\operatorname{Re}(\lambda)+|\lambda|^2\delta)\\ &+\operatorname{Re}(\lambda)\frac{(2\mu_{\max}-\gamma)^2}{2\gamma}(2\operatorname{Re}(\lambda)+|\lambda|^2\mu_{\max})\\ =&\left(1+2\operatorname{Re}(\lambda)\mu_{\max}-\frac{2}{\gamma}\operatorname{Re}(\lambda)\mu_{\max}^2\right)\left(2\operatorname{Re}(\lambda)-\frac{2}{\gamma}\operatorname{Re}(\lambda)\delta\right)\\ &+\operatorname{Re}(\lambda)\frac{(2\mu_{\max}-\gamma)^2}{2\gamma}\left(2\operatorname{Re}(\lambda)+-\frac{2}{\gamma}\operatorname{Re}(\lambda)\mu_{\max}\right)\\ &>0,\end{aligned}$$

which is equivalent to

$$\begin{aligned}&\left(1+2\operatorname{Re}(\lambda)\mu_{\max}-\frac{2}{\gamma}\operatorname{Re}(\lambda)\mu_{\max}^2\right)\left(1-\frac{\delta}{\gamma}\right)\\ &+\frac{(2\mu_{\max}-\gamma)^2}{2\gamma}\left(\operatorname{Re}(\lambda)+-\frac{1}{\gamma}\operatorname{Re}(\lambda)\mu_{\max}\right)<0.\end{aligned}$$

The LHS can be written as

$$\begin{aligned}&\left(1-\frac{\delta}{\gamma}\right)\\ &+\operatorname{Re}(\lambda)\left(1-\frac{\mu_{\max}}{\gamma}\right)\left[2\mu_{\max}\left(1-\frac{\delta}{\gamma}\right)+\frac{(2\mu_{\max}-\gamma)^2}{2\gamma}\right]\end{aligned} \tag{5.8}$$

which is a linear function in $\operatorname{Re}(\lambda)$. Since $1-\delta/\gamma < 0$, it follows that (5.8) is negative for $\operatorname{Re}(\lambda) \in (-\varepsilon, 0)$. □

5.1.2 *Examples*

5.1.2.1 $P[a,b]$ Let $\mathbb{T} = P[a,b]$, the pulse time scale, which is the repeated pattern of a continuous interval of length a followed by a gap of length b. Since $\mathbb{T}$ is periodic with period $a+b$,

$$\delta = \frac{1}{a+b}\int_0^{a+b} \mu(t)\Delta t = \frac{1}{a+b}\left(\int_0^a 0dt + b^2\right) = \frac{b^2}{b+a}.$$

The stability region $\mathcal{S}$ for $P_{[a,b]}$, along with the osculating circle $\mathcal{H}_\delta$ is shown in Figure 5.2 for various values of a and b.

5.1.2.2 $\mathbb{T}_{1,2}^{n.p.}$ Define $\mathbb{T}_{1,2}^{n.p.} := \{t_n\}_{n=0}^\infty$ with $t_0 = 0$ and

$$\{\mu(t_n)\}_{n=0}^\infty = \{1,2,1,1,2,2,1,1,1,2,2,2,\ldots\}.$$

Note that the the graininesses of 1 and 2 occur equally often as $n \to \infty$, but the time scale is not periodic. Working from (5.1), since

$$\frac{\int_{t_0}^{t_N} \mu(t)\Delta t}{t_N - t_0} = \frac{\sum_{n=0}^N \mu(t_n)^2}{\sum_{n=0}^N \mu(t_n)} \begin{cases} = 5/3, & N = k(k+1), \\ < 5/3, & \text{otherwise}, \end{cases}$$

we conclude $\delta = 5/3$ on $\mathbb{T}_{1,2}^{n.p.}$

5.1.2.3 Repeated Cantor Sets Doan *et al.* [17] gave an example of the osculating circle at the origin of the region of exponential stability when $\mathbb{T}$ consists of repeated copies of the Cantor ternary set. They conjectured the osculating circle has a radius of 7, but left the proof as an interesting open problem. The methods outlined here enable us to show analytically that indeed the osculating circle has a radius of 7.

