ABSTRACT

A General Linear Systems Theory on Time Scales:
Transforms, Stability, and Control

Billy Joe Jackson, Ph.D.

Advisor: John M. Davis, Ph.D.

In this work, we examine linear systems theory in the arbitrary time scale setting by considering Laplace transforms, stability, controllability, observability, and realizability. In particular, we revisit the definition of the Laplace transform given by Bohner and Peterson in [10]. We provide sufficient conditions for a given function to be transformable, as well as an inversion formula for the transform. Sufficient conditions for the inverse transform to exist are provided, and uniqueness of this inverse function is discussed. Convolution under the transform is then considered. In particular, we develop an analogue of the Convolution Theorem for arbitrary time scales and discuss the algebraic properties of the convolution. This naturally leads to an algebraic identity for the convolution operator, which is a time scale analogue of the Dirac delta distribution.

Next, we investigate applications of the transform to linear time invariant systems and before discussing linear time varying systems. The focus is on fundamental notions of linear system control such as controllability, observability, and realizability. Sufficient conditions for a system to possess each of these properties are given in the time varying case, while these same criteria often become necessary and sufficient in the time invariant case. Finally, several notions of stability are discussed, and linear state feedback is explored.

A General Linear Systems Theory on Time Scales:
Transforms, Stability, and Control

by

Billy Joe Jackson, B.S., M.A.

A Dissertation

Approved by the Department of Mathematics

Lance L. Littlejohn

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

John M. Davis, Ph.D., Chairperson

Johnny Henderson, Ph.D.

Robert J. Marks II, Ph.D.

Frank H. Mathis, Ph.D.

Qin Sheng, Ph.D.

Accepted by the Graduate School
August 2007

Larry Lyon

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

My sincere gratitude goes to John Davis for his time, consideration, and effort that he has put into helping me produce this dissertation. He has been a great mentor and I shall always treasure the last three years that we have spent together as advisor and student.

Second, I would also like to thank Ian Gravagne and Bob Marks for their contributions to this work. Their valuable input and insight have made this work so much more than it would have been without their ideas.

CHAPTER ONE

A Review of the Time Scales Calculus

1.1 Differential Calculus

The time scales calculus was first introduced in the Ph.D. thesis of Stefan Hilger in 1988 (see [32] and [31]). He begins by defining a time scale to be an arbitrary closed subset of the reals, where $\mathbb{R}$ is given the standard topology. There is no reason to assume that a time scale be unbounded from above, but we shall make this blanket assumption since all of the results in the following chapters deal with time scales of this type.

Let $\mathbb{T}$ be a time scale. Define the *forward jump operator* $\sigma : \mathbb{T} \to \mathbb{R}$ by

$$\sigma(t) := \inf\{s \in \mathbb{T} : s > t\},$$

and the *backward jump operator* $\rho : \mathbb{T} \to \mathbb{R}$ by

$$\rho(t) := \sup\{s \in \mathbb{T} : s < t\}.$$

Here, $\inf \emptyset = \sup \mathbb{T}$ and $\sup \emptyset = \inf \mathbb{T}$. We define the *graininess function* as the distance function between successive points in the time scale, and we denote it by $\mu : \mathbb{T} \to \mathbb{R}$, where $\mu(t) = \sigma(t) - t$.

The *forward derivative* or *delta derivative* of $f : \mathbb{T} \to \mathbb{R}$ (provided it exists) is then defined as the number $f^{\Delta}(t)$ with the property that given any $\epsilon > 0$, there exists a neighborhood U of t (i.e., $U = (t-\delta, t+\delta) \cap \mathbb{T}$ for some $\delta > 0$) such that

$$\left|[f(\sigma(t)) - f(s)] - f^{\Delta}(t)[\sigma(t) - s]\right| \le \epsilon|\sigma(s) - t| \text{ for all } s \in U.$$

It is worth noting that this definition is equivalent to the limit

$$f^{\Delta}(t) = \lim_{s \to t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} = \frac{f(\sigma(t)) - f(t)}{\mu(t)},$$

as long as the difference quotient is interpreted in the limit sense if $\sigma(t) \to t$. Points $t \in \mathbb{T}$ with $\mu(t) = 0$ are termed *right-dense*, while points with $\mu(t) > 0$ are called *right-scattered*. This framework includes two very important cases: namely $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. It is easy to verify that in these two cases, we have $\sigma(t) = t$ and $\sigma(t) = t+1$, respectively. In each case, the corresponding derivatives are $f^{\Delta}(t) = f'(t)$ and $f^{\Delta}(t) = \Delta f(t)$, where $\Delta f(t) := f(t+1) - f(t)$ denotes the usual forward difference operator. Thus, in this sense, any result concerning the time scales calculus will unify the two classically studied cases.

However, there are many other time scales that can now be studied as well. For example, any *hybrid* set can now also be examined. That is, any set containing points some of which are right-dense with others right-scattered easily fit within the framework. Thus, Hilger's calculus will extend the results beyond the cases $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. The time scale

$$\mathbb{P}_{a,b} := \bigcup_{k=0}^{\infty} [k(a+b), k(a+b)+a]$$

is an example of a hybrid set since

$$\sigma(t) = \begin{cases} t, & \text{if } \; t \in \bigcup_{k=0}^{\infty} [k(a+b), k(a+b)+a), \\ t+b, & \text{if } \; t \in \bigcup_{k=0}^{\infty} \{k(a+b)+a\}, \end{cases}$$

and therefore

$$\mu(t) = \begin{cases} 0, & \text{if } \; t \in \bigcup_{k=0}^{\infty} [k(a+b), k(a+b)+a), \\ b, & \text{if } \; t \in \bigcup_{k=0}^{\infty} \{k(a+b)+a\}. \end{cases}$$

Besides hybrid sets, Hilger's framework also handles sets which are nonuniformly spaced unlike $\mathbb{T} = \mathbb{Z}$. Indeed, one of the most important examples of a time scale besides $\mathbb{R}$ and $\mathbb{Z}$ is the set $\mathbb{T} = \overline{q^{\mathbb{Z}}}$, which is defined for $q > 0$ as

$$\mathbb{T} = \overline{q^{\mathbb{Z}}} := \left\{ q^n : n \in \mathbb{Z} \right\} \cup \{0\}.$$

This set is commonly referred to as the *quantum time scale*, and it has a quantum calculus that has been studied in the literature (see [7], [6], and [14]). Notice that this time scale does indeed exhibit hybrid features since $t = 0$ is right dense and every other point is right scattered and nonuniformly spaced.

In what follows, we shall give the most pertinent results from the time scales calculus that will be needed in later chapters. We will focus on the matrix case, as the scalar case will of course be a special case of the matrix case. A complete treatment of the calculus can be found in Bohner and Peterson in [8] and [9].

We begin by developing the geometry of the Hilger complex plane. For $\mu(t) > 0$, the *Hilger complex numbers*, the *Hilger real axis*, the *Hilger alternating axis*, and the *Hilger imaginary circle* (or simply the *Hilger circle*) are all respectively defined as

$$\begin{aligned}
\mathbb{C}_\mu &:= \left\{ z \in \mathbb{C} : z \neq -\frac{1}{\mu(t)} \right\} \\
\mathbb{R}_\mu &:= \left\{ z \in \mathbb{C}_\mu : z \in \mathbb{R} \text{ and } z > -\frac{1}{\mu(t)} \right\}, \\
\mathbb{A}_\mu &:= \left\{ z \in \mathbb{C}_\mu : z \in \mathbb{R} \text{ and } z < -\frac{1}{\mu(t)} \right\}, \\
\mathbb{H}_\mu &:= \left\{ z \in \mathbb{C}_\mu : \left| z + \frac{1}{\mu(t)} \right| = \frac{1}{\mu(t)} \right\},
\end{aligned}$$

and for $\mu(t) = 0$, define $\mathbb{C}_0 := \mathbb{C}, \mathbb{R}_0 := \mathbb{R}, \mathbb{H}_0 := i\mathbb{R}$, and $\mathbb{A}_0 := \emptyset$. For any $z \in \mathbb{C}_\mu$, the *Hilger real part* of z is given by

$$\mathrm{Re}_\mu(z) := \frac{|z\mu(t) + 1| - 1}{\mu(t)},$$

while the *Hilger imaginary part* of z is defined as

$$\mathrm{Im}_\mu(z) := \frac{\mathrm{Arg}(z\mu(t) + 1)}{\mu(t)},$$

where $\mathrm{Arg}(z)$ is the principal argument of z. For $\mu(t) = 0$, define $\mathrm{Re}_\mu(z) = \mathrm{Re}(z)$ and $\mathrm{Im}_\mu(z) = \mathrm{Im}(z)$. The *Hilger purely imaginary number* $\mathring{\imath}\omega$ is defined for $-\frac{\pi}{\mu(t)} \leq$

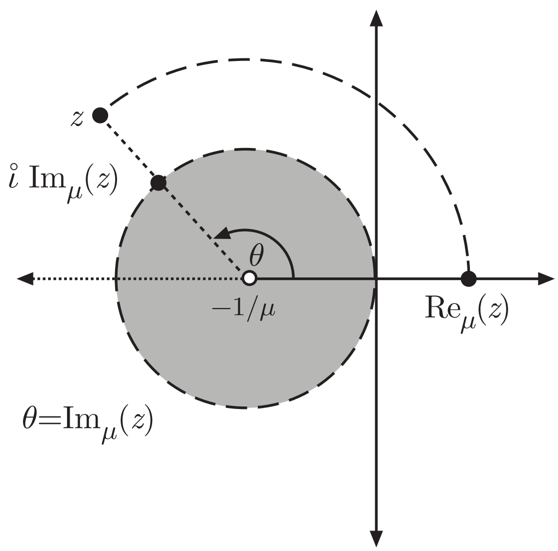

Figure 1.1: The Hilger Complex Plane. Points exterior to the circle have positive Hilger real part, while points interior have negative Hilger real part. Points on the circle have zero Hilger real part, and are thus called the Hilger purely imaginary numbers.

$\omega \leq \frac{\pi}{\mu(t)}$ and is given by

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

For $\mu(t) = 0$, define $\mathring{\iota}\omega := i\omega$. Hilger's description of the complex plane is shown in Figure 1.1.

The set $\mathbb{C}_\mu$ is endowed with a group structure if we define the circle plus addition on $\mathbb{C}_\mu$ as

$$a \oplus b := a + b + ab\mu(t).$$

In fact, $(\mathbb{C}_\mu, \oplus, \ominus)$ is an Abelian group, with

$$\ominus z := -\frac{z}{1 + \mu(t)z}.$$

We then have the following:

Theorem 1.1 ([8]). *For* $z \in \mathbb{C}_\mu$, *we have* $z = \mathrm{Re}_\mu(z) \oplus \mathring{\iota}\mathrm{Im}_\mu(z)$.

Next, we review a few basics of the time scales (matrix) calculus. We use the notation $A^\sigma(t)$ to denote the matrix of functions $A(\sigma(t))$. A matrix is *right-dense* *continuous* (abbreviated rd-continuous) if every entry of A is right-dense continuous. In the scalar case, we say a function is right-dense continuous if it is continuous at right-dense points of $\mathbb{T}$ and its left-sided limits exist as finite numbers at left-dense points in $\mathbb{T}$. The set of all such matrices is denoted by

$$C_{\text{rd}} = C_{\text{rd}}(\mathbb{T}) = C_{\text{rd}}(\mathbb{T}, \mathbb{R}^{m\times n}).$$

$A(t)$ is a delta differentiable matrix if each entry of A is delta differentiable, in which case we define

$$A^{\Delta}(t) = (a_{ij}^{\Delta})_{1\le i\le m, 1\le j\le n}, \quad \text{where} \quad A = (a_{ij})_{1\le i\le m, 1\le j\le n}.$$

We shall need to make frequent use of the next identity, so we state it as a theorem:

Theorem 1.2 ([8]). *If A is differentiable at $t \in \mathbb{T}$, then $A^{\sigma}(t) = A(t) + \mu(t)A^{\Delta}(t)$.*

The following theorem establishes the delta derivative as a linear operator and the analogue of the product rule on an arbitrary time scale:

Theorem 1.3 ([8]). *Suppose A and B are delta differentiable $n \times n$-matrix-valued functions. Then*

(i) $(A + B)^{\Delta} = A^{\Delta} + B^{\Delta}$;

(ii) $(\alpha A)^{\Delta} = \alpha A^{\Delta}$ *if* α *is constant;*

(iii) $(AB)^{\Delta} = A^{\Delta}B^{\sigma} + AB^{\Delta} = A^{\sigma}B^{\Delta} + A^{\Delta}B$;

(iv) $(A^{-1})^{\Delta} = -(A^{\sigma})^{-1}A^{\Delta}A^{-1} = -A^{-1}A^{\Delta}(A^{\sigma})^{-1}$ *if* AA^{σ} *is invertible;*

(v) $(AB^{-1})^{\Delta} = (A^{\Delta} - AB^{-1}B^{\Delta})(B^{\sigma})^{-1} = (A^{\Delta} - (AB^{-1})^{\sigma}B^{\Delta})B^{-1}$ *if* BB^{σ} *is invertible.*

In this work, we wish to study solutions of systems of *dynamic equations* on the (unbounded) time scale $\mathbb{T}$. That is, we wish to examine solutions of the equation

$$y^{\Delta}(t) = A(t)y(t) + f(t), \quad y(t_0) = y_0, \quad t, t_0 \in \mathbb{T}, \quad y_0 \in \mathbb{R}^{n\times m}, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}).$$

To do this, we will need to make appropriate assumptions on the matrix $A(t)$. We say the matrix $A(t)$ is *regressive* if the matrix $I + \mu(t)A(t)$ is invertible for all t. In the scalar case, a function $f(t)$ is *positively regressive* if f is regressive and $1 + \mu(t)f(t) > 0$ for all $t \in \mathbb{T}$. The collection of all regressive matrices is denoted by

$$\mathcal{R} = \mathcal{R}(\mathbb{T}) = \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}),$$

while the positively regressive functions are denoted by

$$\mathcal{R}^+ = \mathcal{R}^+(\mathbb{T}) = \mathcal{R}^+(\mathbb{T}, \mathbb{R}).$$

We define the operation $\oplus$ on $\mathcal{R}$ as

$$(A \oplus B)(t) = A(t) + B(t) + \mu(t)A(t)B(t) \text{ for all } t \in \mathbb{T},$$

and the operation $\ominus$ by

$$(\ominus A)(t) = -A(t)[I + \mu(t)A(t)]^{-1} \text{ for all } t \in \mathbb{T}.$$

With these operations defined on $\mathcal{R}$, we have the following:

Theorem 1.4 ([8]). *$(\mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}), \oplus, \ominus)$ is a group. Furthermore, in the scalar case, $(\mathcal{R}^+(\mathbb{T}, \mathbb{R}), \oplus, \ominus)$ is a subgroup of the regressive group.*

Theorem 1.5 ([8]). *Let $A \in \mathcal{R}$ be an $n \times n$ matrix-valued function on $\mathbb{T}$ and suppose that $f : \mathbb{T} \to \mathbb{R}^n$ is rd-continuous. Let $t_0 \in \mathbb{T}$ and $y_0 \in \mathbb{R}^n$. Then the initial value problem (IVP)*

$$y^\Delta(t) = A(t)y(t) + f(t), \;\; y(t_0) = y_0,$$

has a unique solution $y : \mathbb{T} \to \mathbb{R}^n$.

Thus, the preceding theorem establishes existence and uniqueness of solutions to our dynamic equation. If A is time varying, we shall denote the solution of

$$Y^\Delta(t) = A(t)Y(t), \;\; Y(t_0) = I,$$

as $Y(t) = \Phi_A(t,t_0)$, while if A is time invariant, we denote the solution of the system as $e_A(t,t_0)$. There are important distinctions between the two notations, as $\Phi_A(t,t_0) \equiv e_A(t,t_0)$ if and only if $A(t) \equiv A$ is constant. The usefulness of these two characterizations can easily be seen on $\mathbb{T} = \mathbb{R}$, in which case $e_A(t,t_0) = e^{A(t-t_0)}$. It is known that on $\mathbb{R}$, there are many differences between solutions of autonomous systems versus nonautonomous systems which amount to the differences between $\Phi_A(t,t_0)$ and $e^{A(t-t_0)}$ in general.

In the scalar case, for $p \in \mathcal{R}$, the solution of

$$y^{\Delta}(t) = p(t)y(t), \ \ y(t_0) = 1,$$

is denoted by $y(t) = e_p(t,t_0)$. Hilger proved that the closed form of $e_p(t,t_0)$ is given by

$$e_p(t,t_0) = \exp\left(\int_{t_0}^{t} \frac{\operatorname{Log}(1+\mu(\tau)p(\tau))}{\mu(\tau)}\Delta\tau\right),$$

where the Δt in the integral is used to denote that this is a time scale integral, whose treatment follows. Before discussing the integral, however, it is worth stating the following two theorems from Bohner and Peterson [8] which discuss the sign of the exponential function in the scalar case. In particular, note that the theorems tell us that $e_p(t,t_0)$ is positive if p is positively regressive, a fact which we will make use of later in the discussion on stability.

Theorem 1.6 ([8]). *Assume $p \in \mathcal{R}$ and $t_0 \in \mathbb{T}$.*

(i) If $1+\mu p > 0$ on $\mathbb{T}$, then $e_p(t,t_0) > 0$ for all $t \in \mathbb{T}$.

(ii) If $1+\mu p < 0$ on $\mathbb{T}$, then $e_p(t,t_0) = \alpha(t,t_0)(-1)^{n_t}$ for all $t \in \mathbb{T}$, where

$$\alpha(t,t_0) := \exp\left(\int_{t_0}^{t} \frac{\log|1+\mu(\tau)p(\tau)|}{\mu(\tau)}\Delta\tau\right) > 0$$

and

$$n_t = \begin{cases} |[t_0,t)|, & \text{if } \ t \geq t_0, \\ |[t,t_0)|, & \text{if } \ t < t_0. \end{cases}$$

Note that $1 + \mu(t)p(t) < 0$ for all $t \in \mathbb{T}$ implies that $\mathbb{T}$ contains no right-dense points. Hence, $|[t_0, t)|$, which represents the number of points in the indicated interval, will be finite.

Theorem 1.7 (Sign of the Exponential Function, [8]). *Let $p \in \mathcal{R}$ and $t_0 \in \mathbb{T}$.*

(i) If $p \in \mathcal{R}^+$, then $e_p(t, t_0) > 0$ for all $t \in \mathbb{T}$.

(ii) If $1 + \mu(t)p(t) < 0$ for some $t \in \mathbb{T}$, then

$$e_p(t, t_0)e_p(\sigma(t), t_0) < 0.$$

(iii) If $1 + \mu(t)p(t) < 0$ for all $t \in \mathbb{T}$, then $e_p(t, t_0)$ changes sign at every point $t \in \mathbb{T}$.

(iv) Assume there exist sets $T = \{t_k : k \in \mathbb{N}\} \subset \mathbb{T}$ and $S = \{s_k : k \in \mathbb{N}\} \subset \mathbb{T}$ with

$$\ldots < s_2 < s_1 < t_0 \leq t_1 < t_2 < \ldots$$

such that $1 + \mu(t)p(t) < 0$ for all $t \in S \cup T$ and $1 + \mu(t)p(t) > 0$ for all $t \in \mathbb{T} - \{S \cup T\}$. Furthermore, if $|T| = \infty$, then $\lim_{n\to\infty} t_n = \infty$, and if $|S| = \infty$, then $\lim_{n\to\infty} s_n = -\infty$. If $T \neq \emptyset$ and $S \neq \emptyset$, then

$$e_p(\cdot, t_0) > 0 \text{ on } [\sigma(s_1), t_1).$$

If $|T| = \infty$, then

$$(-1)^k e_p(\cdot, t_0) > 0 \text{ on } [\sigma(t_k), t_{k+1}] \text{ for all } k \in \mathbb{N}.$$

If $|T| = N \in \mathbb{N}$, then

$$(-1)^k e_p(\cdot, t_0) > 0 \text{ on } [\sigma(t_k), t_{k+1}] \text{ for all } 1 \leq k \leq N - 1$$

and

$$(-1)^N e_p(\cdot, t_0) > 0 \text{ on } [\sigma(t_N), \infty).$$

If $T = \emptyset$ and $S \neq \emptyset$, then

$$e_p(\cdot, t_0) > 0 \text{ on } [\sigma(s_1), \infty).$$

If $|S| = \infty$, then

$$(-1)^k e_p(\cdot, t_0) > 0 \text{ on } [\sigma(s_{k+1}), s_k] \text{ for all } k \in \mathbb{N}.$$

If $|S| = M \in \mathbb{N}$, then

$$(-1)^k e_p(\cdot, t_0) > 0 \text{ on } [\sigma(s_{k+1}), s_k] \text{ for all } 1 \leq k \leq M-1$$

and

$$(-1)^M e_p(t, t_0) > 0 \text{ on } (-\infty, s_M].$$

If $S = \emptyset$ and $T \neq \emptyset$, then

$$e_p(\cdot, t_0) > 0 \text{ on } (-\infty, t_1].$$

In particular, the exponential function $e_p(\cdot, t_0)$ is a real-valued function that is never equal to zero but can be negative.

1.2 Integral Calculus

To talk about solutions of the equation, we need to develop an integral process that will act as inverse of the Δ-differential operator. There are many different ways in which one can integrate: for example, we can use any one of the Cauchy, Riemann, or Lebesgue integrals, among others. Of most importance to this work is the Lebesgue integral; we will present its development in the time scale setting here. We follow the construction given by Guseinov in [26].

Denote by $\mathcal{F}_1$ the family of all left closed and right open intervals of $\mathbb{T}$ of the form

$$[a, b) = \{t \in \mathbb{T} : a \leq t < b\},$$

with $a, b \in \mathbb{T}$ and $a \leq b$. We understand $[a, a)$ to be the empty set. The collection $\mathcal{F}_1$ forms a semiring of subsets of $\mathbb{T}$. Let $m_1 : \mathcal{F}_1 \to [0, \infty]$ be the set function defined on $\mathcal{F}_1$ that assigns to each interval $[a, b)$ its length:

$$m_1([a, b)) = b - a.$$

m_1 then becomes a countably additive measure on $\mathcal{F}_1$. We denote the Carathéodory extension of the set function m_1 associated with $\mathcal{F}_1$ by μ_Δ, and we call μ_Δ the *Lebesgue* Δ*-measure* on $\mathbb{T}$.

It is worth examining the Carathéodory extension μ_Δ of m_1. We begin by generating an outer measure m_1^* on the collection of all subsets of $\mathbb{T}$ as follows. Let E be any subset of $\mathbb{T}$. If there exists at least one finite or countable system of intervals $V_j \in \mathcal{F}_1$ for $j \in \mathbb{N}$ such that $E \subset \cup_j V_j$, then we set

$$m_1^* = \inf \sum_j m_1(V_j),$$

where the infimum is taken over all such Vitali coverings of E by a finite system or countable system of intervals $V_j \in \mathcal{F}_1$. If there is no such covering of E, then we set $m_1^*(E) = \infty$.

Next, we define the family $\mathcal{M}(m_1^*)$ of all m_1^*-measurable subsets of $\mathbb{T}$. A subset A of $\mathbb{T}$ is said to be m_1^*-measurable (or Δ-measurable), if

$$m_1^*(E) = m_1^*(E \cap A) + m_1^*(E \cap A^C),$$

holds for all $E \subset \mathbb{T}$, where $A^C = \mathbb{T} - A$ denotes the complement of A. Note that the collection $\mathcal{M}(m_1^*)$ of all m_1^*-measurable subsets of $\mathbb{T}$ is a σ-algebra. Finally, we take the restriction of m_1^* to $\mathcal{M}(m_1^*)$, which we denote by μ_Δ. This measure is then a countably additive measure on $\mathcal{M}(m_1^*)$.

Note that if the time scale is finite, then the time scale will have infinite μ_Δ measure from Guseinov's construction. As a consequence, every time scale will have infinite measure. We will revisit this issue in the next chapter when we discuss the uniqueness of the inverse. For now however, we examine some consequences of this definition:

Theorem 1.8 ([26]). *For each* $t_0 \in \mathbb{T} - \{\max \mathbb{T}\}$, *the single point set* $\{t_0\}$ *is* Δ-*measurable, and its* Δ-*measure is given by*

$$\mu_\Delta(\{t_0\}) = \sigma(t_0) - t_0 = \mu(t_0).$$

Theorem 1.9 ([26]). *If* $a, b \in \mathbb{T}$ *and* $a \leq b$, *then*

$$\mu_\Delta([a,b)) = b - a \text{ and } \mu_\Delta((a,b)) = b - \sigma(a).$$

If $a, b \in \mathbb{T} - \{\max \mathbb{T}\}$ *and* $a \leq b$, *then*

$$\mu_\Delta((a,b]) = \sigma(b) - \sigma(a) \text{ and } \mu_\Delta([a,b]) = \sigma(b) - a.$$

The Lebesgue integral associated with the measure μ_Δ on $\mathbb{T}$ is called the *Lebesgue* Δ-*integral*, and for a measurable set $E \subset \mathbb{T}$ and a measurable function $f : E \to \mathbb{R}$, the integral of f over E is denoted by

$$\int_E f(t)\Delta t.$$

Thus, in terms of the measure theory involved, all of the standard theorems of general Lebesgue integration theory (including the dominated convergence theorem) hold also for the Lebesgue Δ-integral.

The next theorem connects the Riemann and Lebesgue integrals. Although we will not give the treatment of the Riemann Δ-integral here, the interested reader can find this treatment in [9] or in [26].

Theorem 1.10 ([26]). *Let* $[a,b]$ *be a closed bounded interval in* $\mathbb{T}$ *and let* f *be a bounded real-valued function defined on* $[a,b]$. *If* f *is Riemann* Δ-*integrable from* a *to* b, *then* f *is Lebesgue* Δ-*integrable on* $[a,b)$, *and*

$$\mathrm{R}\int_a^b f(t)\Delta t = \mathrm{L}\int_{[a,b)} f(t)\Delta t,$$

where R *and* L *indicate the Riemann and Lebesgue integrals, respectively.*

For completeness, we now state the fundamental theorem of calculus for the time scale setting. The utility of the preceding theorem then becomes clear, since when they all exist, the Lebesgue integral agrees with the Riemann integral which in turn agrees with the Cauchy integral.

Theorem 1.11 (Fundamental Theorem of Calculus, Part I, [26]). *Let f be a function which is Δ-integrable from a to b. For $t \in [a, b]$, define*

$$F(t) = \int_a^t f(\tau)\Delta\tau.$$

Then F is continuous on $[a, b]$. If $t_0 \in [a, b)$ and if f is continuous at t_0 provided t_0 is right-dense, then F is Δ-differentiable at t_0 and

$$F^\Delta(t_0) = f(t_0).$$

Theorem 1.12 (Fundamental Theorem of Calculus, Part II, [26]). *Let f be a Δ-integrable function on $[a, b]$. If f has a Δ-antiderivative $F : [a, b] \to \mathbb{R}$, then*

$$\int_a^b f(t)\Delta t = F(b) - F(a).$$

For computational purposes, it is often easiest to compute the Riemann integral, and so it is worth knowing necessary and sufficient conditions under which the Riemann Δ-integral will exist. The familiar condition is as follows:

Theorem 1.13 ([26]). *Let f be a bounded function defined on the finite closed interval $[a, b]$ of $\mathbb{T}$. Then f is Riemann Δ-integrable from a to b if and only if the set of all right-dense points of $[a, b)$ at which f is discontinuous is a set of Δ-measure zero.*

We conclude our remarks on the time scale integral with two theorems. The first one gives is the familiar parts formula and the second is an analogue of the Leibniz rule in a special case.

Theorem 1.14 (Integration by Parts, [8]). *Let u and v be continuous functions on $[a,b]$ that are Δ-differentiable on $[a,b)$. If u^Δ and v^Δ are integrable from a to b, then*

$$\int_a^b u^\Delta(t)v(t)\Delta t + \int_a^b u^\sigma(t)v^\Delta(t)\Delta t = u(b)v(b) - u(a)v(a).$$

Theorem 1.15 (Leibniz Rule, [8]). *If f and f^{Δ_t} are continuous, then we have the following:*

$$\left[\int_a^t f(t,s)\Delta s\right]^{\Delta_t} = f(\sigma(t),t) + \int_a^t f^{\Delta_t}(t,s)\Delta s.$$

1.3 The Time Scale Exponential Function

With integration now defined, we can examine the exponential in both the scalar and matrix cases for the time scales $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. On $\mathbb{R}$, where $\mu \equiv 0$, the Δ-integral is the usual Lebesgue integral, so that in the scalar case,

$$e_p(t,t_0) = \lim_{\mu \to 0^+} \exp\left(\int_{t_0}^t \frac{\log(1+\mu(\tau)p(\tau))}{\mu(\tau)}\Delta\tau\right) = \exp\left(\int_{t_0}^t p(\tau)d\tau\right),$$

while in the matrix case for A constant, $e_A(t,t_0) = e^{A(t-t_0)}$. On $\mathbb{Z}$, where $\mu \equiv 1$, the scalar exponential is

$$e_p(t,t_0) = \exp\left(\int_{t_0}^t \log(1+p(\tau))\Delta\tau\right) = \exp\left(\sum_{t_0}^{t-1} \log(1+p(\tau))\right) = \prod_{\tau=t_0}^{t-1}(1+p(\tau)),$$

and in the matrix case for A constant, $e_A(t,t_0) = (I+A)^{(t-t_0)}$. Notice that in the scalar case, if α is constant, then $e_\alpha(t,t_0) = e^{(\alpha(t-t_0))}$ and $e_\alpha(t,t_0) = (1+\alpha)^{(t-t_0)}$, respectively.

We now return to the fundamental problem of finding solutions to dynamic equations on time scales. We begin with a theorem on properties of the system transition matrix $\Phi_A(t,t_0)$. (Note that these properties also hold for $e_A(t,t_0)$ since this matrix is simply the transition matrix when A is time invariant.) In what follows, for a matrix A, A^* denotes the conjugate transpose of A.

Theorem 1.16 ([8]). *If $A, B \in \mathcal{R}$ are matrix-valued functions on $\mathbb{T}$, then*

(i) $\Phi_0(t,s) \equiv I$ and $\Phi_A(t,t) \equiv I$;

(ii) $\Phi_A(\sigma(t),s) = (I + \mu(t)A(t))\Phi_A(t,s)$;

(iii) $\Phi_A^{-1}(t,s) = \Phi^*_{\ominus A^*}(t,s)$;

(iv) $\Phi_A(t,s) = \Phi_A^{-1}(s,t) = \Phi^*_{\ominus A^*}(s,t)$;

(v) $\Phi_A(t,s)\Phi_A(s,r) = \Phi_A(t,r)$;

(vi) $\Phi_A(t,s)\Phi_B(t,s) = \Phi_{A\oplus B}(t,s)$ *if* $\Phi_A(t,s)$ *and* $B(t)$ *commute.*

Theorem 1.17 ([8]). *If $A \in \mathcal{R}$ and $a, b, c \in \mathbb{T}$, then*

$$[\Phi_A(c,\cdot)]^\Delta = -[\Phi_A(c,\cdot)]^\sigma A$$

and

$$\int_a^b \Phi_A(c,\sigma(t))A(t)\Delta t = \Phi_A(c,a) - \Phi_A(c,b).$$

With this foundation, we can present a useful result for solving first order dynamic IVPs:

Theorem 1.18 (Variation of Constants, [8]). *Let $A \in \mathcal{R}$ be an $n \times n$-matrix-valued function on $\mathbb{T}$ and suppose that $f : \mathbb{T} \to \mathbb{R}^n$ is rd-continuous. Let $t_0 \in \mathbb{T}$ and $y_0 \in \mathbb{R}^n$. Then the initial value problem*

$$y^\Delta(t) = A(t)y(t) + f(t), \;\; y(t_0) = y_0,$$

has a unique solution $y : \mathbb{T} \to \mathbb{R}^n$. Moreover, this solution is given by

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

The next concept that we will use frequently in later chapters is the notion of *exponential stability.* We will begin with the results obtained by Pötzsche, Siegmund, and Wirth in [38]. In their work, they define exponential stability as follows:

Definition 1.1 ([38]). For $t, t_0 \in \mathbb{T}$ and $x_0 \in \mathbb{R}^n$, the system

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R}^{n \times n}), \tag{1.1}$$

is said to be

(i) *exponentially stable* if there exists a constant $\alpha > 0$ such that for every $t_0 \in \mathbb{T}$ there exists a $K = K(t_0) \geq 1$ with

$$||\Phi_A(t, t_0)|| \leq Ke^{-\alpha(t-t_0)}, \text{ for } t \geq t_0,$$

(ii) *uniformly exponentially stable* if K can be chosen independently of t_0 in the definition of exponential stability,

(iii) *robustly exponentially stable* if there is an $\epsilon > 0$ such that the exponential stability of (1.1) implies the exponential stability of $x^{\Delta}(t) = B(t)x(t)$ for any rd-continuous $B : \mathbb{T} \to \mathbb{K}^{n \times n}$ with $\sup_{t \in \mathbb{T}} ||B(t) - A(t)|| \leq \epsilon$. (Here, $\mathbb{K}$ is the complex or real field.) In particular, if A is constant, we call (1.1) *robustly exponentially stable* if for all matrices B in a suitable neighborhood of A the corresponding system is exponentially stable.

