

NUMERICAL INTEGRATION FOR TIME SCALES

by

BREE ETTINGER

(Under the direction of Edward Azoff)

ABSTRACT

Time Scale theory unifies and extends real and discrete analysis. Euler's Method is an algorithm for constructing a linear approximation to the initial value problem

$$\frac{\delta y}{\delta t} = f(t, y) \quad a \leq t \leq b \text{ and } y(a) = \alpha.$$

Although Euler's Method is seldom used in practice, it lays the ground work for developing higher order approximation schemes. In this thesis, we use Euler's Method to develop a numerical integration scheme for time scales. First we explore our numerical integration scheme on an isolated time scale, and then extend our results to an arbitrary time scale. We also give error approximations for these schemes as well as examples.

INDEX WORDS: Time scales, Euler's Method, Numerical Integration

NUMERICAL INTEGRATION FOR TIME SCALES

by

BREE ETTINGER

BS, The University of Georgia, 2002

A Thesis Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2004

© 2004

Bree Ettinger

All Rights Reserved

NUMERICAL INTEGRATION FOR TIME SCALES

by

BREE ETTINGER

Approved:

Major Professor: Edward Azoff

Committee: Malcolm Adams
Joan Hoffacker
Paul Wenston

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2004

ACKNOWLEDGMENTS

First and foremost I would like to thank Dr. Joan Hoffacker for being a wonderful mentor and friend. She put in many long hours into this thesis with me proofreading and discussing ideas. I would also like to thank Dr. Ed Azoff for being a wonderful advisor and always encouraging me to ride my bike. Thank you to Dr. Malcom Adams, Dr. Ed Azoff, Dr. Joan Hoffacker and Dr. Paul Weston for being on my committee. I would like to thank my fellow graduate students for their guidance especially Alan Thomas. Last but not least I would like to thank my friends and family for always being supportive. A special thanks to Jamie Saye, my mom and my dad for being understanding when I had to work nights, weekends and holidays.

TABLE OF CONTENTS

	Page
ACKNOWLEDEGMENTS	iv
CHAPTER	
1 INTRODUCTION	1
1.1 BASIC NOTATION AND DEFINITIONS	2
1.2 DIFFERENTIATION	11
1.3 INTEGRATION	19
1.4 POLYNOMIALS AND TAYLOR'S FORMULA	39
1.5 THE EXPONENTIAL FUNCTION	49
1.6 EULER'S METHOD	55
2 NUMERICAL INTEGRATION FOR TIME SCALES	66
2.1 NUMERICAL INTEGRATION ON ISOLATED TIME SCALES	66
2.2 NUMERICAL INTEGRATION ON ARBITRARY TIME SCALES	73
2.3 CONCLUSION	86
BIBLIOGRAPHY	88

CHAPTER 1

INTRODUCTION

In 1988, Stefan Hilger introduced time scale calculus in his Ph.D. dissertation [6]. His motivation was to settle the discrepancies between continuous and discrete analysis. For example, time scale theory allows one to prove results for both difference equations and differential equations simultaneously. Before time scale theory, one would prove a result for a differential equation defined on the real numbers $\mathbb{R}$ and a similar result for a difference equation defined on the integers $\mathbb{Z}$. Now using time scale theory one can prove a general result for an equation defined on any closed subset of the reals, which we call a time scale $\mathbb{T}$. Not only does time scale theory unify continuous and discrete analysis, but it also extends our understanding to other sets, such as $q^{\mathbb{N}_0} = \{q^n : n \in \mathbb{N}_0\}$, where $q > 1$ is fixed. As one may expect, many results can be proven for arbitrary time scales but more can always be said about a specific time scale. Although time scale calculus is rich in theory very few applications have been developed so far.

Euler's Method approximates a piecewise linear solution to the continuous initial value problem

$$\frac{\delta y}{\delta t} = f(t, y) \quad a \leq t \leq b, \quad y(a) = \alpha,$$

by evaluating

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + hf(t_i, w_i), \end{aligned}$$

at points t_i in the interval $[a, b]$ that are all a step size h apart. This set of evenly spaced points is a closed subset of the reals which in turn is a time scale. Thus Euler's Method

uses an isolated time scale to approximate a solution that is defined on all of the reals. Therefore it is natural to want to extend Euler's Method to time scales. Euler's Method is not usually used in practice; however it provides a conceptual basis for developing higher order approximations.

In this thesis, we use Euler's Method as a basis for developing a numerical integration scheme for time scales. We will see for discrete time scales our numerical integration scheme finds the exact solution. When we consider an arbitrary time scale we will see that real intervals and dense points in our time scale introduce error and left scattered and isolated points do not contribute to the error.

In this chapter, we introduce time scale calculus and Euler's Method. Most of the results in this section are from the two books on time scales by Bohner and Peterson, [1] and [2]. For background on Euler's Method we will use Burden and Faires [3]. Some of the proofs in these references were left as exercises. We provide proofs here. Others were presented, in which case we have added some details and explanations.