As $\mathbb{T}$ is periodic with period 1,

$$\delta = \int_0^1 \mu(t)\Delta t.$$

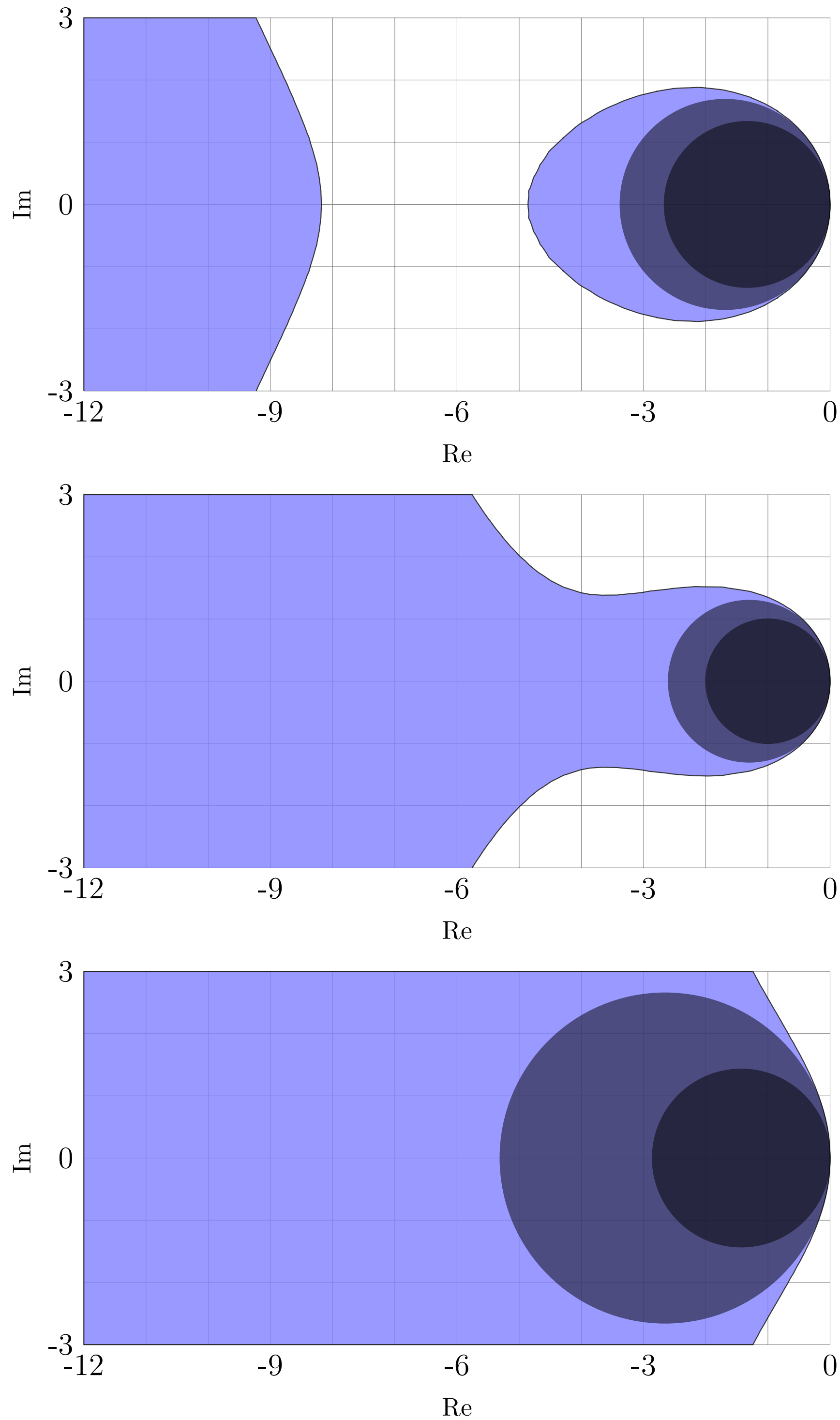

Figure 5.2: The region of exponential stability, $\mathcal{H}_\delta$, and $\mathcal{H}_{\mu_{\max}}$ for $P_{[a,b]}$ with various values of a and b. Top: $a = 1/5$ and $b = 3/4$; Middle: $a = 3/10$ and $b = 1$; Bottom: $a = 6/10$ and $b = 7/10$.