Pötzsche, Siegmund, and Wirth show that the three definitions are in fact necessary: that is, the three notions do not have to coincide with each other even in the time invariant case. They then proceed to prove a very powerful theorem in time scale stability theory:

Theorem 1.19 ([38]). *Let $\mathbb{T}$ be a time scale which is unbounded above and let $\lambda \in \mathbb{C}$. The scalar equation*

$$x^{\Delta}(t) = \lambda x(t), \quad x(t_0) = x_0, \quad \lambda \in \mathbb{C},$$

is exponentially stable if and only if one of the following conditions is satisfied for arbitrary $t_0 \in \mathbb{T}$:

(i) $\gamma(\lambda) := \limsup_{T\to\infty} \frac{1}{T-t_0}\int_{t_0}^{T} \lim_{s\searrow\mu(t)} \frac{\log|1+s\lambda|}{s}\Delta t < 0,$

(ii) $\forall T \in \mathbb{T} : \exists t \in \mathbb{T}$ *with* $t > T$ *such that* $1+\mu(t)\lambda = 0$*,*

where we use the convention $\log 0 = -\infty$ *in (i).*

In light of this theorem, they define the *set of exponential stability.*

Definition 1.2 ([38]). Given a time scale $\mathbb{T}$ which is unbounded above, we define for arbitrary $t_0 \in \mathbb{T}$,

$$\mathcal{S}_{\mathbb{C}}(\mathbb{T}) := \{\lambda \in \mathbb{C} : \limsup_{T\to\infty} \frac{1}{T-t_0}\int_{t_0}^{T} \lim_{s\searrow\mu(t)} \frac{\log|1+s\lambda|}{s}\Delta t < 0\},$$

and

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

Then the *set of exponential stability* for the time scale $\mathbb{T}$ is defined by

$$\mathcal{S}(\mathbb{T}) = \mathcal{S}_{\mathbb{C}}(\mathbb{T}) \cup \mathcal{S}_{\mathbb{R}}(\mathbb{T}).$$

This set can often be very difficult to compute for an arbitrary time scale $\mathbb{T}$. Thus, while in theory the theorem is strong, in practice it has limitations. With this difficulty in mind, Hoffacker and Gard proved in [22] that there is a particularly nice subset of $\mathcal{S}(\mathbb{T})$ to work with regardless of the time scale. We call this region the *Hilger circle*, denote it by $\mathbb{H}$, and define the region as

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

(Observe that in this region, the Hilger real part of z is negative.) Strictly speaking, this definition is an abuse of the language as the set $\mathbb{H}$ is the interior of the Hilger circle previously defined. It is worth noting if we choose $\lambda \in \mathcal{R}^+$, then by Theorem 1.6 and Theorem 1.7, the exponential will decay in a positive, monotonic manner to the zero state. If we choose $\lambda \in (\mathbb{H}_{\min} - \mathcal{R}^+) \cap \mathbb{R}$, where $\mathbb{H}_{\min}$ is the smallest Hilger circle (that is, the Hilger circle corresponding to $\mu_{\max}$), then the exponential will be real-valued, alternate in sign, and tend to the zero state. Every other $\lambda \in \mathbb{H}$ will cause the exponential to be complex-valued in general and go to the zero state as $t \to \infty$.

There are a few things worth noting here. First, the set $\mathcal{S}_{\mathbb{R}}(\mathbb{T})$ is really the set of nonregressivity for the exponential, and since we will only concern ourselves with the regressive case, we only need to focus on $\mathcal{S}_{\mathbb{C}}(\mathbb{T})$. Second, Pötzsche, Siegmund, and Wirth with this theorem have shown that elements of the stability set really only need to lie in the Hilger circle on average and not necessarily for all time t. Third, notice that the members of $\mathcal{S}(\mathbb{C})$ have negative real part of necessity since their Hilger real part is negative on average. (If the real part of λ were positive, then it would be impossible for the Hilger real part to be negative on average.) Finally, although not obvious, the stability region in general can be disconnected, but Pötzsche, Siegmund, and Wirth do show that connected components are in fact simply connected (see [38]).

Pötzsche, Siegmund, and Wirth next define a function $\lambda : \mathbb{T} \to \mathbb{R}$ to be *uniformly regressive* if there exists a $\gamma > 0$ such that

$$\gamma^{-1} \geq |1 + \mu(t)\lambda(t)| \text{ for } t \in \mathbb{T}.$$

This leads to the following result.

Theorem 1.20 ([38]). *Let $\mathbb{T}$ be a time scale that is unbounded above and let $A \in \mathbb{K}^{n\times n}$ be regressive. Then the following hold:*

(i) If the system (1.1) is exponentially stable, then $\operatorname{spec}(A) \subset \mathcal{S}_{\mathbb{C}}(\mathbb{T})$.

(ii) If all eigenvalues λ of A are uniformly regressive, and if $\operatorname{spec}(A) \subset \mathcal{S}_{\mathbb{C}}(\mathbb{T})$, *then (1.1) is exponentially stable.*

As is evidenced by the preceding discussion, Pötszhe, Siegmund, and Wirth only deal with the autonomous case, in which they show that eigenvalue placement is sufficient for exponential stability of the system. DaCunha, in [15], was more interested in the regressive time varying system

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}), \tag{1.2}$$

where eigenvalue placement in general is not enough to guarantee stability. However, he takes as his definition of exponential stability the special case of the stability set given by Pötzsche, Siegmund, and Wirth in which all eigenvalues are positively regressive, a restriction which always places the eigenvalues in the Hilger circle $\mathbb{H}$. This is evidenced in the following definition taken from [15]:

Definition 1.3 (DaCunha, [15]). The regressive time varying system

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}),$$

is called *uniformly exponentially stable* if there exist constants $\gamma, \lambda > 0$ with $-\lambda \in \mathcal{R}^+$ such that for any t_0 and $x(t_0)$, the corresponding solution satisfies

$$||x(t)|| \leq ||x(t_0)||\gamma e_{-\lambda}(t, t_0), \;\; t \geq t_0.$$

Thus, DaCunha's definition of exponential stability is much weaker than that of Pötzsche, Siegmund, and Wirth in the sense that the latter implies the former but not conversely. However, DaCunha's definition allows for many standard theorems concerning stability to follow through into the time scale case. DaCunha uses the second method of Lyapunov to gain his results. He begins by examining the derivative of the scalar function $||x(t)||^2$ which plays the role of the energy of the system. Upon taking the delta derivative of this function, DaCunha investigates the existence of a regressive symmetric matrix $Q(t)$ that will make the symmetric form $(xQ(t)x^T)^{\Delta}$ negative definite in turn giving stability of the system. This line of reasoning yields the following theorem.

Theorem 1.21 (Lyapunov Stability Criterion I, [15]). *The regressive time varying linear dynamic system*

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}),$$

is uniformly exponentially stable if there exists a symmetric matrix $Q(t) \in C^1_{\mathrm{rd}}(\mathbb{T}, \mathbb{R}^{n\times n})$ *such that for all* $t \in \mathbb{T}$

(i) $\eta I \leq Q(t) \leq \rho I$,

(ii) $A^T(t)Q(t) + (I + \mu(t)A^T(t))(Q^{\Delta}(t) + Q(t)A(t) + \mu(t)Q^{\Delta}(t)A(t)) \leq -\nu I$,

where $\nu, \eta, \rho > 0$ *and* $-\frac{\nu}{\rho} \in \mathcal{R}^+$.

Although DaCunha did not address the following versions of the previous theorem, their proofs are the same.

Theorem 1.22 (Lyapunov Stability Criterion II). *The regressive time varying linear dynamic system*

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}),$$

is uniformly exponentially stable if there exists a symmetric matrix $Q(t) \in C^1_{\mathrm{rd}}(\mathbb{T}, \mathbb{R}^{n\times n})$ *such that for all* $t \in \mathbb{T}$

(i) $\eta I \leq Q(t) \leq \rho I$,

(ii) $A^T(t)Q(\sigma(t))[I + \mu(t)A(t)] + Q^{\Delta}(t)[I + \mu(t)A(t)] + Q(t)A(t) \leq -\nu I$,

where $\nu, \eta, \rho > 0$ *and* $-\frac{\nu}{\rho} \in \mathcal{R}^+$.

Theorem 1.23 (Lyapunov Stability Criterion III). *The regressive time varying linear dynamic system*

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}),$$

is uniformly exponentially stable if there exists a symmetric matrix $Q(t) \in C^1_{\mathrm{rd}}(\mathbb{T}, \mathbb{R}^{n\times n})$ *such that for all* $t \in \mathbb{T}$

(i) $\eta I \leq Q(t) \leq \rho I$,

(ii) $\left[(I+\mu(t)A^T(t))Q(\sigma(t))(I+\mu(t)A(t))-Q(t)\right]/\mu(t) \leq -\nu I$,

where $\nu, \eta, \rho > 0$ *and* $-\frac{\nu}{\rho} \in \mathcal{R}^+$.

It is this last version of the result that we will need in the section on control. Note that if $\mu \to 0^+$, then in the limit, the expression

$$\left[(I+\mu(t)A^T(t))Q(\sigma(t))(I+\mu(t)A(t))-Q(t)\right]/\mu(t),$$

reduces to

$$A^T(t)Q(t)+Q(t)A(t)+Q'(t),$$

a familiar expression for stability in the continuous case. If $\mu \equiv 1$, the expression becomes

$$(I+A^T(t))Q(t+1)(I+A(t))-Q(t),$$

which is a shifted version of the corresponding result in the discrete case. Thus, the result does unify the two cases and extends them to other time scales as well.

We will also need the following results from DaCunha in [15] to prove various theorems in the chapter on control.

Theorem 1.24 ([15]). *The regressive dynamic equation*

$$x^\Delta(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n\times n}),$$

is uniformly exponentially stable if and only if there exist $\lambda, \gamma > 0$ *with* $-\lambda \in \mathcal{R}^+$ *such that*

$$||\Phi_A(t,t_0)|| \leq \gamma e_{-\lambda}(t,t_0),$$

for all $t \geq t_0$ *with* $t, t_0 \in \mathbb{T}$.

Theorem 1.25 ([15]). *Suppose there exists a constant α such that for all $t \in \mathbb{T}$, $||A(t)|| \leq \alpha$. Then the regressive linear state equation*

$$x^{\Delta}(t) = A(t)x(t), \quad x(t_0) = x_0, \quad A(t) \in \mathcal{R}(\mathbb{T}, \mathbb{R}^{n \times n}),$$

is uniformly exponentially stable if and only if there exists a finite $\beta > 0$ such that

$$\int_{\tau}^{t} ||\Phi_A(t, \sigma(s))||\Delta s \leq \beta,$$

for all t, τ with $t \geq \sigma(\tau)$.

The last two theorems of DaCunha that we need can be found in [17]. The first result relies on the time scale polynomials $h_k(t, 0)$. These functions are defined recursively for $k \in \mathbb{N}_0$ by

$$\begin{aligned} h_0(t,s) &\equiv 1 \text{ for all } s,t \in \mathbb{T}, \\ h_{k+1}(t,s) &= \int_s^t h_k(\tau,s)\Delta\tau \text{ for all } s,t \in \mathbb{T}. \end{aligned}$$

Theorem 1.26 ([17]). *Suppose that A is a constant matrix. Then the transition matrix for (1.1) is*

$$\Phi_A(t, t_0) \equiv e_A(t, t_0),$$

where the matrix exponential is defined by the power series

$$e_A(t, t_0) = \sum_{i=0}^{\infty} A^i h_i(t, t_0),$$

which converges uniformly on $[-T, T]_{\mathbb{T}}$ for any $T > 0$.

Theorem 1.27 ([17]). *For the system (1.1) with A constant, there exist scalar functions $\gamma_0(t, t_0), \ldots, \gamma_{n-1}(t, t_0) \in C^{\infty}_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$ such that the unique solution has representation*

$$e_A(t, t_0) = \sum_{i=0}^{n-1} A^i \gamma_i(t, t_0).$$

1.4 Other Time Scale Functions

We conclude this chapter by introducing a few of the elementary functions on arbitrary time scales and the dynamic equations that they solve. We have already discussed the exponential function and the first order IVP that it solves. We gave the definition of the polynomials $h_k(t, t_0)$ above. The critical thing to observe about these functions is that for each $k \in \mathbb{N}$, $h_k(t, t_0)$ is the unique solution to the IVP

$$y^{\Delta}(t) = h_{k-1}(t, t_0), \;\; y(t_0) = 0.$$

The *time scale trigonometric functions* $\sin_\alpha(t, t_0)$ and $\cos_\alpha(t, t_0)$, for $\alpha \in C_{\mathrm{rd}}$ and $\mu\alpha^2 \in \mathcal{R}$ are defined as

$$\cos_\alpha(t, t_0) = \frac{e_{i\alpha}(t, t_0) + e_{-i\alpha}(t, t_0)}{2} \;\text{ and }\; \sin_\alpha(t, t_0) = \frac{e_{i\alpha}(t, t_0) - e_{-i\alpha}(t, t_0)}{2i}.$$

Bohner and Peterson in [8] show that these functions form a fundamental solution set of the second order dynamic equation

$$y^{\Delta\Delta}(t) + \alpha^2 y(t) = 0.$$

As a consequence of the definition of the two functions, *Euler's formula* remains true on the arbitrary time scale: that is,

$$e_{i\alpha}(t, t_0) = \cos_\alpha(t, t_0) + i \sin_\alpha(t, t_0).$$

Further, the following theorem holds.

Theorem 1.28 ([8]). *Let $p \in C_{\mathrm{rd}}$. If $\mu p^2 \in \mathcal{R}$, then we have*

$$\cos_p^{\Delta}(t, t_0) = -p(t)\sin_p(t, t_0) \;\text{ and }\; \sin_p^{\Delta}(t, t_0) = p(t)\cos_p(t, t_0),$$

and

$$\cos_p^2(t, t_0) + \sin_p^2(t, t_0) = e_{\mu p^2}(t, t_0).$$

The *time scale hyperbolic functions* $\sinh_\alpha(t, t_0)$ and $\cosh_\alpha(t, t_0)$, for $\alpha \in C_{\mathrm{rd}}$ and $\mu\alpha^2 \in \mathcal{R}$ are defined as

$$\cosh_\alpha(t, t_0) = \frac{e_\alpha(t, t_0) + e_{-\alpha}(t, t_0)}{2} \text{ and } \sin_\alpha(t, t_0) = \frac{e_\alpha(t, t_0) - e_{-\alpha}(t, t_0)}{2}.$$

The hyperbolic functions will form a fundamental solution set for the equation

$$y^{\Delta\Delta}(t) - \alpha^2 y(t) = 0.$$

Their calculus is as follows.

Theorem 1.29 ([8]). *Let $p \in C_{\mathrm{rd}}$. If $-\mu p^2 \in \mathcal{R}$, then we have*

$$\cosh_p^\Delta(t, t_0) = p(t)\sinh_p(t, t_0) \text{ and } \sinh_p^\Delta(t, t_0) = p(t)\cosh_p(t, t_0),$$

and

$$\cosh_p^2(t, t_0) - \sinh_p^2(t, t_0) = e_{-\mu p^2}(t, t_0).$$

With the necessary time scale preliminaries now established, we are in position to begin with the first focus of this dissertation which is the Laplace transform.

CHAPTER TWO

Laplace Transform

2.1 An Overview of the Cases $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$

The Laplace transform is a tool that has been to study differential equations for almost two centuries, although when Laplace first used the transform he did not use it in this fashion. (It is not commonly known that it was actually Euler who discovered it, and Lagrange who fitted the integral into probability theory for which Laplace used it.) The tool has become quite popular in engineering because of its ease of use and utility in understanding and manipulating LTI systems. Indeed, the transform will "algebratize" the problem in the sense that it allows the analyst to understand sophisticated phenomena occurring by examining what is happening in the frequency domain through investigating the system's transfer function rather than thinking about solutions in the state space, which is given in the time domain. This allows one to design systems with favorable attributes such as stability in the frequency domain by simply effecting things such as pole placement.

In the continuous or (in the engineering vernacular) *analogue* case, the (unilateral) Laplace transform of $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$\mathcal{L}\{f\}(z) = F(z) = \int_0^\infty e^{-zt} f(t)dt, \tag{2.1}$$

where $z \in \mathbb{C}$ is chosen so that the integral converges absolutely. It is known (see for example [5]) that a sufficient condition for the integral to converge is for f to be piecewise continuous of exponential order with constant $c > 0$, in which case if we choose any $z \in \mathbb{C}$ with $\mathrm{Re}(z) > c$, then the integral will converge absolutely. If we use the definition of the transform and integration by parts, then it is easy to show that

$$\mathcal{L}\{f'\}(z) = z\mathcal{L}\{f\}(z) - f(0),$$

and using an induction argument,

$$\mathcal{L}\{f^{(n)}\}(z) = z^n\mathcal{L}\{f\}(z) - z^{n-1}f(0) - z^{n-2}f'(0) - \cdots - f^{(n-1)}(0).$$

Of course, it is this structure that allows one to solve an n-th order differential equation with constant coefficients since the transform is a linear operator. Transforms of many of the elementary functions commonly studied in calculus are readily computed, and so if one is justified in associating a function with its transform, then one can obtain a solution to the equation in question quite easily.

Surprisingly, the question of the uniqueness of the inverse took a century to answer. It was not until 1916 when Thomas John I'Anson Bromwich in [11] arrived at his now famous result after making the methods of Oliver Heaviside (see [27], [28], and [29]) rigorous in terms of the Laplace integral that this problem was solved when the right meaning of "uniqueness" was given. In particular, Bromwich was familiar with Lebesgue's work in measure theory and the integral, and so Bromwich knew that for f and g differing on a set of measure zero, then it is possible for f and g to have the same transform. However, if we redefine uniqueness to mean that $f = g$ almost everywhere, then Bromwich showed that for any analytic function $F(z)$ in the half-plane $\mathrm{Re}(z) > c$ for some real $c > 0$, we have that if the integral

$$\mathcal{L}^{-1}\{F\} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{zt}F(z)dz$$

converges absolutely, then there exists some real-valued piecewise continuous function $f : \mathbb{R} \to \mathbb{R}$ with $f(t)$ of exponential order with constant $c > 0$ having Laplace transform of $F(z)$. (For the reader interested in the history of the transform, see [30] and the references contained therein.)

In this setting, Bromwich indeed showed that it is permissible to associate a function with its transform. The inversion process is a linear operator, and so in algebraic terms there is one fundamental question remaining: what function has transform $F(z)G(z)$? The answer is clearly not the product of f and g, since we know that the integral does not evolve in this fashion. The answer to the question lies in the convolution product $f * g$, which is given by

$$(f * g)(t) = \int_0^t f(\tau)g(t-\tau)d\tau.$$

It is straightforward to prove that this product is commutative and associative. It is also well known that the product has an identity element vested in the Dirac delta functional, which as its name suggests, lives in the space of distributions rather than the space of functions.

By the late 1940s/early 1950s, the electrical engineering community began investigating a discrete or *digital* version of the (unilateral) Laplace transform; that is, when the domain of the function is not $\mathbb{R}$ but rather $\mathbb{Z}$. The first appearance of the transform in its modern form was given by Jury in [34]. This transform is now commonly known as the (unilateral) Z*-transform* and is given by

$$Z\{f\}(z) = F(z) = \sum_{k=0}^{\infty} f(k)z^{-k}.$$

The Z-transform is a linear operator as well, and the shifts of functions have transforms that are related to the transforms of the unshifted functions, much in the same manner as the Laplace transforms of derivatives are related the transforms of the original functions in the continuous case. Indeed,

$$\begin{aligned} Z[f(k-1)] &= z^{-1}Z[f(k)], \\ Z[f(k+1)] &= zZ[f(k)] - zf(0). \end{aligned}$$

It is well known (see for example [36]) that if $|f(t)| \leq K\alpha^t$ for $K, \alpha > 0$ constant, then the Z-transform of f will exist. By using the shifting property of the transform previously mentioned, we can solve constant coefficient linear difference equations, much in the same manner as we do in using the Laplace transform to solve differential equations. As for uniqueness in the discrete case, for the unilateral transform it can be shown that if f and g have the same transform, then $f = g$. The Bromwich inversion formula in this case becomes

$$f(k) = Z^{-1}\{F(z)\} = \frac{1}{2\pi i}\oint_C F(z)z^{k-1}dz,$$

where C is a counterclockwise path encircling the origin and lies entirely in the region of convergence of $F(z)$. (If f is bounded as above, then the region of convergence will be $|z| > \alpha$.)

Again, it is worth asking what function has transform $F(z)G(z)$. The answer is found in the convolution product. For $f, g : \mathbb{N}_0 \to \mathbb{R}$, their convolution product is defined as

$$(f * g)(k) = \sum_{j=0}^{k} f(k-j)h(j).$$

We then have

$$Z[f * g] = Z[f]Z[g].$$

The convolution product on $\mathbb{Z}$ is commutative and associative as well, but the product has an identity that is a function given by the Kronecker delta function.

Furthermore, there are some striking similarities between the two transforms. For utility purposes, tables are commonly used in both cases to associate a function with its inverse instead of using Bromwich's inversion integrals, which can be complicated to compute in general. However, even though there are nice relationships between the two transforms, the tables for "corresponding" functions can be quite different. For example, the Laplace transform of the function $f(t) = \cos(\alpha t)$ is $F(z) = \frac{z}{z^2+\alpha^2}$. The same function has a z-transform of $Z(z) = \frac{1-z^{-1}\cos(\alpha)}{1-2z^{-1}-\cos(\alpha)+z^{-2}}$. Thus, the same function has drastically different transforms corresponding to the two different domains.

A natural question here is: Does there exist one transform that will give rise to the usual Laplace transform when the domain is $\mathbb{R}$ and to the Z-transform when the domain is $\mathbb{Z}$? That is, is it possible to *unify* these two cases? Secondly, we would certainly like a transform that would work on any domain that we choose between these extremes, so that we can *extend* the transform as well to an arbitrary time scale.

2.2 Introduction to the Arbitrary Time Scale Setting

In their initial work on the subject, Bohner and Peterson [10] define the Laplace transform of the time scale function f as follows:

Definition 2.1 ([10]). For $f : \mathbb{T} \to \mathbb{R}$, the *Laplace transform of* f, denoted by $\mathcal{L}\{f\}$ or $F(z)$, is given by

$$\mathcal{L}\{f\}(z) = F(z) = \int_0^\infty f(t) g^\sigma(t)\, \Delta t, \tag{2.2}$$

where $g(t) = e_{\ominus z}(t, 0)$.

We will assume that $\mathbb{T}$ is unbounded above and $0 \in \mathbb{T}$. Bohner and Peterson go on to state that the transform is defined for some appropriate collection of complex numbers $z \in \mathbb{C}$ for which the integral converges and give no inversion formula for the transform. Instead, they give the transform of the elementary functions and show the uniqueness of their transforms by using the uniqueness of solutions to ordinary dynamic equations. Hence, the method is actually a formal one.

We wish to justify the method. Thus, the goal of this chapter is first to quantify a subset of the complex plane for which the integral converges; that is, establish a *region of convergence* in the complex plane for (2.2). Furthermore, we provide a relatively simple inversion formula and show that inverses are uniquely determined by it. Finally, a notion of convolution arises and its algebraic structure is explored. In this process, the identity element is determined to be the appropriate analogue of the Dirac delta functional.

Before we begin with the justification, it is worth examining the definition for the cases $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. If $\mathbb{T} = \mathbb{R}$, then we know that $\sigma(t) = t$, $e_{\ominus z}(t, 0) = e^{-zt}$

and the delta integral is the usual continuous Lebesgue integral. Thus, we have that

$$\mathcal{L}_{\mathbb{T}}\{f\} = \int_0^\infty e^\sigma_{\ominus z}(t,0)\Delta t = \int_0^\infty e^{-zt} f(t)dt = \mathcal{L}_{\mathbb{R}}\{f\},$$

so that for $\mathbb{T} = \mathbb{R}$ the time scale version agrees with the usual version of the Laplace transform.

For $\mathbb{T} = \mathbb{Z}$, we have $\sigma(t) = t+1$, $e_{\ominus z}(t,0) = (1+z)^{-t}$, and the delta integral in this case is simply a sum. Therefore, it follows that

$$\mathcal{L}_{\mathbb{T}}\{f\} = \int_0^\infty e^\sigma_{\ominus z}(t,0) f(t)\Delta t = \sum_{t=0}^{\infty} (1+z)^{-(t+1)} f(t) = \frac{Z[f](z+1)}{z+1},$$

so that for $\mathbb{T} = \mathbb{Z}$, the time scale Laplace transform is not exactly the Z-transform of f, but a shift of it. These computations show that if we can justify the method in general, we will have accomplished our goal of unifying the two cases. However, we can also extend the results to any other time scale $\mathbb{T}$ (with bounded graininess) that we wish as our justification will not rely upon the underlying domain.

One question that arises is: Does the time scale Laplace transform preserve an algebraic structure on derivatives? The answer to this question is affirmative since an application of the time scales parts formula (Theorem 1.14) reveals

$$\mathcal{L}\{f^{\Delta^n}\} = \int_0^\infty e^\sigma_{\ominus z}(t,0) f^{\Delta^n}(t)\Delta t = z^n \mathcal{L}\{f\}(z) - \sum_{j=0}^{n-1} z^j f^{\Delta^{n-j-1}}(0),$$

provided the integral converges. For completeness, we state the state and prove the preceding result for the case $n = 1$, which can be found in [8] and [10]. An induction argument then shows the result for arbitrary $n \in \mathbb{N}$. (For the theorem, we need for x^Δ to be *regulated*. A time scale function is said to be regulated if all limits exist as finite numbers at left-dense and right-dense points.) To prove the result, we shall need a lemma whose proof we shall not give here, but it can be found in the aforementioned works.

Lemma 2.1 ([10]). *If $z \in \mathbb{C}$ is regressive, then*

$$e^\sigma_{\ominus z}(t,0) = -\frac{(\ominus z)(t)}{z} e_{\ominus z}(t,0).$$

Theorem 2.1 (Transform of Derivatives, [10]). *Assume* $x : \mathbb{T} \to \mathbb{C}$ *is such that* x^Δ *is regulated. Then*

$$\mathcal{L}\{x^\Delta\}(z) = z\mathcal{L}\{x\}(z) - x(0)$$

for those regressive $z \in \mathbb{C}$ *satisfying*

$$\lim_{t\to\infty} \{x(t)e_{\ominus z}(t, 0)\} = 0.$$

Proof. As previously mentioned, by the parts formula and Lemma 2.1,

$$\begin{aligned}
\mathcal{L}\{x^\Delta\}(z) &= \int_0^\infty x^\Delta(t)e^\sigma_\ominus(t, 0)\Delta t \\
&= [x(t)e_{\ominus z}(t, 0)]_{t=0}^{t\to\infty} - \int_0^\infty x(t)(\ominus z)(t)e_{\ominus z}(t, 0)\Delta t \\
&= -x(0) + z\int_0^\infty x(t)e^\sigma_{\ominus z}(t, 0)\Delta t \\
&= z\mathcal{L}\{x\}(z) - x(0),
\end{aligned}$$

provided the integral converges, which will happen when x^Δ is regulated. □

It is most likely puzzling at first glance as to why the same algebraic structure is preserved for the shifted version of the transform on $\mathbb{Z}$ compared with the usual Z-transform in this domain. The answer lies in the fact that the Z-transform is usually applied to the *recursive* form of the equation rather than the *difference* form. The time scale analysis always deals with the difference form, and so any transform that we define should take this information into account, as Bohner and Peterson's definition does for the Laplace transform.

We begin by giving a sufficient condition that characterizes those functions that are transformable. We have already seen that on $\mathbb{R}$, the functions of exponential order are a sufficient class for this purpose, while on $\mathbb{Z}$, a sufficient class is given by the class of functions $f(t)$ with $|f(t)| \leq K\alpha^t$. Thus, in both of the known cases, functions that are bounded by exponential functions are transformable. As the next definition and theorem show, this is in fact true in general.

Definition 2.2. The function $f : \mathbb{T} \to \mathbb{R}$ is said to be of *exponential type I* if there exists constants $M, c > 0$ such that $|f(t)| \leq Me^{ct}$. Furthermore, f is said to be of *exponential type II* if there exists constants $M, c > 0$ such that $|f(t)| \leq Me_c(t, 0)$.

The time scale exponential function itself is type II. In their work on the stability of the time scale exponential function, Pötzsche, Siegmund, and Wirth [38] show that the time scale polynomials $h_k(t, 0)$ are type I. It turns out that as an easy application of Theorem 1.26, one can show that type II functions are in fact type I.

Recall that throughout this work, we will assume that $\mathbb{T}$ is a time scale that is unbounded above with bounded graininess, that is, $\mu_{\min} \leq \mu(t) \leq \mu_{\max} < \infty$ for all $t \in \mathbb{T}$. We set $\mu_{\min} = \mu_*$ and $\mu_{\max} = \mu^*$.

To give an appropriate domain for the transform, which of course is tied to the region of convergence (ROC) of the integral in (2.2), for any $c > 0$ define the set

$$D = \{z \in \mathbb{C} : \mathrm{Re}_\mu(z) > \mathrm{Re}_\mu(c) \text{ for all } t \in \mathbb{T}\}.$$

Notice that this set is nonempty since the collection of all $z \in \mathbb{C}$ with $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(c)$ is a nonempty subset of D. (This last statement follows since for fixed z, the function $f(\mu) = \mathrm{Re}_\mu(z)$ is an increasing function of μ. In particular, the set of all complex z with $\mathrm{Re}(z) > c$ is a subset of the collection of all z with $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(c)$.) Note that if $\mu_* = 0$, then this set is a right half plane; see Figure 2.1. In fact, there are a couple of equivalent formulations of the set D. Recall from Chapter 1 that under the assumption that $\mathbb{T}$ has bounded graininess, Pötzsche, Siegmund, and Wirth [38] as well as Hoffacker and Gard [22] show that by choosing $\lambda \in \mathbb{H}$, where $\mathbb{H}$ denotes the *Hilger circle* given by

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

we obtain $\lim_{t\to\infty} e_\lambda(t, 0) = 0$. This limit condition will play a crucial role in the analysis of our transform. (Note, however, that the expression for the transform

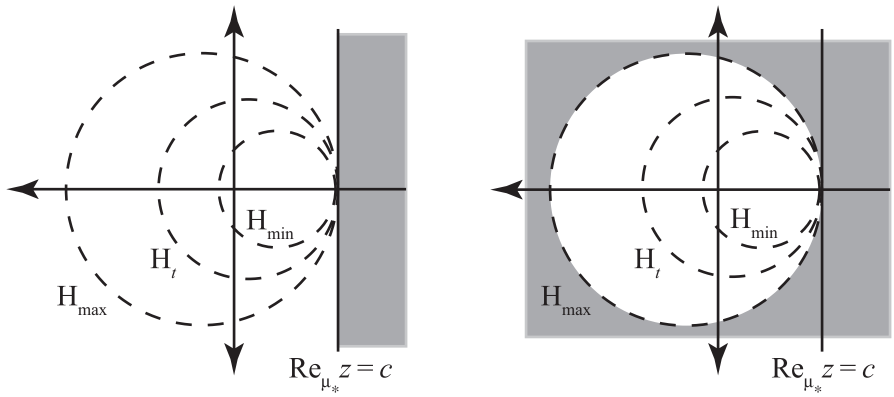

Figure 2.1: The region of convergence is shaded. On the left, the $\mu_* = 0$ case. On the right, the $\mu_* \neq 0$ case. In the latter, note our proof of the inversion formula is only valid for $\operatorname{Re} z > c$, i.e. the right half plane bounded by this abscissa of convergence even though the region of convergence is clearly a superset of this right half plane.

has a slight complication since the function $\ominus z$ is time varying and not constant. Fortunately, this is only a minor problem to overcome as we shall see.) One can characterize D in the following ways.

$$\begin{aligned} D &= \{z \in \mathbb{C} : \operatorname{Re}_\mu(z) > \operatorname{Re}_\mu(c) \text{ for all } t \in \mathbb{T}\} \\ &= \{z \in \mathbb{C} : z \in \overline{\mathbb{H}}^{\mathsf{C}}_{\max} \text{ and } z \text{ satisfies } \operatorname{Re}_{\mu_*}(z) > \operatorname{Re}_{\mu_*}(c)\} \\ &= \{z \in \mathbb{C} : \ominus z \in \mathbb{H} \text{ and } \operatorname{Re}_\mu(z) > \operatorname{Re}_\mu(c) \text{ for all } t \in \mathbb{T}\}, \end{aligned}$$

where $\overline{\mathbb{H}}^{\mathsf{C}}_{\max}$ denotes the complement of the closure of largest Hilger circle corresponding to μ_*; see Figure 2.2. This last equality is included to highlight the connection between the Hilger circle and the region of convergence. Furthermore, if $z \in D$, then $\ominus z \in \mathbb{H}_{\min} \subset \mathbb{H}_t$ since for all $z \in D$, $\ominus z$ satisfies the inequality

$$\left| \ominus z + \frac{1}{\mu^*} \right| < \frac{1}{\mu^*}.$$

The choice of D is not arbitrary. To make the integral converge, we must choose a region in which the exponential decays faster than the function being transformed grows. If f is of exponential type, then our claimed D is precisely the region in which this happens, as the next theorem establishes.

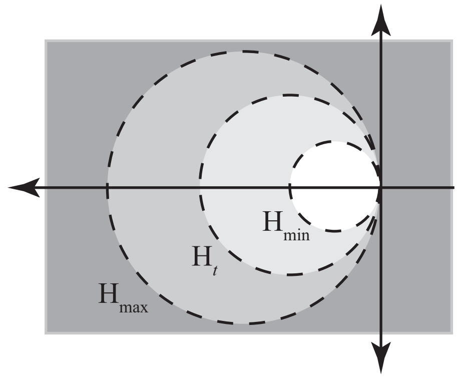

Figure 2.2: Time varying Hilger circles. The largest, $\mathbb{H}_{\max}$ has center $-1/\mu_{\min}$ while the smallest, $\mathbb{H}_{\min}$ has center $-1/\mu_{\max}$. In general, the Hilger circle at time t is denoted by $\mathbb{H}_t$ and has center $1/\mu(t)$. The exterior of each circle is shaded representing the corresponding regions of convergence (with respect to the transform).

Theorem 2.2 (Domain of the Laplace Transform). *The integral* $\int_0^\infty e^\sigma_{\ominus z}(t,0)f(t)\ \Delta t$ *converges absolutely for* $z \in D$ *if* $f(t)$ *is of exponential type II with exponential constant* c.