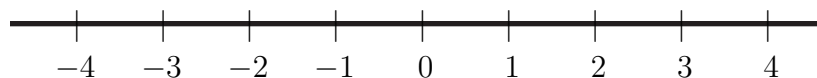
In Chapter 2, we develop a numerical integration scheme for time scales by using Euler's Method. First we consider a time scale consisting of isolated points then we extend our results to an arbitrary time scale. For both cases we develop error approximations and provide examples.

1.1 BASIC NOTATION AND DEFINITIONS

Definition 1.1 A *time scale*, denoted $\mathbb{T}$, is a nonempty closed subset of the real numbers.

A few common examples are:

- (i) the real numbers $\mathbb{R}$



(ii) the integers $\mathbb{Z}$



(iii) $h\mathbb{Z} = \{hz : z \in \mathbb{Z}\}$, where h is a fixed positive real number. For example $2\mathbb{Z}$ is the even integers



(iv) the natural numbers $\mathbb{N} = \{1, 2, \dots\}$

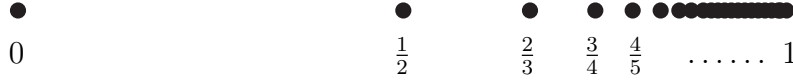


(v) $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$

(vi) $q^{\mathbb{N}_0} = \{q^n : n \in \mathbb{N}_0\}$, where $q > 1$ is fixed. For example $2^{\mathbb{N}_0}$ is the set of powers of 2



(vii) $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$.



It is usual to assume that the time scales are endowed with the standard relative topology from the real numbers.

Definition 1.2 For $t \in \mathbb{T}$ the *forward jump operator* $\sigma : \mathbb{T} \rightarrow \mathbb{T}$ is defined by

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

the *backward jump operator* $\rho : \mathbb{T} \rightarrow \mathbb{T}$ is defined by

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

and the *graininess function* $\mu : \mathbb{T} \rightarrow [0, \infty)$ is defined by

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

The convention here is to let $\inf \emptyset = \sup \mathbb{T}$ and $\sup \emptyset = \inf \mathbb{T}$.

Example 1.3 Let $\mathbb{T} = \mathbb{Z}$. Consider the point $t \in \mathbb{Z}$. The next greatest point is $t + 1$, thus $\sigma(t) = t + 1$. Similarly the greatest previous point is $t - 1$ and $\rho(t) = t - 1$. The space between any two points has a constant graininess $\mu(t) = \sigma(t) - t = t + 1 - t = 1$

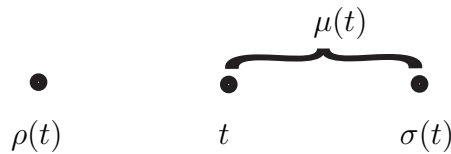


Figure 1.1: $\rho(t)$, $\sigma(t)$ and $\mu(t)$ for t on $\mathbb{Z}$

However, if we let $\mathbb{T} = \mathbb{R}$ and consider $t \in \mathbb{R}$, then $\sigma(t) = t$ and $\rho(t) = t$. Thus, $\mathbb{R}$ has a constant graininess $\mu(t) = 0$

Now let us consider a time scale that does not have a constant graininess. Let $\mathbb{T} = (-\infty, 0] \cup \mathbb{N}$



Figure 1.2: $\mathbb{T} = (-\infty, 0] \cup \mathbb{N}$

For a point $t \in \mathbb{T}$ there are two cases. If $t < 0$ then $\sigma(t) = t$ and $\mu(t) = 0$. If $t \geq 0$ then $\sigma(t) = t + 1$ and $\mu(t) = 1$.

Let us take a closer look at the different points in the previous example $\mathbb{T} = (-\infty, 0] \cup \mathbb{N}$. For any point $t \in \mathbb{T}$ such that $t < 0$, we have $\sigma(t) = t$ and $\rho(t) = t$. This point is called dense. Yet, when $t > 0$, we have $\sigma(t) > t$ and $\rho(t) < t$ so this point is called isolated. The most interesting point in this time scale is $t = 0$, if we look to the right of this point $\sigma(t) > t$ so the point is isolated to the right. Yet, if we look to the left of t , $\rho(t) = t$, so the point is dense to the left. In fact the point $t = 0$ is called right scattered and left dense. The following definition clarifies the situation.

Definition 1.4 If $\sigma(t) > t$, then we say that t is *right scattered*. If $\rho(t) < t$ we say that t is *left scattered*. If a point is both right scattered and left scattered, then it is called *isolated*. If $t < \sup \mathbb{T}$ and $\sigma(t) = t$, then t is called *right dense*. If $t > \inf \mathbb{T}$ and $\rho(t) = t$, then t is called *left dense*. Points that are both right dense and left dense are called *dense*.

t_1	left dense and right scattered
t_2	right dense and left scattered
t_3	dense
t_4	isolated

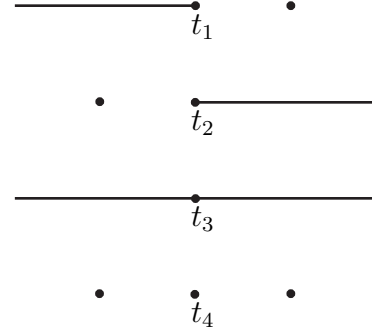


Figure 1.3: Classification of Points

We need define a subset, $\mathbb{T}^\kappa$ of a given timescale such that every point in the timescale will have a defined forward jump operator. For example, if $\mathbb{T}$ has a left scattered maximum then we need to ignore that point so that the forward jump operator, σ , is defined. Thus we have the following definition.