Since the Cantor ternary set has 2^n gaps of length $3^{-(n+1)}$, we see the curvature of the osculating circle is

$$\delta = \int_0^1 \mu(t)\Delta t = \sum_{n=0}^{\infty} \frac{2^n}{3^{n+1}} \frac{1}{3^{n+1}} = \frac{1}{9} \sum_{n=0}^{\infty} \left(\frac{2}{9}\right)^n = \frac{1}{9}\left(\frac{1}{1-2/9}\right) = \frac{1}{7}.$$

Thus, the osculating circle of $\mathcal{S}$ at the origin has curvature 1/7 and hence has a radius of 7. The geometry of the stability regions is shown in Figure 5.3.

5.1.2.4 The Bounded Graininess Assumption is Neccesary With this example, we show that the bounded graininess assumption is necessary for the conclusion of Theorem 5.2.

Define $\mathbb{T} = \bigcup_{n=0}^{\infty} \left[\frac{n(n+1)(2n+1)}{6}, \frac{n(n+1)(2n+1)}{6} + n^2 + n\right]$ and

$$f(T) := \frac{1}{T} \int_0^T \mu(t)\Delta t.$$

For each $n \in \mathbb{N}$, $f(n(n+1)(2n+1)/6) = 1$, whereas $f(T) \leq 1$ for every $T \in \mathbb{T}$. Therefore, $\delta = \limsup_{T\to\infty} f(T) = 1$. The structure of the time scale as well as select values of F are shown in Figure 5.4. Moreover, it is easy to check that $\mathcal{S} = \mathbb{C}^-$. Therefore, if the time scale has unbounded graininess, then $\mathcal{H}_\delta$ is not neccesarily the osculating circle of $\mathcal{S}$ at the origin.

5.1.3 Uniform Exponential Stability

A survey of the literature reveals that little is known about the region of uniform exponential stability $\mathcal{US}$ in general. The region is known for special classes of time scales, such as periodic time scales [17]. Doan *et al.* [17] showed that $\mathcal{H}_{\mu_{\max}}$ is always a subset of $\mathcal{US}$. Since $\mathcal{H}_\delta$ satisfies

$$\mathcal{H}_{\mu_{\max}} \subset \mathcal{H}_\delta \subset \mathcal{S},$$

it is tempting to conjecture that $\mathcal{H}_\delta$ is the region of uniform exponential stability. This is not the case, as the following example demonstrates.

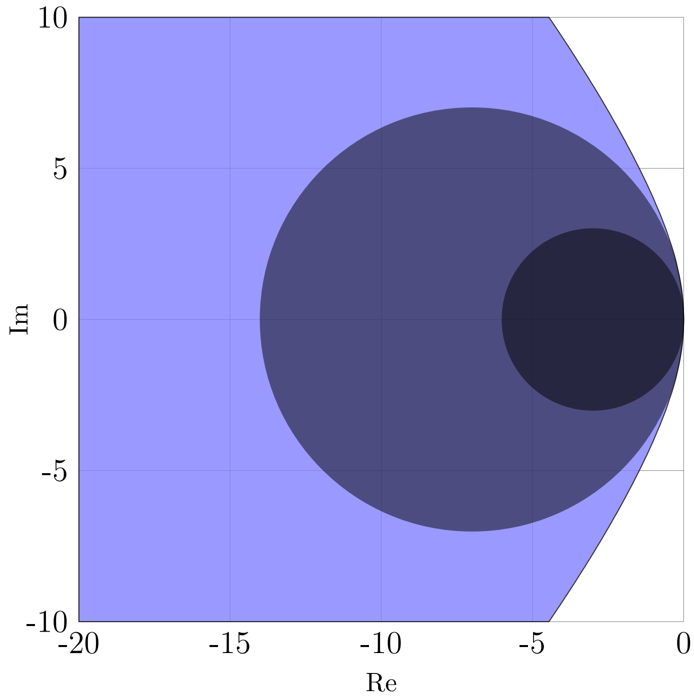

Figure 5.3. $\mathcal{S}$, $\mathcal{H}_\delta$ and $\mathcal{H}_{\mu_{\max}}$ for the Cantor ternary set.