Proof. For $z \in D$, we have

$$\begin{aligned}
\left|\int_0^\infty e^\sigma_{\ominus z}(t,0)f(t)\,\Delta t\right| &\leq \int_0^\infty \left|e^\sigma_{\ominus z}(t,0)f(t)\right|\ \Delta t \\
&\leq M\int_0^\infty \left|\frac{1}{1+z\mu(t)}\right| \left|e_{\ominus z}(t,0)e_c(t,0)\right|\ \Delta t \\
&\leq \frac{M}{|1+\mu_* z|}\int_0^\infty \left|e_{\ominus z \oplus c}(t,0)\right|\ \Delta t \\
&= \frac{M}{|1+\mu_* z|}\int_0^\infty \exp\left(\int_0^t \frac{\log|1+\mu(\tau)(\ominus z\oplus c)|}{\mu(\tau)}\Delta\tau\right)\Delta t \\
&= \frac{M}{|1+\mu_* z|}\int_0^\infty \exp\left(\int_0^t \frac{\log\left|\frac{1+\mu(\tau)c}{1+\mu(\tau)z}\right|}{\mu(\tau)}\ \Delta\tau\right)\ \Delta t \\
&\leq \frac{M}{1+\mu_* c}\int_0^\infty e^{-\alpha t}dt \\
&\leq \frac{M}{\alpha},
\end{aligned}$$

where $\alpha = \left|\dfrac{\log\left|\frac{1+\mu_* c}{1+\mu^* z}\right|}{\mu^*}\right|$. ∎

The same estimates used in the proof of the preceding theorem can be used to show that if $f(t)$ is of exponential type II with constant c and $\mathrm{Re}_\mu(z) > \mathrm{Re}_\mu(c)$, then $\lim_{t\to\infty} e_{\ominus z}(t,0)f(t) = 0$.

With a ROC for the integral now defined, next we examine transforms of some of the elementary functions and their corresponding regions of convergence. We start with the exponential function itself. (Note that the computations that follow can all be found in [8] and [10]).

Example 2.1. Obviously, $e_\alpha(t,t_0)$ is of type II with a corresponding ROC for the integral given by $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(|\alpha|)$. Using the group property of the exponential function in the scalar case yields

$$\begin{aligned}
\mathcal{L}\{e_\alpha(t,0)\} &= \int_0^\infty e^\sigma_{\ominus z}(t,0)e_\alpha(t,0)\Delta t \\
&= \int_0^\infty \frac{1}{1+\mu(t)z} e_{\alpha\ominus z}(t,0)\Delta t \\
&= \frac{1}{\alpha - z}\int_0^\infty \frac{\alpha - z}{1+\mu(t)z} e_{\alpha\ominus z}(t,0)\Delta t \\
&= \frac{1}{\alpha - z}\int_0^\infty (\alpha\ominus z)(t) e_{\alpha\ominus z}(t,0)\Delta t \\
&= \frac{1}{\alpha - z}[e_{\alpha\ominus z}(t,0)]_{t=0}^{t\to\infty} \\
&= \frac{1}{z-\alpha},
\end{aligned}$$

with the integral converging in $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(|\alpha|)$.

Example 2.2. The transform of the function $f(t) \equiv 1$ can be found by using Lemma 2.1, once we note that this function is of exponential type II with ROC $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(0) = 0$. With this information, it then follows that

$$\begin{aligned}
\mathcal{L}\{1\}(z) &= \int_0^\infty 1\cdot e^\sigma_{\ominus z}(t,0)\Delta t \\
&= -\frac{1}{z}\int_0^\infty (\ominus z)(t)e_{\ominus z}(t,0)\Delta t \\
&= -\frac{1}{z}[e_{\ominus z}(t,0)]_{t=0}^{t\to\infty} \\
&= \frac{1}{z}, \quad \mathrm{Re}_{\mu_*}(z) > 0.
\end{aligned}$$

Before examining the transform of the time scale polynomials $h_k(t,0)$, we will need a theorem that relates the transform of the integral of a function with the transform of the function itself. To this end, the proof of Theorem 2.1 shows that the ROC of the derivative f^{Δ} will be the same as the ROC for f. Thus, the ROC of f^{Δ^n} will be the same as the ROC of $f^{\Delta^{n-1}}$ for all n, which implies that f and all of its derivatives will have the same ROC. Thus, we obtain the following.

Theorem 2.3 ([10]). *Assume $x : \mathbb{T} \to \mathbb{C}$ is regulated and transformable with region of convergence $\mathcal{C} \subset \mathbb{C}$. If*

$$X(t) := \int_0^t x(\tau)\Delta\tau,$$

for $t \in \mathbb{T}$, then

$$\mathcal{L}\{X\}(z) = \frac{1}{z}\mathcal{L}\{x\}(z),$$

for those regressive $z \neq 0 \in \mathcal{C}$.

Proof. Using the parts formula and Lemma 2.1, we obtain

$$\begin{aligned}
\mathcal{L}\{X\}(z) &= \int_0^\infty X(t)e^{\sigma}_{\ominus z}(t,0)\Delta t \\
&= -\frac{1}{z}\int_0^\infty X(t)(\ominus z)(t)e_{\ominus z}(t,0)\Delta t \\
&= -\frac{1}{z}\left\{ -X(0) - \int_0^\infty x(t)e^{\sigma}_{\ominus z}(t,0)\Delta t \right\} \\
&= \frac{1}{z}\mathcal{L}\{x\}(z),
\end{aligned}$$

provided $z \neq 0 \in \mathcal{C}$. □

We will return to this result when we talk about convolution and give an easier proof of it.

Example 2.3. For the time scale polynomials $h_k(t,0), k \in \mathbb{N}_0$, we have

$$\mathcal{L}\{h_k(t,0)\}(z) = \frac{1}{z^{k+1}}, \tag{2.3}$$

for all regressive z in the region $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(0) = 0$. This ROC comes from noting that for $t \geq 0, h_k(t,0)$ is of exponential type II with constant c for all $c > 0$. This can be shown either by using the time scales version of L'Hospital's Rule (see [8]), or by using Theorem 1.26 in the scalar case. Using the latter, we see that for $c > 0$,

$$c^k h_k(t,0) \leq \sum_{i=0}^{\infty} c^i h_i(t,0) = e_c(t,0),$$

so that $h_k(t,0) \leq c^{-k} e_c(t,0)$. We prove this claim by induction. First note that for z in the ROC,

$$\lim_{t\to\infty} \{h_i(t,0)e_{\ominus z}(t,0)\} = 0,$$

for $0 \leq i \leq k$. We showed in Example 2.2 that the claim is true for $i = 0$. Now assume that $1 \leq i < k$ and (2.3) holds with i replaced by $i-1$. Then by Theorem 2.3,

$$\begin{aligned}
\mathcal{L}\{h_i(t,0)\}(z) &= \mathcal{L}\left\{\int_0^t h_{i-1}(\tau,0)\Delta\tau\right\}(z) \\
&= \frac{1}{z}\mathcal{L}\{h_{i-1}(t,0)\}(z) \\
&= \frac{1}{z^{i+1}}.
\end{aligned}$$

Thus, the claim follows for all z in the ROC given by $\mathrm{Re}_{\mu_*}(z) > 0$.

Example 2.4. Since we know the transform of $e_\alpha(t,0)$ by Example 2.1, we can use the linearity of the transform to find the transforms of the trigonometric and hyperbolic functions defined in Chapter 1. Our earlier remarks concerning the ROC of a function and the ROC of its derivative being equivalent imply that we only need to find the ROC for the functions $f(t) = \cos_\alpha(t,0)$ and $f(t) = \cosh_\alpha(t,0)$ since by Theorem 1.28 and Theorem 1.29, their derivatives are just the functions $f(t) = -\alpha\sin_\alpha(t,0)$ and $f(t) = \alpha\sinh_\alpha(t,0)$, respectively. Thus, we first note that an immediate consequence of Theorem 1.26 in the scalar case is that $|e_\alpha(t,0)| \leq e_{|\alpha|}(t,0)$. Hence,

$$|\cosh_\alpha(t,0)| = \left|\frac{e_\alpha(t,0) + e_{-\alpha}(t,0)}{2}\right| \leq e_{|\alpha|}(t,0),$$

and likewise

$$|\cos_\alpha(t,0)| = \left|\frac{e_{i\alpha}(t,0)+e_{-i\alpha}(t,0)}{2}\right| \leq e_{|\alpha|}(t,0).$$

Consequently, all four trigonometric and hyperbolic functions are of exponential type II with constant $|\alpha|$. Using the linearity of the transform, we see

$$\begin{aligned}
\mathcal{L}\{\cosh_\alpha(t,0)\}(z) &= \mathcal{L}\left\{\frac{e_\alpha(t,0)+e_{-\alpha}(t,0)}{2}\right\}(z) \\
&= \frac{1}{2}\left(\frac{1}{z-\alpha}+\frac{1}{z+\alpha}\right) \\
&= \frac{z}{z^2-\alpha^2}, \\
\mathcal{L}\{\cos_\alpha(t,0)\}(z) &= \mathcal{L}\left\{\frac{e_{i\alpha}(t,0)+e_{-i\alpha}(t,0)}{2}\right\} \\
&= \frac{1}{2}\left(\frac{1}{z-i\alpha}+\frac{1}{z+i\alpha}\right) \\
&= \frac{z}{z^2+\alpha^2}, \\
\mathcal{L}\{\sinh_\alpha(t,0)\}(z) &= \mathcal{L}\left\{\frac{e_\alpha(t,0)-e_{-\alpha}(t,0)}{2}\right\}(z) \\
&= \frac{1}{2}\left(\frac{1}{z-\alpha}-\frac{1}{z+\alpha}\right) \\
&= \frac{\alpha}{z^2-\alpha^2}, \\
\mathcal{L}\{\sin_\alpha(t,0)\}(z) &= \mathcal{L}\left\{\frac{e_{i\alpha}(t,0)-e_{-i\alpha}(t,0)}{2i}\right\}(z) \\
&= \frac{1}{2i}\left(\frac{1}{z-i\alpha}-\frac{1}{z+i\alpha}\right) \\
&= \frac{\alpha}{z^2+\alpha^2},
\end{aligned}$$

where all integrals converge in $\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(|\alpha|)$.

Example 2.5. We conclude our examples here with the transform of the *unit step function* or, as it is sometimes called, the *Heaviside function*. It is defined for $a > 0 \in \mathbb{T}$ by

$$u_a(t) = \begin{cases} 0, & \text{if } t \in \mathbb{T} \cap (-\infty, a), \\ 1, & \text{if } t \in \mathbb{T} \cap [a, \infty). \end{cases}$$

The unit step function is obviously of exponential type II with constant c for all $c > 0$. Thus, the transform is

$$\int_0^\infty u_a(t)e_{\ominus z}(\sigma(t),0)\Delta t = \int_a^\infty e_{\ominus z}(\sigma(t),0)\Delta t = \frac{e_{\ominus z}(a,0)}{z},$$

which holds for all z in the ROC which is given by $\mathrm{Re}_{\mu_*}(z) > 0$.

We conclude this section by summarizing the transforms in tabular form given in Table 2.1. There are a couple of things to notice here. First, at this point we are still lacking uniqueness of the transform so that right now we are *not* justified in reading the table as an association between the function and its transform. All we can do at this point is read the table as saying that for the given functions, applying the transform to those functions gives the results stated in the table. Second, notice that the table works for the given functions independent of the time scale involved. That is, we have one table for the cases $\mathbb{R}$ and $\mathbb{Z}$ rather than the two we get by using the usual Laplace and Z-transforms. We need no other table for any other time scale, and this is an idea we will revisit and try to understand in functional terms when we discuss the inversion formula. We give the explicit representations of the functions on $\mathbb{R}$ and $\mathbb{Z}$ as they are easily computed in these cases. Note that in doing so, we are demonstrating that these functions are *not* the same functions on all time scales.

2.2.1 Properties of the Transform

As we look towards an inverse for the transform, we would like to know which functions are the transform of some function. To answer this question, the following properties are needed. The reader familiar with the cases $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$ will note striking similarities of the corresponding result in each of these cases, both of which are special cases of this more general result.

Table 2.1. Laplace transforms of functions on $\mathbb{T}$ and their ROC.

$x_{\mathbb{T}}(t)$	$x_{\mathbb{R}}(t)$	$x_{\mathbb{Z}}(t)$	$\mathcal{L}\{x\}(z)$	ROC
1	1	1	$\frac{1}{z}$	$\mathrm{Re}_{\mu_*}(z) > 0$
t	t	t	$\frac{1}{z^2}$	$\mathrm{Re}_{\mu_*}(z) > 0$
$h_k(t,0),\ k \geq 0$	$\frac{t^k}{k!}$	$\binom{t}{k}$	$\frac{1}{z^{k+1}}$	$\mathrm{Re}_{\mu_*}(z) > 0$
$e_\alpha(t,0)$	$e^{\alpha t}$	$(1+\alpha)^t$	$\frac{1}{z-\alpha}$	$\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(\|\alpha\|)$
$\cosh_\alpha(t,0)$	$\cosh(\alpha t)$	$\frac{(1+\alpha)^t+(1-\alpha)^t}{2}$	$\frac{z}{z^2-\alpha^2}$	$\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(\|\alpha\|)$
$\sinh_\alpha(t,0)$	$\sinh(\alpha t)$	$\frac{(1+\alpha)^t-(1-\alpha)^t}{2}$	$\frac{\alpha}{z^2-\alpha^2}$	$\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(\|\alpha\|)$
$\cos_\alpha(t,0)$	$\cos(\alpha t)$	$\frac{(1+i\alpha)^t+(1-i\alpha)^t}{2}$	$\frac{z}{z^2+\alpha^2}$	$\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(\|\alpha\|)$
$\sin_\alpha(t,0)$	$\sin(\alpha t)$	$\frac{(1+i\alpha)^t-(1-i\alpha)^t}{2i}$	$\frac{\alpha}{z^2+\alpha^2}$	$\mathrm{Re}_{\mu_*}(z) > \mathrm{Re}_{\mu_*}(\|\alpha\|)$
$u_a(t), a \in \mathbb{T}$ (unit step)	$u_a(t)$	$u_a(t)$	$\frac{e_{\ominus z}(a,0)}{z}$	$\mathrm{Re}_{\mu_*}(z) > 0$

Theorem 2.4. *Let F denote the generalized Laplace transform for $f : \mathbb{T} \to \mathbb{R}$.*

(1) $F(z)$ is analytic in $\mathrm{Re}_\mu(z) > \mathrm{Re}_\mu(c)$.

(2) $F(z)$ is bounded in $\mathrm{Re}_\mu(z) > \mathrm{Re}_\mu(c)$.

(3) $\lim_{|z|\to\infty} F(z) = 0$.

Proof. For the first, we see

$$\begin{aligned}
\frac{d}{dz}\mathcal{L}\{f\}(z) &= \frac{d}{dz}\int_0^\infty e^\sigma_{\ominus z}(t,0)f(t)\,\Delta t \\
&= \int_0^\infty \frac{d}{dz}\left(\frac{1}{1+\mu(t)z}\exp\left(\int_0^t \frac{\log\left(\frac{1}{1+\mu(\tau)z}\right)}{\mu(\tau)}\Delta\tau\right)\right) f(t)\,\Delta t \\
&= \int_0^\infty \left(\int_0^t \frac{-1}{1+\mu(\tau)z}\Delta\tau - \frac{\mu(t)}{1+\mu(t)z}\right) e^\sigma_{\ominus z}(t,0)f(t)\,\Delta t \\
&= -\int_0^\infty \left(\int_0^{\sigma(t)} \frac{1}{1+\mu(\tau)z}\Delta\tau\right) e^\sigma_{\ominus z}(t,0)f(t)\Delta t \\
&= -\mathcal{L}\{gf\}(z),
\end{aligned}$$

where $g(t) = \int_0^{\sigma(t)} \frac{1}{1+\mu(\tau)z}\Delta\tau$. The second equation follows from the Lebesgue Dominated Convergence Theorem. Note that on $\mathbb{R}$, this calculation shows that we get the familiar formula that derivatives of the transform correspond to multiplication by powers of t in the function. On $\mathbb{Z}$, the calculations show that (in the shifted version) derivatives of the transform correspond to multiplication by powers of $t+1$ in the function.

The second claim is an immediate consequence of the preceding theorem since it shows $|F(z)| < \frac{M}{\alpha}$.

As for the third, a direct calculation yields

$$\lim_{|z|\to\infty} F(z) = \lim_{|z|\to\infty}\int_0^\infty e^\sigma_{\ominus z}(t,0)f(t)\,\Delta t = \int_0^\infty \lim_{|z|\to\infty} e^\sigma_{\ominus z}(t,0)f(t)\,\Delta t = 0,$$

which follows again from the Lebesgue Dominated Convergence Theorem. $\square$

Theorem 2.5 (Initial and Final Values). *Let $f : \mathbb{T} \to \mathbb{R}$ have generalized Laplace transform $F(z)$. Then $f(0) = \lim_{z\to\infty} zF(z)$ and $\lim_{t\to\infty} f(t) = \lim_{z\to 0} zF(z)$ when the limits exist.*

Proof.

$$\begin{aligned}\mathcal{L}\{f^{\Delta}(t)\} &= \int_0^{\infty} f^{\Delta}(t)e^{\sigma}_{\ominus z}(t,0)\,\Delta t \\ &= f(t)e_{\ominus z}(t,0)\Big|_{t=0}^{\infty} - \int_0^{\infty} f(t)(\ominus z)(t)e_{\ominus z}(t,0)\,\Delta t \\ &= zF(z) - f(0).\end{aligned}$$

Now $z \to \infty$ above yields $\lim_{z\to\infty}\int_0^{\infty} f^{\Delta}(t)e^{\sigma}_{\ominus z}(t,0)\,\Delta t = 0 = \lim_{z\to\infty}\left[zF(z) - f(0)\right]$, i.e., $f(0) = \lim_{z\to\infty} zF(z)$.

On the other hand, $z \to 0$ yields

$$\lim_{z\to 0}\int_0^{\infty} f^{\Delta}(t)e^{\sigma}_{\ominus z}(t,0)\,\Delta t = \int_0^{\infty} f^{\Delta}(t)\,\Delta t = \lim_{t\to\infty} f(t) - f(0) = \lim_{z\to 0}\left[zF(z) - f(0)\right],$$

i.e., $\lim_{t\to\infty} f(t) = \lim_{z\to 0} zF(z)$. □

2.2.2 Inversion Formula

Using Theorem 2.4 we can establish an inversion formula for the transform. As is the case with $\mathbb{T} = \mathbb{R}$, these properties are not sufficient to guarantee that $F(z)$ is the transform of some continuous function $f(t)$, but they are necessary as we have just seen. For sufficiency, we have the following:

Theorem 2.6 (Inversion of the Transform). *Let $F_{\mathbb{T}}(z)$ be a complex valued function of a complex variable that satisfies the following.*

(1) $F_{\mathbb{T}}(z)$ *is analytic in the region* $\mathrm{Re}_{\mu}(z) > \mathrm{Re}_{\mu}(c)$.

(2) $F_{\mathbb{T}}(z) \to 0$ *uniformly as* $|z| \to \infty$ *in the region* $\mathrm{Re}_{\mu}(z) > \mathrm{Re}_{\mu}(c)$.

(3) $F_{\mathbb{T}}(z)$ *has finitely many regressive poles of finite order* $\{z_1, z_2, \ldots z_n\}$.

Further, let $F_{\mathbb{R}}(z)$ be the transform of the function $f_{\mathbb{R}}(t)$ that corresponds to the transform $F_{\mathbb{T}}(z)$ of $f_{\mathbb{T}}(t)$. If $\int_{c-i\infty}^{c+i\infty} |F_{\mathbb{R}}(z)|dz < \infty$, then

$$f_{\mathbb{T}}(t) = \sum_{k=1}^{n} \mathrm{Res}_{z=z_k}\, e_z(t,0)F_{\mathbb{T}}(z),$$

has transform $F_{\mathbb{T}}(z)$ for all z with $\mathrm{Re}(z) > c$.

Proof. The proof follows from the commutative diagram between the appropriate function spaces in Figure 2.3.

$$\begin{tikzcd}
C_{\text{p-eo}}(\mathbb{R},\mathbb{R}) \arrow[r, shift left, "\mathcal{L}_{\mathbb{R}}"] \arrow[d, shift right, "\theta"'] & \mathcal{C} \arrow[l, shift left, "\mathcal{L}_{\mathbb{R}}^{-1}"] \arrow[d, shift right, "\gamma"'] \\
C_{\text{prd-e2}}(\mathbb{T},\mathbb{R}) \arrow[r, shift left, "\mathcal{L}_{\mathbb{T}}"] \arrow[u, shift right, "\theta^{-1}"'] & \mathcal{D} \arrow[l, shift left, "\mathcal{L}_{\mathbb{T}}^{-1} = \theta \circ \mathcal{L}_{\mathbb{R}}^{-1} \circ \gamma^{-1}"] \arrow[u, shift right, "\gamma^{-1}"']
\end{tikzcd}$$

Figure 2.3. Commutative diagram between the function spaces.

Define the sets

$$\begin{aligned}
\mathcal{C} &:= \{F_{\mathbb{R}}(z) : F_{\mathbb{R}}(z) = G(z)e^{-z\tau}\}, \\
\mathcal{D} &:= \{F_{\mathbb{T}}(z) : F_{\mathbb{T}}(z) = G(z)e_{\ominus z}(\tau,0)\},
\end{aligned}$$

for G a rational function in z and for τ an appropriate constant. Let $C_{\text{p-eo}}(\mathbb{R},\mathbb{R})$ denote the space of piecewise continuous functions of exponential order, and $C_{\text{prd-e2}}(\mathbb{T},\mathbb{R})$ denote the space of piecewise right dense continuous functions of exponential type II.

We now examine the maps between these spaces shown in Figure 2.3. Each of θ, γ, θ^{-1}, γ^{-1} maps functions involving the continuous exponential to the time scale exponential and vice versa. For example, γ maps the function $F_{\mathbb{R}}(z) = \frac{e^{-za}}{z}$ to the function $F_{\mathbb{T}}(z) = \frac{e_{\ominus z}(a,0)}{z}$, while γ^{-1} maps $F_{\mathbb{T}}(z)$ back to $F_{\mathbb{R}}(z)$ in the obvious manner. If the representation of $F_{\mathbb{T}}(z)$ is independent of the exponential (that is, $\tau = 0$), then γ and its inverse will act as the identity. For example,

$$\gamma\left(\frac{1}{z^2+1}\right) = \gamma^{-1}\left(\frac{1}{z^2+1}\right) = \frac{1}{z^2+1}.$$

θ sends the continuous exponential function to the time scale exponential function in the following manner: if we write $f_{\mathbb{R}}(t) \in C_{\text{p-eo}}(\mathbb{R},\mathbb{R})$ as

$$f_{\mathbb{R}}(t) = \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} e^{zt} F_{\mathbb{R}}(z),$$

then

$$\theta(f_{\mathbb{R}}(t)) = \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} e_z(t,0) F_{\mathbb{T}}(z).$$

To go from $F_{\mathbb{R}}(z)$ to $F_{\mathbb{T}}(z)$, we simply switch expressions involving the continuous exponential in $F_{\mathbb{R}}(z)$ with the time scale exponential giving $F_{\mathbb{T}}(z)$ as was done for γ and its inverse. θ^{-1} will then act on the collection of all $g \in C_{\text{prd-e2}}(\mathbb{T},\mathbb{R})$ such that g can be written in the form

$$g(t) = \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} e_z(t,0) G_{\mathbb{T}}(z),$$

as

$$\theta^{-1}(g(t)) = \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} e^{zt} G_{\mathbb{R}}(z).$$

For example, for the unit step function $f_{\mathbb{R}}(t) = u_a(t)$, we know from the continuous result that we may write the step function as

$$f_{\mathbb{R}}(t) = u_a(t) = \operatorname{Res}_{z=0} e^{zt} \cdot \frac{e^{-az}}{z},$$

so that if $a \in \mathbb{T}$, then

$$\theta(u_a(t)) = \operatorname{Res}_{z=0} e_z(t,0) \frac{e_{\ominus z}(a,0)}{z}.$$

With these operators defined on these spaces, the claim in the theorem follows.

For a given time scale Laplace transform $F_{\mathbb{T}}(z)$, we begin by mapping to $F_{\mathbb{R}}(z)$ via γ^{-1}. The hypotheses on $F_{\mathbb{T}}(z)$ and $F_{\mathbb{R}}(z)$ are enough to guarantee the inverse of $F_{\mathbb{R}}(z)$ exists for all z with $\mathrm{Re}(z) > c$ (see [5]), and is given by

$$f_{\mathbb{R}}(t) = \sum_{k=1}^{n} \mathrm{Res}_{z=z_k}\, e^{zt} F_{\mathbb{R}}(z).$$

Apply θ to $f_{\mathbb{R}}(t)$ to retrieve the time scale function

$$f_{\mathbb{T}}(t) = \sum_{k=1}^{n} \mathrm{Res}_{z=z_k}\, e_z(t,0) F_{\mathbb{T}}(z),$$

whereby $(\gamma \circ \mathcal{L}_{\mathbb{R}} \circ \theta^{-1})(f_{\mathbb{T}}(t)) = F_{\mathbb{T}}(z)$ as claimed. $\square$

Before looking at a few examples, some remarks are in order. First, it is reasonable to ask if there is a contour in the complex plane around which it is possible to integrate to obtain the same results that we have obtained here through a more operational approach. *At present, we leave this as a very interesting although nontrivial open problem.* It is well known that there are such contours when $\mathbb{T} = \mathbb{R}$ or $\mathbb{T} = \mathbb{Z}$. In fact, it can easily be shown that if $\mathbb{T}$ is completely discrete, then if we choose any circle in the region of convergence which encloses all of the singularities of $F(z)$, we will obtain the inversion formula. However, in general, we do not know whether or not there exists a contour which gives the formula, and if so, what it actually is.

Second, it is possible to use the technique we have presented here to define and find inverses for any time scale transform. These would include the Fourier, Mellin, and many other transforms. Once the inverse is known for $\mathbb{T} = \mathbb{R}$ and the appropriate time scale integral is developed to give the correct transform analogues for any $\mathbb{T}$, the diagram becomes completed and the inversion formula for any $\mathbb{T}$ is readily obtained.

Finally, notice in our construction, for any transformable function $f_{\mathbb{T}}(t)$, there is a *shadow* function $f_{\mathbb{R}}(t)$. That is, to determine the appropriate *time scale analogue* of the function $f_{\mathbb{R}}(t)$ in terms of the transform, we use the diagram to map its Laplace transform on $\mathbb{R}$ to its Laplace transform on $\mathbb{T}$.

Before looking at the examples, it is worth noting that even with the inversion formula, we are still *not* justified in viewing Table 2.1 as association between functions and their transforms as we still have not established uniqueness of the inverse. However, we will do this in the next section.

Example 2.6. Let $F(z) = \frac{1}{z-\alpha}$. $F(z)$ is obviously analytic in $\mathrm{Re}_\mu(z) > \mathrm{Re}_\mu(|\alpha|)$, and it certainly tends to zero uniformly in this region. As the function $\frac{1}{z-\alpha}$ is independent of the exponential function, we see that the function $F_{\mathbb{R}}$ that corresponds to $F = F_{\mathbb{T}}$ is simply $F_{\mathbb{R}} = F$. The integral

$$\int_{|\alpha|-i\infty}^{|\alpha|+i\infty} \frac{1}{z-\alpha} dz$$

converges absolutely in $\mathrm{Re}(z) > |\alpha|$, and so we have an inverse of F for $z \in \mathrm{Re}(z) > |\alpha|$ and all regressive α in $\mathbb{C}$ given by

$$f(t) = \mathrm{Res}_{z=\alpha} \frac{e_z(t,0)}{z-\alpha} = e_\alpha(t,0).$$

Example 2.7 (New Representation for $h_k(t,0)$). Suppose $F(z) = \frac{1}{z^2}$. For this F, we have that F is analytic in $\mathrm{Re}_\mu(z) > 0$ and again tends to zero uniformly in this region. The correspondence in this case is given by $F_{\mathbb{R}} = F$, and since the integral

$$\int_{-i\infty}^{i\infty} \frac{1}{z^2} dz,$$

converges absolutely, F has an inverse for all z with $\mathrm{Re}(z) > 0$ given by

$$\mathcal{L}^{-1}\{F\} = f(t) = \mathrm{Res}_{z=0} \frac{e_z(t,0)}{z^2} = e_z(t,0) \int_0^t \frac{1}{1+\mu(\tau)z} \Delta\tau \bigg|_{z=0} = t.$$

Likewise, for $F(z) = \frac{1}{z^3}$, we have that F satisfies all of the conditions in the same region as above, with the complex integral involved converging absolutely. Thus, F

again has an inverse for all z with $\mathrm{Re}(z) > 0$ given by

$$\begin{aligned}
\mathcal{L}^{-1}\{F\} &= f(t) = \mathrm{Res}_{z=0} \frac{e_z(t,0)}{z^3} \\
&= \left. \frac{e_z(t,0)\left(\left(\int_0^t \frac{1}{1+\mu(\tau)z}\Delta\tau\right)^2 - \int_0^t \frac{\mu(\tau)}{1+\mu(\tau)z}\Delta\tau\right)}{2}\right|_{z=0} \\
&= \frac{t^2 - \int_0^t \mu(\tau)\Delta\tau}{2} = h_2(t,0).
\end{aligned}$$

The last equality is justified since the function

$$f(t) = \frac{t^2 - \int_0^t \mu(\tau)\Delta\tau}{2},$$

is the unique solution to the initial value problem $f^{\Delta}(t) = h_1(t,0),\ f(0) = 0$. In a similar manner, we can use an induction argument coupled with Theorem 2.6 to show that the inverse of $F(z) = \frac{1}{z^{k+1}}$, for k a positive integer, is $h_k(t,0)$.

Example 2.8. Now suppose F is one of the following: $\frac{z}{z^2-\alpha^2}, \frac{\alpha}{z^2-\alpha^2}, \frac{z}{z^2+\alpha^2}, \frac{\alpha}{z^2+\alpha^2}$. Each of these functions is analytic in the region $\mathrm{Re}_\mu(z) > \mathrm{Re}_\mu(|\alpha|)$ and approach zero as $|z| \to \infty$ in this region. $F_{\mathbb{R}} = F$ in each of these cases, and in each case the integral

$$\int_{|\alpha|-i\infty}^{|\alpha|+i\infty} F_{\mathbb{R}}(z)dz,$$

converges absolutely, so that each F has an inverse. If we use the linearity of the inverse operator and Example 2.6, then each inverse is given by

$$\begin{aligned}
f_1(t) &= \mathcal{L}^{-1}\{\frac{z}{z^2-\alpha^2}\} \\
&= \frac{1}{2}\left(\mathcal{L}^{-1}\left\{\frac{1}{z-\alpha}\right\} + \mathcal{L}^{-1}\left\{\frac{1}{z+\alpha}\right\}\right) \\
&= \frac{1}{2}\left(e_\alpha(t,0) + e_{-\alpha}(t,0)\right) \\
&= \cosh_\alpha(t,0),
\end{aligned}$$

$$\begin{aligned}
f_2(t) &= \mathcal{L}^{-1}\{\frac{\alpha}{z^2-\alpha^2}\} \\
&= \frac{1}{2}\left(\mathcal{L}^{-1}\left\{\frac{1}{z-\alpha}\right\} - \mathcal{L}^{-1}\left\{\frac{1}{z+\alpha}\right\}\right) \\
&= \frac{1}{2}\left(e_\alpha(t,0) - e_{-\alpha}(t,0)\right) \\
&= \mathrm{sinh}_\alpha(t,0), \\
f_3(t) &= \mathcal{L}^{-1}\{\frac{z}{z^2+\alpha^2}\} \\
&= \frac{1}{2}\left(\mathcal{L}^{-1}\left\{\frac{1}{z-i\alpha}\right\} + \mathcal{L}^{-1}\left\{\frac{1}{z+i\alpha}\right\}\right) \\
&= \frac{1}{2}\left(e_{i\alpha}(t,0) + e_{-i\alpha}(t,0)\right) \\
&= \cos_\alpha(t,0), \\
f_4(t) &= \mathcal{L}^{-1}\{\frac{\alpha}{z^2+\alpha^2}\} \\
&= \frac{1}{2i}\left(\mathcal{L}^{-1}\left\{\frac{1}{z-i\alpha}\right\} - \mathcal{L}^{-1}\left\{\frac{1}{z+i\alpha}\right\}\right) \\
&= \frac{1}{2i}\left(e_{i\alpha}(t,0) - e_{-i\alpha}(t,0)\right) \\
&= \sin_\alpha(t,0),
\end{aligned}$$

for all regressive $\alpha \in \mathbb{R}$ and $z \in \mathbb{C}$ with $\mathrm{Re}(z) > |\alpha|$.

Thus, the preceding examples show that in terms of the transform, the elementary functions defined in Chapter 1 all become the appropriate *time scale analogues* or *shadows* of their continuous counterparts.

Example 2.9. Suppose we would like to determine the appropriate shadow of the function

$$f_{\mathbb{R}}(t) = \frac{1}{2a} t \sin at, \quad a > 0.$$

At first glance, a reasonable guess for its time scale analogue might be the function

$$f_{\mathbb{T}}(t) = \frac{1}{2a} h_1(t,0) \sin_a(t,0).$$

However, closer inspection shows that this guess is in fact incorrect. To see this, note that the Laplace transform of $f_{\mathbb{R}}(t)$ is

$$F_{\mathbb{R}}(z) = \frac{z}{(z^2+a^2)^2}.$$

To find the the proper analogue $f_{\mathbb{T}}(t)$ of $f_{\mathbb{R}}(t)$, we search for a time scale function with the same transform as $f_{\mathbb{R}}(t)$. To do this, note that

$$\begin{aligned}
f_{\mathbb{T}}(t) &= \mathcal{L}_{\mathbb{T}}^{-1}\{F\} \\
&= \mathcal{L}_{\mathbb{T}}^{-1}\left\{\frac{z}{(z^2+a^2)^2}\right\} \\
&= \sum_{k=1}^{2} \mathrm{Res}_{z=z_k} \frac{ze_z(t,0)}{(z^2+a^2)^2} \\
&= \frac{1}{2a}\sin_a(t,0)\int_0^t \frac{1}{1+(\mu(\tau)a)^2}\Delta\tau - \frac{1}{2}\cos_a(t,0)\int_0^t \frac{\mu(\tau)}{1+(\mu(\tau)a)^2}\Delta\tau,
\end{aligned}$$

so that in general the correct analogue involves the cosine function as well.