Definition 1.5 The set $\mathbb{T}^\kappa$ is defined as

$$\mathbb{T}^\kappa := \begin{cases} \mathbb{T} \setminus (\rho(\sup \mathbb{T}), \sup \mathbb{T}] & \text{if } \sup \mathbb{T} < \infty \\ \mathbb{T} & \text{if } \sup \mathbb{T} = \infty \end{cases}.$$

Whenever we consider an interval I in the time scale setting we actually want $I \cap \mathbb{T}$. Next, we will look at the *induction principle*. It is a useful tool we will need for later proofs.

Theorem 1.6 (Induction Principle). *Let $t_0 \in \mathbb{T}$ and assume that*

$$\{S(t) : t \in [t, \infty)\}$$

is a family of statements satisfying:

(i) *The statement $S(t_0)$ is true.*

(ii) *If $t \in [t_0, \infty)$ is right scattered and $S(t)$ is true, then $S(\sigma(t))$ is also true.*

(iii) If $t \in [t_0, \infty)$ is right dense and $S(t)$ is true, then there is a neighborhood U of t such that $S(s)$ is true for all $s \in U \cap (t, \infty)$.

(iv) If $t \in (t_0, \infty)$ is left dense and $S(s)$ is true for all $s \in [t_0, t)$, then $S(t)$ is true.

Then $S(t)$ is true for all $t \in [t_0, \infty)$.

Proof. Let $S^* := \{t \in [t_0, \infty) : S(t) \text{ is not true}\}$. Assume $S^* \neq \emptyset$. Since S^* is non-empty and $\mathbb{T}$ is closed then

$$\inf S^* =: t^* \in \mathbb{T}.$$

Let $S(t^*)$ be true. If $t^* = t_0$, then $S(t^*)$ is true by (i) and $t^* \notin S^*$. If t is left dense then $t^* \neq t_0$ and $\rho(t^*) = t^*$, thus $S(t^*)$ is true by (iv). If t is right scattered then $\rho(t^*) \leq t^*$, and $S(t^*)$ is true from (ii). Thus for any case $t^* \notin S^*$. If t^* is not right scattered, and $t^* \neq \max \mathbb{T}$ then t^* is right dense and (iii) is true. Thus any case leads us to a contradiction, and $S(t^*) = \emptyset$. $\square$

The Induction Principle is used to show a statement, $S(t)$ is true for $t \in [t_0, \infty)$. Now we introduce a dual version to prove a statement $S(t)$ is true for $t \in (-\infty, t_0]$.

Theorem 1.7 (Dual Induction Principle). *Let $t_0 \in \mathbb{T}$ and assume that*

$$\{S(t) : t \in (-\infty, t_0]\}$$

is a family of statements satisfying:

(i) *The statement $S(t_0)$ is true.*

(ii) *If $t \in (-\infty, t_0]$ is left scattered and $S(t)$ is true, then $S(\rho(t))$ is also true.*

(iii) *If $t \in (-\infty, t_0]$ is left dense and $S(t)$ is true, then there is a neighborhood U of t such that $S(r)$ is true for all $r \in U \cap (-\infty, t)$.*

(iv) *If $t \in (-\infty, t_0]$ is right dense and $S(r)$ is true for all $r \in (t, t_0]$, then $S(t)$ is true.*

Then $S(t)$ is true for all $t \in (-\infty, t_0]$.

The proof of the Dual Induction Principle is similar to that of the Induction Principle. Before we can get into classes of functions we need a notion of continuity for time scales.

Definition 1.8 A function $f : \mathbb{T} \rightarrow \mathbb{R}$ is called *regulated* provided that its right sided limits exist at all right dense points in $\mathbb{T}$ and its left sided limits exist at all left dense points in $\mathbb{T}$, that is only jump discontinuities are possible.

Example 1.9 To discuss regulated functions we will consider a function that is not regulated on $\mathbb{R}$. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by

$$f(t) := \begin{cases} \frac{1}{t} & t \neq 0 \\ 0 & t = 0 \end{cases}.$$

While f is continuous for all $t \in \mathbb{R} \setminus \{0\}$, neither the left or right sided limit exists at 0 so it is not regulated. However, if we consider f on $\mathbb{Z}$ then f is regulated only because $\mathbb{Z}$ does not have any dense points.

Definition 1.10 A function $f : \mathbb{T} \rightarrow \mathbb{R}$ is called *right dense continuous* provided it is continuous at right dense points in $\mathbb{T}$ and its left sided limits exist at left dense points in $\mathbb{T}$. Note that if a function is regulated and continuous from the right it is right dense continuous. We often denote the set of right dense continuous functions by C_{rd} .

Example 1.11 Let $\mathbb{T} = \{1 + \frac{1}{n} : n \in \mathbb{N}\} \cup \{2 - \frac{1}{n} : n \in \mathbb{N}\}$.