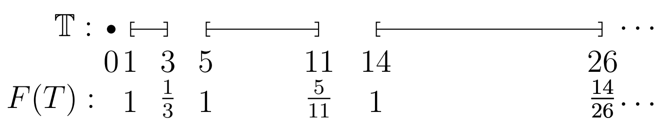

Figure 5.4. The time scale in Section *5.1.2.4 along with $f(T)$ at select time scales points.*

Example 1. Consider once more $\mathbb{T} = \mathbb{T}_{1,2}^{n.p.}$. We already showed $\delta = 5/3$ for this time scale. However, $\mathcal{H}_{5/3} \not\subset \mathcal{US}$. To verify this, consider $x^\Delta = -11/10x$. Now, $-11/10 \in \mathcal{H}_\delta$ but, given $M > 0$, we can choose an initial point t_k such that t_k proceeds $N > \ln(M)/\ln(12/10)$ points of graininess 2. Then

$$\begin{aligned}|e_{-11/10}(t_{k+N}, t_k)| &= |(1 + 2(-11/10))|^N |x(t_k)| \\ &= (12/10)^N |x(t_k)| \\ &> M|x(t_k)|,\end{aligned}$$

and hence $x^\Delta = -11/10x$ is not uniformly exponentially stable on $\mathbb{T}$ even though $-11/10 \in \mathcal{H}_\delta$.

This last example naturally leads us to ask: under what conditions is $\mathcal{H}_\delta$ a subset of the region of uniform exponential stability? The last example certainly shows that long strings of graininesses greater than δ can destroy uniform exponential stability. Arbitrarily large graininesses in the tail of the time scale have the same effect. The next condition rules out these types of examples.

Definition 5.1. A time scale $\mathbb{T}$ is *mean-stationary* provided there exists $K \geq 0$ such that for any $u, v \in \mathbb{T}$,

$$\int_u^v (\mu(t) - \delta)\Delta t \leq K.$$

Intuitively, this condition means that no matter where we look in the time scale, the time scale average value of the graininess in a window is within a fixed number of the true average. Arbitrarily large graininesses in the tail or arbitrarily long sequences of graininesses above δ violate the condition. Both of these conditions can cause the local average value of μ to be arbitrarily larger than the global average value of μ, δ.

The next lemma shows the definition of mean-stationarity can be cast in terms of rates of convergence.

Lemma 5.5. $\mathbb{T}$ *is mean-stationary if and only if*

$$\frac{\int_{t_0}^T \mu(t)\Delta t}{T - t_0} - \delta = \mathcal{O}\left(\frac{1}{T - t_0}\right) \text{ as } T \to \infty,$$

where δ is as in (5.1).

The next result reveals that in fact a broad class of time scales is mean-stationary.

Lemma 5.6. *Let $\mathbb{T}$ be a periodic time scale. Then $\mathbb{T}$ is mean-stationary.*

Proof. Let $\mathbb{T}$ be periodic with period T. Let $u, v \in \mathbb{T}$ with $u \leq v$. Since $\mathbb{T}$ is periodic, by [33, Lemma 9 (ii)]

$$\delta = \frac{1}{T}\int_u^{u+T} \mu(t)\Delta t.$$

Let

$$K = \max_{\substack{r \le s \le r+T, \\ r,s \in \mathbb{T}}} \int_r^s \mu(t)\Delta t = \max_{\substack{u \le s \le u+T \\ s \in \mathbb{T}}} \int_u^s \mu(t)\Delta t.$$

Write $v = u + kT + R$ with $R < T$ using the division algorithm. Then

$$\begin{aligned}
\int_u^v (\mu(t) - \delta)\Delta t = & \int_u^{u+kT} (\mu(t) - \delta)\Delta t + \int_{u+kT}^v (\mu(t) - \delta)\Delta t \\
= & \int_u^{u+kT} \mu(t)\Delta t - \delta kT + \int_{u+kT}^v (\mu(t) - \delta)\Delta t \\
= & 0 + \int_{u+kT}^v (\mu(t) - \delta)\Delta t \\
\le & K.
\end{aligned}$$

□

In Doan *et al.* [17], the authors show that if $\mathbb{T}$ is periodic, then $\mathcal{S} = \mathcal{US}$. Therefore, when $\mathbb{T}$ is periodic, $\mathcal{H}_\delta \subset \mathcal{US}$. The next result extends this by showing that $\mathcal{H}_\delta \subset \mathcal{US}$ if $\mathbb{T}$ is mean-stationary.