Example 2.10. A useful Laplace transform property is the ability to compute the matrix exponential $e_A(t,0)$ when A is a constant matrix. This is a property that we will heavily exploit in the next chapter when discussing control. As in the discrete and continuous cases, $e_A(t,0)$ solves the initial value problem $Y^{\Delta} = AY$, $Y(0) = I$. Transforming yields $z\mathcal{L}\{Y\} - Y(0) = A\mathcal{L}\{Y\}$, so that $\mathcal{L}\{Y\} = (zI-A)^{-1}$, or equivalently $Y = e_A(t,0) = \mathcal{L}^{-1}\{(zI-A)^{-1}\}$.

For example, let $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$. Then

$$(zI-A)^{-1} = \begin{bmatrix} \frac{1}{(z-2)} & \frac{1}{(z-2)(z-3)} \\ 0 & \frac{1}{(z-3)} \end{bmatrix},$$

so that

$$e_A(t,0) = \begin{bmatrix} e_2(t,0) & e_3(t,0) - e_2(t,0) \\ 0 & e_3(t,0) \end{bmatrix}.$$

Example 2.11. Consider the function $F(z) = e_{\ominus z}(\sigma(a), 0)$. This function has no regressive poles, and hence Theorem 2.6 cannot be applied. If a is right scattered, the Hilger Delta function which has representation

$$\delta_a^{\mathbb{H}}(t) = \begin{cases} \frac{1}{\mu(a)}, & t = a, \\ 0, & t \neq a, \end{cases}$$

has $F(z)$ as a transform, while if a is right dense, the Dirac delta functional has $F(z)$ as a transform as we shall see later.

2.2.3 Uniqueness of the Inverse

If two functions f and g have the same (unilateral) transform, then are f and g necessarily the same function? We have already seen that on $\mathbb{R}$, the answer to this question is affirmative if we define the equality in an almost everywhere (a.e.) sense, whereas on $\mathbb{Z}$ the answer is affirmative. Of course, the biggest difference between these two sets is that one only contains right-dense points, while the other only contains right-scattered points. Thus, this might lead one to conjecture that in the time scale case, it is necessary to consider points that are scattered and dense separately. This is in fact the case as we show. Thus, the answer to our question on uniqueness in the time scales case is affirmative when $f = g$ a.e. for our definition of a.e. on a time scale. Of course, in order to do so, we must first clarify what is meant by *a.e. on a time scale.*

Recall from Chapter 1 that in his initial work on the time scale Lebesgue integral, Guseinov [26] defines the Carathéodory extension of the set function that assigns each time scale interval its length to be the Lebesgue Δ-measure on $\mathbb{T}$. To construct an outer measure, Guseinov does a Vitali covering of subsets of $\mathbb{T}$ by finite or countable systems of intervals of $\mathbb{T}$, and then naturally defines the outer measure to be the infimum of the sums of the lengths of the intervals that cover the subsets. If there is no such covering, he defines the outer measure of the set to be infinite.

Thus, any time scale will have infinite Δ-measure. However, Guseinov points out that his choosing the Δ-measure to be infinite is merely to preserve the monotonicity of the outer measure. The monotonicity will also be preserved if for a subset E of $\mathbb{T}$ that cannot be covered, we define the outer measure of E to be the outer measure of the maximal coverable subset of E, call it F, plus some positive extended real number c chosen independently of E. For his purposes, Guseinov chooses $c = \infty$, but for our purposes, it is convenient to choose $c = 0$.

By doing this, the Lebesgue Δ-measure μ_Δ can be decomposed nicely. For any subset E of $\mathbb{T}$, decompose E as $E = D \cup S$, where

$$D = \{t \in \mathbb{T} : t \text{ is right dense}\}, \quad S = \{t \in \mathbb{T} : t \text{ is right scattered}\}.$$

Since $c = 0$ above, we may write

$$\mu_\Delta(E) = m(E \cap D) + c(E \cap S),$$

where $m(D)$ denotes the usual Lebesgue measure of the set D and $c(S)$ is the measure given by $c(S) = \sum_{s \in S} \mu(s)$.

Notice that with this decomposition, the sets of measure zero can only consist of right dense points since $\mu_\Delta(E) = 0$ if and only if $m(E \cap D) = 0$ and $E \cap S = \emptyset$. Thus, to show that a property holds *almost everywhere on a time scale*, it is necessary to show that the property holds for every right scattered point in the time scale, and that the set of right dense points for which the property fails has Lebesgue measure zero.

We are now in a position to prove the uniqueness theorem.

Theorem 2.7 (Uniqueness of the Inverse). *If the functions $f : \mathbb{T} \to \mathbb{R}$ and $g : \mathbb{T} \to \mathbb{R}$ have the same Laplace transform, then $f = g$ a.e. on $\mathbb{T}$.*

Proof. Suppose

$$\int_0^\infty e^\sigma_{\ominus z}(t, 0) f(t)\, \Delta t = \int_0^\infty e^\sigma_{\ominus z}(t, 0) g(t)\, \Delta t,$$

so that the function $h = f - g$ has transform of zero. That is, if we let F denote the inversion operator given in Theorem 2.6 and G denote the transform operator, then $h \in \ker G$. But, it follows that $(F \circ G)(h) = F(0) = 0 = h$, where we note that the function $h(t) = 0$ a.e. also satisfies the equations above. Thus, $f = g$ a.e. on $\mathbb{T}$. $\square$

With uniqueness up to sets of measure zero now established, notice that *we are now justified in using Table 2.1 to associate a given function with its transform*, as long as we agree that any other function that differs from those given in the table on a set of measure zero will also have the same transform. Indeed, this agrees with the cases $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$ with which we are well familiar and have already discussed.

2.2.4 Frequency Shifting

For $f \in C_{\text{prd-e2}}(\mathbb{T}, \mathbb{R})$ with exponential type II constant $c > 0$, define the function $f_\mu(a, t)$ by

$$f_\mu(a,t) := \sum \mathrm{Res}_{z=z_k} e_{\frac{z}{1+\mu a}}(t,0)F(z),$$

where $F(z)$ is the transform of $f(t)$.

Theorem 2.8 (Frequency Shifting). *For $f \in C_{prd\text{-}e2}(\mathbb{T}, \mathbb{R})$ with exponential type II constant $c > 0$,*

$$\mathcal{L}\{e_a(t,0)f_\mu(a,t)\}(z) = F(z-a),$$

where $F(z)$ denotes the transform of $f(t)$.

Proof. First, note that if $F(z)$ has poles $z_0, z_1, \dots, z_{n-1}$, then $F(z-a)$ has poles $z_0 + a, z_1 + a, \dots, z_{n-1} + a$. Thus,

$$\begin{aligned}
\mathcal{L}^{-1}\{F(z-a)\} &= \sum \mathrm{Res}_{z=z_k+a}\ e_z(t,0)F(z-a) \\
&= \sum \mathrm{Res}_{z=z_k}\ e_{z+a}(t,0)F(z) \\
&= \sum \mathrm{Res}_{z=z_k}\ e_{\frac{z}{1+\mu a}\oplus a}(t,0)F(z) \\
&= e_a(t,0)f_\mu(a,t),
\end{aligned}$$

an expression which equivalently says that

$$\mathcal{L}\{e_a(t,0)f_\mu(a,t)\}(z) = F(z-a).$$

□

We now consider a few examples of uses of Theorem 2.8. First, note that if $\mathbb{T}=\mathbb{R}$, then $\mu \equiv 0$ and $f_0(a,t) = f(t)$, and so on $\mathbb{R}$, the frequency shifting theorem says

$$\mathcal{L}\{e_a(t,0)f_\mu(a,t)\} = \mathcal{L}\{e^{at}f(t)\} = F(z-a),$$

as expected.

Next, notice for the exponential function, the theorem yields

$$\mathcal{L}\{e_{\alpha+\beta}(t,0)\} = \mathcal{L}\{e_\alpha(t,0)e_{\frac{\beta}{1+\mu\alpha}}(t,0)\} = \frac{1}{z-(\alpha+\beta)},$$

which was expected from Table 2.1.

Next, we examine the time scale trigonometric functions. We know

$$\mathcal{L}\{\sin_\beta(t,0)\} = \frac{\beta}{z^2+\beta^2}, \quad \mathcal{L}\{\cos_\beta(t,0)\} = \frac{z}{z^2+\beta^2},$$

so that each of these complex functions has simple poles at $z = \{\beta i, -\beta i\}$. Thus, for $f(t) = \cos_\beta(t,0)$, $f_\mu(\alpha,t) = \cos_{\frac{\beta}{1+\mu\alpha}}(t,0)$, and for $f(t) = \sin_\beta(t,0)$, $f_\mu(\alpha,t) = \sin_{\frac{\beta}{1+\mu\alpha}}(t,0)$. Applying Theorem 2.8,

$$\begin{aligned}\mathcal{L}\{e_\alpha(t,0)\sin_{\frac{\beta}{1+\mu\alpha}}(t,0)\} &= \frac{\beta}{(z-\alpha)^2+\beta^2},\\ \mathcal{L}\{e_\alpha(t,0)\cos_{\frac{\beta}{1+\mu\alpha}}(t,0)\} &= \frac{z-\alpha}{(z-\alpha)^2+\beta^2}.\end{aligned}$$

Bohner and Peterson show this in [8, 10] by using uniqueness of solutions to certain IVPs.

Likewise, for the time scale hyperbolic trig functions,

$$\mathcal{L}\{\sinh_\beta(t,0)\} = \frac{\beta}{z^2-\beta^2}, \quad \mathcal{L}\{\cosh_\beta(t,0)\} = \frac{z}{z^2-\beta^2},$$

so that each of these complex functions has simple poles at $z = \{\beta, -\beta\}$. Thus, for $f(t) = \cosh_\beta(t,0)$, $f_\mu(\alpha,t) = \cosh_{\frac{\beta}{1+\mu\alpha}}(t,0)$, and for $f(t) = \sinh_\beta(t,0)$, $f_\mu(\alpha,t) = \sinh_{\frac{\beta}{1+\mu\alpha}}(t,0)$. Applying Theorem 2.8,

$$\mathcal{L}\{e_\alpha(t,0)\sinh_{\frac{\beta}{1+\mu\alpha}}(t,0)\} = \frac{\beta}{(z-\alpha)^2-\beta^2},$$
$$\mathcal{L}\{e_\alpha(t,0)\cosh_{\frac{\beta}{1+\mu\alpha}}(t,0)\} = \frac{z-\alpha}{(z-\alpha)^2-\beta^2}.$$

Bohner and Peterson also showed this in [8, 10] by using uniqueness of solutions to certain IVPs.

The preceding examples are special because the poles are simple in each of these cases, a fact which probably led Bohner and Peterson to their results. However, they never talk about the analogue of shifting in the frequency domain for the function $f(t) = t^k/k!$ on $\mathbb{R}$. The reason for this is that the poles for the transform of this function, and thus the time scale polynomials $h_k(t,0)$ as well, are *not simple* anymore. Indeed, the poles are of order $k+1$ for $k \geq 0 \in \mathbb{N}_0$. For example, on $\mathbb{T} = \mathbb{R}$, the function $f(t) = te^{\alpha t}$ has transform $F(z-\alpha) = \frac{1}{(z-\alpha)^2}$, and so it is quite natural to wonder what plays the role of this function in general for arbitrary $\mathbb{T}$. Using Theorem 2.8, we note $F(z) = \frac{1}{z^2}$ has a pole of order 2 at $z = 0$, and since the residue calculations in this case involve the derivative of the integral $\int_0^t \frac{\log(1+\mu(\tau)\left(\frac{z}{(1+\mu(\tau)\alpha)}\right)}{\mu(\tau)}\Delta\tau$ with respect to z, we see $f_\mu(\alpha,t) = \int_0^t \frac{1}{1+\mu(\tau)\alpha}\Delta\tau$. Thus, another application of Theorem 2.8 yields

$$\mathcal{L}\left\{e_a(t,0)\int_0^t \frac{1}{1+\mu(\tau)\alpha}\Delta\tau\right\} = \frac{1}{(z-a)^2}.$$

2.3 Convolution

We wish to know what function in the time domain is associated with the product $F(z)G(z)$ in the z domain. We know that on $\mathbb{R}$ and $\mathbb{Z}$, these functions are the convolution products. Thus, in this section, we turn our attention to the convolution on an arbitrary time scale $\mathbb{T}$ via the Laplace transform.

We begin by noting that on $\mathbb{R}$ and $\mathbb{Z}$, the convolution is defined by

$$(f * g)(t) = \int_0^t f(t-s)g(s)\Delta s.$$

However, on these two time scales, the difference $t - s \in \mathbb{T}$ whenever $s, t \in \mathbb{T}$. This need not be true on a general $\mathbb{T}$, so that we cannot define the generalized convolution within this framework. Instead, we need to define an analogue of the delay or shift of a function on $\mathbb{T}$. Motivated by this, we have the following:

Definition 2.3. The *delay* or *shift* of the function $x(t)$ with $x \in C_{\text{prd-e2}}(\mathbb{T}, \mathbb{R})$ by $\sigma(\tau) \in \mathbb{T}$, denoted by $x(t, \sigma(\tau))$, is given by

$$u_{\sigma(\tau)}(t)x(t,\sigma(\tau)) = \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} X(z)e_z(t,\sigma(\tau)).$$

Here, recall that $u_\xi(t) : \mathbb{T} \to \mathbb{R}$ is the time scale unit step function activated at time $t = \xi \in \mathbb{T}$.

Notice that $u_{\sigma(\tau)}(t)x(t,\sigma(\tau))$ has transform $X(z)e^{\sigma}_{\ominus z}(\tau, 0)$. Indeed,

$$\begin{aligned} u_{\sigma(\tau)}(t)x(t,\sigma(\tau)) &= \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} X(z)e_z(t,\sigma(\tau)) \\ &= \sum_{k=1}^{n} \operatorname{Res}_{z=z_k} \left[X(z)e^{\sigma}_{\ominus z}(\tau,0)\right] e_z(t,0) \\ &= \mathcal{L}^{-1}\{X(z)e^{\sigma}_{\ominus z}(\tau,0)\}. \end{aligned}$$

This allows us to use the term delay or shift to describe $x(t,\sigma(\tau))$, since on $\mathbb{T} = \mathbb{R}$, the transformed function $X(z)e^{-z\tau}$ corresponds to the function $u_\tau(t)x(t-\tau)$. This definition also allows us to give an analogue of the time shifting theorem in general.

Theorem 2.9 (Time Shifting). *For* $g(t) \in C_{\text{prd-e2}}(\mathbb{T}, \mathbb{R})$,

$$\mathcal{L}\{u_{\sigma(s)}(t)g(t,\sigma(s))\} = \int_0^\infty e^{\sigma}_{\ominus z}(t,0)u_{\sigma(s)}(t)g(t,\sigma(s))\Delta t = G(z)e^{\sigma}_{\ominus z}(s,0),$$

with the integral converging in the ROC of g.

With the delay operator now defined, we define the convolution of any two arbitrary transformable time scale functions.

Definition 2.4. The convolution of the functions $f(t)$ and $g(t)$, denoted $f * g$, with $f, g \in C_{\text{prd-e2}}(\mathbb{T}, \mathbb{R})$ is given by

$$(f * g)(t) = \int_0^t f(\tau)g(t, \sigma(\tau))\Delta\tau.$$

Before stating the and proving the Convolution Theorem, it is worth noting that on $\mathbb{R}$ and $\mathbb{Z}$, the shift $g(t, \sigma(s))$ of the function $g(t)$ is given by $g(t-s)$. Thus, this tells us that the convolution product reduces to the familiar one that is known for each of $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$.

Theorem 2.10 (Convolution Theorem). *The transform of a convolution product that is absolutely integrable is the product of the transforms, with the Laplace integral converging in the region $Re_{\mu_*}(z) > Re_{\mu_*}(\hat{c})$, where $\hat{c} = \max\{c_f, c_g\}$ and c_f and c_g are the exponential constants corresponding to f and g, respectively.*

Proof. If we assume absolute integrability of all functions involved, then by the delay property of the transform previously mentioned, we obtain

$$\begin{aligned}
\mathcal{L}\{f * g\} &= \int_0^\infty e^\sigma_{\ominus z}(t, 0)\left[(f * g)(t)\right]\Delta t \\
&= \int_0^\infty \left[\int_0^t f(\tau)g(t, \sigma(\tau))\Delta\tau\right] e^\sigma_{\ominus z}(t, 0)\,\Delta t \\
&= \int_0^\infty f(\tau)\left[\int_{\sigma(\tau)}^\infty g(t, \sigma(\tau))e^\sigma_{\ominus z}(t, 0)\,\Delta t\right]\Delta\tau \\
&= \int_0^\infty f(\tau)\mathcal{L}\{u_{\sigma(\tau)}(t)g(t, \sigma(\tau))\}\Delta\tau \\
&= \int_0^\infty f(\tau)\left[G(z)e^\sigma_{\ominus z}(\tau, 0)\right]\Delta\tau \\
&= \int_0^\infty f(\tau)e^\sigma_{\ominus z}(\tau, 0)\Delta\tau G(z) \\
&= F(z)G(z).
\end{aligned}$$

For the convergence of the integral, note that for f and g of exponential type II with constants c_f and c_g, respectively, we have

$$\begin{aligned}
|(f*g)(t)| &= \left|\int_0^t f(\tau)g(t,\sigma(\tau))\Delta\tau\right| \\
&\leq \int_0^t |f(\tau)|\cdot|g(t,\sigma(\tau))|\Delta\tau \\
&\leq Me_{c_g}(t,0)\int_0^t e_{c_f}(\tau,0)e_{c_g}(0,\sigma(\tau))\Delta\tau \\
&\leq Me_{c_g}(t,0)\int_0^t e_{c_f}(\tau,0)e_{\ominus c_g}(\tau,0)\Delta\tau \\
&\leq \frac{M}{|c_f-c_g|}e_{c_g}(t,0)\left(e_{c_f\ominus c_g}(t,0)-1\right) \\
&\leq \frac{M}{|c_f-c_g|}\left(e_{c_f}(t,0)+e_{c_g}(t,0)\right) \\
&\leq \frac{2M}{|c_f-c_g|}e_{\hat{c}}(t,0),
\end{aligned}$$

so that $f*g$ is of exponential type II with constant $\hat{c}$. $\square$

Example 2.12. Suppose $f(t)=f(t,0)$ is one of the elementary functions: that is, $f(t)$ is one of $h_k(t,0)$, $e_a(t,0)$, $\sin_a(t,0)$, $\cos_a(t,0)$, $\cosh_a(t,0)$, or $\sinh_a(t,0)$. Direct residue calculations show that the delay $f(t,\sigma(\tau))$ of each of these functions is given by $h_k(t,\sigma(\tau))$, $e_a(t,\sigma(\tau))$, $\sin_a(t,\sigma(\tau))$, $\cos_a(t,\sigma(\tau))$, $\cosh_a(t,\sigma(\tau))$, and $\sinh_a(t,\sigma(\tau))$. We wish to build a table of convolution products of these functions since their products commonly arise in the solutions of dynamic equations. We will demonstrate the computations involved for one of the products; the rest are similar.

Consider the product $e_\alpha(t,0)*e_\beta(t,0)$. By definitions of the convolution product and the delay, we have

$$\begin{aligned}
e_\alpha(t,0)*e_\beta(t,0) &= \int_0^t e_\alpha(\tau,0)e_\beta(t,\sigma(\tau))\Delta\tau \\
&= e_\beta(t,0)\int_0^t e_\alpha(\tau,0)e_\beta(0,\sigma(\tau))\Delta\tau \\
&= e_\beta(t,0)\int_0^t \frac{1}{1+\mu\beta}e_{\alpha\ominus\beta}(\tau,0)\Delta\tau
\end{aligned}$$

$$
\begin{aligned}
&= \frac{e_\beta(t,0)}{\beta-\alpha}\left[1 - e_{\alpha\ominus\beta}(t,0)\right] \\
&= \frac{1}{\beta-\alpha}\left[e_\beta(t,0) - e_\alpha(t,0)\right].
\end{aligned}
$$

The remaining products can be found in Table 2.2.

Example 2.13. Consider $g(t,\sigma(\tau)) = 1$. In this instance, we see that for any $f \in C_{\text{prd-e2}}(\mathbb{T},\mathbb{R})$, the transform of $h(t) = \int_0^t f(\tau)\Delta\tau$ is given by

$$
\mathcal{L}\{h\} = \mathcal{L}\{f * 1\} = F(z)\mathcal{L}\{1\} = \frac{F(z)}{z},
$$

with the integral converging in the ROC of f. Recall that this is a result which we obtained by (a notably more lengthy) direct calculation earlier.

Example 2.14. Let

$$
f(t) = \frac{1}{2\alpha}\sin_\alpha(t,0)\int_0^t \frac{1}{1+(\mu(\tau)\alpha)^2}\Delta\tau - \frac{1}{2}\cos_\alpha(t,0)\int_0^t \frac{\mu(\tau)}{1+(\mu(\tau)\alpha)^2}\Delta\tau,
$$

for $\alpha > 0$, and $g(t) = e_\beta(t,0)$. We wish to compute $(f * g)(t)$. By the Convolution Theorem, we know that the convolution product has transform $F(z)G(z)$. By Example 2.9,

$$
F(z) = \frac{z}{(z^2+\alpha^2)^2}.
$$

Thus, we have

$$
\begin{aligned}
&(f * g)(t) = \mathcal{L}^{-1}\{F(z)G(z)\} \\
&= \sum_{k=1}^{3} \mathrm{Res}_{z=z_k} \frac{z e_z(t,0)}{(z^2+\alpha^2)^2(z-\beta)} \\
&= \frac{1}{2\alpha^2(\alpha^2+\beta^2)(\alpha^4+\beta^2)}\left(\int_0^t \frac{(-\beta\mu+1)(\alpha^2)(\alpha^4+\beta^2)}{1+(\mu\alpha)^2}\Delta\tau + (\alpha-1)\beta^3 + 2\alpha^3\beta\right)\cos_\alpha(t,0) \\
&+ \frac{1}{2\alpha(\alpha^2+\beta^2)(\alpha^4+\beta^2)}\left(\int_0^t \frac{\beta(\alpha^4+\beta^2)+\alpha^2(\alpha^4+\beta^2)\mu}{1+(\mu\alpha)^2}\Delta\tau + \alpha^4 + (-\alpha^2+\alpha+1)\beta^2\right) \\
&\qquad \times \sin_\alpha(t,0) + \frac{\beta}{(\alpha^2+\beta^2)^2}e_\beta(t,0).
\end{aligned}
$$

Table 2.2. Convolutions of some of the elementary functions on $\mathbb{T}$.

$f(t)$	$g(t)$	$(f*g)(t)$
$e_\alpha(t,0)$	$e_\beta(t,0), \alpha \neq \beta$	$\frac{1}{\beta-\alpha}\left[e_\beta(t,0) - e_\alpha(t,0)\right]$
	$e_\alpha(t,0)$	$e_\alpha(t,0)\int_0^t \frac{1}{1+\alpha\mu(s)}\Delta s$
	$h_k(t,0), \alpha \neq 0$	$\frac{1}{\alpha^{k+1}}e_\alpha(t,0) - \sum_{j=0}^{k}\frac{1}{\alpha^{k+1-j}}h_j(t,0)$
	$\sin_\beta(t,0), \alpha^2+\beta^2>0$	$\frac{\beta e_\alpha(t,0)}{\alpha^2+\beta^2} - \frac{\alpha\sin_\beta(t,0)}{\alpha^2+\beta^2} - \frac{\beta\cos_\beta(t,0)}{\alpha^2+\beta^2}$
	$\cos_\beta(t,0)$	$\frac{\alpha e_\alpha(t,0)}{\alpha^2+\beta^2} + \frac{\beta\sin_\beta(t,0)}{\alpha^2+\beta^2} - \frac{\alpha\cos_\beta(t,0)}{\alpha^2+\beta^2}$
$\sin_\alpha(t,0)$	$\sin_\beta(t,0), \alpha\neq 0, \alpha\neq\beta$	$\frac{\alpha}{\alpha^2-\beta^2}\sin_\beta(t,0) - \frac{\beta}{\alpha^2-\beta^2}\sin_\alpha(t,0)$
	$\cos_\beta(t,0), \alpha\neq 0, \alpha\neq\beta$	$\frac{\alpha}{\alpha^2-\beta^2}\cos_\beta(t,0) - \frac{\alpha}{\alpha^2-\beta^2}\cos_\alpha(t,0)$
	$\sin_\alpha(t,0), \alpha\neq 0$	$\frac{1}{\alpha}\sin_\alpha(t,0) - \frac{1}{2}t\cos_\alpha(t,0)$
	$h_k(t,0), k$ even	$(-1)^{(k+1)(k+2)/2}\frac{1}{\alpha^{k+1}}\cos_\alpha(t,0)$ $+\sum_{j=0}^{k/2}(-1)^j\frac{h_{k-2j}(t,0)}{\alpha^{2j+1}}$
	$h_k(t,0), k$ odd	$(-1)^{(k+1)(k+2)/2}\frac{1}{\alpha^{k+1}}\sin_\alpha(t,0)$ $+\sum_{j=0}^{(k-1)/2}(-1)^j\frac{h_{k-2j}(t,0)}{\alpha^{2j+1}}$
$\cos_\alpha(t,0)$	$\cos_\beta(t,0), \alpha\neq 0, \alpha\neq\beta$	$-\frac{\beta}{\alpha^2-\beta^2}\sin_\beta(t,0) + \frac{\alpha}{\alpha^2-\beta^2}\sin_\alpha(t,0)$
	$\cos_\alpha(t,0)$	$\frac{1}{\alpha}\sin_\alpha(t,0) + \frac{1}{2}t\cos_\alpha(t,0)$
	$h_k(t,0), k$ even	$(-1)^{(k+1)(k+2)/2}\frac{1}{\alpha^{k+1}}\sin_\alpha(t,0)$ $+\sum_{j=0}^{(k-2)/2}(-1)^j\frac{h_{k-2j-1}(t,0)}{\alpha^{2j+1}}$
	$h_k(t,0), k$ odd	$(-1)^{(k+1)(k+2)/2}\frac{1}{\alpha^{k+1}}\cos_\alpha(t,0)$ $+\sum_{j=0}^{(k-1)/2}(-1)^j\frac{h_{k-2j-1}(t,0)}{\alpha^{2j+1}}$
$h_k(t,0)$	$h_j(t,0)$	$h_{k+j+1}(t,0)$

It is worth noting that the convolution product is both commutative and associative. Indeed, the products $f * g$ and $g * f$ have the same transform as do the products $f * (g * h)$ and $(f * g) * h$, and since the inverse is unique, the functions defined by these products must agree almost everywhere.

At first glance, one may think that the identity is vested in the Hilger delta. Unfortunately, this is not the case. It can be shown that any identity for the convolution will of necessity have transform 1 by the Convolution Theorem. But the Hilger delta does not have transform 1 since its transform is just the exponential. Thus, we develop an analogue of the Dirac delta in the next section in order to establish an identity element.

2.4 The Dirac Delta Functional

Let $f, g : \mathbb{T} \to \mathbb{R}$ be given functions with $f(x)$ having unit area. To define the delta functional, we construct the following functional. Let $C_c^\infty(\mathbb{T})$ denote the $C^\infty(\mathbb{T})$ functions with compact support. For $g^\sigma \in C_c^\infty(\mathbb{T})$ and for all $\epsilon > 0$, define the functional $F : C_c^\infty(\mathbb{T}, \mathbb{R}) \times \mathbb{T} \to \mathbb{R}$ by

$$F(g^\sigma, a) := \begin{cases} \int_0^\infty \delta_a^{\mathbb{H}}(x) g^\rho(\sigma(x))\, \Delta x, & \text{if } a \text{ is right scattered,} \\ \lim_{\epsilon \to 0} \int_0^\infty \frac{1}{\epsilon} f(\frac{x}{\epsilon}) g(\sigma(x))\, \Delta x, & \text{if } a \text{ is right dense.} \end{cases}$$

The *time scale Dirac delta functional* is then given by the symmetric form

$$\langle \delta_a^{\mathbb{T}}, g^\sigma \rangle := F(g^\sigma, a).$$

To demonstrate how this functional acts on functions from $C_c^\infty(\mathbb{T})$, let $g : \mathbb{T} \to \mathbb{R}$ be such that $g^\sigma \in C_c^\infty(\mathbb{T})$. If a is right dense, consider

$$f(x) = \begin{cases} \frac{1}{\epsilon}, & \text{if } a \le x \le a + \epsilon, \\ 0, & \text{else,} \end{cases}$$

with the understanding that any sequence of ϵ's we choose will be under the restriction that $a + \epsilon \in \mathbb{T}$ for each ϵ in the sequence. Then for $h(x) = g^\sigma(x)$, we have (for any antiderivative $H(x)$ of $h(x)$),

$$\begin{aligned} F(g^\sigma, a) &= \lim_{\epsilon\to 0}\int_0^\infty f(x)h(x)\,\Delta x \\ &= \lim_{\epsilon\to 0}\frac{\int_a^{a+\epsilon} h(x)\,\Delta x}{\epsilon} \\ &= \lim_{\epsilon\to 0}\frac{H(a+\epsilon)-H(a)}{\epsilon} \\ &= H^\Delta(a) = h(a) = g(\sigma(a)) = g(a), \end{aligned}$$

in the sense of weak limits. If a is right scattered, then

$$F(g^\sigma, a) = \int_0^\infty \delta_a^{\mathbb{H}}(x)g^\rho(\sigma(x))\,\Delta x = g^\rho(\sigma(a)) = g(a).$$

Thus, in functional terms, the time scale Dirac delta functional acts as

$$\langle g^\sigma, \delta_a^{\mathbb{T}}\rangle = g(a),$$

independently of the time scale involved. Also, if $g(t) = e_{\ominus z}(t,0)$, then $\langle \delta_a^{\mathbb{T}}, g^\sigma\rangle = e_{\ominus z}(a,0)$, so that for $a=0$, the Dirac delta functional $\delta_0^{\mathbb{T}}(t)$ has Laplace transform of 1, thereby producing an identity element for the convolution. However, the present definition of the delay operator only holds for functions. It is necessary to extend this definition for the Dirac delta functional. To maintain consistency with the delta function's action on $g(t) = e_{\ominus z}^\sigma(t,0)$, it follows that for any $t\in\mathbb{T}$, the *shift of* $\delta_a^{\mathbb{T}}(\tau)$ is given by

$$\delta_a^{\mathbb{T}}(t,\sigma(\tau)) := \delta_t^{\mathbb{T}}(\sigma(\tau)).$$

With this definition, our claim holds:

$$\begin{aligned} (\delta_0^{\mathbb{T}} * g)(t_0) &= \int_0^{t_0} \delta_0^{\mathbb{T}}(\tau)g(t,\sigma(\tau))\,\Delta\tau \\ &= g(t_0,0) \\ &= g(t_0) \\ &= \int_0^{t_0} g(\tau)\delta_{t_0}^{\mathbb{T}}(\sigma(\tau))\,\Delta\tau \\ &= \int_0^{t_0} g(\tau)\delta_0^{\mathbb{T}}(t_0,\sigma(\tau))\,\Delta\tau \\ &= (g*\delta_0^{\mathbb{T}})(t_0). \end{aligned}$$

$$\begin{tikzcd}
(C_c^\infty(\mathbb{R},\mathbb{R}))^* \arrow[r, shift left, "\mathcal{L}_\mathbb{R}"] \arrow[d, shift right, "I"'] & T_f(\mathbb{R}) \arrow[l, shift left, "\mathcal{L}_\mathbb{R}^{-1}"] \arrow[d, shift right, "\gamma"'] \\
(C_c^\infty(\mathbb{T},\mathbb{R}))^* \arrow[r, shift left, "\mathcal{L}_\mathbb{T}"] \arrow[u, shift right, "I"'] & T_f(\mathbb{T}) \arrow[l, shift left, "\mathcal{L}_\mathbb{T}^{-1} = I \circ \mathcal{L}_\mathbb{R}^{-1} \circ \gamma^{-1}"] \arrow[u, shift right, "\gamma^{-1}"']
\end{tikzcd}$$

Figure 2.4. Commutative diagram in the dual spaces.

In other words, when we perform the convolution $\langle g, \delta_0^{\mathbb{T}}(\sigma(t))\rangle$, we must give meaning to this symbol and do so by defining $\delta_0^{\mathbb{T}}(\sigma(t))$ to be the Kronecker delta when t is right scattered and the usual Dirac delta if t is right dense. While this *ad hoc* approach does not address convolution with an arbitrary (shifted) distribution on the right, this will suffice (at least for now) since our eye is on solving generalizations of canonical partial dynamic equations which will involve the Dirac delta distribution.

We now turn to uniqueness of the inverse transform of the Dirac delta. We would like an analogue of the diagram given in the proof of Theorem 2.6 where we move from the space of continuous functions to its dual space (or at least restricted to the dual space of those functions that are infinitely differentiable with compact support). As frequently happens when we move from a space to its dual space, the diagram also becomes the dual of the original one. Indeed, the diagram in terms of the dual space is shown in Figure 2.4.

The mappings γ and γ^{-1} in the dual diagram act on the transform spaces just as in Figure 2.3. It is worth comparing Figures 2.3 and 2.4 since $X \subset X^*$. So how do we reconcile the two diagrams? We first must clarify what is meant by the identity map between the two dual spaces. The Dirac delta is defined in terms of its action on any function: in both dual spaces, $\langle \delta_a^{\mathbb{T}}, g^\sigma \rangle = g(a)$. To make this agree with the map for the function spaces, note that for $f \in C_{\text{prd-e2}}(\mathbb{T},\mathbb{R})$ and $g \in C_{\text{p-eo}}(\mathbb{R},\mathbb{R})$, the action of $f(t)$ on $h^\sigma(t) = e^\sigma_{\ominus z}(t,0)$ should be the same as the action of $g(t)$ on $\tilde{h}(t) = e^{-zt}$. For example, the function on $\mathbb{T}$ that acts on $h^\sigma(t)$ with a result of $\frac{1}{z^2+1}$ is the function $f(t) = \sin_1(t,0)$, while the function on $\mathbb{R}$ that acts on $\tilde{h}(t)$ with the same result is $g(t) = \sin(t)$. As another example, the function on $\mathbb{T}$ that acts on $h^\sigma(t)$ with a result of $\frac{h(\tau)}{z}$ is the time scale unit step function, while on $\mathbb{R}$, the function that results in $\frac{\tilde{h}(\tau)}{z}$ is the continuous step function. Thus, in terms of the dual spaces, the map on the left hand side of the diagram is the identity map, while in the function spaces this identity map maps g into f^σ by the switching of the exponentials between the two domains.