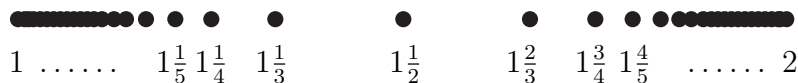


Figure 1.4: $\mathbb{T} = \{1 + \frac{1}{n} : n \in \mathbb{N}\} \cup \{2 - \frac{1}{n} : n \in \mathbb{N}\}$

Define $f : \mathbb{T} \rightarrow [1, 2]$ by

$$f(t) := \begin{cases} t & t \neq 2 \\ 1 & t = 2 \end{cases}.$$

This function is continuous at isolated points in $\mathbb{T}$. The two interesting points are the right dense point 1 and the left dense point 2. In addition f is continuous at 1 because the right sided limit of f at 1 exists and is equal to $f(1)$. However f is not continuous at 2, but the left sided limit of f exists at 2. So f is right dense continuous. Even though f is not continuous on $\mathbb{T}$ it is right dense continuous.

Theorem 1.12. *Let $f : \mathbb{T} \rightarrow \mathbb{R}$ and $g : \mathbb{T} \rightarrow \mathbb{T}$. Then*

- (i) *If f is continuous, then f is right dense continuous.*
- (ii) *If f is continuous and g is regulated or right dense continuous, then $f \circ g$ is respectively regulated or right dense continuous.*
- (iii) *If f is right dense continuous, then it is regulated.*

Proof. (i) This is obvious since f is continuous it is continuous in both directions and therefore continuous from the right. Similarly the limit from the left exists and is finite since f is continuous.

- (ii) Assume f is continuous and g is right dense continuous. Let x_0 be a right dense point, since g is continuous at x_0 . Let $c = g(x_0)$ then for any given $\epsilon > 0$ there exists a $\delta_1 > 0$ such that if

$$|t - c| < \delta_1,$$

then

$$|f(t) - f(c)| < \epsilon.$$

There also exists a $\delta_2 > 0$ for a given $\epsilon > 0$ such that when

$$|x - x_0| < \delta_2$$

implies that

$$|g(x) - g(x_0)| < \delta_1.$$

Let $\delta = \min\{\delta_1, \delta_2\}$ and substitute $t = g(x)$ and $c = g(x_0)$. Then

$$|x - x_0| < \delta$$

implies

$$|g(x) - g(x_0)| < \delta_1$$

and thus

$$|f(g(x)) - f(g(x_0))| < \epsilon.$$

So, $f \circ g$ is continuous at right dense points. Now, let x_1 be a left dense point. Again let $c = g(x_0)$ then for any $\epsilon > 0$ there exists a $\delta_1 > 0$ such that for

$$|t - c| < \delta_1,$$

then

$$|f(t) - f(c)| < \epsilon.$$

Since g is right dense continuous the limit L at x_1 exists and is finite. For any $\epsilon > 0$ there exists a $\delta_2 > 0$ such that

$$|x - x_1| < \delta_2$$

implies

$$|g(x) - L| < \epsilon.$$

Let $\delta = \min\{\delta_1, \delta_2\}$ and substitute $t = g(x)$ and $c = L$. Then

$$|x - x_0| < \delta$$

implies

$$|g(x) - L| < \delta_1$$

and thus

$$|f(g(x)) - f(L)| < \epsilon.$$

Therefore the limit of left dense points exists and $f \circ g$ is right dense continuous. The proof for regulated is similar.

(iii) Assume f is right dense continuous. Then by the definition of right dense continuous we know that all of its left sided limits exist and are finite. So we only need to show that the right sided limits of right dense points exist and are finite. Take $L = f(x_0)$ and let x_0 be a right dense point then; f is continuous at x_0 . Thus for a point x to the right of x_0 for any $\epsilon > 0$, there exists a $\delta > 0$ such that when $|x - x_0| < \delta$ we have

$$|f(x) - L| = |f(x) - f(x_0)| < \epsilon.$$

Thus the limit at right dense points exists and is finite.

□

1.2 DIFFERENTIATION

Definition 1.13 A function $f : \mathbb{T} \rightarrow \mathbb{R}$ is said to be Δ -differentiable at $t \in \mathbb{T}^\kappa$ provided

$$f^\Delta(t) := \lim_{s \rightarrow t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s}, \quad s \in \mathbb{T} \setminus \{\sigma(t)\}$$

exists. We call $f^\Delta(t)$ the Δ -derivative of f at t and say that the function f is Δ -differentiable on $\mathbb{T}^\kappa$ if the limit exists and is finite for all $t \in \mathbb{T}^\kappa$. $f^\Delta : \mathbb{T}^\kappa \rightarrow \mathbb{R}$ is called the Δ -derivative of f on $\mathbb{T}^\kappa$.

This definition of the Δ -derivative is similar to the usual definition of the derivative. When a point is right scattered, it is the slope of the line through the points $(t, f(t))$ and $(\sigma(t), f(\sigma(t)))$. There exists a backwards derivative called the ∇ -derivative.