Theorem 5.3. *Let $\mathbb{T}$ be mean-stationary. Then $\mathcal{H}_\delta \subset \mathcal{US}$. Therefore, $\mathcal{H}_\delta$ is the osculating circle to the region of uniform exponential stability at the origin.*

Proof. Let $\lambda = x + iy \in \mathcal{H}_\delta$. Then $\alpha := x + |\lambda|^2\delta/2 < 0$. Since $\mathbb{T}$ is mean-stationary, there exists $K \ge 0$ such that for any $t, t_0 \in \mathbb{T}$, $\int_{t_0}^t (\mu(t) - \delta)\Delta t \le K$. Then

$$\begin{aligned}
|e_\lambda(t, t_0)| \le & \exp\left(\int_{t_0}^t \lim_{s \to \mu(t)^+} \frac{\ln((1 + sx)^2 + (sy)^2)}{2s} \Delta t \right) \\
\le & \exp\left(\int_{t_0}^t x + |\lambda|^2 \mu(t)/2 \Delta t \right) \\
= & \exp\left((x + |\lambda|^2\delta/2)(t - t_0) \right) \exp\left(\frac{|\lambda|^2}{2} \int_{t_0}^t (\mu(t) - \delta)\Delta t \right) \\
= & \exp\left(\alpha(t - t_0) \right) \exp\left(K|\lambda|^2/2 \right).
\end{aligned}$$

□

Of course, the last result is of no use if the class of mean-stationary time scales coincides with the class of periodic time scales. This final example shows this is not the case.

5.1.4 A Mean-Stationary, Non-Periodic Time Scale

Define the time scale $\mathbb{T} = \{0, t_1, t_2, \ldots\}$ in the following way:

$$\mu(0) = 1; \quad \mu(t_{2n}) = \mu(t_n); \quad \mu(t_{2n+1}) = 3 - \mu(t_n).$$

The sequence of graininesses begins

$$\{\mu(t_n)\} = \{1, 2, 2, 1, 2, 1, 1, 2, 2, 1, 1, 2, 1, 2, 2, 1, \ldots\}.$$

The generating laws for the time scale define a sequence of graininesses equal to the celebrated Thue-Morse sequence on the symbols 1 and 2. The Thue-Morse sequence is not periodic, and therefore the time scale is not periodic. It is easy to show that $\delta = 5/3$ and that $\int_u^v \mu(t) - \delta\Delta t \leq 4/3$ for every $u, v \in \mathbb{T}$ with $u < v$. Therefore, $\mathbb{T}$ is mean-stationary, but not periodic.

5.2 "Optimal" Control on Deterministic Time Scales

Using Corollary 4.2, we are able to find a constant feedback control law that stabilizes the system and that minimizes the expected cost functional. The ubiquity of the constant $\mathrm{E}[\mu^2]/\mathrm{E}[\mu]$ in (STSARE) leads us to conjecture that solving the equation

$$0 = Q + A^T V + VA + \delta A^T VA - (I + \delta A)^T VB(R + \delta B^T VB)^{-1} B^T V(I + \delta A) \quad \text{(TSARE)}$$

will yield a feedback gain $K = -(R + \delta B^T VB)^{-1} B^T V(I + \delta A)$ which will minimize the cost functional

$$J(x, u) = \limsup_{T\to\infty} \int_{t_0}^{T} x^T(t)Qx(t) + u^T(t)Ru(t)\Delta t. \tag{5.9}$$

Indeed, as the solution V of (STSARE) yields an optimal control K which stabilizes $x^\Delta = (A + BK)x$ in the mean-square sense, it follows that

$$\mathrm{spec}(A + BK) \subset \mathcal{H}_{\mathrm{E}[\mu^2]/\mathrm{E}[\mu]}.$$

Similarly, the solution V of (TSARE) yields the control law gain matrix

$$K = -(R + \delta B^T V B)^{-1} B^T V (I + \delta A).$$

Since all we have done is replace $\mathrm{E}[\mu^2]/\mathrm{E}[\mu]$ by δ, we must have $\mathrm{spec}(A+BK) \subset \mathcal{H}_\delta$. Therefore, as long as the time scale is mean stationary or if $A + BK$ is uniformly regressive, then K is a stabilizing feedback.