The preceding discussion provides the basis for the uniqueness of the transform of the delta functional. Through the diagram we see that if another functional had the same transform, then there would be two such functionals over $C_c^\infty(\mathbb{R},\mathbb{R})$ which had the same transform, which we know to be false.

Next, we examine more properties of the Dirac delta. We have already noted that the functional has transform 1 (as desired) since $\langle \delta_0^{\mathbb{T}}, e^\sigma_{\ominus z}(t,0)\rangle = 1$. Second, the transform of the derivative of the delta functional is familiar:

$$\begin{aligned}
\mathcal{L}\{\delta_0^{\mathbb{T}^\Delta}\} &= \int_0^\infty \delta_0^{\mathbb{T}^\Delta} e^\sigma_{\ominus z}(t,0)\,\Delta t \\
&= \delta_0^{\mathbb{T}} e_{\ominus z}(t,0)\Big|_{t=0}^{\infty} - \int_0^\infty \ominus z e_{\ominus z}(t,0)\delta_0^{\mathbb{T}}\,\Delta t \\
&= z\int_0^\infty e^\sigma_{\ominus z}(t,0)\delta_0^{\mathbb{T}}\,\Delta t \\
&= z\mathcal{L}\{\delta_0^{\mathbb{T}}\} \\
&= z,
\end{aligned}$$

so that $\mathcal{L}\{\delta_0^{\mathbb{T}^\Delta}\}$ is the same as it on $\mathbb{R}$.

Finally, just as on $\mathbb{R}$, the derivative of the Heaviside function is still the Dirac delta:

$$\begin{aligned}\langle H^\Delta, g^\sigma\rangle &= \int_0^\infty H^\Delta(t) g^\sigma(t)\,\Delta t \\ &= H(t)g(t)\Big|_{t=0}^{\infty} - \int_0^\infty H(t) g^\Delta(t)\,\Delta t \\ &= -\int_0^\infty g^\Delta(t)\,\Delta t \\ &= g(0) \\ &= \langle \delta_0^{\mathbb{T}}, g^\sigma\rangle.\end{aligned}$$

2.5 Applications to Green's Function Analysis

We now demonstrate a powerful use of the Dirac delta applied to the Green's function analysis ubiquitous in the study of boundary value problems.

Consider the operator $L : S \to C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$ given by

$$Lx(t) := (px^\Delta)^\Delta(t) + q(t)x^\sigma(t),$$

where $p, q \in C_{\mathrm{rd}}$, $p(t) \neq 0$ for all $t \in \mathbb{T}$, and

$$S := \{x \in C^1(\mathbb{T}, \mathbb{R}) : (p\,x^\Delta)^\Delta \in C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})\}.$$

Bohner and Peterson [8] showed that if the homogeneous boundary value problem

$$Lx = 0, \qquad \alpha x(a) - \beta x^\Delta(a) = \gamma x(\sigma^2(b)) + \delta x^\Delta(\sigma(b)) = 0,$$

has only the trivial solution, then the nonhomogeneous boundary value problem

$$Lx = h(\sigma(t)), \qquad \alpha x(a) - \beta x^\Delta(a) = A,\ \gamma(\sigma^2(b)) + \delta x^\Delta(\sigma(b)) = B, \tag{2.4}$$

where $h^\sigma \in C_{\mathrm{rd}}$, and A and B are given constants, has a unique solution. If ϕ is the solution of

$$L\phi = 0, \qquad \phi(a) = \beta,\ \phi^\Delta(a) = \alpha,$$

and ψ is the solution of

$$L\psi = 0, \qquad \psi(\sigma^2(b)) = \delta,\ \psi^\Delta(\sigma(b)) = -\gamma,$$

then the solution of (2.4) can be written in the form

$$x(t) := \int_a^{\sigma(b)} G(t,\sigma(s))h(\sigma(s))\,\Delta s,$$

where $G(t,\sigma(s))$ is the Green's function for the boundary value problem and is given by

$$G(t,\sigma(s)) = \begin{cases} \frac{1}{c}\phi(t)\psi(\sigma(s)), & \text{if } t \le s, \\ \frac{1}{c}\psi(t)\phi(\sigma(s)), & \text{if } t \ge \sigma(s), \end{cases}$$

where $c := p(t)W(\phi,\psi)(t)$ is a constant. They further show that the Green's function is symmetric.

For any $s \in [a,\sigma(b)]$, if $h(\sigma(t)) = \delta^{\mathbb{T}}_{\sigma(t)}(s)$, then

$$x(t) = \int_a^{\sigma(b)} \delta^{\mathbb{T}}_{\sigma(s)}(s)G(t,\sigma(s))\,\Delta s = G(t,\sigma(s)),$$

and therefore,

$$L(G(t,\sigma(s)) = \delta^{\mathbb{T}}_{\sigma(t)}(s).$$

In fact,

$$Lx = \int_a^{\sigma(b)} L(G(t,\sigma(s)))h(\sigma(s))\,\Delta s = \int_a^{\sigma(b)} \delta^{\mathbb{T}}_{\sigma(t)}(s)h(\sigma(s))\,\Delta s = h(\sigma(t)).$$

CHAPTER THREE

Linear Systems Theory on Time Scales

3.1 Controllability

We now turn our attention to the fundamental notions of controllability, observability, realizability, and stability commonly dealt with in control theory. Our focus here is the time scale setting, but our definitions coincide with those that appear in the literature for $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. Throughout this chapter, $A(t) \in \mathbb{R}^{n\times n}, B(t) \in \mathbb{R}^{n\times m}, C(t) \in \mathbb{R}^{p\times n}$, and $D(t) \in \mathbb{R}^{p\times m}$. We will assume the systems in question are regressive, a restriction which in turn implies that the matrix $I + \mu(t)A(t)$ is invertible. Hence, on $\mathbb{Z}$, the transition matrix will always be invertible, and so we are justified in talking about controllability rather than reachability which is common (see [12], [19], and [40]) since the transition matrix in general need not be invertible for $\mathbb{T} = \mathbb{Z}$.

We begin with the time varying case, and then proceed to treat the time invariant case, in which we will get stronger results than from the latter. We remark here that some of the statements that follow in this chapter can be found in [3], [4], and [20], but we have found serious errors throughout these works which are not present in this chapter.

3.1.1 Time Varying Case

When the term *controllability* is used to discuss dynamical systems, it means that the solutions of the dynamic equations involved can be driven to some desired final state in finite time. Thus, mathematically, this notion is defined as follows.

Definition 3.1. Let $A(t) \in \mathbb{R}^{n\times n}, B(t) \in \mathbb{R}^{n\times m}, C(t) \in \mathbb{R}^{p\times n}$, and $D(t) \in \mathbb{R}^{p\times m}$ all be rd-continuous matrix functions defined on $\mathbb{T}$. Here, $p, m \leq n$. The regressive linear state equation

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \;\; x(t_0) = x_0, \qquad (3.1) \\ y(t) &= C(t)x(t) + D(t)u(t), \end{aligned}$$

is called *controllable* on $[t_0, t_f]$ if given any initial state x_0 there exists a rd-continuous input signal $u(t)$ such that the corresponding solution of the system satisfies $x(t_f) = x_f$.

We begin by giving a necessary and sufficient condition for a regressive linear dynamic system to be controllable. The result is an analogue of the corresponding known results for $\mathbb{R}$ and $\mathbb{Z}$, for which the theorem reduces to the classical results in each case.

Theorem 3.1 (Controllability Gramian Condition). *The regressive linear state equation*

$$x^{\Delta}(t) = A(t)x(t) + B(t)u(t), \;\; x(t_0) = x_0, \qquad (3.2)$$

$$y(t) = C(t)x(t) + D(t)u(t), \qquad (3.3)$$

is controllable on $[t_0, t_f]$ *if and only if the* $n \times n$ *controllability Gramian matrix*

$$\mathcal{G}_C(t_0, t_f) = \int_{t_0}^{t_f} \Phi_A(t_0, \sigma(t)) B(t) B^T(t) \Phi_A^T(t_0, \sigma(t)) \; \Delta t,$$

is invertible.

Proof. Suppose $\mathcal{G}_C(t_0, t_f)$ is invertible. Then, given x_0 and x_f, we can choose the input signal $u(t)$ as

$$u(t) = -B^T(t)\Phi_A^T(t_0, \sigma(t))\mathcal{G}_C^{-1}(t_0, t_f)(x_0 - \Phi_A(t_0, t_f)x_f), \quad t \in [t_0, t_f),$$

and extend $u(t)$ continuously for all other values of t. The corresponding solution of the system at $t = t_f$ can be written as

$$\begin{aligned}
x(t_f) &= \Phi_A(t_f, t_0)x_0 + \int_{t_0}^{t_f} \Phi_A(t_f, \sigma(t))B(t)u(t)\Delta t \\
&= \Phi_A(t_f, t_0)x_0 \\
&- \int_{t_0}^{t_f} \Phi_A(t_f, \sigma(t))B(t)B^T(t)\Phi_A^T(t_0, \sigma(t))\mathcal{G}_C^{-1}(t_0, t_f)(x_0 - \Phi_A(t_0, t_f)x_f)\ \Delta t, \\
&= \Phi_A(t_f, t_0)x_0 \\
&- \Phi_A(t_f, t_0)\int_{t_0}^{t_f} \Phi_A(t_0, \sigma(t))B(t)B^T(t)\Phi_A^T(t_0, \sigma(t))\ \Delta t\, \mathcal{G}_C^{-1}(t_0, t_f)(x_0 - \Phi_A(t_0, t_f)x_f) \\
&= \Phi_A(t_f, t_0)x_0 - (\Phi_A(t_f, t_0)x_0 - x_f) \\
&= x_f,
\end{aligned}$$

so that the state equation is controllable on $[t_0, t_f]$.

For the converse, suppose that the state equation is controllable, but for the sake of a contradiction, assume that the matrix $\mathcal{G}_C(t_0, t_f)$ is not invertible. If $\mathcal{G}_C(t_0, t_f)$ is not invertible, then there exists a vector $x_a \neq 0$ such that

$$0 = x_a^T \mathcal{G}_C(t_0, t_f)x_a = \int_{t_0}^{t_f} x_a^T \Phi_A(t_0, \sigma(t))B(t)B^T(t)\Phi_A^T(t_0, \sigma(t))x_a\ \Delta t. \tag{3.4}$$

But the function in this expression is the nonnegative continuous function $||x_a^T \Phi_A(t_0, \sigma(t))B(t)||^2$, and so it follows that

$$x_a^T \Phi_A(t_0, \sigma(t))B(t) = 0, \ \ t \in [t_0, t_f). \tag{3.5}$$

However, the state equation is controllable on $[t_0, t_f]$, and so choosing $x_0 = x_a + \Phi_A(t_0, t_f)x_f$, there exists an input signal $u_a(t)$ such that

$$x_f = \Phi_A(t_f, t_0)x_0 + \int_{t_0}^{t_f} \Phi_A(t_f, \sigma(t))B(t)u_a(t)\ \Delta t,$$

which is equivalent to the equation

$$x_a = -\int_{t_0}^{t_f} \Phi_A(t_0, \sigma(t))B(t)u_a(t)\ \Delta t.$$

Multiplying through by x_a^T and using (3.4) and (3.5) yields $x_a^T x_a = 0$, a contradiction. Thus, the matrix $\mathcal{G}_C(t_0, t_f)$ is invertible. $\square$

The controllability Gramian is symmetric and positive semidefinite. Thus, the preceding theorem states that (3.1) is controllable on $[t_0, t_f]$ if and only if the Gramian is positive definite. A system that is not controllable on $[t_0, t_f]$ may become so when either t_f is increased or t_0 is decreased. Likewise, a system that is controllable on $[t_0, t_f]$ may become uncontrollable if t_0 is increased and/or t_f is decreased.

Although the preceding theorem is strong in theory, in practice it has limitations. Indeed, computing the Gramian requires explicit knowledge of the transition matrix which is generally not known and can even be difficult to approximate in some cases. Thus, an easier sufficient condition to check is given by the following definition and theorem.

Definition 3.2. If $\mathbb{T}$ is a time scale such that μ is sufficiently differentiable with the indicated derivatives and rd-continuity existing, define the sequence of $n \times m$ matrix functions

$$\begin{aligned} K_0(t) &= B(t) \\ K_{j+1}(t) &= (I + \mu(\sigma(t))A(\sigma(t)))^{-1}K_j^{\Delta}(t) \\ &- \big[(I + \mu(\sigma(t))A(\sigma(t)))^{-1}(\mu^{\Delta}(t)A(\sigma(t)) \\ &+ \mu(t)A^{\Delta}(t))(I + \mu(t)A(t))^{-1} + A(t)(I + \mu(t)A(t))^{-1}\big]\, K_j(t), \qquad j = 0, 1, 2, \ldots \end{aligned}$$

A straightforward induction proof will show that for all t, s, we have that

$$\frac{\partial^j}{\Delta s^j}\left[\Phi_A(\sigma(t), \sigma(s))B(s)\right] = \Phi_A(\sigma(t), \sigma(s))K_j(s), \qquad j = 0, 1, \ldots$$

Indeed, note that the claim trivially holds for $j = 0$, and for $j = n$, our inductive hypothesis gives

$$\frac{\partial^n}{\Delta s^n}\left[\Phi_A(\sigma(t), \sigma(s))B(s)\right] = \Phi_A(\sigma(t), \sigma(s))K_n(s).$$

Hence,

$$
\begin{aligned}
&\frac{\partial^{n+1}}{\Delta s^{n+1}}\left[\Phi_A(\sigma(t),\sigma(s))B(s)\right] \\
&= \frac{\partial}{\Delta s}\left[\Phi_A(\sigma(t),\sigma(s))K_n(s)\right] \\
&= \Phi_A(\sigma(t),\sigma(s))[(I+\mu(\sigma(s))A(\sigma(s)))^{-1}K_n^{\Delta}(s) \\
&-[(I+\mu(\sigma(s))A(\sigma(s)))^{-1}(\mu^{\Delta}(s)A(\sigma(s))+\mu(s)A^{\Delta}(s))(I+\mu(s)A(s))^{-1} \\
&+A(s)(I+\mu(s)A(s))^{-1}]K_n(s)] \\
&= \Phi_A(\sigma(t),\sigma(s))K_{n+1}(s).
\end{aligned}
$$

Evaluation of these matrices at $s = t$ yields a nice relationship between these matrices and the matrices given in the definition above:

$$
K_j(t) = \frac{\partial^j}{\Delta s^j}\left[\Phi_A(\sigma(t),\sigma(s))B(s)\right]\Bigg|_{s=t}, \qquad j = 0, 1, 2, \ldots
$$

This relationship enables us to establish the following result.

Theorem 3.2 (Controllability Rank Theorem). *Suppose q is a positive integer such that, for $t \in [t_0, t_f]$, $B(t)$ is q-times rd-continuously differentiable and both of $\mu(t)$ and $A(t)$ are $(q-1)$-times rd-continuously differentiable. Then the regressive linear system*

$$
\begin{aligned}
x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \;\; x(t_0) = x_0, \qquad (3.6)\\
y(t) &= C(t)x(t) + D(t)u(t),
\end{aligned}
$$

is controllable on $[t_0, t_f]$ if for some $t_c \in [t_0, t_f)$, we have

$$
\operatorname{rank}\begin{bmatrix} K_0(t_c) & K_1(t_c) & \ldots & K_q(t_c) \end{bmatrix} = n,
$$

where

$$
K_j(t) = \frac{\partial^j}{\Delta s^j}\left[\Phi_A(\sigma(t),\sigma(s))B(s)\right]\Bigg|_{s=t}, \qquad j = 0, 1, \ldots, q.
$$

Proof. Suppose there is some $t_c \in [t_0, t_f)$ such that the rank condition holds. For the sake of a contradiction, suppose that the state equation is not controllable on $[t_0, t_f]$. Then the controllability Gramian $\mathcal{G}_C(t_0, t_f)$ is not invertible and, as in the proof of Theorem 3.1, there exists a nonzero $n \times 1$ vector x_a such that

$$x_a^T \Phi_A(t_0, \sigma(t))B(t) = 0, \quad t \in [t_0, t_f). \tag{3.7}$$

If we choose the nonzero vector x_b so that $x_b = \Phi_A^T(t_0, \sigma(t_c))x_a$, then (3.7) yields

$$x_b^T \Phi_A(\sigma(t_c), \sigma(t))B(t) = 0, \quad t \in [t_0, t_f).$$

In particular, at $t = t_c$, we have $x_b^T K_0(t_c) = 0$. Differentiating (3.7) with respect to t,

$$x_b^T \Phi_A(\sigma(t_c), \sigma(t))K_1(t) = 0, \quad t \in [t_0, t_f),$$

so that $x_b^T K_1(t_c) = 0$. In general,

$$\frac{d^j}{\Delta t^j}\left[x_b^T \Phi_A^T(\sigma(t_c), \sigma(t))B(t)\right]\Bigg|_{t=t_c} = x_b^T K_j(t_c) = 0, \qquad j = 0, 1, \ldots, q.$$

Thus,

$$x_b^T \begin{bmatrix} K_0(t_c) & K_1(t_c) & \ldots & K_q(t_c) \end{bmatrix} = 0,$$

which contradicts the linear independence of the rows guaranteed by the rank condition. Hence, the equation is controllable on $[t_0, t_f]$. □

On $\mathbb{R}$, the collection of matrices $K_j(t)$ is such that each member is the jth derivative of the matrix $\Phi_A(\sigma(t), \sigma(s))B(s) = \Phi_A(t, s)B(s)$. This agrees with the literature in the continuous case (see [39] for example). However, while still tractable, in general the collection $K_j(t)$ is nontrivial to compute. The mechanics are more involved even on $\mathbb{Z}$, which is still a very "tame" time scale. Thus, the complications of the extension of the usual theory become self evident.

Furthermore, the preceding theorem shows that if the rank condition holds for some q and some $t_c \in [t_0, t_f)$, then the linear state equation is controllable on *any* interval $[t_0, t_f]$ containing t_c. This strong conclusion partly explains why the condition is only a sufficient one.

3.1.2 Time Invariant Case

We now turn our attention to establishing results concerning the controllability of regressive linear time invariant systems. The Laplace transform presented in Chapter 2 will provide us with results that are not available in the time varying case. We begin with an analogue of Theorem 3.2 that gives a necessary and sufficient condition for controllability. We first need the following lemma.

Lemma 3.1. *Given* $A, B \in \mathbb{R}^{n\times n}$, *and* $u = u_{x_0}(t_f, \sigma(s)) \in \mathbb{R}^{n\times 1}$ *an arbitrary rd-continuous function, then*

$$\operatorname{span}\left\{ \int_{t_0}^{t_f} e_A(s,t_0)Bu_{x_0}(t_f,\sigma(s))\Delta s \right\} = \operatorname{span}\{B, AB, \ldots, A^{n-1}B\}. \tag{3.8}$$

Proof. Let $\{\gamma_k(t,t_0)\}_{k=0}^{n-1}$ be the collection of functions that decompose the exponential matrix as guaranteed by Theorem 1.27. This collection forms a linearly independent set since it can be taken as the solution set of an n-th order system of linear ODEs. Apply the Gram-Schmidt process to generate an orthonormal collection $\{\hat{\gamma}_k(t,t_0)\}_{k=0}^{n-1}$. The two collections are related by

$$\begin{bmatrix} \gamma_0(t,t_0) & \gamma_1(t,t_0) & \ldots & \gamma_{n-1}(t,t_0) \end{bmatrix} \tag{3.9}$$

$$= \begin{bmatrix} \hat{\gamma}_0(t,t_0) & \hat{\gamma}_1(t,t_0) & \ldots & \hat{\gamma}_{n-1}(t,t_0) \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ 0 & p_{22} & \ldots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & p_{nn} \end{bmatrix}, \tag{3.10}$$

where the matrix on the right is the triangular matrix obtained from the QR factorization of the vector consisting of the functions $\{\gamma_k(t,t_0)\}_{k=0}^{n-1}$ on the left.

Using the QR factorization, we can write the matrix exponential as

$$\begin{aligned} e_A(t,t_0) &= \sum_{k=0}^{n-1} \gamma_k(t,t_0)A^k \\ &= \sum_{k=0}^{n-1} \begin{bmatrix} \hat{\gamma_0}(t,t_0) & \hat{\gamma_1}(t,t_0) & \ldots & \hat{\gamma}_{n-1}(t,t_0) \end{bmatrix} \cdot p_k A^k, \end{aligned}$$

where p_k is the k-th column vector of the triangular matrix R. It is worth recalling here that the entries on the diagonal of this matrix are norms of nonzero vectors and are thus nonzero and positive. That is, $p_{ii} > 0$ for all i.

Rewriting the integral from (3.8),

$$\begin{aligned} & \int_{t_0}^{t_f} e_A(s,t_0)Bu_{x_0}(t_f,\sigma(s))\Delta s \\ = & \int_{t_0}^{t_f} \sum_{k=0}^{n-1} \gamma_k(s,t_0)A^k B u_{x_0}(t_f,\sigma(s))\Delta s \\ = & \sum_{k=0}^{n-1} A^k B \int_{t_0}^{t_f} \gamma_k(s,t_0)u_{x_0}(t_f,\sigma(s))\Delta s \\ = & \sum_{k=0}^{n-1} A^k B \int_{t_0}^{t_f} \begin{bmatrix} \hat{\gamma_0}(s,t_0) & \hat{\gamma_1}(s,t_0) & \ldots & \hat{\gamma_{n-1}}(s,t_0) \end{bmatrix} \cdot p_k u_{x_0}(t_f,\sigma(s))\Delta s. \end{aligned}$$

Let

$$y_k = \int_{t_0}^{t_f} \begin{bmatrix} \hat{\gamma_0}(s,t_0) & \hat{\gamma_1}(s,t_0) & \ldots & \hat{\gamma_{n-1}}(s,t_0) \end{bmatrix} \cdot p_k u_{x_0}(t_f,\sigma(s))\Delta s, \qquad k = 0,1,\ldots,n-1.$$

We will show that $\text{span}\{y_0, y_1, \ldots, y_{n-1}\} = \mathbb{R}^n$. That is, there exists some $u \in C_{\text{rd}}(\mathbb{R}^{n\times 1})$ so that for any arbitrary but fixed collection of vectors $\{z_0, z_1, \ldots, z_{n-1}\} \in \mathbb{R}^{n\times 1}$, the system

$$\int_{t_0}^{t_f} p_{11}\hat{\gamma_0}(s,t_0)u_{x_0}(t_f,\sigma(s))\Delta s := z_0 = \begin{bmatrix} z_{00} \\ z_{01} \\ \vdots \\ z_{0(n-1)} \end{bmatrix}$$

$$\int_{t_0}^{t_f} \left(\hat{\gamma_0}(s,t_0)p_{12} + \hat{\gamma_1}(s,t_0)p_{22}\right) u_{x_0}(t_f,\sigma(s))\Delta s := z_1 = \begin{bmatrix} z_{10} \\ z_{11} \\ \vdots \\ z_{1(n-1)} \end{bmatrix}$$

$$\vdots$$

$$\int_{t_0}^{t_f} \left(\hat{\gamma_0}(s,t_0)p_{1n} + \ldots + \hat{\gamma_{n-1}}(s,t_0)p_{nn}\right) u_{x_0}(t_f,\sigma(s))\Delta s := z_{n-1} = \begin{bmatrix} z_{(n-1)0} \\ z_{(n-1)1} \\ \vdots \\ z_{(n-1)(n-1)} \end{bmatrix}$$

has a solution.

To accomplish this, we use the fact that the collection $\hat{\gamma_k}(s,t_0)$ is orthonormal and search for a solution of the form

$$u_{x_0}(t_f,\sigma(s)) = (u_j) = \left(\sum_{i=0}^{n-1} \beta_i^j \hat{\gamma_i}(s,t_0)\right).$$

Starting with u_0, the equations become

$$\begin{aligned}
&\int_{t_0}^{t_f} \hat{\gamma_0}p_{11}\left(\sum_{i=0}^{n-1} \beta_i^0 \hat{\gamma_i}\right)\Delta s = z_{00} \\
&\int_{t_0}^{t_f} \left(\hat{\gamma_0}p_{12} + \hat{\gamma_1}p_{22}\right)\left(\sum_{i=0}^{n-1} \beta_i^0 \hat{\gamma_i}\right)\Delta s = z_{10} \\
&\vdots \\
&\int_{t_0}^{t_f} \left(\hat{\gamma_0}p_{1n} + \hat{\gamma_1}p_{2n} + \ldots + \hat{\gamma_{n-1}}p_{nn}\right)\left(\sum_{i=0}^{n-1} \beta_i^0 \hat{\gamma_i}\right)\Delta s = z_{(n-1)0}.
\end{aligned}$$

Since the system $\hat{\gamma_k}$ is orthonormal, we can simplify the equations above using the fact that the integral of cross terms $\hat{\gamma_i}\hat{\gamma_j}$, $i \neq j$, is zero. After doing so, the system becomes a lower triangular system that can be solved by forward substitution. (The observation that $p_{ii} \neq 0$ is crucial here, since this is exactly what allows us to solve the system.) For example, the first equation becomes

$$\begin{aligned}\int_{t_0}^{t_f} \hat{\gamma}_0 p_{11} \left(\sum_{i=0}^{n-1} \beta_i^0 \hat{\gamma}_i \right) \Delta s &= \int_{t_0}^{t_f} \hat{\gamma}_0{}^2 \beta_0^0 p_{11} \Delta s \\ &= \beta_0^0 p_{11} \\ &= z_{01},\end{aligned}$$

so that $\beta_0^0 = \frac{z_{00}}{p_{11}}$. Using this value for β_0^0 in the second equation,

$$\begin{aligned}\int_{t_0}^{t_f} (\hat{\gamma}_0 p_{12} + \hat{\gamma}_1 p_{22}) \left(\sum_{i=0}^{n-1} \beta_i^0 \hat{\gamma}_i \right) \Delta s &= \int_{t-0}^{t_f} (\hat{\gamma}_0 p_{12} + \hat{\gamma}_1 p_{22}) \left(\beta_0^0 \hat{\gamma}_0 + \beta_1^0 \hat{\gamma}_1 \right) \Delta s \\ &= \frac{p_{12}}{p_{11}} z_{01} + \beta_1^1 p_{22} \\ &= z_{10},\end{aligned}$$

so that $\beta_1^0 = \frac{1}{p_{22}} z_{11} - \frac{p_{12}}{p_{11} p_{22}} z_{01}$. We can continue solving the system in like manner by using forward substitutions to find β_j^0 for all $j = 0, 1, \ldots, n-1$, which will in turn yield $u_0 = \sum_{i=0}^{n-1} \beta_i^0 \hat{\gamma}_i$. Repeating this process for $u_1, u_2, \ldots, u_{n-1}$, we find the correct linear combinations of $\hat{\gamma}_k$ to solve the system, and so the claim follows. $\square$

Theorem 3.3 (Kalman Controllability Rank Condition). *The regressive linear time invariant system*

$$\begin{aligned}x^\Delta(t) &= Ax(t) + Bu(t), \;\; x(t_0) = x_0, \\ y(t) &= Cx(t) + Du(t),\end{aligned} \tag{3.11}$$

is controllable on $[t_0, t_f]$ *if and only if the* $n \times nm$ *controllability matrix*

$$\begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix},$$

satisfies

$$\operatorname{rank} \begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix} = n.$$

Proof. Suppose the system is controllable, but for the sake of a contradiction that the rank condition fails. Then there exists an $n \times 1$ vector x_a such that $x_a^T A^k B = 0, k = 0, \ldots, n-1$. Now, there are two cases to consider: either $x_a^T x_f = 0$ or $x_a^T x_f \neq 0$.

Suppose $x_a^T x_f \neq 0$. Then for any t, the solution at time t is given by

$$\begin{aligned} x(t) &= \int_{t_0}^{t} e_A(t,\sigma(s))Bu_{x_0}(s)\ \Delta s\ + e_A(t,t_0)x_0 \\ &= e_A(t,0) * Bu(t) + e_A(t,0)x_0 \\ &= Bu(t) * e_A(t,0) + e_A(t,0)x_0 \\ &= \int_{t_0}^{t} e_A(s,t_0)Bu_{x_0}(t,\sigma(s))\ \Delta s\ + e_A(t,t_0)x_0, \end{aligned}$$

where we observe that the solution is a convolution and is commutative. Choose initial state $x_0 = By$, where y is arbitrary. Then, again by commutativity of the convolution and Theorem 1.27, we see

$$\begin{aligned} x_a^T x(t) &= x_a^T \int_{t_0}^{t} e_A(s,t_0)Bu_{x_0}(t,\sigma(s))\ \Delta s + x_a^T e_A(t,t_0)x_0 \\ &= \int_{t_0}^{t} \sum_{k=0}^{n-1} \gamma_k(s,t_0)x_a^T A^k Bu_{x_0}(t,\sigma(s))\ \Delta s + \sum_{k=0}^{n-1} \gamma_k(t,t_0)x_a^T A^k By \\ &= 0, \end{aligned}$$

so that $x_a^T x(t) = 0$ for all t, which is a contradiction since $x_a^T x(t_f) = x_a^T x_f \neq 0$.

Now suppose $x_a^T x_f = 0$. This time, we choose initial state $x_0 = e_A^{-1}(t_f,t_0)x_a$. Similar to the equation above,

$$\begin{aligned} x_a^T x(t) &= \int_{t_0}^{t} \sum_{k=0}^{n-1} \gamma_k(s,t_0)x_a^T A^k Bu_{x_0}(t,\sigma(s))\ \Delta s + x_a^T e_A(t,t_0)e_A^{-1}(t_f,t_0)x_a \\ &= x_a^T e_A(t,t_0)e_A^{-1}(t_f,t_0)x_a. \end{aligned}$$

In particular, at $t = t_f$, $x_a^T x(t_f) = ||x_a||^2 \neq 0$, another contradiction.

Thus in either case we arrive at a contradiction, and so controllability implies the rank condition.

Conversely, suppose that the system is not controllable. Then there exists an initial state $x_0 \in \mathbb{R}^{n\times 1}$ such that for all input signals $u(t) \in \mathbb{R}^{m\times 1}$, we have that

$x(t_f) \neq x_f$. Again, it follows from the commutativity of the convolution that

$$\begin{aligned} x_f \neq x(t_f) &= \int_{t_0}^{t_f} e_A(t_f, \sigma(s)) B u_{x_0}(s) \ \Delta s + e_A(t_f, t_0) x_0 \\ &= \int_{t_0}^{t_f} e_A(s, t_0) B u_{x_0}(t_f, \sigma(s)) \ \Delta s + e_A(t_f, t_0) x_0 \\ &= \int_{t_0}^{t_f} \sum_{k=0}^{n-1} \gamma_k(s, t_0) A^k B u_{x_0}(t_f, \sigma(s)) \ \Delta s \ + e_A(t_f, t_0) x_0. \end{aligned}$$

In particular,

$$\sum_{k=0}^{n-1} A^k B \int_{t_0}^{t_f} \gamma_k(s, t_0) u_{x_0}(t_f, \sigma(s)) \ \Delta s \neq x_f - e_A(t_f, t_0) x_0.$$

Notice that the last equation holds if and only if there is no linear combination of the matrices $A^k B$ for $k = 0, 1, \ldots, n-1$, which satisfy

$$\sum_{k=0}^{n-1} A^k B \alpha_k = x_f - e_A(t_f, t_0) x_0.$$

The fact that there is no such linear combination follows from Lemma 3.1 once we realize that an argument similar to the one given in the proof of this result holds if $m < n$. Thus, the matrix

$$\begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix},$$

cannot have rank n, and so we have shown that if the matrix has rank n, then it is controllable by contraposition. $\square$

The preceding theorem is commonly called the Kalman rank condition after R.E. Kalman who first proved it in 1960 (see [33] and [35]) for the cases $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. Thus, in our analysis, we have again unified the two cases: however, once again we have also extended the result to the arbitrary time scale with bounded graininess. However, it is important to point out that the proof here is not the one that Kalman gave, which is the one classically used for $\mathbb{R}$ and $\mathbb{Z}$ (see [39] for example). In these two special cases, an observation about the particular form of the matrix exponential on $\mathbb{R}$ and $\mathbb{Z}$ allows one to arrive at the result in a more straightforward manner.

Example 3.1. Consider the system

$$
\begin{aligned}
x^{\Delta}(t) &= \begin{bmatrix} -\frac{8}{45} & \frac{1}{30} \\ -\frac{1}{45} & -\frac{1}{10} \end{bmatrix} x(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u(t), \quad x(0) = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \\
y(t) &= \begin{bmatrix} 3 & 4 \end{bmatrix} x(t).
\end{aligned}
$$

It is straightforward to verify that

$$
\operatorname{rank} \begin{bmatrix} B & AB \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 2 & -\frac{29}{90} \\ 1 & -\frac{13}{90} \end{bmatrix} = 2,
$$

so that the state equation is controllable by Theorem 3.3.

The next theorem establishes that there is a state variable change in the time invariant case that demonstrates the "controllable part" of the state equation.