Example 1.14 Consider the function $f : \mathbb{T} \rightarrow \mathbb{R}$ defined by $f(t) = t^2$ for all $t \in \mathbb{T}$ then for $\sigma(t) \neq s$,

$$\begin{aligned} f^\Delta(t) &= \lim_{s \rightarrow t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{(\sigma(t))^2 - s^2}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{(\sigma(t) - s)(\sigma(t) + s)}{\sigma(t) - s} \\ &= \sigma(t) + t. \end{aligned}$$

We should point out that the definition stipulates s cannot equal $\sigma(t)$ but s can equal t . In addition when $\mathbb{T} = \mathbb{R}$ our solution checks with the usual definition of derivative. However when $t \in \mathbb{Z}$, $f^\Delta(t) = 2t + 1$, indicating that polynomials on time scales are more complicated than one might expect.

Theorem 1.15. *Let $f : \mathbb{T} \rightarrow \mathbb{R}$ and let $t \in \mathbb{T}^\kappa$. If f is Δ -differentiable at t , then:*

$$(i) \quad f(\sigma(t)) = f(t) + \mu(t)f^\Delta(t).$$

$$(ii) \quad f \text{ is continuous at } t.$$

Proof. (i) Let us assume f is differentiable at a point $t \in \mathbb{T}$. When t is right dense $\mu(t) = 0$ and $\sigma(t) = t$, thus

$$f(\sigma(t)) = f(t) = f(t) + \mu(t)f^\Delta(t).$$

Now, consider the case when t is right scattered. Then, $\sigma(t) \neq t$ and f Delta-differentiable at t , and the derivative at t is

$$\begin{aligned} f^\Delta(t) &= \frac{f(\sigma(t)) - f(t)}{\sigma(t) - t} \\ &= \frac{f(\sigma(t)) - f(t)}{\mu(t)}. \end{aligned}$$

So

$$f(\sigma(t)) - f(t) = \mu(t)f^\Delta(t)$$

or

$$f(\sigma(t)) = f(t) + \mu(t)f^\Delta(t).$$

(ii) First note that for any $s \in \mathbb{T}^\kappa$ we can rewrite

$$\sigma(t) - s = (\sigma(t) - t) + (t - s) = \mu(t) + (t - s). \quad (1.1)$$

Let $0 < \epsilon < 1$, and define $\epsilon^* = \epsilon[1 + |f^\Delta(t)| + \mu(t)]^{-1}$. Then $0 < \epsilon^* < 1$. Now by the definition of the Δ -derivative, given $0 < \epsilon < 1$ there exists a $\delta > 0$ such that when $|t - s| < \delta$, $s \neq \sigma(t)$, we have

$$\left| \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} - f^\Delta(t) \right| < \epsilon^*.$$

This implies

$$\left| \frac{f(\sigma(t)) - f(s) - (\sigma(t) - s)f^\Delta(t)}{\sigma(t) - s} \right| < \epsilon^*,$$

and so

$$|f(\sigma(t)) - f(s) - (\sigma(t) - s)f^\Delta(t)| < \epsilon^*|\sigma(t) - s|. \quad (1.2)$$

Next, we will use (1.1) and (1.2) to show $|f(t) - f(s)| < \epsilon$. Let $|t - s| < \min\{\epsilon^*, \delta\}$.

Then

$$\begin{aligned} |f(t) - f(s)| &= |f(t) - f(s) + f(\sigma(t)) - f(\sigma(t)) + f^\Delta(t)(\sigma(t) - s) \\ &\quad - f^\Delta(t)(\sigma(t) - s)|. \end{aligned}$$

Using (1.1) this can be rewritten as

$$\begin{aligned} &|f(t) - f(s) + f(\sigma(t)) - f(\sigma(t)) + f^\Delta(t)[\mu(t) + (t - s)] - f^\Delta(t)(\sigma(t) - s)| \\ &= |f(t) - f(s) + f(\sigma(t)) - f(\sigma(t)) + f^\Delta(t)\mu(t) + f^\Delta(t)(t - s) \\ &\quad - f^\Delta(t)(\sigma(t) - s)|. \end{aligned}$$

By (i) we have $f(\sigma(t)) - f(t) - \mu(t)f^\Delta(t) = 0$. Therefore

$$\begin{aligned}
& |\{f(\sigma(t)) - f(s) - f^\Delta(t)(\sigma(t) - s)\} + f^\Delta(t)(t - s)| \\
& \leq \epsilon^*|\sigma(t) - s| + |t - s||f^\Delta(t)| \\
& = \epsilon^*|\mu(t) + (t - s)| + |t - s||f^\Delta(t)| \\
& \leq \epsilon^*\mu(t) + \epsilon'|t - s| + |t - s||f^\Delta(t)| \\
& < \epsilon^*\mu(t) + \epsilon^* + \epsilon^*|f^\Delta(t)| \\
& = \epsilon^*[1 + |f^\Delta(t)| + \mu(t)] \\
& = \epsilon.
\end{aligned}$$

Hence $|f(t) - f(s)| < \epsilon$. □

Theorem 1.15 clearly reduces to the standard results in the real case. We should note that (i) is useful if t is right scattered. If t is right dense, it only says $f(t) = f(t)$. Next, we will develop some rules for Δ -differentiation.