While the control Kx generated by solving (TSARE) indeed stabilizes the system, it does not, in general, minimize the cost functional (5.9).

Our interpretation of stochastic time scales helps us reconcile why solving (TSARE) does not yield an *optimal* control with respect to (5.9). For a fixed dynamic equation $x^\Delta = Ax + Bu$, and fixed Q and R, there are infinitely many time scales which have the same average graininess δ. For each of these time scales, the feedback gain K generated from solving (TSARE) will be the same, yet for each time scale, the corresponding system will evolve in a different manner. We can view the choice of K from solving (TSARE) as a reasonable (but not optimal) choice to work across every time scale with the same average graininess δ.

BIBLIOGRAPHY

[1] Ravi A., M. Bohner, D. O'Regan, and A. Peterson, *Dynamic equations on time scales: a survey*, Journal of Computational and Applied Mathematics **141** (2002), no. 12, 1 – 26, Dynamic Equations on Time Scales.

[2] H. Abou-Kandil, *Matrix Riccati Equations: In Control and Systems Theory*, Springer, 2003.

[3] P.J. Antsaklis and A.N. Michel, *Linear systems*, 2006.

[4] D.H.J. Baert and A.A.K. Vervaet, *A new method for the measurement of the double layer capacitance for the estimation of battery capacity*, The 25th International Telecommunications Energy Conference, Oct 2003, pp. 733–738.

[5] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific optimization and computation series, no. v. 2, Athena Scientific, 2007.

[6] R.R. Bitmead, *Convergence properties of LMS adaptive estimators with unbounded dependent inputs*, 1981 20th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes, Dec 1981, pp. 607–612.

[7] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser Boston, Inc., Boston, MA, 2001.

[8] M. Bohner and N. Wintz, *The linear quadratic regulator on time scales*, Int. J. Difference Equ **5** (2010), no. 2, 149–174.

[9] W.L. Brogan, *Modern Control Theory*, third ed., Prentice-Hall, Englewood Cliffs, NJ, 1991.

[10] M. Chen and G.A. Rincon-Mora, *Accurate electrical battery model capable of predicting runtime and I-V performance*, IEEE Transactions on Energy Conversion **21** (2006), no. 2, 504–511.

[11] F. Codeca, S.M. Savaresi, and G. Rizzoni, *On battery state of charge estimation: A new mixed algorithm*, 2008 IEEE International Conference on Control Applications, Sept 2008, pp. 102–107.

[12] J.J. DaCunha, *Lyapunov stability and Floquet theory for nonautonomous linear dynamic systems on time scales*, Ph.D. thesis, Baylor University, 2004.

[13] T. Damm, *Rational matrix equations in stochastic control*, Lecture Notes in Control and Information Sciences, vol. 297, Springer-Verlag, Berlin, 2004.

[14] J.M. Davis, I.A. Gravagne, R.J. Marks, and B.J. Jackson, *Regions of exponential stability for LTI systems on nonuniform discrete domains*, 2011 IEEE 43rd Southeastern Symposium on System Theory, March 2011, pp. 37–42.

[15] M.H.A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall mathematics series, Chapman and Hall, 1977.

[16] T. S. Doan, A. Kalauch, and S. Siegmund, *Exponential stability of linear time-invariant systems on time scales*, Nonlinear Dynamics and Systems Theory **9** (2009), no. 1, 37–50.

[17] T.S. Doan, A. Kalauch, S. Siegmund, and F.R. Wirth, *Stability radii for positive linear time-invariant systems on time scales*, Systems & Control Letters **59** (2010), no. 34, 173 – 179.

[18] G. B. Eisenbarth, *Quadratic Lyapunov theory for dynamic linear switched systems*, Ph.D. thesis, Baylor University, 2013.

[19] M. D. Fragoso, O. L. V. Costa, and C. E. de Souza, *A new approach to linearly perturbed Riccati equations arising in stochastic control*, Applied Mathematics and Optimization **37** (1998), no. 1, 99–126.

[20] T. Gard and J. Hoffacker, *Asymptotic behavior of natural growth on time scales*, Dynamic Systems and Applicationsc **12** (2002).