Theorem 3.4. *Suppose the controllability matrix for the regressive linear time invariant state equation*

$$
\begin{aligned}
x^{\Delta}(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\
y(t) &= Cx(t),
\end{aligned}
$$

satisfies

$$
\operatorname{rank} \begin{bmatrix} B & AB & \ldots A^{n-1}B \end{bmatrix} = q,
$$

where $0 < q < n$. *Then there exists an invertible matrix* P *such that*

$$
P^{-1}AP = \begin{bmatrix} \hat{A_{11}} & \hat{A_{12}} \\ 0_{(n-q)\times q} & \hat{A_{22}} \end{bmatrix}, \qquad P^{-1}B = \begin{bmatrix} \hat{B_{11}} \\ 0_{(n-q)\times m} \end{bmatrix},
$$

where $\hat{A_{11}}$ *is* $q \times q$, $\hat{B_{11}}$ *is* $q \times m$, *and*

$$
\operatorname{rank} \begin{bmatrix} \hat{B_{11}} & \hat{A_{11}}\hat{B_{11}} & \ldots & \hat{A_{11}}^{q-1}\hat{B_{11}} \end{bmatrix} = q.
$$

Proof. We begin constructing P by choosing q linearly independent columns $p_1, p_2, \ldots, p_q$, from the controllability matrix for the system. Then choose $p_{q+1}, \ldots, p_n$ as $n \times 1$ vectors so that

$$P = \begin{bmatrix} p_1 & \ldots & p_q & p_{q+1} & \ldots & p_n \end{bmatrix}$$

is invertible. Define G so that $PG = B$. Writing the j-th column of B as a linear combination of the linearly independent columns of P given by $p_1, p_2, \ldots, p_q$, we find that the last $n - q$ entries of the j-th column of G must be zero. This argument holds for $j = 1, \ldots, m$, and so $G = P^{-1}B$ does indeed have the desired form.

Now set $F = P^{-1}AP$, yielding

$$PF = \begin{bmatrix} Ap_1 & Ap_2 \ldots Ap_n \end{bmatrix}.$$

The column vectors $Ap_1, \ldots, Ap_q$ can be written as linear combinations of $p_1, \ldots, p_n$ since each column of $A^k B, k \geq 0$ can be written as a linear combination of these vectors. As for G above, the first q columns of F must have zeros as the last $n - q$ entries. Thus, $P^{-1}AP$ has the desired form. Multiply the rank-q controllability matrix by P^{-1} to obtain

$$\begin{aligned} P^{-1} \begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix} &= \begin{bmatrix} P^{-1}B & P^{-1}AB & \ldots & P^{-1}A^{n-1}B \end{bmatrix} \\ &= \begin{bmatrix} G & FG & \ldots & F^{n-1}G \end{bmatrix} \\ &= \begin{bmatrix} \hat{B_{11}} & \hat{A_{11}}\hat{B_{11}} & \ldots & \hat{A_{11}}^{n-1}\hat{B_{11}} \\ 0 & 0 & \ldots & 0 \end{bmatrix}. \end{aligned}$$

Applying the Cayley-Hamilton theorem gives

$$\text{rank} \begin{bmatrix} \hat{B_{11}} & \hat{A_{11}}\hat{B_{11}} & \ldots & \hat{A_{11}}^{q-1}\hat{B_{11}} \end{bmatrix} = q.$$

□

Theorem 3.5. *The regressive linear time invariant state equation*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \;\; x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is controllable if and only if for every scalar λ *the only complex* $n \times 1$ *vector* p *that satisfies* $p^T A = \lambda p^T, p^T B = 0$ *is* $p = 0$.

Proof. For necessity, note that if there exists $p \neq 0$ and a complex λ such that the equation given is satisfied, then it follows that

$$\begin{aligned} p^T \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} &= \begin{bmatrix} p^T B & p^T AB & \dots & p^T A^{n-1} B \end{bmatrix} \\ &= \begin{bmatrix} p^T B & \lambda p^T B & \dots & \lambda^{n-1} p^T B \end{bmatrix}, \end{aligned}$$

so that the n rows rows of the controllability matrix are linearly dependent, and hence the system is not controllable.

For sufficiency, suppose that the state equation is not controllable. Then by Theorem 3.4, there exists an invertible P such that

$$P^{-1}AP = \begin{bmatrix} \hat{A_{11}} & \hat{A_{12}} \\ 0_{(n-q)\times q} & \hat{A_{22}} \end{bmatrix}, \qquad P^{-1}B = \begin{bmatrix} \hat{B_{11}} \\ 0_{(n-q)\times m} \end{bmatrix},$$

with $0 < q < n$. Let $p^T = \begin{bmatrix} 0_{1\times q} & p_q^T \end{bmatrix} P^{-1}$, where p_q is a left eigenvector for $\hat{A_{22}}$. Thus, for some complex scalar λ, $p_q^T \hat{A_{22}} = \lambda p_q^T, p_q \neq 0$. Then $p \neq 0$, and

$$\begin{aligned} p^T B &= \begin{bmatrix} 0 & p_q^T \end{bmatrix} \begin{bmatrix} \hat{B_{11}} \\ 0 \end{bmatrix} = 0 \\ p^T A &= \begin{bmatrix} 0 & p_q^T \end{bmatrix} \begin{bmatrix} \hat{A_{11}} & \hat{A_{12}} \\ 0 & \hat{A_{22}} \end{bmatrix} P^{-1} = \begin{bmatrix} 0 & \lambda p_q^T \end{bmatrix} P^{-1} = \lambda p^T. \end{aligned}$$

Thus, the claim follows. $\square$

We can paraphrase the preceding result as saying that in a controllable time invariant system, A can have no left eigenvectors that are orthogonal to the columns of B. This fact can then be used to prove the next theorem.

Theorem 3.6. *The regressive linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \;\; x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is controllable if and only if

$$\operatorname{rank} \begin{bmatrix} zI - A & B \end{bmatrix} = n$$

for every complex scalar z.

Proof. By Theorem 3.5, the state equation is not controllable if and only if there exists a nonzero complex $n \times 1$ vector p and complex scalar λ such that

$$p^T \begin{bmatrix} \lambda I - A & B \end{bmatrix}, \; p \neq 0.$$

But this condition is equivalent to saying that

$$\operatorname{rank} \begin{bmatrix} \lambda I - A & B \end{bmatrix} < n,$$

so the claim follows. □

3.2 Observability

The next notion from linear systems theory that we explore is observability. As before, we will treat the time varying case first followed by the time invariant case.

3.2.1 Time Varying Case

In linear systems theory, when the term *observability* is used, it refers to the effect that the state vector has on the output of the state equation. As such, the concept is unchanged by considering simply the response of the system to zero input. Motivated by this, we define the following.

Definition 3.3. The regressive linear state equation

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t), \;\; x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is called *observable* on $[t_0, t_f]$ if any initial state $x(t_0) = x_0$ is uniquely determined by the corresponding response $y(t)$ for $t \in [t_0, t_f)$.

The notions of controllability and observability can be thought of as dual to one another. Thus, any theorem that we obtain for controllability should have an analogue in terms of observability. Thus, we begin by formulating observability in terms of the *Gramian*.

Theorem 3.7 (Observability Gramian Condition). *The regressive linear system*

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t), \;\; x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is observable on $[t_0, t_f]$ *if and only if the* $n \times n$ *observability Gramian matrix*

$$\mathcal{G}_O(t_0, t_f) = \int_{t_0}^{t_f} \Phi_A^T(t, t_0) C^T(t) C(t) \Phi_A(t, t_0) \; \Delta t$$

is invertible.

Proof. If we multiply the solution expression

$$y(t) = C(t)\Phi_A(t, t_0)x_0,$$

on both sides by $\Phi_A^T(t, t_0)C(t)$ and integrate, we obtain

$$\int_{t_0}^{t_f} \Phi_A^T(t, t_0) C^T(t) y(t) \Delta t = \mathcal{G}_O(t_0, t_f) x_0.$$

The left side of this equation is determined by $y(t)$ for $t \in [t_0, t_f)$, and thus this equation is a linear algebraic equation in x_0. If $\mathcal{G}_O(t_0, t_f)$ is invertible, then x_0 is uniquely determined.

Conversely, if $\mathcal{G}_O(t_0, t_f)$ is not invertible, then there exists a nonzero vector x_a such that $\mathcal{G}_O(t_0, t_f)x_a = 0$. But then $x_a^T \mathcal{G}_O(t_0, t_f)x_a = 0$, so that

$$C(t)\Phi_A(t, t_0)x_a = 0, \quad t \in [t_0, t_f).$$

Thus, $x(t_0) = x_0 + x_a$ yields the same zero-input response for the system as $x(t_0) = x_0$, and so the system is not observable on $[t_0, t_f]$. □

The observability Gramian, like the controllability Gramian, is symmetric positive semidefinite. It is positive definite if and only if the state equation is observable.

Once again we see that the Gramian condition is not very practical as it requires explicit knowledge of the transition matrix. Thus, we present a sufficient condition that is easier to check for observability. As before, observability and controllability can be considered dual notions to one another, and as such, proofs of corresponding results are often similar if not the same. Thus, any result for which we do not give the proof in observability has a proof that mirrors the proof of the result for controllability.

Definition 3.4. If $\mathbb{T}$ is a time scale such that μ is sufficiently differentiable with the indicated derivatives and rd-continuity existing, define the sequence of $p \times n$ matrix functions

$$\begin{aligned} L_0(t) &= C(t), \\ L_{j+1}(t) &= L_j(t)A(t) + L_j^{\Delta}(t)(I + \mu(t)A(t)), \quad j = 0, 1, 2, \dots \ . \end{aligned}$$

As in the case of controllability, an induction argument shows that

$$L_j(t) = \frac{\partial^j}{\Delta t^j}\left[C(t)\Phi_A(t, s)\right]\Bigg|_{s=t}.$$

With this, an argument similar to the one before shows the following:

Theorem 3.8 (Observability Rank Condition). *Suppose q is a positive integer such that, for $t \in [t_0, t_f]$, $C(t)$ is q-times rd-continuously differentiable and both of $\mu(t)$*

and $A(t)$ are $(q-1)$-times rd-continuously differentiable. Then the regressive linear system

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t), \;\; x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned} \tag{3.12}$$

is observable on $[t_0, t_f]$ if for some $t_c \in [t_0, t_f)$, we have

$$\operatorname{rank} \begin{bmatrix} L_0(t_c) \\ L_1(t_c) \\ \vdots \\ L_q(t_c) \end{bmatrix} = n,$$

where

$$L_j(t) = \frac{\partial^j}{\Delta s^j} \left[C(t)\Phi_A(t,s) \right] \Bigg|_{s=t}, \quad j = 0, 1, \dots, q.$$

3.2.2 Time Invariant Case

Like controllability, observability has equivalent conditions that become necessary and sufficient in the time invariant case. The first is the Kalman rank condition, whose statement and proof follow.

Theorem 3.9 (Kalman Observability Rank Condition). *The autonomous linear regressive system*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is observable on $[t_0, t_f]$ if and only if the $nm \times n$ observability matrix

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

satisfies

$$\operatorname{rank}\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n.$$

Proof. Again, we show that the rank condition fails if and only if the observability Gramian is not invertible. Thus, suppose that the rank condition fails. Then there exists a nonzero $n \times 1$ vector x_a such that

$$CA^k x_a = 0, \qquad k = 0, \ldots, n-1.$$

This implies, using Theorem 1.27, that

$$\begin{aligned} \mathcal{G}_O(t_0, t_f)x_a &= \int_{t_0}^{t_f} e_A^T(t, t_0)C^T C e_A(t, t_0)x_a \,\Delta t \\ &= \int_{t_0}^{t_f} e_A^T(t, t_0)C^T \sum_{k=0}^{n-1} \gamma_k(t, t_0) CA^k x_a \,\Delta t \\ &= 0, \end{aligned}$$

so that the Gramian is not invertible.

Conversely, suppose that the Gramian is not invertible. Then there exists nonzero x_a such that

$$x_a^T \mathcal{G}_O(t_0, t_f)x_a = 0.$$

As argued before, this then implies

$$Ce_A(t, t_0)x_a = 0, \;\; t \in [t_0, t_f).$$

At $t = t_0$, we obtain $Cx_a = 0$, and differentiating k times and evaluating the result at $t = t_0$ gives

$$CA^k x_a = 0, \;\; k = 0, \ldots, n-1.$$

Thus,

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x_a = 0,$$

so that the rank condition fails. $\square$

The proof of the preceding result demonstrates an important point about controllability and observability in the arbitrary time scale setting: namely, proofs of similar results for the two notions are often similar, but can sometimes be very different. (Comparing the proof of the Kalman condition for controllability with the one for observability shows this stark contrast.)

The following example makes use of Theorem 3.9.

Example 3.2. Consider the system

$$\begin{aligned} x^{\Delta}(t) &= \begin{bmatrix} -\frac{8}{45} & \frac{1}{30} \\ -\frac{1}{45} & -\frac{1}{10} \end{bmatrix} x(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u(t), \quad x(0) = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \\ y(t) &= \begin{bmatrix} 3 & 4 \end{bmatrix} x(t). \end{aligned}$$

From Example 3.1, recall that the system is controllable. We claim the system is also observable. This follows from Theorem 3.9 since

$$\operatorname{rank} \begin{bmatrix} C \\ CA \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 3 & 4 \\ -\frac{28}{45} & -\frac{3}{10} \end{bmatrix} = 2.$$

The following three theorems concerning observability have proofs that mirror their controllability counterparts, and so will not be given here.

Theorem 3.10. *Suppose the observability matrix for the regressive linear time invariant state equation*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

satisfies

$$\operatorname{rank}\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = \ell,$$

where $0 < \ell < n$*. Then there exists an invertible* $n \times n$ *matrix* Q *such that*

$$Q^{-1}AQ = \begin{bmatrix} \hat{A_{11}} & 0 \\ \hat{A_{21}} & \hat{A_{22}} \end{bmatrix}, \qquad CQ = \begin{bmatrix} \hat{C_{11}} & 0 \end{bmatrix},$$

where $\hat{A_{11}}$ *is* $\ell \times \ell$*,* $\hat{C_{11}}$ *is* $p \times \ell$*, and*

$$\operatorname{rank}\begin{bmatrix} \hat{C_{11}} \\ \hat{C_{11}}\hat{A_{11}} \\ \vdots \\ \hat{C_{11}}\hat{A_{11}^{l-1}} \end{bmatrix} = \ell.$$

The state variable change Theorem 3.10 is constructed by choosing $n-\ell$ vectors in the nullspace of the observability matrix, and preceding them by ℓ vectors that yield a set of n linearly independent vectors.

Theorem 3.11. *The regressive time invariant linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \;\; x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is observable if and only if for every complex scalar λ*, the only complex* $n \times 1$ *vector* p *that satisfies* $Ap = \lambda p, Cp = 0$ *is* $p = 0$*.*

Again, Theorem 3.11 can be restated as saying that in an observable time invariant system, A can have no right eigenvectors that are orthogonal to the rows of C.

Theorem 3.12. *The regressive time invariant linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is observable if and only if

$$\operatorname{rank} \begin{bmatrix} C \\ zI - A \end{bmatrix} = n,$$

for every complex scalar z.

3.3 Realizability

The term *realizability* in linear systems theory refers to the ability to characterize a known output in terms of a linear system with some input. A precise definition of the concept follows.

Definition 3.5. The regressive linear state equation

$$\begin{aligned} x^{\Delta} &= A(t)x(t) + B(t)u(t), \qquad x(t_0) = 0, \\ y(t) &= C(t)x(t), \end{aligned}$$

of dimension n is called a *realization* of the weighting pattern $G(t, \sigma(s))$ if

$$G(t, \sigma(s)) = C(t)\Phi_A(t, \sigma(s))B(s),$$

for all t, s. If a realization of this system exists, then the weighting pattern is called *realizable*. The system is called a *minimal realization* if no realization of $G(t, \sigma(s))$ with dimension less than n exists.

Notice that for the system

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \;\; x(t_0) = 0, \\ y(t) &= C(t)x(t) + D(t)u(t), \end{aligned} \tag{3.13}$$

the output signal $y(t)$ corresponding to a given input $u(t)$ and weighting pattern $G(t, \sigma(s)) = C(t)\Phi_A(t, \sigma(s))B(s)$ is given by

$$y(t) = \int_{t_0}^{t} G(t, \sigma(s))u(s)\ \Delta s + D(t)u(t), \ \ t \geq t_0.$$

When there exists a realization of a particular weighting response $G(t, \sigma(s)$, there will in fact exist many since a change of state variables will leave the weighting pattern unchanged. Also, there can be many different realizations of the same weighting pattern that all have different dimensions. This is why we are careful to distinguish between realizations and minimal realizations in our definition.

We now give equivalent conditions for realizability: as before, we begin with the time variant case and then proceed to the time invariant case.

3.3.1 Time Varying Case

The next theorem gives a characterization of realizable systems in general.

Theorem 3.13 (Factorization of $G(t, \sigma(s))$). *The weighting pattern $G(t, \sigma(s))$ is realizable if and only if there exist a rd-continuous matrix $H(t)$ that is of dimension $q \times n$ and a rd-continuous matrix $F(t)$ of dimension $n \times r$ such that $G(t, \sigma(s)) = H(t)F(\sigma(s))$ for all t, s.*

Proof. Suppose there exist of the matrices $H(t)$ and $F(t)$ with $G(t, \sigma(s)) = H(t)F(\sigma(s))$. Then the system

$$\begin{aligned} x^{\Delta}(t) &= F(t)u(t), \\ y(t) &= H(t)x(t), \end{aligned}$$

is a realization of $G(t, \sigma(s))$ since the transition matrix of the zero system is the identity.

Conversely, suppose that $G(t, \sigma(s))$ is realizable. We may assume that the system

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t), \end{aligned}$$

is one such realization. Since the system is regressive, we may write

$$G(t, \sigma(s)) = C(t)\Phi_A(t, \sigma(s))B(s) = C(t)\Phi_A(t, 0)\Phi_A(0, \sigma(s))B(s),$$

and so by choosing $H(t) = C(t)\Phi_A(t, 0)$ and $F(t) = \Phi_A(0, \sigma(t))B(t)$, the claim follows. $\square$

Although the preceding theorem gives a basic condition for realization of linear systems, often in practice it is not very useful because writing the weighting pattern in its factored form can be very difficult. Also, as the next example demonstrates, the realization given by the factored form can often be undesirable in certain aspects.

Example 3.3. Suppose $\mathbb{T}$ is a time scale with $0 \leq \mu \leq 2$. Under this assumption, $-\frac{1}{4} \in \mathcal{R}^+(\mathbb{T})$. Then the weighting pattern

$$G(t, \sigma(s)) = e_{-1/4}(t, \sigma(s)),$$

has the factorization

$$G(t, \sigma(s)) = e_{-1/4}(t, \sigma(s)) = e_{-1/4}(t, 0)e_{\ominus(-1/4)}(\sigma(s), 0).$$

By the previous theorem, a one-dimensional realization of G is

$$\begin{aligned} x^{\Delta}(t) &= e_{\ominus(-1/4)}(t, 0)u(t), \\ y(t) &= e_{-1/4}(t, 0)x(t). \end{aligned}$$

This state equation has an unbounded coefficient (note that $e_{\ominus(-1/4)}(t,0) = e_{1/(4-\mu)}(t,0)$ is unbounded since $1/(4-\mu) > 0$) and is not uniformly exponentially stable. However, the one-dimensional realization of G given by

$$\begin{aligned} x^{\Delta}(t) &= -\frac{1}{4}x(t) + u(t), \\ y(t) &= x(t), \end{aligned}$$

does have bounded coefficients and is uniformly exponentially stable. Thus, this realization is the more desirable of the two realizations because of this fact.

Before examining minimal realizations, some remarks are in order. First, note that inverse and σ operators commute:

$$\begin{aligned} P^{-1}(\sigma(t)) &= P^{-1}(t) + \mu(t)(P^{-1}(t))^{\Delta} \\ &= P^{-1}(t) + \mu(t)(-P(\sigma(t)))^{-1}P^{\Delta}(t)P^{-1}(t) \\ &= P^{-1}(t) - (P(\sigma(t))^{-1}(P(\sigma(t)) - P(t))P^{-1}(t) \\ &= (P(\sigma(t)))^{-1}. \end{aligned}$$

Second, it is possible to do a variable change on the system

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)x(t), \\ y(t) &= C(t)x(t), \end{aligned}$$

so that the coefficient of $x^{\Delta}(t)$ in the new system is zero, while at the same time preserving realizability of the system under the change of variables.

Indeed, set $z(t) = P^{-1}(t)x(t)$ and note that $P(t) = \Phi_A(t, t_0)$ satisfies

$$(P(\sigma(t)))^{-1}A(t)P(t) - (P(\sigma(t)))^{-1}P^{\Delta}(t) = 0.$$

If we make this substitution, then the system becomes

$$\begin{aligned} z^{\Delta}(t) &= P^{-1}(\sigma(t))B(t)u(t), \\ y(t) &= C(t)P(t)z(t). \end{aligned}$$

Thus, in terms of realizability, we may assume without loss of generality that $A(t) \equiv 0$ by changing the system to the form given above. We shall make frequent use of this fact when proving the results that follow.

It is important to know when a given realization is minimal. The following theorem gives a necessary and sufficient condition for this in terms of controllability and observability.

Theorem 3.14 (Characterization of Minimal Realizations). *Suppose the regressive linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)x(t), \\ y(t) &= C(t)x(t), \end{aligned}$$

is a realization of the weighting pattern $G(t, \sigma(s))$. Then this realization is minimal if and only if for some t_0 and $t_f > t_0$ the state equation is both controllable and observable on $[t_0, t_f]$.

Proof. As argued above, we may assume without loss of generality that $A(t) \equiv 0$. Suppose the $n-$dimensional realization given is not minimal. Then there is a lower-dimension realization of $G(t, \sigma(s))$ having form

$$\begin{aligned} z^{\Delta}(t) &= R(t)u(t), \\ y(t) &= S(t)z(t), \end{aligned} \tag{3.14}$$

with the dimension of $z(t)$ being $n_z < n$. Writing the weighting pattern in terms of both realizations produces $C(t)B(s) = S(t)R(s)$ for all t, s. Thus,

$$C^T(t)C(t)B(s)B^T(s) = C^T S(t)R(s)B^T(s),$$

for all t, s. For any t_0 and $t_f > t_0$, it is possible to integrate this expression with respect to t and then with respect to s to obtain

$$\mathcal{G}_0(t_0, t_f)\mathcal{G}_C(t_0, t_f) = \int_{t_0}^{t_f} C^T(t)S(t)\ \Delta t \int_{t_0}^{t_f} R(s)B^T(s)\ \Delta s.$$

The right hand side of this equation is the product of an $n \times n_z$ matrix and an $n_z \times n$ matrix, and as such, it cannot have full rank since the dimension of the space spanned by the product is at most $n_z < n$. Therefore, $\mathcal{G}_O(t_0, t_f)$ and $\mathcal{G}_C(t_0, t_f)$ cannot be simultaneously invertible. The argument is independent of the t_0 and t_f chosen, and so sufficiency is established.

Conversely, suppose that the given state equation is a minimal realization of the weighting pattern $G(t, \sigma(s))$, with $A(t) \equiv 0$. We begin by showing that if either

$$\mathcal{G}_C(t_0, t_f) = \int_{t_0}^{t_f} B(t)B^T(t)\ \Delta t,$$

or

$$\mathcal{G}_O(t_0, t_f) = \int_{t_0}^{t_f} C^T(t)C(t)\ \Delta t,$$

is singular for all t_0 and t_f, then minimality is violated. Thus, there exist intervals $[t_0^a, t_f^a]$ and $[t_0^b, t_f^b]$ such that $\mathcal{G}_C(t_0^a, t_f^a)$ and $\mathcal{G}_O(t_0^b, t_f^b)$ are both invertible. If we let $t_0 = \min\{t_0^a, t_0^b\}$ and $t_f = \max\{t_f^a, t_f^b\}$, then the positive definiteness of the observability and controllability Gramians yield that both $\mathcal{G}_C(t_0, t_f)$ and $\mathcal{G}_O(t_0, t_f)$ are invertible.

To show this, we begin by supposing that for every interval $[t_0, t_f]$ the matrix $\mathcal{G}_C(t_0, t_f)$ is not invertible. Then, given t_0 and t_f there exists an $n \times 1$ vector $x = x(t_0, t_f)$ such that

$$0 = x^T \mathcal{G}_C(t_0, t_f) x = \int_{t_0}^{t_f} B(t)B^T(t)x\ \Delta t.$$

Thus, $x^T B(t) = 0$ for $t \in [t_0, t_f)$.

We claim that there exists at least one such vector x that is independent of t_0 and t_f. To this end, note that if $\mathbb{T}$ is unbounded from above and below, then for each positive integer k there exists an $n \times 1$ vector x_k with

$$||x_k|| = 1;\ \ x_k^T B(t) = 0, \quad t \in [-k, k].$$

Thus, $\{x_k\}_{k=1}^{\infty}$ is a bounded sequence of $n \times 1$ vectors and by the Bolzano-Wierstrauss Theorem, it has a convergent subsequence since $\mathbb{T}$ is closed. We label this convergent subsequence by $\{x_{k_j}\}_{j=1}^{\infty}$ and denote its limit by $x_0 = \lim_{j\to\infty} x_{k_j}$. Note $x_0^T B(t) = 0$ for all t, since for any given time t_a, there exists a positive integer J_a such that $t_a \in [-k_j, k_j]$ for all $j \geq J_a$, which in turn implies $x_{k_j}^T B(t_a) = 0$ for all $j \geq J_a$. Hence, x_0^T satisfies $x_0^T B(t_a) = 0$.

Now let P^{-1} be a constant, invertible, $n \times n$ matrix with bottom row x_0^T. Using P^{-1} as a change of state variables gives another minimal realization of the weighting pattern, with coefficient matrices

$$P^{-1}B(t) = \begin{bmatrix} \hat{B}_1(t) \\ 0_{1\times m} \end{bmatrix}, \qquad C(t)P = \begin{bmatrix} \hat{C}_1(t) & \hat{C}_2(t) \end{bmatrix},$$

where $\hat{B}_1(t)$ is $(n-1)\times m$, and $\hat{C}_1(t)$ is $p \times (n-1)$. Then a straightforward calculation shows $G(t, \sigma(s)) = \hat{C}_1(t)\hat{B}_1(\sigma(s))$ so that the linear state equation

$$\begin{aligned} z^{\Delta}(t) &= \hat{B}_1(t)u(t), \\ y(t) &= \hat{C}_1(t)z(t), \end{aligned}$$

is a realization for $G(t, \sigma(s))$ of dimension $n-1$. This contradicts the minimality of the original n-dimensional realization. Thus, there must be at least one t_0^a and one $t_f^a > t_0^a$ such that $\mathcal{G}_C(t_0^a, t_f^a)$ is invertible.

A similar argument shows that there exists at least one t_0^b and one $t_f^b > t_0^b$ such that $\mathcal{G}_O(t_0^b, t_f^b)$ is invertible. Taking $t_0 = \min\{t_0^a, t_0^b\}$ and $t_f = \max\{t_f^a, t_f^b\}$ shows that the minimal realization of the state equation is both controllable and observable on $[t_0, t_f]$. □

3.3.2 Time Invariant Case

We now restrict ourselves to the time invariant case and use a Laplace transform approach to establish our results. Instead of considering the time-domain description of the input-output behavior given by

$$y(t) = \int_0^t G(t, \sigma(s))u(s) \;\Delta s,$$

we examine the corresponding behavior in the z-domain. Laplace transforming the equation above and using the convolution theorem yields $Y(z) = G(z)U(z)$. The question is: Given a transfer function $G(z)$, when does there exist a time invariant form of the state equation such that

$$C(zI - A)^{-1}B = G(z),$$

and when is this realization minimal?

To answer this, we begin by characterizing time invariant realizations. In what follows, a *strictly-proper rational function* of z is a rational function of z such that the degree of the numerator is strictly less than the degree of the denominator.

Theorem 3.15. *The $p \times q$ transfer function $G(z)$ admits a time invariant realization of the regressive system*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned}$$

if and only if each entry of $G(z)$ is a strictly-proper rational function of z.

Proof. If $G(z)$ has a time invariant realization, then G has the form $G(z) = C(zI - A)^{-1}B$. We showed in Chapter 2 that for each Laplace transformable function $f(t)$, $F(z) \to 0$ as $z \to \infty$, which in turn implies that $F(z)$ is a strictly-proper rational function in z. Thus, the matrix $(zI - A)^{-1}$ is a matrix of strictly-proper rational functions, and $G(z)$ is a matrix of strictly-proper rational functions since linear combinations of such functions are themselves strictly-proper and rational.

Conversely, suppose that each entry $G_{ij}(z)$ in the matrix is strictly-proper and rational. Without loss of generality, we can assume that each polynomial in the denominator is monic (i.e. has leading coefficient of 1). Suppose

$$d(z) = z^r + d_{r-1}z^{r-1} + \ldots + d_0,$$

is the least common multiple of the polynomials in the denominators. Then $d(z)G(z)$ can be decomposed as a polynomial in z with $p \times q$ constant coefficient matrices, so that

$$d(z)G(z) = P_{r-1}z^{r-1} + \ldots + P_1 z + P_0.$$

We claim that the qr-dimensional matrices given by

$$A = \begin{bmatrix} 0_q & I_q & \ldots & 0_q \\ 0_q & 0_q & \ldots & 0_q \\ \vdots & \vdots & \vdots & \vdots \\ 0_q & 0_q & \ldots & 0_q \\ -d_0 I_q & -d_1 I_q & \ldots & -d_{r-1} I_q \end{bmatrix}, \qquad B = \begin{bmatrix} 0_q \\ 0_q \\ \vdots \\ 0_q \\ I_q \end{bmatrix}, \qquad C = \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_{r-1} \end{bmatrix},$$

form a realization of $G(z)$. To see this, let

$$R(z) = (zI - A)^{-1} B,$$

and partition the $qr \times q$ matrix $R(z)$ into r blocks $R_1(z), R_2(z), \ldots, R_r(z)$, each of size $q \times q$. Multiplying $R(z)$ by $(zI - A)$ and writing the result in terms of submatrices gives rise to the relations

$$R_{i+1}(z) = zR_i, \;\; i = 1, \ldots, r-1, \tag{3.15}$$

and

$$zR_r(z) + d_0 R_1(z) + d_1 R_2(z) + \ldots + d_{r-1} R_r(z) = I_q. \tag{3.16}$$

Using (3.15) to rewrite (3.16) in terms of $R_1(z)$ gives

$$R_1(z) = \frac{1}{d(z)} I_q,$$

and thus from (3.15) again, we have

$$R(z) = \frac{1}{d(z)} \begin{bmatrix} I_q \\ zI_q \\ \vdots \\ z^{r-1} I_q \end{bmatrix}.$$

Multiplying by C yields

$$C(zI-A)^{-1}B = \frac{1}{d(z)}\left(P_0 + zP_1 + \ldots + z^{r-1}P_{r-1}\right) \quad = \quad G(z),$$

which is a realization of $G(z)$. Thus, the claim follows. $\square$

The realizations that are minimal are characterized in the following theorem. The result is repeated here for completeness sake in this easier case of Theorem 3.14.

Theorem 3.16. *Suppose the time invariant regressive linear-state equation*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bx(t), \\ y(t) &= Cx(t), \end{aligned}$$

is a realization of the transfer function $G(z)$. Then this state equation is a minimal realization of $G(z)$ if and only if it is both controllable and observable.

Proof. Suppose the state equation is a realization of $G(z)$ that is not minimal. Then there is a realization of $G(z)$ given by

$$\begin{aligned} z^{\Delta}(t) &= Pz(t) + Qz(t), \\ y(t) &= Rz(t), \end{aligned}$$

with dimension $n_z < n$. Thus,

$$Ce_A(t,0)B = Re_P(t,0)Q, \quad t \geq 0.$$

Repeated differentiation with respect to t, followed by evaluation at $t = 0$ yields

$$CA^kB = RF^kQ, \;\; k = 0, 1, \ldots.$$

Rewriting this information in matrix form for $k = 0, 1, \ldots, 2n-2$, we see

$$\begin{bmatrix} CB & CAB & \ldots & CA^{n-1}B \\ \vdots & \vdots & \vdots & \vdots \\ CA^{n-1}B & CA^nB & \ldots & CA^{2n-2}B \end{bmatrix} = \begin{bmatrix} RQ & RPQ & \ldots & RP^{n-1}Q \\ \vdots & \vdots & \vdots & \vdots \\ RP^{n-1}Q & RP^nQ & \ldots & RP^{2n-2}Q \end{bmatrix},$$

and can be rewritten as

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} = \begin{bmatrix} R \\ RP \\ \vdots \\ RP^{n-1} \end{bmatrix} \begin{bmatrix} Q & PQ & \dots & P^{n-1}Q \end{bmatrix}.$$

However, since the right hand side of the equation is the product of an $n_z p \times n_z$ and an $n_z \times n_z m$ matrix, the rank of the product can be no greater than n_z. Thus, $n_z < n$, which implies that that the realization given in the statement of the theorem cannot be both controllable and observable. Therefore, by contraposition a controllable and observable realization must be minimal.