Theorem 1.16. *Let $f, g : \mathbb{T} \rightarrow \mathbb{R}$ be Δ -differentiable at $t \in \mathbb{T}^\kappa$. Then the following hold:*

(i) *The sum $f + g : \mathbb{T} \rightarrow \mathbb{R}$ is Δ -differentiable at t with*

$$(f + g)^\Delta(t) = f^\Delta(t) + g^\Delta(t).$$

(ii) *For any constant $\alpha \in \mathbb{R}$, $\alpha f : \mathbb{T} \rightarrow \mathbb{R}$ is Δ -differentiable at t with*

$$(\alpha f)^\Delta(t) = \alpha f^\Delta(t).$$

(iii) *The product $fg : \mathbb{T} \rightarrow \mathbb{R}$ is Δ -differentiable at t with*

$$(fg)^\Delta(t) = f^\Delta(t)g(\sigma(t)) + f(t)g^\Delta(t) = f^\Delta(t)g(t) + f(\sigma(t))g^\Delta(t).$$

(iv) *If $f(t)f(\sigma(t)) \neq 0$, then $\frac{1}{f}$ is differentiable at t and*

$$\left(\frac{1}{f}\right)^\Delta(t) = -\frac{f^\Delta(t)}{f(t)f(\sigma(t))}$$

(v) If $g(t)g(\sigma(t)) \neq 0$ then $\frac{f}{g}$ is differentiable at t and

$$\left(\frac{f}{g}\right)^\Delta(t) = \frac{f^\Delta(t)g(t) - f(t)g^\Delta(t)}{g(t)g(\sigma(t))}$$

Proof. (i) Let $f, g : \mathbb{T} \rightarrow \mathbb{R}$

$$\begin{aligned} (f+g)^\Delta(t) &= \lim_{s \rightarrow t} \frac{(f+g)(\sigma(t)) - (f+g)(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{f(\sigma(t)) + g(\sigma(t)) - f(s) - g(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} + \lim_{s \rightarrow t} \frac{g(\sigma(t)) - g(s)}{\sigma(t) - s} \\ &= f^\Delta(t) + g^\Delta(t). \end{aligned}$$

Therefore $(f+g)^\Delta(t) = f^\Delta(t) + g^\Delta(t)$.

(ii) Let $\alpha \in \mathbb{R}$ and consider

$$\begin{aligned} (\alpha f)^\Delta(t) &= \lim_{s \rightarrow t} \frac{\alpha f(\sigma(t)) - \alpha f(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{\alpha(f(\sigma(t)) - f(s))}{\sigma(t) - s} \\ &= \alpha \lim_{s \rightarrow t} \frac{(f(\sigma(t)) - f(s))}{\sigma(t) - s} \\ &= \alpha f^\Delta(t). \end{aligned}$$

Thus for any constant $\alpha \in \mathbb{R}$, $(\alpha f)^\Delta(t) = \alpha f^\Delta(t)$.

(iii) Let $f, g : \mathbb{T} \rightarrow \mathbb{R}$

$$\begin{aligned} (fg)^\Delta(t) &= \lim_{s \rightarrow t} \frac{(fg)(\sigma(t)) - (fg)(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{f(\sigma(t))g(\sigma(t)) - f(s)g(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} \frac{f(\sigma(t))g(\sigma(t)) - f(\sigma(t))g(s) + f(\sigma(t))g(s) - f(s)g(s)}{\sigma(t) - s} \\ &= \lim_{s \rightarrow t} f(\sigma(t)) \frac{g(\sigma(t)) - g(s)}{\sigma(t) - s} + \lim_{s \rightarrow t} g(s) \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} \\ &= g(t)f^\Delta(t) + f(\sigma(t))g^\Delta(t). \end{aligned}$$

Thus $(fg)^\Delta(t) = f^\Delta(t)g(t) + f(\sigma(t))g^\Delta(t)$. To get the other equality we would add and subtract $f(s)g(\sigma(t))$ instead of $f(\sigma(t))g(s)$.

(iv) By definition

$$\begin{aligned}
 \left(\frac{1}{f}\right)^\Delta(t) &= \lim_{s \rightarrow t} \frac{\frac{1}{f(\sigma(t))} - \frac{1}{f(s)}}{\sigma(t) - s} \\
 &= \lim_{s \rightarrow t} \frac{\frac{f(s) - f(\sigma(t))}{f(s)f(\sigma(t))}}{\sigma(t) - s} \\
 &= \lim_{s \rightarrow t} \left(\frac{-(f(\sigma(t)) - f(s))}{\sigma(t) - s} \right) \left(\frac{1}{f(s)f(\sigma(t))} \right) \\
 &= -\frac{f^\Delta}{f(t)f(\sigma(t))}
 \end{aligned}$$

which exists as long as $f(t)f(\sigma(t)) \neq 0$.

(v) Now by (ii) and (iv)

$$\begin{aligned}
 \left(\frac{f}{g}\right)^\Delta(t) &= \left(f \cdot \frac{1}{g}\right)^\Delta(t) \\
 &= f(t) \left(\frac{1}{g}\right)^\Delta(t) + f^\Delta(t) \frac{1}{g(\sigma(t))} \\
 &= -f(t) \frac{g^\Delta(t)}{g(t)g(\sigma(t))} + f^\Delta(t) \frac{1}{g(\sigma(t))} \\
 &= \frac{f^\Delta(t)g(t) - f(t)g^\Delta(t)}{g(t)g(\sigma(t))}
 \end{aligned}$$

which exists as long as $g(t)g(\sigma(t)) \neq 0$. □

We should observe that (i) and (ii) are the same as the real case. Although (iii) is a little different it still reduces to the real case since for the real case $\sigma(t) = t$.