[21] I.A. Gravagne, J.M. Davis, and D.R. Poulsen, *Time scale-based observer design for battery state-of-charge estimation*, 2012 IEEE 44th Southeastern Symposium on System Theory, March 2012, pp. 12–17.

[22] A. Gray, *Curvature of curves in the plane*, "Drawing Plane Curves with Assigned Curvature," and "Drawing Space Curves with Assigned Curvature." **1** (1997), no. 5, 222–224.

[23] Abderrezak Hammouche, Eckhard Karden, and Rik W. De Doncker, *Monitoring state-of-charge of Ni-MH and Ni-Cd batteries using impedance spectroscopy*, Journal of Power Sources **127** (2004), no. 12, 105 – 111, Eighth Ulmer Electrochemische Tage.

[24] C. Heij, A. Ran, and F. van Schagen, *Introduction to mathematical systems theory*.

[25] S. Hilger, *Ein maßkettenkalkül mit anwendung auf zentrumsmannigfaltigkeiten*, Ph.D. thesis, Universität Würzburg, 1988.

[26] B. J. Jackson, J. M. Davis, I. A. Gravagne, and R. J. Marks I. I., *Linear state feedback stabilisation on time scales*, Int. J. Dynamical Systems and Differential Equations **3** (2011), 163–177.

[27] R.E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME–Journal of Basic Engineering **82** (1960), no. Series D, 35–45.

[28] L. Koralov and Y.G. Sinai, *Theory of Probability and Random Processes*, Universitext, Springer Berlin Heidelberg, 2007.

[29] H.J. Kushner, *Introduction to stochastic control*, Holt, Rinehart and Winston New York, 1971.

[30] V. Lakshmikantham, S. Sivasundaram, and B. Kaymakcalan, *Dynamic Systems on Measure Chains*, Mathematics and Its Applications (Kluwer Academic Publishers), V. 370, Springer, 1996.

[31] T. Morozan, *Stabilization of some stochastic discrete–time control systems*, Stochastic Analysis and Applications **1** (1983), no. 1, 89–116.

[32] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Communications and signal processing, McGraw-Hill, 1991.

[33] C. Pötzsche, S. Siegmund, and F. Wirth, *A spectral characterization of exponential stability for linear time-invariant systems on time scales*, Discrete and Continuous Dynamical Systems **9** (2003), no. 5, 1223–1242.

[34] D.R. Poulsen, J.M. Davis, and I.A. Gravagne, *Observer based feedback controllers on stochastic time scales*, 2013 IEEE 45th Southeastern Symposium on System Theory, March 2013, pp. 104–107.

[35] ______, *Stochastic time scales: Quadratic Lyapunov functions and probabilistic regions of stability*, 2013 IEEE 45th Southeastern Symposium on System Theory, March 2013, pp. 98–103.

[36] D.R. Poulsen, J.M. Davis, I.A. Gravagne, and R.J. Marks, *On the stability of mu-varying dynamic equations on stochastically generated time scales*, 2012 IEEE 44th Southeastern Symposium on System Theory, March 2012, pp. 18–23.

[37] A. A. Ramos, *Stability of hybrid dynamic systems: Analysis and design*, Ph.D. thesis, Baylor University, 2009.

[38] A. C. M. Ran and M. C. B. Reurings, *A fixed point theorem in partially ordered sets and some applications to matrix equations*, Proc. Amer. Math. Soc. **132** (2004), no. 5, 1435–1443 (electronic).

[39] S. Rodrigues, N. Munichandraiah, and A.K. Shukla, *A review of state-of-charge indication of batteries by means of A.C. impedance measurements*, Journal of Power Sources **87** (2000), no. 12, 12 – 20.

[40] W. Wonham, *On a matrix Riccati equation of stochastic control*, SIAM Journal on Control **6** (1968), no. 4, 681–697.

[41] X.Y. Xiong, H. Vander Poorten, and M. Crappe, *Impedance parameters of Ni/Cd batteries - individual electrode characteristics. Application to modelling and state of charge determinations*, Electrochemica Acta **41** (1996), no. 7, 1267 – 1275.

[42] J. Zhang and C. Xia, *Application of statistical parameter identification in battery management system*, Intelligent Control and Automation (WCICA), 2010 8th World Congress on, July 2010, pp. 5938–5942.