Conversely, suppose the state equation given in the statement of the theorem is a minimal realization that is not controllable. Then there exists an $n \times 1$ vector $y \neq 0$ such that

$$y^T \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} = 0,$$

which implies $y^T A^k B = 0$ for all $k \geq 0$ by the Cayley-Hamilton theorem. For P^{-1} an invertible $n \times n$ matrix with bottom row y^T, then a variable change of $z(t) = P^{-1}x(t)$ produces the state equations

$$\begin{aligned} z^\Delta(t) &= \hat{A}z(t) + \hat{B}u(t), \\ y(t) &= \hat{C}z(t), \end{aligned}$$

which is also an n-dimensional minimal realization of $G(z)$. Partition the coefficient matrices of the state equation above as

$$\hat{A} = P^{-1}AP = \begin{bmatrix} \hat{A_{11}} & \hat{A_{12}} \\ \hat{A_{21}} & \hat{A_{22}} \end{bmatrix}, \ \hat{B} = P^{-1}B = \begin{bmatrix} \hat{B}_1 \\ 0 \end{bmatrix}, \ \hat{C} = CP = \begin{bmatrix} \hat{C}_1 & CA \end{bmatrix},$$

where $\hat{A_{11}}$ is $(n-1) \times (n-1)$, $\hat{B}_1$ is $(n-1) \times 1$, and $\hat{C}_1$ is $1 \times (n-1)$. From these partitions, it follows from the construction of P that $\hat{A}\hat{B} = P^{-1}AB$ has the form

$$\hat{A}\hat{B} = \begin{bmatrix} \hat{A_{11}}\hat{B_1} \\ \hat{A_{21}}\hat{B_1} \end{bmatrix} = \begin{bmatrix} \hat{A_{11}}\hat{B_1} \\ 0 \end{bmatrix}.$$

Since the bottom row of $P^{-1}A^kB$ is zero for all $k \geq 0$,

$$\hat{A}^k\hat{B} = \begin{bmatrix} \hat{A_{11}}^k\hat{B_1} \\ 0 \end{bmatrix}, \quad k \geq 0.$$

But, $\hat{A_{11}}, \hat{B_1}, \hat{C_1}$ give an $(n-1)$-dimensional realization of $G(z)$ since

$$\begin{aligned} \hat{C}e_{\hat{A}}(t,0)\hat{B} &= \begin{bmatrix} \hat{C_1} & \hat{C_2} \end{bmatrix} \sum_{k=0}^{\infty} \hat{A}^k\hat{B}h_k(t,0) \\ &= \begin{bmatrix} \hat{C_1} & \hat{C_2} \end{bmatrix} \sum_{k=0}^{\infty} \begin{bmatrix} \hat{A_{11}}^k\hat{B_1} \\ 0 \end{bmatrix} h_k(t,0) \\ &= \hat{C_1}e_{\hat{A_{11}}}(t,0)\hat{B_1}, \end{aligned}$$

so that the state equation in the statement of the theorem is in fact not minimal, a contradiction. A similar argument holds if the system is assumed not to be observable. $\square$

We now illustrate Theorem 3.15 and Theorem 3.16 with an example.

Example 3.4. Consider the transfer function $G(z) = \frac{9(37+300z)}{5+75z+270z^2}$. $G(z)$ admits a time invariant realization by Theorem 3.15 since $G(z)$ is a strictly-proper rational function of z. The form of $G(z)$ indicates that we should look for a 2-dimensional realization with a single input and single output. We can write

$$G(z) = \begin{bmatrix} 3 & 4 \end{bmatrix} \left(zI - \begin{bmatrix} -\frac{8}{45} & \frac{1}{30} \\ -\frac{1}{45} & -\frac{1}{10} \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

so that a time invariant realization of $G(z)$ is given by

$$\begin{aligned} x^{\Delta}(t) &= \begin{bmatrix} -\frac{8}{45} & \frac{1}{30} \\ -\frac{1}{45} & -\frac{1}{10} \end{bmatrix} x(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u(t), \qquad x(0) = x_0, \\ y(t) &= \begin{bmatrix} 3 & 4 \end{bmatrix} x(t). \end{aligned}$$

We showed in Example 3.1 that this realization is in fact controllable, and we showed in Example 3.2 that it is also observable. Thus, Theorem 3.16 guarantees that this realization of $G(z)$ is minimal.

3.4 Stability

We complete our foray into linear systems theory by considering stability. Pötzsche, Siegmund, and Wirth deal with exponential stability in [38]. DaCunha also deals with this concept under a different definition in [15, 16] and emphasizes the time varying case. We begin by revisiting exponential stability in the time invariant case and then proceed to another notion of stability commonly used in linear systems theory.

3.4.1 Exponential Stability in the Time Invariant Case

We start this section by revisiting the notion of exponential stability. We are interested in both the time invariant and time varying cases separately since it is often possible to obtain stronger results in the time invariant case.

We have already noted that if A is constant, then $\Phi_A(t, t_0) = e_A(t, t_0)$. In what follows, we will consider autonomous systems with $t_0 = 0$ in order to talk about the Laplace transform.

Recall from Chapter 1 that DaCunha in [16] defines *uniform exponential stability* as follows.

Definition 3.6 (DaCunha, [16]). The regressive time varying system

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

is called *uniformly exponentially stable* if there exist constants $\gamma, \lambda > 0$ with $-\lambda \in \mathcal{R}^+$ such that for any t_0 and $x(t_0)$, the corresponding solution satisfies

$$||x(t)|| \leq ||x(t_0)||\gamma e_{-\lambda}(t, t_0), \quad t \geq t_0.$$

With this definition of exponential stability, we can prove the next theorem.

Theorem 3.17. *The autonomous equation*

$$x^{\Delta}(t) = Ax(t), \quad x(0) = x_0,$$

is uniformly exponentially stable if and only if

$$\int_0^{\infty} ||e_A(t, 0)|| \; \Delta t \leq \beta,$$

for some $\beta > 0$.

Proof. For necessity, note that if the system is uniformly exponentially stable, then by Theorem 1.24, we have that

$$\begin{aligned} \int_0^{\infty} ||e_A(t, 0)|| \; \Delta t &\leq \int_0^{\infty} \gamma e_{-\lambda}(t, 0) \; \Delta t \\ &= \frac{\gamma}{\lambda}, \end{aligned}$$

so that the claim follows by choosing $\beta = \frac{\gamma}{\lambda}$.

For sufficiency, assume the integral condition holds but for the sake of contradiction that the system is not exponentially stable. Then, again by Theorem 1.24, we know that for all $\lambda, \gamma > 0$ with $-\lambda \in \mathcal{R}^+$, we have that

$$||e_A(t, 0)|| > \gamma e_{-\lambda}(t, 0).$$

Computing the integral gives

$$\begin{aligned} \int_0^{\infty} ||e_A(t, 0)|| \; \Delta t &> \int_0^{\infty} \gamma e_{-\lambda}(t, 0) \; \Delta t \\ &= \frac{\gamma}{-\lambda} e_{-\lambda}(t, 0)\Big|_0^{\infty} \\ &= \frac{\gamma}{\lambda}. \end{aligned}$$

In particular, if we choose $\gamma > \beta\lambda$, then

$$\int_0^\infty ||e_A(t,0)||\ \Delta t > \frac{\beta\lambda}{\lambda} = \beta,$$

a contradiction. $\square$

Now consider the system

$$x^\Delta(t) = Ax(t),\ \ x(0) = I.$$

Transforming this system yields

$$X(z) = (zI - A)^{-1},$$

which is the transform of $e_A(t,0)$. We know that this result is unique as we argued in the previous chapter. Note that this matrix contains only strictly-proper rational functions of z since we have the formula

$$(zI - A)^{-1} = \frac{\operatorname{adj}(zI - A)}{\det(zI - A)}.$$

Specifically, $\det(zI-A)$ is a degree-n polynomial in z, while each entry of $\operatorname{adj}(zI-A)$ is a polynomial of degree at most $n-1$. Suppose

$$\det(zI - A) = (z-\lambda_1)^{\psi_1}(z-\lambda_2)^{\psi_2}\dots(z-\lambda_m)^{\psi_m},$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the distinct eigenvalues of the $n \times n$ matrix A, with corresponding multiplicities $\psi_1, \psi_2, \dots, \psi_m$. Decomposing $(zI-A)^{-1}$ in terms of partial fractions gives

$$(zI - A)^{-1} = \sum_{k=1}^{m}\sum_{j=1}^{\psi_k} W_{kj}\frac{1}{(z-\lambda_k)^j},$$

where each W_{kj} is an $n \times n$ matrix of partial fraction expansion coefficients given by

$$W_{kj} = \frac{1}{(\psi_k - j)!}\frac{d^{\psi_k - j}}{dz^{\psi_k - j}}\left[(z-\lambda_k)^{\psi_k}(zI-A)^{-1}\right]\Bigg|_{z=\lambda_k}.$$

If we now take the inverse Laplace transform of $(zI - A)^{-1}$ in the form given above, we obtain the representation

$$e_A(t,0) = \sum_{k=1}^{m} \sum_{j=1}^{\psi_k} W_{kj} \frac{f_{j-1}(\mu, \lambda_k)}{(j-1)!} e_{\lambda_k}(t,0),$$

where $f_j(\mu, \lambda_k)$ is the sequence of functions obtained from the residue calculations of the jth derivative in the inversion formula. For example, the first few terms in the sequence are

$$\begin{aligned}
f_0(\mu, \lambda_k) &= 1, \\
f_1(\mu, \lambda_k) &= \int_0^t \frac{1}{1+\mu\lambda_k}\, \Delta\tau, \\
f_2(\mu, \lambda_k) &= \left(\int_0^t \frac{1}{1+\mu\lambda_k}\, \Delta\tau \right)^2 - \int_0^t \frac{\mu}{(1+\mu\lambda_k)^2}\, \Delta\tau, \\
f_3(\mu, \lambda_k) &= \left(\int_0^t \frac{1}{1+\mu\lambda_k}\, \Delta\tau \right)^3 - 3 \int_0^t \frac{\mu}{(1+\mu\lambda_k)^2}\, \Delta\tau \int_0^t \frac{1}{1+\mu\lambda_k}\, \Delta\tau \\
&+ \int_0^t \frac{2\mu^2}{(1+\mu\lambda_k)^3}\, \Delta\tau, \\
&\vdots
\end{aligned}$$

Notice that if μ is bounded, then each $f_j(\mu, \lambda_k)$ can be bounded by a "regular" polynomial of degree j in t, call it $a_j(t)$. That is, f_j can be bounded by functions of the form $a_j(t) = a_j x^j + a_{j-1} x^{j-1} \cdots + a_0$. This observation will play a key role in the next theorem. Pötzsche, Siegmund, and Wirth do prove this result in [38], but our proof differs from theirs somewhat in that we use the transform to obtain it, while they use other techniques. Note, however, that in the theorem we do use their definition of exponential stability rather than the one given by DaCunha. For completeness, we remind the reader by restating their definition here.

Definition 3.7 (Pötzsche, Siegmund, Wirth, [38]). For $t, t_0 \in \mathbb{T}$ and $x_0 \in \mathbb{R}^n$, the system

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

is said to be *uniformly exponentially stable* if there exists a constant $\alpha > 0$ such that for every $t_0 \in \mathbb{T}$ there exists a $K \geq 1$ with

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

with K being chosen independently of t_0.

Recall that DaCunha's definition of uniform exponential stability of a system will imply that the system is uniformly exponential stable if we use Pötzsche, Siegmund, and Wirth's definition of the concept, but the converse need not be true in general. Thus, DaCunha's definition is weaker in this sense.

Theorem 3.18. *The autonomous system*

$$x^{\Delta}(t) = Ax(t), \quad x(0) = x_0,$$

is exponentially stable if and only if all eigenvalues of A live in $\mathcal{S}(\mathbb{C})$, the stability region of $\mathbb{T}$, which has bounded graininess.

Proof. Suppose the eigenvalue condition holds. Then, appealing to Theorem 3.17 and writing the exponential in the explicit form given above in terms of the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, we obtain

$$\begin{aligned}
\int_0^{\infty} ||e_A(t,0)|| \; \Delta t &= \int_0^{\infty} ||\sum_{k=1}^{m} \sum_{j=1}^{\psi_k} W_{kj} \frac{f_{j-1}(\mu, \lambda_k)}{(j-1)!} e_{\lambda_k}(t,0)|| \; \Delta t \\
&\leq \sum_{k=1}^{m} \sum_{j=1}^{\psi_k} ||W_{kj}|| \int_0^{\infty} \left| \frac{f_{j-1}(\mu, \lambda_k)}{(j-1)!} e_{\lambda_k}(t,0) \right| \; \Delta t \\
&\leq \sum_{k=1}^{m} \sum_{j=1}^{\psi_k} ||W_{kj}|| \int_0^{\infty} |a_{j-1}(t) e_{\lambda_k}(t,0)| \; \Delta t \\
&\leq \sum_{k=1}^{m} \sum_{j=1}^{\psi_k} ||W_{kj}|| \int_0^{\infty} a_{j-1}(t) e^{-\alpha t} \; \Delta t \\
&\leq \sum_{k=1}^{m} \sum_{j=1}^{\psi_k} ||W_{kj}|| \int_0^{\infty} a_{j-1}(t) e^{-\alpha t} \; dt \\
&< \infty.
\end{aligned}$$

Notice that the last three lines hold by appealing to Definition 3.7. Thus, by Theorem 3.17 the system is exponentially stable.

Now, for the sake of a contradiction, assume that the eigenvalue condition fails. Let λ be an eigenvalue of A with associated eigenvector v, with $\lambda \notin \mathcal{S}(\mathbb{C})$. Direct calculation shows that the solution of the system

$$x^{\Delta} = Ax, \ \ x(0) = v,$$

is given by $x(t) = e_{\lambda}(t, 0)v$. According to Pötzsche, Siegmund, and Wirth, if $\lambda \notin \mathcal{S}(\mathbb{C})$, then

$$\lim_{t \to \infty} e_{\lambda}(t, 0) \neq 0,$$

so that we arrive at a contradiction. □

3.4.2 BIBO Stability in the Time Varying Case

Besides exponential stability, the concept of bounded-input, bounded-output stability is also a useful property for a system to have. As its name suggests, the notion is one that compares the supremum of the output signal with the supremum of the input signal. Thus, we can define the term as follows.

Definition 3.8. The regressive time varying linear state equation

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \ \ x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is said to be *uniformly bounded-input, bounded-output stable* if there exists a finite constant η such that for any t_0 and any input signal $u(t)$ the corresponding zero-state response satisfies

$$\sup_{t \geq t_0} ||y(t)|| \leq \eta \sup_{t \geq t_0} ||u(t)||.$$

Note that we use the word 'uniform' to stress that the same η works for all t_0 and all input signals. We wish to know when a system is BIBO stable, and so a characterization of BIBO stability follows.

Theorem 3.19. *The regressive time varying linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \ \ x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is uniformly bounded-input, bounded-output stable if and only if there exists a finite constant ρ such that for all t, τ with $t \geq \tau$,

$$\int_{\tau}^{t} ||G(t, \sigma(s))|| \ \Delta s \leq \rho.$$

Proof. Assume such a ρ exists. Then for any t_0 and any input signal, the corresponding zero-state response of the state equation satisfies

$$\begin{aligned} ||y(t)|| &= \left|\left| \int_{t_0}^{t} C(t)\Phi_A(t, \sigma(s))B(s)u(s) \, \Delta s \right|\right| \\ &\leq \int_{t_0}^{t} ||G(t, \sigma(s))|| \ ||u(s)|| \, \Delta s, \quad t \geq t_0. \end{aligned}$$

Replacing $||u(s)||$ by its supremum over $s \geq t_0$, and using the integral condition, we obtain

$$\begin{aligned} ||y(t)|| &\leq \int_{t_0}^{t} ||G(t, \sigma(s))|| \Delta s \sup_{t \geq t_0} ||u(t)|| \\ &\leq \rho \sup_{t \geq t_0} ||u(t)||, \quad t \geq t_0. \end{aligned}$$

Thus, taking the supremum of the left hand side of the inequality over $t \geq t_0$, the system is BIBO stable if we choose $\eta = \rho$.

Conversely, suppose the state equation is uniformly BIBO stable. Then there exists a constant η so that, in particular, the zero-state response for any t_0 and any input signal such that $\sup_{t \geq t_0} ||u(t)|| \leq 1$ satisfies $\sup_{t \geq t_0} ||y(t)|| \leq \eta$. For the sake of a contradiction, suppose no finite ρ exists that satisfies the integral condition. Then for any given $\rho > 0$, there exist τ_ρ and $t_\rho > \tau_\rho$ such that

$$\int_{\tau_\rho}^{t_\rho} ||G(t_\rho, \sigma(s))|| \Delta s > \rho.$$

In particular, if $\rho = \eta$, this implies that there exist τ_η, with $t_\eta > \tau_\eta$, and indices i, j such that the i, j-entry of the impulse response satisfies

$$\int_{\tau_\eta}^{t_\eta} |G_{ij}(t_\eta, \sigma(s))| \Delta s > \eta.$$

With $t_0 = \tau_\eta$ consider the $m \times 1$ input signal $u(t)$ defined for $t \geq t_0$ as follows: set $u(t) = 0$ for $t > t_\eta$, and for $t \in [t_0, t_\eta]$ set every component of $u(t)$ to zero except for the j-th component given by the piecewise continuous signal

$$u_j(t) = \begin{cases} 1, & G_{ij}(t_\eta, \sigma(t)) > 0, \\ 0, & G_{ij}(t_\eta, \sigma(t)) = 0 \ , \quad t \in [t_0, t_\eta], \\ -1, & G_{ij}(t_\eta, \sigma(t)) < 0. \end{cases}$$

This input signal satisfies $||u(t)|| \leq 1$ for all $t \geq t_0$, but because of the integral condition above, the i-th component of the corresponding zero-state response satisfies

$$\begin{aligned} y_i(t_\eta) &= \int_{t_0}^{t_\eta} G_{ij}(t_\eta, \sigma(s)) u_j(s) \Delta s \\ &= \int_{t_0}^{t_\eta} |G_{ij}(t_\eta, \sigma(s))| \Delta s \\ &> \eta. \end{aligned}$$

Since $||y(t_\eta)|| \geq |y_i(t_\eta)|$, we arrive at a contradiction that completes the proof. $\square$

We now wish to give conditions under which the notions of exponential stability and BIBO stability are equivalent. To this end, we begin with the following.

Theorem 3.20. *Suppose the regressive time varying linear state equation*

$$\begin{aligned} x^\Delta(t) &= A(t)x(t) + B(t)u(t), \ \ x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is uniformly exponentially stable, and there exist constants β and γ such that

$$||B(t)|| \leq \beta \ \text{ and } \ ||C(t)|| \leq \alpha$$

for all t. Then the state equation is also uniformly bounded-input, bounded-output stable.

Proof. Using the bound implied by uniform exponential stability, we have

$$\begin{aligned}
\int_{\tau}^{t} ||G(t,\sigma(s))||\Delta s &\leq \int_{\tau}^{t} ||C(t)||\ ||\Phi_A(t,\sigma(s))||\ ||B(s)||\Delta s \\
&\leq \alpha\beta \int_{\tau}^{t} ||\Phi_A(t,\sigma(s))||\Delta s \\
&\leq \alpha\beta \int_{\tau}^{t} \gamma e_{-\lambda}(t,\sigma(s))\Delta s \\
&\leq \frac{\alpha\beta\gamma}{\lambda} \int_{\tau}^{t} \frac{\lambda}{1-\mu(s)\lambda} e_{\lambda/(1-\mu\lambda)}(s,t)\Delta s \\
&= \frac{\alpha\beta\gamma}{\lambda}\left(1 - e_{-\lambda}(t,\tau)\right) \\
&\leq \frac{\alpha\beta\gamma}{\lambda}.
\end{aligned}$$

By Theorem 3.19, the state equation is also bounded-input, bounded-output stable.

□

The following example illustrates the use of Theorem 3.20.

Example 3.5. Let $\mathbb{T}$ be a time scale with $0 \leq \mu < \frac{1}{2}$. Consider the system

$$\begin{aligned}
x^{\Delta}(t) &= \begin{bmatrix} -2 & 1 \\ -1 & -\sin(t) - 2 \end{bmatrix} x(t) + \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} u(t), \quad x(t_0) = x_0, \\
y(t) &= \begin{bmatrix} 1 & e_{-1}(t,0) \end{bmatrix} x(t),
\end{aligned}$$

where here, $\sin(t)$ and $\cos(t)$ are the *usual* trigonometric functions and not their time scale counterparts. DaCunha, in [15], shows that the system is uniformly exponentially stable by applying Theorem 1.21 with the choice $Q(t) = I$. For $t \geq 0$, we have $||B(t)|| = \sqrt{\cos^2(t) + \sin^2(t)} = 1$ and $||C(t)|| = \sqrt{1 + (e_{-1}(t,0))^2} \leq \sqrt{2}$ since $p = -1 \in \mathcal{R}^+$ from our assumption on $\mathbb{T}$, and hence, by Theorem 3.20, the state equation is also uniformly bounded-input, bounded-output stable.

For the converse of the previous theorem, it is known on $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$ that stronger hypotheses than simply having the system be BIBO stable are necessary to establish exponential stability (see [1], [2], and [37]). At present, we lack an analogue of this result for an arbitrary time scale in the time varying system case. We will see that the time invariant case does allow for the equivalence of the two notions in the general time scale case under certain conditions.

3.4.3 BIBO Stability in the Time Invariant Case

We need to extend the definition of BIBO stability to the time invariant case, but for reasons that will soon become apparent, we will need to modify the definition slightly.

Definition 3.9. For any shift $u(t, \sigma(s))$ of the transformable function $u(t)$, the time invariant system

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \;\; x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is said to be *uniformly bounded-input, bounded-output stable* if there exists a finite constant η such that the corresponding zero-state response satisfies

$$\sup_{t \geq 0} ||y(t)|| \leq \eta \sup_{t \geq 0} \sup_{s \geq 0} ||u(t, \sigma(s))||.$$

Note that Definitions 3.8 and 3.9 are different: one deals with the time varying case and the other with the time invariant case. The modified definition in the time invariant case says that the output stays bounded over all shifts of the input.

Theorem 3.21. *The regressive linear time invariant system*

$$\begin{aligned} x^{\Delta}(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is bounded-input, bounded-output stable if and only if there exists a finite $\beta > 0$ *such that*

$$\int_0^\infty ||G(t)||\Delta t \leq \beta.$$

Proof. Suppose we have the existence of the claimed $\beta > 0$. For any time t, we have

$$\begin{aligned} y(t) &= \int_0^t Ce_A(t, \sigma(s))Bu(s)\Delta s \\ &= \int_0^t Ce_A(s, 0)Bu(t, \sigma(s))\Delta s, \end{aligned}$$

since $y(t)$ is a convolution, so that

$$\begin{aligned} ||y(t)|| &\leq ||C|| \int_0^t ||e_A(s, 0)||\ ||B|| \sup_{0\leq s\leq t} ||u(t, \sigma(s))||\Delta s \\ &\leq ||C|| \int_0^\infty ||e_A(s, 0)||\Delta s\ ||B|| \sup_{s\geq 0} ||u(t, \sigma(s))||. \end{aligned}$$

Therefore,

$$\sup_{t\geq 0} ||y(t)|| \leq ||C|| \int_0^\infty ||e_A(s, 0)||\Delta s\ ||B|| \sup_{t\geq 0} \sup_{s\geq 0} ||u(t, \sigma(s))||.$$

If we choose $\eta = ||C||\, \beta\, ||B||$, the claim follows.

Conversely, suppose that the system is bounded-input bounded-output stable, but for the sake of a contradiction that the integral is unbounded. Then,

$$\sup_{t\geq 0} ||y(t)|| \leq \eta \sup_{t\geq 0} \sup_{s\geq 0} ||u(t, \sigma(s))||,$$

and

$$\int_0^\infty ||G(t)||\Delta t > \beta, \text{ for all } \beta > 0.$$

In particular, there exist indices i, j such that

$$\int_0^\infty |G_{ij}(t)|\Delta t > \beta.$$

Choose $u(t, \sigma(s))$ in the following manner: set $u_k(t, \sigma(s)) = 0$ for all $k \neq j$, and define $u_j(t,\sigma(s))$ by

$$u_j(t,\sigma(s)) = \begin{cases} 1, & \text{if } G_{ij}(s) > 0, \\ 0, & \text{if } G_{ij}(s) = 0, \\ -1, & \text{if } G_{ij}(s) < 0, \end{cases}$$

and choose $\beta > \eta > 0$. Notice that $\sup\limits_{t\geq 0}\sup\limits_{s\geq 0} ||u(t,\sigma(s)|| \leq 1$, so that $\sup\limits_{t\geq 0} ||y(t)|| \leq \eta$. However,

$$\begin{aligned} \sup_{t\geq 0} ||y(t)|| &= \sup_{t\geq 0} || \int_0^t G(s)u(t,\sigma(s))\Delta s|| \\ &= \sup_{t\geq 0} || \int_0^t G_j(s)\cdot u_j(s)\Delta s|| \\ &\geq \sup_{t\geq 0} \int_0^t |G_{ij}(s)|\Delta s \\ &= \int_0^\infty |G_{ij}(s)|\Delta s > \beta > \eta, \end{aligned}$$

which is a contradiction. Thus, the claim follows. $\square$

The next theorem demonstrates the equivalence of exponential and BIBO stability in the time invariant case. Recall that this is a notion we currently lack in the time varying case.

Theorem 3.22 (Equivalence of BIBO and Exponential Stability). *Suppose the linear regressive time invariant state equation*

$$\begin{aligned} x^\Delta(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is controllable and observable. Then the system is uniformly bounded-input, bounded output stable if and only if it is exponentially stable.

Proof. If the system is exponentially stable, then by Theorem 3.17,

$$\int_0^\infty ||Ce_A(t,0)B||\ \Delta t \leq ||C||\,||B|| \int_0^\infty ||e_A(t,0)||\ \Delta t \leq \eta.$$

Conversely, suppose the system is uniformly bounded-input, bounded output stable. Then

$$\int_0^\infty ||Ce_A(t,0)B||\ \Delta t < \infty,$$

which implies

$$\lim_{t\to\infty} Ce_A(t,0)B = 0. \tag{3.17}$$

Using the representation of the matrix exponential given earlier in terms of the transform, we may write

$$Ce_A(t,0)B = \sum_{k=1}^{m}\sum_{j=1}^{\psi_k} N_{kj}\frac{f_{j-1}(\mu,\lambda_k)}{(j-1)!}e_{\lambda_k}(t,0), \tag{3.18}$$

where the λ_k are the distinct eigenvalues of A, the N_{kj} are constant matrices, and the $f_j(\mu,\lambda_k)$ are the terms from the residue calculations. In this form,

$$\frac{d}{\Delta t}Ce_A(t,0)B = \sum_{k=1}^{m}\left(N_{k1}\lambda_k + \sum_{j=2}^{\psi_k}\left(\frac{f_{j-1}^{\Delta}(\mu,\lambda_k)(1+\mu(t)\lambda_k)}{(j-2)!} + \frac{\lambda_k f_{j-1}(\mu,\lambda_k)}{(j-1)!}\right)\right)e_{\lambda_k}(t,0).$$

If this function does not to go to zero as $t\to\infty$, then using (3.18), we could compare this result with (3.17) to obtain a contradiction. Thus,

$$\lim_{t\to\infty}\left(\frac{d}{\Delta t}Ce_A(t,0)B\right) = \lim_{t\to\infty} CAe_A(t,0)B = \lim_{t\to\infty} Ce_A(t,0)AB = 0,$$

where the last equation holds by noting that if A is constant, then A and $e_A(t,0)$ commute. Similarly, it can easily be shown that any order time derivative of the exponential goes to zero as $t\to\infty$. Thus,

$$\lim_{t\to\infty} CA^ie_A(t,0)A^jB = 0, \quad i,j = 0,1,\ldots$$

It then follows that

$$\lim_{t\to\infty}\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1}\end{bmatrix} e_A(t,0)\begin{bmatrix} B & AB & \ldots & A^{n-1}B\end{bmatrix} = 0. \tag{3.19}$$

But, the system is controllable and observable, and so we can form invertible matrices $\mathcal{G}_C^a$ and $\mathcal{G}_O^a$ by choosing n independent columns of the controllability matrix and n independent rows of the observability matrix, respectively. Then, by (3.19), $\lim_{t\to\infty} \mathcal{G}_O^a e_A(t,0)\mathcal{G}_C^a = 0$. Hence, $\lim_{t\to\infty} e_A(t,0) = 0$ and so the exponential stability follows from the arguments given in Theorem 3.18. $\square$

We make use of the preceding theorem in the following example.

Example 3.6. Suppose $\mathbb{T}$ is a time scale with $0 \le \mu \le 4$. The system

$$\begin{aligned} x^\Delta(t) &= \begin{bmatrix} -\frac{8}{45} & \frac{1}{30} \\ -\frac{1}{45} & -\frac{1}{10} \end{bmatrix} x(t) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u(t), \quad x(0) = x_0, \\ y(t) &= \begin{bmatrix} 3 & 4 \end{bmatrix} x(t). \end{aligned}$$

is controllable by Example 3.1 and observable by Example 3.2. The eigenvalues of A are $\lambda_1 = -\frac{1}{9}$ and $\lambda_2 = -\frac{1}{6}$. Note that the assumption on $\mathbb{T}$ implies $\lambda_1, \lambda_2 \in \mathcal{S}(\mathbb{C})$, the stability region of $\mathbb{T}$. Thus, by Theorem 3.18, the system is exponentially stable. Theorem 3.22 then says that the system is also BIBO stable.

As we have seen, the Laplace transform can be a useful tool for analyzing stability in the time invariant case. With this in mind, we desire a theorem that determines if a system is BIBO stable by examining its transfer function. The following theorem does this.

Theorem 3.23. *The regressive linear time invariant system*

$$\begin{aligned} x^\Delta(t) &= Ax(t) + Bu(t), \;\; x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

is bounded-input, bounded-output stable if and only if all poles of the transfer function $G(z) = C(zI - A)^{-1}B$ *are contained in* $S(\mathbb{C})$.

Proof. If each entry of $G(z)$ has poles that lie in $S(\mathbb{C})$, then the partial fraction decomposition of $G(z)$ discussed earlier shows that each entry of $G(t)$ has a sum of "polynomial-multiplied exponential terms" form. Since the exponentials will all have subscripts living in the stability region,

$$\int_0^\infty ||G(t)||\Delta t < \infty,$$

and so the system is bounded-input, bounded-output stable.

Conversely, if

$$\int_0^\infty ||G(t)||\Delta t < \infty,$$

then the exponential terms in any entry of $G(t)$ must have subscripts that lie in the stability region by using a standard contradiction argument. Thus, every entry of $G(z)$ must have poles that lie in the stability region. $\square$

3.5 Linear Feedback

In this section, we examine linear feedback in systems. In particular, we focus on *state feedback* and leave *output feedback* as an area of future research outside the scope of this dissertation.

We begin by defining several ubiquitous terms in standard linear systems theory. The *open loop state equation* is given by

$$\begin{aligned} x^\Delta(t) &= A(t)x(t) + B(t)u(t), \ \ x(t_0) = x_0, \\ y(t) &= C(t)x(t). \end{aligned}$$

This equation is called *open loop* because the controller computes its input into the system using only the current state and its model of the system.

In linear control, *linear state feedback* replaces the input $u(t)$ by an expression of the form $u(t) = K(t)x(t) + N(t)r(t)$, where $r(t)$ represents a new input signal, and $K(t) \in \mathbb{R}^{m\times n}$, $N(t) \in R^{m\times m}$ are rd-continuous. Thus, substituting the linear feedback into the original equation yields the *closed loop state equation* given by

$$\begin{aligned} x^{\Delta}(t) &= [A(t) + B(t)K(t)]\, x(t) + B(t)N(t)r(t), \;\; x(t_0) = x_0, \\ y(t) &= C(t)x(t). \end{aligned}$$

This equation is termed *closed loop* because the outputs of the system are fed back to the inputs of the controller. That is, process inputs have have an effect on process outputs which is measured in some way and then processed by the controller; the resulting control signal is used as an input to the process, closing the loop (see [21]).

For *linear output feedback*, we choose $u(t)$ as

$$u(t) = L(t)y(t) + N(t)r(t).$$

In this case, the new resulting state equation is

$$\begin{aligned} x^{\Delta}(t) &= [A(t) + B(t)L(t)C(t)]\, x(t) + B(t)N(t)r(t), \;\; x(t_0) = x_0, \\ y(t) &= C(t)x(t). \end{aligned}$$

In what follows, we will need to use the inverse of the matrix $I - F(z)$, where $F(z)$ is a matrix of strictly-proper rational functions of z. Invertibility follows from noting that the function $\det [I - F(s)]$ is a rational function of z, and it must be a nonzero rational function since $||F(z)|| \to 0$ as $|z| \to \infty$. Thus, $[I - F(z)]^{-1}$ exists for all but a finite number of values for z, and it is a matrix of rational functions.

We begin by comparing the transition matrices of the open-loop and closed-loop equations.

Theorem 3.24 (Equivalence of Transition Matrices). *If $\Phi_A(t, \tau)$ is the transition matrix for the corresponding open-loop state equation*

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \;\; x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

and $\Phi_{A+BK}(t,\tau)$ *is the transition matrix for the corresponding closed-loop equation*

$$\begin{aligned} x^{\Delta}(t) &= [A(t)+B(t)K(t)]\,x(t)+B(t)N(t)r(t), \;\; x(t_0)=x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

resulting from linear state feedback, then

$$\Phi_{A+BK}(t,\tau)=\Phi_A(t,\tau)+\int_{\tau}^{t}\Phi_A(t,\sigma(s))B(s)K(s)\Phi_{A+BK}(s,\tau)\Delta s.$$

If the open-loop equation and state feedback are both time-invariant, then the Laplace transform of the closed-loop matrix exponential can be expressed in terms of the Laplace transform of the open-loop matrix exponential as

$$(zI-A-BK)^{-1}=\left[I-(zI-A)^{-1}BK\right]^{-1}(zI-A)^{-1}.$$

Proof. For the first claim, let τ be arbitrary but fixed. Evaluation of the right hand side of the equation at $t=\tau$ yields the identity matrix. Differentiating the right side of the equation with respect to t yields

$$\begin{aligned} &\frac{d}{\Delta t}\left[\Phi_A(t,\tau)+\int_{\tau}^{t}\Phi_A(t,\sigma(s))B(s)K(s)\Phi_{A+BK}(s,\tau)\Delta s\right] \\ = \;& A(t)\Phi_A(t,\tau) \\ + \;& A(t)\int_{\tau}^{t}\Phi_A(t,\sigma(s))B(s)K(s)\Phi_{A+BK}(s,\tau)\Delta s+\Phi_A(\sigma(t),\sigma(t))B(t)K(t)\Phi_{A+BK}(t,\tau) \\ = \;& A(t)\left[\Phi_A(t,\tau)+\int_{\tau}^{t}\Phi_A(t,\sigma(s))B(s)K(s)\Phi_{A+BK}(s,\tau)\Delta s\right]+B(t)K(t)\Phi_{A+BK}(t,\tau). \end{aligned}$$

That is, the right hand side of the last equation satisfies the matrix differential equation that uniquely determines $\Phi_{A+BK}(t,\tau)$ for any τ.