At this time we would like to develop the chain rule. Recall, for functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ the chain rule is $(f \circ g)'(t) = f'(g(t))g'(t)$. However, for arbitrary time scales no such rule exists. We will see this in the following example from Bohner and Peterson [1].

Example 1.17 Let $\mathbb{T} = \mathbb{Z}$ and let $f, g : \mathbb{Z} \rightarrow \mathbb{Z}$ be defined by $f(t) = g(t) = t^2 = t \cdot t$. First, we should note that the Δ -derivative is

$$f^\Delta(t) = g^\Delta(t) = t + \sigma(t) = t + t + 1 = 2t + 1.$$

Next observe that

$$(f \circ g)(t) = (t^2)^2 = t^4 = f(t)g(t).$$

Now by using product rule in Theorem 1.16, we get

$$\begin{aligned} (f \circ g)^\Delta(t) &= (f(t)g(t))^\Delta \\ &= f^\Delta(t)g(t) + f(\sigma(t))g^\Delta(t) \\ &= (2t+1)t^2 + (\sigma(t))^2(2t+1) \\ &= (2t+1)t^2 + (t+1)^2(2t+1) \\ &= 2t^3 + t^2 + (t^2 + 2t + 1)(2t+1) \\ &= 2t^3 + t^2 + 2t^3 + 4t^2 + 2t + t^2 + 2t + 1 \\ &= 4t^3 + 6t^2 + 4t + 1. \end{aligned}$$

If we compute by using the usual chain rule, we get

$$\begin{aligned} f^\Delta(g(t))g^\Delta(t) &= (2t^2 + 1)(2t + 1) \\ &= 4t^3 + 2t^2 + 2t + 1. \end{aligned}$$

If we assume that $(f \circ g)^\Delta(t) = f^\Delta(g(t))g^\Delta(t)$, then we have

$$\begin{aligned} 4t^3 + 6t^2 + 4t + 1 &= 4t^3 + 2t^2 + 2t + 1 \\ 4t^2 + 2t &= 0. \end{aligned}$$

Which implies

$$t \in \left\{0, -\frac{1}{2}\right\}.$$

Thus $(f \circ g)^\Delta(t) = f^\Delta(g(t))g^\Delta(t)$ only holds for $t = 0$ in $\mathbb{Z}$.

In Bohner and Peterson [1], there are a few chain rules for time scales, all of which are weaker than the chain rule for real numbers.

Theorem 1.18 (Chain Rule). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be continuously differentiable and suppose $g : \mathbb{T} \rightarrow \mathbb{R}$ is Δ -differentiable. Then $f \circ g : \mathbb{T} \rightarrow \mathbb{R}$ is Δ -differentiable and the following*

formula holds

$$(f \circ g)^\Delta(t) = \left\{ \int_0^1 f'(g(t) + h\mu(t)g^\Delta(t))dh \right\} g^\Delta(t).$$

Proof. First, let us apply the substitution rule from calculus to find

$$\begin{aligned} f(g(\sigma(t))) - f(g(s)) &= \int_{g(s)}^{g(\sigma(t))} f'(\tau) d\tau \\ &= [g(\sigma(t)) - g(s)] \int_0^1 f'(hg(\sigma(t)) + (1-h)g(s)) dh \end{aligned}$$

Let $t \in \mathbb{T}^\kappa$ and $\epsilon > 0$ be given. Define

$$\epsilon^* = \frac{\epsilon}{1 + 2 \int_0^1 |f'(hg(\sigma(t)) + (1-h)g(t))| dh}.$$

The Δ -differentiability of g at t , implies there exists a $\delta_1 > 0$ such that when $|t - s| < \delta_1$

$$|g(\sigma(t)) - g(s) - g^\Delta(t)(\sigma(t) - s)| \leq \epsilon^* |\sigma(t) - s|$$

for all $s \in \mathbb{T}$. In addition, we know f' is continuous on $\mathbb{R}$. In fact we know f' is uniformly continuous on compact subsets of $\mathbb{R}$. By Theorem 1.12, we also know g is continuous and differentiable therefore there exists a $\delta_2 > 0$ such that when $|t - s| < \delta_2$,

$$|f'(hg(\sigma(t)) + (1-h)g(s)) - f'(hg(\sigma(t)) + (1-h)g(t))| \leq \frac{\epsilon}{2(\epsilon^* + |g^\Delta(t)|)}.$$