For the time invariant case, with $\tau=0$, the equation becomes

$$e_{A+BK}(t,0)=e_A(t,0)+\int_{0}^{t}e_A(t,\sigma(s))BKe_{A+BK}(s,0)\Delta s.$$

Transforming each side and recognizing the right hand side as a convolution produces

$$(zI-A-BK)^{-1}=(zI-A)^{-1}+(zI-A)^{-1}BK(zI-A-BK)^{-1},$$

an expression that is easily rewritten to give the claimed form. $\square$

Theorem 3.25. *If $G(t, \sigma(s))$ is the weighting pattern of the regressive open-loop time invariant state equation, and $\hat{G}(t, \sigma(s))$ is the weighting pattern of the regressive closed-loop time invariant state equation resulting from static output feedback, then the transfer functions of the two state equations are related by*

$$\hat{G}(z) = [I - G(z)L]^{-1} G(z)N.$$

Proof. The preceding theorem with $\tau = 0$ yields

$$e_{A+BK}(t, 0) = e_A(t, 0) + \int_0^t e_A(t, \sigma(s))BKe_{A+BK}(s, 0)\Delta s.$$

Replace K with LC to reflect output feedback, and then premultiply by C and postmultiply by BN to obtain

$$Ce_{A+BK}(t, 0)BN = Ce_A(t, 0)BN + \int_0^t Ce_A(t, \sigma(s))BLCe_{A+BK}(s, 0)BN\Delta s,$$

or equivalently,

$$\hat{G}(t, 0) = G(t, 0)N + \int_0^t G(t, \sigma(s))L\hat{G}(s, 0)\Delta s.$$

Again, recognizing the right hand side as a convolution and transforming yields

$$\hat{G}(z) = G(z)N + G(z)L\hat{G}(z),$$

from which the claim follows immediately. □

We would now like to consider when it is in fact possible to stabilize the system in question, and how to do so. To answer this, we first need a couple of lemmas.

Definition 3.10. The regressive linear state equation

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \qquad x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is called *uniformly exponentially stable with rate* $\lambda > 0$, where $-\lambda \in \mathcal{R}^+$, if there exists a constant $\gamma > 0$ such that for any t_0 and x_0 the corresponding solution satisfies

$$||x(t)|| \leq \gamma e_{-\lambda}(t, t_0)||x_0||, \ \ t \geq t_0.$$

Lemma 3.2. *The Hilger circle* $\mathbb{H}$ *is closed under the operation* $\oplus$ *for all* $t \in \mathbb{T}$.

Proof. Let $\alpha \in \mathbb{C}$ be such that $|\alpha| < 1$. Then for a given graininess μ, the number $a = \frac{\alpha-1}{\mu} \in \mathbb{H}$. Similarly, let $\beta \in \mathbb{C}$ be such that $|\beta| < 1$, so that $b = \frac{\beta-1}{\mu} \in \mathbb{H}$. We set

$$c := a \oplus b = a + b + \mu ab.$$

Now, $c \in \mathbb{H}$ if there exists a $\gamma \in \mathbb{C}$ such that $|\gamma| < 1$ with $c = \frac{\gamma-1}{\mu}$. We claim that the choice $\gamma = \alpha\beta$ will suffice, from which the claim follows immediately.

Indeed, with this choice of γ, we have that

$$\frac{\gamma-1}{\mu} = \frac{\alpha-1}{\mu} + \frac{\beta-1}{\mu} + \mu\frac{\alpha-1}{\mu}\frac{\beta-1}{\mu},$$

and since $|\gamma| = |\alpha| \cdot |\beta| < 1$, the claim follows. □

Lemma 3.3 (Stability Under State Variable Change). *The regressive linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= A(t)x(t) + B(t)u(t), \qquad x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

is uniformly exponentially stable with rate $\frac{\lambda+\alpha}{1+\mu^*\alpha}$, *where* $\lambda, \alpha > 0$ *such that* $-\lambda \in \mathcal{R}^+$, *if the linear state equation*

$$z^{\Delta}(t) = [A(t)(1 + \mu\alpha) + \alpha I]z(t),$$

is uniformly exponentially stable with rate λ.

Proof. By direct calculation, $x(t)$ satisfies

$$x^{\Delta}(t) = A(t)x(t), \ \ x(t_0) = x_0,$$

if and only if $z(t) = e_{\alpha}(t, t_0)x(t)$ satisfies

$$z^{\Delta}(t) = [A(t)(1 + \mu\alpha) + \alpha I]z(t), \ \ z(t_0) = x_0. \tag{3.20}$$

Now assume there exists a $\gamma > 0$ such that for any x_0 and t_0 the solution of (3.20) satisfies

$$||z(t)|| \le \gamma e_{-\lambda}(t, t_0)||x_0||, \ \ t \ge t_0.$$

Then, substituting for $z(t)$ yields

$$||e_{\alpha}(t, t_0)x(t)|| = e_{\alpha}(t, t_0)||x(t)|| \le \gamma e_{-\lambda}(t, t_0)||x_0||,$$

so that

$$||x(t)|| \le \gamma e_{-\lambda \ominus \alpha}(t, t_0) \le \gamma e_{-(\lambda+\alpha)/(1+\mu^*\alpha)}(t, t_0).$$

An application of Lemma 3.2 then gives the result. $\square$

In order to achieve the desired stabilization result, we need to define a weighted version of the controllability Gramian defined earlier as

$$\mathcal{G}_C(t_0, t_f) = \int_{t_0}^{t_f} \Phi_A(t_0, \sigma(s))B(s)B^T(s)\Phi_A^T(t, \sigma(s))\Delta s. \tag{3.21}$$

To this end, for $\alpha > 0$ define the matrix $\mathcal{G}_{C_\alpha}(t_0, t_f)$ by

$$\mathcal{G}_{C_\alpha}(t_0, t_f) \ \ = \ \ \int_{t_0}^{t_f} (e_{\alpha}(t_0, s))^4 \Phi_A(t_0, \sigma(s))B(s)B^T(s)\Phi_A^T(t_0, \sigma(s))\Delta s. \tag{3.22}$$

We are now in position to prove the following major result of Chapter 3.

Theorem 3.26 (Gramian Exponential Stability Criterion). *Let $\mathbb{T}$ be a time scale with bounded graininess. For the regressive linear state equation*

$$\begin{aligned} x^{\Delta}(t) &= A(t)x + B(t)u(t), \ \ x(t_0) = x_0, \\ y(t) &= C(t)x(t), \end{aligned}$$

suppose there exist positive constants ϵ_1, ϵ_2 and a strictly increasing function $\mathcal{C} : \mathbb{T} \to \mathbb{T}$ such that $0 < \mathcal{C}(t) - t \leq M$ holds for some constant $0 < M < \infty$ and all $t \in \mathbb{T}$ with

$$\epsilon_1 I \leq \mathcal{G}_C(t, \mathcal{C}(t)) \leq \epsilon_2 I, \tag{3.23}$$

for all t. Then given a positive constant α, the state feedback gain

$$K(t) = -B^T(t)(I + \mu(t)A^T(t))^{-1}\mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t)), \tag{3.24}$$

is such that the resulting closed-loop state equation is uniformly exponentially stable with rate α.

Proof. We first note that for $N = \sup_{t \in \mathbb{T}} \dfrac{\log(1 + \mu(t)\alpha)}{\mu(t)}$, we have $0 < N < \infty$ since $\mathbb{T}$ has bounded graininess. Thus,

$$\begin{aligned} e_\alpha(t, \mathcal{C}(t)) &= \exp\left(-\int_t^{\mathcal{C}(t)} \frac{\log(1 + \mu(s)\alpha)}{\mu(s)} \Delta s\right) \\ &\geq \exp\left(-\int_t^{\mathcal{C}(t)} N \Delta s\right) \\ &= e^{-N(\mathcal{C}(t) - t)} \\ &\geq e^{-MN}. \end{aligned}$$

Comparing the quadratic forms $x^T \mathcal{G}_{C_\alpha}(t, \mathcal{C}(t))x$ and $x^T \mathcal{G}_C(t, \mathcal{C}(t))x$ using their respective definitions (3.21) and (3.22) gives

$$e^{-4MN}\mathcal{G}_C(t, \mathcal{C}(t)) \leq \mathcal{G}_{C_\alpha}(t, \mathcal{C}(t)) \leq \mathcal{G}_C(t, \mathcal{C}(t)),$$

for all t. Thus, (3.23) implies

$$\epsilon_1 e^{-4MN} I \leq \mathcal{G}_{C_\alpha}(t, \mathcal{C}(t)) \leq \epsilon_2 I \tag{3.25}$$

for all t, and so the existence of $\mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t))$ is immediate. Now, we show that the linear state equation

$$z^\Delta(t) = [\hat{A}(t)(1 + \mu(t)\alpha) + \alpha I]z(t), \tag{3.26}$$

where $\hat{A}(t) = A(t) - B(t)B^T(t)(I + \mu(t)A^T(t))\mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t))$, is uniformly exponentially stable by applying Theorem 1.23 with the choice

$$Q(t) = \mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t)). \tag{3.27}$$

Lemma 3.3 then gives the desired result. To apply the theorem, we first note that $Q(t)$ is symmetric and continuously differentiable. Thus, (3.25) implies

$$\frac{1}{\epsilon_2} I \leq Q(t) \leq \frac{e^{4MN}}{\epsilon_1} I, \tag{3.28}$$

for all t. Hence, it only remains to show that there exists $\nu > 0$ such that

$$\frac{\left(\left[I + \mu(t)\left[(1 + \mu(t)\alpha)\hat{A}(t) + \alpha I\right]^T\right] Q(\sigma(t)) \left[I + \mu(t)\left[(1 + \mu(t)\alpha)\hat{A}(t) + \alpha I\right]\right] - Q(t)\right)}{\mu(t)}$$

is less than or equal to $-\nu I$.

We begin with the first term, writing

$$\begin{aligned} & \left[I + \mu(t)\left[(1 + \mu(t)\alpha)\hat{A}(t) + \alpha I\right]^T\right] Q(\sigma(t)) \left[I + \mu(t)\left[(1 + \mu(t)\alpha)\hat{A}(t) + \alpha I\right]\right] \\ = \quad & (1 + \mu(t)\alpha)^2 \left[\left[I + \mu(t)A^T(t)\right] - \mathcal{G}_{C_\alpha}^{-1}(t, C(t)) \left[I + \mu(t)A(t)\right]^{-1} \mu(t)B(t)B^T(t)\right] \\ \cdot \quad & \mathcal{G}_{C_\alpha}^{-1}(\sigma(t), \mathcal{C}(\sigma(t))) \left[\left[I + \mu(t)A(t)\right] - \mu(t)B(t)B^T(t)\left[I + \mu(t)A^T(t)\right]^{-1} \mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t))\right]. \end{aligned}$$

We pause to establish an important identity. Notice that

$$\begin{aligned} & \left[I + \mu(t)A(t)\right] \mathcal{G}_{C_\alpha}(t, \mathcal{C}(t)) \left[I + \mu(t)A^T(t)\right] \\ & = \mu(t)B(t)B^T(t) + \frac{\mathcal{G}_{C_\alpha}(\sigma(t), \mathcal{C}(t))}{(1 + \mu(t)\alpha)^4}. \end{aligned} \tag{3.29}$$

This leads to

$$\begin{aligned} & I - \left[I + \mu(t)A(t)\right]^{-1} \mu(t)B(t)B^T(t)\left[I + \mu(t)A^T(t)\right]^{-1} \mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t)) \\ & = (1 + \mu(t)\alpha)^{-4} \left[I + \mu(t)A(t)\right]^{-1} \mathcal{G}_{C_\alpha}(\sigma(t), \mathcal{C}(t)) \left[I + \mu(t)A^T(t)\right]^{-1} \\ & \quad \cdot \mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t)), \end{aligned} \tag{3.30}$$

which in turn yields

$$\begin{aligned}
&I-\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\,[I+\mu(t)A(t)]^{-1}\,\mu(t)B(t)B^T(t)\left[I+\mu(t)A^T(t)\right]^{-1}\\
&=(1+\mu(t)\alpha)^{-4}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\,[I+\mu(t)A(t)]^{-1}\,\mathcal{G}_{C_\alpha}(\sigma(t),\mathcal{C}(t))\\
&\quad\cdot\left[I+\mu(t)A^T(t)\right]^{-1}. \qquad (3.31)
\end{aligned}$$

The first term can now be rewritten as

$$\begin{aligned}
&(1+\mu(t)\alpha)^2\left[\left[I+\mu(t)A^T(t)\right]-\mathcal{G}_{C_\alpha}^{-1}(t,C(t))\,[I+\mu(t)A(t)]^{-1}\,\mu(t)B(t)B^T(t)\right]\\
\cdot\;&\mathcal{G}_{C_\alpha}^{-1}(\sigma(t),\mathcal{C}(\sigma(t)))\left[[I+\mu(t)A(t)]-\mu(t)B(t)B^T(t)\left[I+\mu(t)A^T(t)\right]^{-1}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\right]\\
=\;&(1+\mu(t)\alpha)^2\left[I-\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\,[I+\mu(t)A(t)]^{-1}\,\mu(t)B(t)B^T(t)\left[I+\mu(t)A^T(t)\right]^{-1}\right]\\
\cdot\;&\left[I+\mu(t)A^T(t)\right]\mathcal{G}_{C_\alpha}^{-1}(\sigma(t),\mathcal{C}(t))\,[I+\mu(t)A(t)]\\
\cdot\;&\left[I-[I+\mu(t)A(t)]^{-1}\,\mu(t)B(t)B^T(t)\left[I+\mu(t)A^T(t)\right]^{-1}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\right]
\end{aligned}$$

Using (3.30) and (3.31), we can now write

$$\begin{aligned}
&\left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}^T(t)+\alpha I\right]\right]Q(\sigma(t))\left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}(t)+\alpha I\right]\right]\\
=\;&(1+\mu(t)\alpha)^{-6}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\,[I+\mu(t)A(t)]^{-1}\,\mathcal{G}_{C_\alpha}(\sigma(t),\mathcal{C}(t))\mathcal{G}_{C_\alpha}^{-1}(\sigma(t),\mathcal{C}(\sigma(t)))\\
\cdot\;&\mathcal{G}_{C_\alpha}(\sigma(t),\mathcal{C}(t))\left[I+\mu(t)A^T(t)\right]^{-1}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t)). \qquad (3.32)
\end{aligned}$$

On the other hand, from the definition of $\mathcal{G}_{C_\alpha}(t,\mathcal{C}(t))$, we have

$$\mathcal{G}_{C_\alpha}(\sigma(t),\mathcal{C}(\sigma(t)))\geq\mathcal{G}_{C_\alpha}(\sigma(t),\mathcal{C}(t)),$$

which in turn implies

$$\mathcal{G}_{C_\alpha}^{-1}(\sigma(t),\mathcal{C}(\sigma(t)))\leq\mathcal{G}_{C_\alpha}^{-1}(\sigma(t),\mathcal{C}(t)).$$

Combining this with (3.32) gives

$$\begin{aligned}
&\left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}^T(t)+\alpha I\right]\right]Q(\sigma(t))\left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}(t)+\alpha I\right]\right]\\
\leq\;&(1+\mu(t)\alpha)^{-6}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t))\left[[I+\mu(t)A(t)]^{-1}\,\mathcal{G}_{C_\alpha}(\sigma(t),\mathcal{C}(t))\left[I+\mu(t)A^T(t)\right]^{-1}\right]\\
\cdot\;&\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t)).
\end{aligned}$$

Applying (3.29) again yields

$$
\begin{aligned}
& \left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}^T(t)+\alpha I\right]\right] Q(\sigma(t)) \left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}(t)+\alpha I\right]\right] \\
\leq\ & (1+\mu(t)\alpha)^{-6}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t)) \\
\cdot\ & \left[(1+\mu(t)\alpha)^4\mathcal{G}_{C_\alpha}(t,\mathcal{C}(t))-(1+\mu(t)\alpha)^4\left[I+\mu(t)A(t)\right]^{-1}\mu(t)B(t)B^T(t)\left[I+\mu(t)A^T(t)\right]^{-1}\right] \\
\cdot\ & \mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t)) \\
\leq\ & (1+\mu(t)\alpha)^{-2}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t)).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& \frac{\left(\left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}^T(t)+\alpha I\right]\right] Q(\sigma(t)) \left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}(t)+\alpha I\right]\right]-Q(t)\right)}{\mu(t)} \\
\leq\ & -\frac{(1+\mu(t)\alpha)^2-1}{\mu(t)(1+\mu(t)\alpha)^2}\mathcal{G}_{C_\alpha}^{-1}(t,\mathcal{C}(t)) \\
\leq\ & -\frac{(1+\mu(t)\alpha)^2-1}{\mu(t)\epsilon_2(1+\mu(t)\alpha)^2}I.
\end{aligned}
$$

Now, the quantity $((1+\mu(t)\alpha)^2-1)\,/\,(\mu(t)\epsilon_2(1+\mu(t)\alpha)^2)$ is certainly not constant, but it can be bounded by a quantity that is (here $\mu^*=\mu_{\max}$):

$$
\frac{(1+\mu(t)\alpha)^2-1}{\mu(t)\epsilon_2(1+\mu(t)\alpha)^2}=\frac{2\alpha+\mu(t)\alpha^2}{\epsilon_2(1+\mu(t)\alpha)^2}\geq\frac{\alpha}{\epsilon_2(1+\mu^*\alpha)^2}.
$$

Thus, if we set $\nu=\alpha/(\epsilon_2(1+\mu^*\alpha)^2)$, then

$$
\frac{\left(\left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}^T(t)+\alpha I\right]\right] Q(\sigma(t)) \left[I+\mu(t)\left[(1+\mu(t)\alpha)\hat{A}(t)+\alpha I\right]\right]-Q(t)\right)}{\mu(t)}
$$

is in fact less than or equal to $-\nu I$. □

Some questions remain regarding Theorem 3.26. What types of functions $\mathcal{C}(t)$, which we term the *compactification operator*, will suffice to meet the hypothesis of the preceding theorem? For $\mathbb{T}=\mathbb{R}$, it is known that $\mathcal{C}(t)=t+\delta$ for any $\delta>0$ is sufficient, while on $\mathbb{T}=\mathbb{Z}$, the function $\mathcal{C}(t)=t+k$ for $k\in\mathbb{N}$ meets the criteria. In fact, our result agrees with the results known in each of these cases: on $\mathbb{R}$, for $\mathcal{C}(t)=t+\delta$ and $K(t)=-B^T(t)\mathcal{G}_{C_\alpha}^{-1}(t,t+\delta)$, it is proved in [13] and [39] that the result holds, as is the case on $\mathbb{Z}$, where $\mathcal{C}(t) = t + k$ and $K(t) = -B^T(t)A^{-T}\mathcal{G}_{C_\alpha}^{-1}(t, t+k)$. (Recall that the time scales analysis deals with the difference form rather than the recursive form, so we do indeed expect to get a shift of the known result on $\mathbb{Z}$.) For the time scale $\mathbb{P}_{a,b}$ defined earlier, a candidate for $\mathcal{C}(t)$ is given by $\mathcal{C}(t) = t + a + b$. If we examine $\mathbb{Z}$ a little more closely, we see that on this time scale the choice $\mathcal{C}(t) = t + k$ is really $\mathcal{C}(t) = \sigma^k(t)$, which leads to the conclusion that a possible choice for a purely discrete time scale $\mathbb{T}$ in general (that is, a time scale with no right dense points) could be $\mathcal{C}(t) = \sigma^k(t)$ for some $k > 0 \in \mathbb{N}$. For a general time scale $\mathbb{T}$ with both right dense and right scattered points, one possible choice for $\mathcal{C}(t)$ is

$$\mathcal{C}(t) = \begin{cases} t + \delta_1, & \text{if } t \text{ is right dense,} \\ \sigma^k(t), & \text{if } \sigma^i(t) \neq t \text{ for all } 0 \leq i \leq k, \\ \sigma^k(t) + \delta_2, & \text{else.} \end{cases}$$

We now demonstrate the preceding theorem with an example.

Example 3.7. Let $p, q \in \mathbb{R}^+$ be constants such that the system

$$\begin{aligned} x^\Delta(t) &= \begin{bmatrix} \frac{9}{10}\frac{(\sqrt{4+e_p(t,0)})^\Delta}{\sqrt{4+e_p(t,0)}} + \frac{1}{10}\frac{(\sqrt{10+e_q(t,0)})^\Delta}{\sqrt{10+e_q(t,0)}} & \frac{3}{10}\frac{(\sqrt{4+e_p(t,0)})^\Delta}{\sqrt{4+e_p(t,0)}} - \frac{3}{10}\frac{(\sqrt{10+e_q(t,0)})^\Delta}{\sqrt{10+e_q(t,0)}} \\ \frac{3}{10}\frac{(\sqrt{4+e_p(t,0)})^\Delta}{\sqrt{4+e_p(t,0)}} - \frac{3}{10}\frac{(\sqrt{10+e_q(t,0)})^\Delta}{\sqrt{10+e_q(t,0)}} & \frac{1}{10}\frac{(\sqrt{4+e_p(t,0)})^\Delta}{\sqrt{4+e_p(t,0)}} + \frac{9}{10}\frac{(\sqrt{10+e_q(t,0)})^\Delta}{\sqrt{10+e_q(t,0)}} \end{bmatrix} x(t) \\ &+ \begin{bmatrix} \frac{\sqrt{10}}{10} & 0 \\ 0 & -\frac{\sqrt{10}}{10} \end{bmatrix} u(t), \\ y(t) &= x(t), \end{aligned}$$

with initial condition

$$x(0) = \begin{bmatrix} 3\sqrt{5} & \sqrt{11} \\ \sqrt{5} & -3\sqrt{11} \end{bmatrix},$$

is regressive.

A direct calculation (i.e. verification) shows that the transition matrix is given by

$$\Phi_A(t,\sigma(s)) = \begin{bmatrix} 3(\sqrt{4+e_p(t,\sigma(s))}) & \sqrt{10+e_q(t,\sigma(s))} \\ \sqrt{4+e_p(t,\sigma(s))} & -3(\sqrt{10+e_q(t,\sigma(s))}) \end{bmatrix}.$$

Now, we have

$$\begin{aligned} &\Phi_A(t,\sigma(s))B(s)B^T(s)\Phi_A^T(t,\sigma(s)) \\ &= \begin{bmatrix} \frac{9}{10}(4+e_p(t,\sigma(s))) + \frac{1}{10}(10+e_q(t,\sigma(s))) & \frac{3}{10}(4+e_p(t,\sigma(s)) - \frac{3}{10}(10+e_q(t,\sigma(s))) \\ \frac{3}{10}(4+e_p(t,\sigma(s)) - \frac{3}{10}(10+e_q(t,\sigma(s))) & \frac{1}{10}(4+e_p(t,\sigma(s))) + \frac{9}{10}(10+e_q(t,\sigma(s))) \end{bmatrix}, \end{aligned}$$

which can be diagonalized as

$$\begin{bmatrix} 3 & -\frac{1}{3} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4+e_p(t,\sigma(s)) & 0 \\ 0 & 10+e_q(t,\sigma(s)) \end{bmatrix} \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{bmatrix}.$$

Thus, $\mathcal{G}_C(t,\mathcal{C}(t))$ can be written as

$$\mathcal{G}_C(t,\mathcal{C}(t)) = \int_t^{\mathcal{C}(t)} \begin{bmatrix} 3 & -\frac{1}{3} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4+e_p(t,\sigma(s)) & 0 \\ 0 & 10+e_q(t,\sigma(s)) \end{bmatrix} \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{bmatrix} \Delta s.$$

For $s \geq t$, the eigenvalues $\lambda_1(t,\sigma(s)) = 4+e_p(t,\sigma(s))$ and $\lambda_2(t,\sigma(s)) = 10+e_q(t,\sigma(s))$ have the bounds $4 \leq \lambda_1(t,\sigma(s)) \leq 5$ and $10 \leq \lambda_2(t,\sigma(s)) \leq 11$, respectively. Thus,

$$\begin{aligned} &\begin{bmatrix} 3 & -\frac{1}{3} \\ 1 & 1 \end{bmatrix} \left(\int_t^{\mathcal{C}(t)} \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \Delta s \right) \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{bmatrix} \\ \leq\ & \mathcal{G}_C(t,\mathcal{C}(t)) \\ \leq\ & \begin{bmatrix} 3 & -\frac{1}{3} \\ 1 & 1 \end{bmatrix} \left(\int_t^{\mathcal{C}(t)} \begin{bmatrix} 11 & 0 \\ 0 & 11 \end{bmatrix} \Delta s \right) \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{bmatrix}, \end{aligned}$$

or equivalently

$$\begin{aligned} &\begin{bmatrix} 3 & -\frac{1}{3} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4(\mathcal{C}(t)-t) & 0 \\ 0 & 4(\mathcal{C}(t)-t) \end{bmatrix} \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{bmatrix} \\ \leq\ & \mathcal{G}_C(t,\mathcal{C}(t)) \\ \leq\ & \begin{bmatrix} 3 & -\frac{1}{3} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 11(\mathcal{C}(t)-t) & 0 \\ 0 & 11(\mathcal{C}(t)-t) \end{bmatrix} \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{bmatrix}. \end{aligned}$$

Thus, if we assume $0 < N \leq \mathcal{C}(t) - t \leq M < \infty$, then

$$4NI \leq \mathcal{G}_C(t, \mathcal{C}(t)) \leq 11MI.$$

By Theorem 3.26, the closed loop equation

$$\begin{aligned} x^{\Delta}(t) &= (A + BK)(t)x(t), \\ y(t) &= x(t), \end{aligned}$$

is uniformly exponentially stable if we choose $\alpha > 0$ and

$$K(t) = -B^T(t)(I + \mu(t)A^T(t))^{-1}\mathcal{G}_{C_\alpha}^{-1}(t, \mathcal{C}(t)).$$

CHAPTER FOUR

Conclusions and Future Directions

In this dissertation, we have examined several different aspects of a linear systems theory in the arbitrary time scale setting. The Laplace transform given by Bohner and Peterson in [10] has been analyzed and given a rigorous foundation. The concepts of controllability, observability, and reachability have been introduced and conditions for a system to possess these properties presented. Exponential stability as defined by DaCunha in [15, 16] has also been discussed further, with particular interest concerning its relation to bounded-input, bounded-output stability, another concept that has been defined in the dissertation. We also examined the effects of linear state feedback in systems and showed that it is possible to stabilize a system whose controllability Gramian is bounded by positive perturbations of the identity regardless of the spacing of the time scale, as long as the time scale does not have arbitrarily large spacing.

The applications of the theory could most likely happen in the area of adaptive control (see [23], [24], and [25]). In particular, in real time communication networks, it would be advantageous to be able to analyze the system without knowing the time scale *a priori*. Indeed, an analysis that allows the time scale to be created "on the fly" would be useful as the times of serious activity on the network are most often unknown and can be difficult to account for. Resources used by the network could be saved and bandwidth limitations maintained if the system does not have to be sampled frequently in an uniform fashion as is currently done. Thus, the utility of our analysis becomes self evident: we are never concerned with the underlying spacing to obtain our results.

As for future directions, there are certain crucial tools in systems analysis still missing. For example, we currently lack a frequency analysis in the general time scale setting. This tool is critical for determining how systems and their dynamics evolve in time, and it is also useful for feedback as it can be used to select appropriate feedback gains to obtain desired physical properties of systems. Another tool that we lack currently is the theory of output feedback and observers. These are necessary to design effective controlled systems that achieve prescribed results. The results established in this dissertation are useful in this regard since we now know how to invert the transform, which provides the basis for associating the usual trigonometric functions with their generalized transforms thereby understanding frequency. The results concerning state feedback give insight as to determining corresponding results for output feedback.

Another area that deserves attention is the control of systems involving boundary value problems (BVPs). To the author's knowledge, at present there is nothing in the literature that deals with this concept in the time scale setting. Thus, it is an area ripe for research.

BIBLIOGRAPHY

[1] B.D.O Anderson and J.B. Moore. New results in linear system theory. *SIAM Journal on Control and Optimization*, 6(1):121–130, 1968.

[2] B.D.O. Anderson and L.M. Silverman. Controllability, observability, and stability of linear systems. *SIAM Journal on Control and Optimization*, 6(1):121–130, 1968.

[3] Z. Bartosiewicz, Ü. Kotta, and E. Pawluszewicz. Equivalence of linear control systems on time scales. *Proceedings of the Estonian Academy of Sciences*, 55:43–52, 2006.

[4] Z. Bartosiewicz and E. Pawluszewicz. Unification of continuous-time and discrete-time systems: the linear case. url: http://katmat.pb.bialystok.pl/mat/barz/mtns2004n.pdf.

[5] L. Berg. *Introduction to the Operational Calculus*. John Wiley & Sons, New York, 1967.

[6] N. Bettaibi, K. Brahim, and A. Fitouhi. The Mellin transform in quantum calculus. *Constructions and Approximations*, 23(3):305–323, 2006.

[7] N. Bettaibi and A. Fitouhi. Wavelet transforms in quantum calculus. *Journal of Nonlinear Mathematical Physics*, 13(4):492–506, 2006.

[8] M. Bohner and A. Peterson. *Dynamic Equations on Time Scales: An Introduction with Applications*. Birkäuser, Boston, 2001.

[9] M. Bohner and A. Peterson. *Advances in Dynamic Equations on Time Scales*. Birkhäuser, Boston, 2003.

[10] M. Bohner and A. Peterson. Laplace transform and Z-transform: unification and extension. *Methods and Applications of Analysis*, 9(1):155–162, 2002.

[11] T. Bromwich. Normal Co-ordinates in Dynamical Systems. *Proceedings of the London Mathematical Society*, 15(2):401–448, 1916.

[12] F.M. Callier and C.A. Desoer. *Linear System Theory*. Springer-Verlag, New York, 1991.

[13] V.H.L. Cheng. A direct way to stabilize continuous-time and discrete-time linear time-varying systems. *IEEE Transactions on Automatic Control*, 24(4):641–643, 1979.

[14] P. Cheung and V. Kac. *Quantum Calculus*. Universitext, Springer-Verlag, New York, 2002.

[15] J.J. DaCunha. *Lyapunov Stability and Floquet Theory for Nonautonomous Linear Dynamic Systems on Time Scales*. Ph.D. thesis, Baylor University, 2004.

[16] J.J. DaCunha. Stability for time varying linear dynamic systems on time scales. J. Comput. Appl. Math. 176(2), 381–410, 2005.

[17] J.J. DaCunha. Transition matrix and generalized exponential via the Peano-Baker series. *Journal of Difference Equations and Applications*, 11(15):1245–1264, 2005.

[18] J.M. Davis, I.A. Gravagne, B.J. Jackson, R.J. Marks II, and A.A. Ramos. The Laplace transform on time scales revisited. *Journal of Mathematical Analysis and Applications*, 332(2): 1291–1306, 2007.

[19] J.C. Engwerda. Control aspects of linear discrete time-varying systems. *International Journal of Control*, 48(4):1631–1658, 1988.

[20] L. Fausett and K. Murty. Controllability, observability, and realizability criteria on time scale dynamical systems. *Nonlinear Studies*, 11(4):627–638, 2004.

[21] G.F. Franklin, J.D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems*. Prentice Hall, New Jersey, 2002.

[22] T. Gard and J. Hoffacker. Asymptotic behavior of natural growth on time scales. *Dynamic Systems and Applications*, 12(1–2):131–148, 2003.

[23] I.A. Gravagne, J.M. Davis, and J.J. DaCunha. A unified approach to discrete and continuous high-gain adaptive controllers using time scales, submitted.

[24] I.A. Gravagne, J.M. Davis, J.J. DaCunha, and R.J. Marks II. Bandwidth reduction for controller area networks using adaptive sampling. *Proc. Int. Conf. Robotics and Automation (ICRA)*, New Orleans, LA, April 2004, pp. 5250-5255.

[25] I.A. Gravagne, J.M. Davis, and R.J. Marks II. How deterministic must a real-time controller be? *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS 2005)*, Alberta, Canada. Aug. 2–6, 2005, pp. 3856–3861.

[26] G. Guseinov. Integration on time scales. *Journal of Mathematical Analysis and Applications*, 285:107–127, 2003.

[27] O. Heaviside. *Electrical Papers*. The Macmillan Company, New York, 1892.

[28] O. Heaviside. *Electromagnetic Theory*. Reprinted, Benn Brothers, London, 1922.

[29] O. Heaviside. On Operators in Physical Mathematics. *Proceedings of the Royal Society of London*, 52:505–529, 1893, 54:105–143, 1893.

[30] T. Higgins. History of the operational calculus as used in electric circuit analysis. *Electrical Engineering*, 68:42–45, 1949.

[31] S. Hilger. Analysis on measure chains—a unified approach to continuous and discrete calculus. *Results Math.*, 18:18–56, 1990.

[32] S. Hilger. *Ein Maßkettenkalkül mit Anwendung auf Zentrumsmannigfaltigkeiten*. Ph.D. thesis, Universität Würzburg, 1988.

[33] Y.C. Ho, R.E. Kalman, and K.S. Narendra. Controllability of linear dynamical systems. *Contr. Diff. Eqns.*, 1:189–213, 1963.

[34] E.I. Jury. *Sampled-Data Control Systems*. John Wiley & Sons, 1958.

[35] R.E. Kalman. Contributions to the theory of optimal control. *Bol. Soc. Mat. Mex.* 5:102–119, 1960.

[36] W.G. Kelley and A.C. Peterson. *Difference Equations: An Introduction with Applications*, 2nd edition. Academic Press, San Diego, 2001.

[37] P.P. Khargonekar and R. Ravi. Exponential and input-output stability are equivalent for linear time-varying systems. *Sadhana*, 18(1):31–37, 1993.

[38] 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(5):1223–1241, 2003.

[39] W. Rugh. *Linear System Theory*, 2nd edition. Prentice Hall, New Jersey, 1996.

[40] L. Weiss. Controllability, realization, and stability of discrete-time systems. *SIAM Journal on Control and Optimization*, 10(2):230–251, 1972.