We should also note that

$$\begin{aligned} &|hg(\sigma(t)) + (1-h)g(s) - (hg(\sigma(t)) + (1-h)g(t))| \\ &= (1-h)|g(s) - g(t)| \\ &\leq |g(s) - g(t)| \end{aligned}$$

holds for all $0 \leq h \leq 1$. Define $\delta = \min\{\delta_1, \delta_2\}$ and for convenience let

$$\alpha = hg(\sigma(t)) + (1-h)g(s)$$

and

$$\beta = hg(\sigma(t)) + (1-h)g(t) = g(t) + h\mu(t)g^\Delta(t).$$

Thus we have

$$\begin{aligned}
& \left| (f \circ g)(\sigma(t)) - (f \circ g)(s) - (\sigma(t) - s)g^\Delta(t) \int_0^1 f'(\beta)dh \right| \\
&= \left| [g(\sigma(t)) - g(s)] \int_0^1 f'(\alpha)dh - (\sigma(t) - s)g^\Delta(t) \int_0^1 f'(\beta)dh \right| \\
&= \left| [g(\sigma(t)) - g(s) - (\sigma(t) - s)g^\Delta(t)] \int_0^1 f'(\alpha)dh + (\sigma(t) - s)g^\Delta(t) \int_0^1 (f'(\alpha) - f'(\beta))dh \right| \\
&\leq |g(\sigma(t)) - g(s) - (\sigma(t) - s)g^\Delta(t)| \int_0^1 |f'(\alpha)|dh + |\sigma(t) - s| |g^\Delta(t)| \int_0^1 |f'(\alpha) - f'(\beta)|dh \\
&\leq \epsilon^* |\sigma(t) - s| \int_0^1 |f'(\alpha)|dh + |\sigma(t) - s| |g^\Delta(t)| \int_0^1 |f'(\alpha) - f'(\beta)|dh \\
&\leq \epsilon^* |\sigma(t) - s| \int_0^1 |f'(\beta)|dh + [\epsilon^* + |g^\Delta(t)|] |\sigma(t) - s| \int_0^1 |f'(\alpha) - f'(\beta)|dh \\
&\leq \frac{\epsilon}{2} |\sigma(t) - s| + \frac{\epsilon}{2} |\sigma(t) - s| \\
&= \epsilon |\sigma(t) - s|,
\end{aligned}$$

thus $f \circ g$ is differentiable at t . The derivative is as claimed in Theorem 1.18. $\square$

1.3 INTEGRATION

As an introduction to integration we will show that antiderivatives exist. To do this we need to discuss pre-differentiable functions and the existence of pre-antiderivatives.

Definition 1.19 A continuous function $f : \mathbb{T} \rightarrow \mathbb{R}$ is called *pre-differentiable* with a region of differentiation D , provided $D \subset \mathbb{T}^\kappa$, $\mathbb{T}^\kappa \setminus D$ is countable and contains no right scattered

elements of $\mathbb{T}$, and f is differentiable at each $t \in D$.

Now, we can develop a Mean Value Theorem for pre-differentiable functions. The Mean Value Theorem will be essential for proving existence and uniqueness theorems for pre-antiderivatives and antiderivatives.

Theorem 1.20 (Mean Value Theorem). *Let f and g be real valued functions defined on $\mathbb{T}$, both pre-differentiable with D . Then*

$$|f^\Delta(t)| \leq g^\Delta(t) \text{ for all } t \in D$$

implies

$$|f(s) - f(r)| \leq g(s) - g(r) \text{ for all } r, s \in \mathbb{T}, r \leq s.$$

Proof. Let $r, s \in \mathbb{T}$ be such that $r \leq s$, and denote $[r, s] \setminus D = \{t_n : n \in \mathbb{N}\}$. Given $\epsilon > 0$, we want to show for all $t \in [r, s]$ the following holds

$$S(t): |f(t) - f(r)| \leq g(t) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right).$$

This is done by way of the Induction Principle, Theorem 1.6.

(i) $S(r)$ is trivial.

$$\begin{aligned} S(r): |f(r) - f(r)| &= 0 \\ &\leq g(r) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\ &= \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right). \end{aligned}$$

(ii) Let t be right scattered and assume that $S(t)$ holds. Then for $t \in D$

$$\begin{aligned}
|f(\sigma(t)) - f(r)| &= |f(t) + \mu(t)f^\Delta(t) - f(r)| \\
&\leq \mu(t)|f^\Delta(t)| + |f(t) - f(r)| \\
&\leq \mu(t)g^\Delta(t) + g(t) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\
&= g(\sigma(t)) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\
&= g(\sigma(t)) - g(r) + \epsilon \left(\sigma(t) - r + \sum_{t_n < \sigma(t)} 2^{-n} \right).
\end{aligned}$$

Therefore $S(\sigma(t))$ is true.

(iii) Suppose $S(t)$ is true and take $t \neq s$ to be right dense. Recall that right dense requires $\sigma(t) = t$. There are two cases to consider: $t \in D$, and $t \notin D$.

First we will consider $t \in D$. Here f and g are differentiable at t , and given $\epsilon > 0$, there exists a $\delta > 0$ such that when $|t - \tau| < \delta$ then

$$|f(t) - f(\tau) - f^\Delta(t)(t - \tau)| \leq \frac{\epsilon}{2}|t - \tau|,$$

and

$$|g(t) - g(\tau) - g^\Delta(t)(t - \tau)| \leq \frac{\epsilon}{2}|t - \tau|.$$

Thus by the triangle inequality we have

$$|f(t) - f(\tau)| \leq \left[f^\Delta(t) + \frac{\epsilon}{2} \right] |t - \tau|$$

and by the definition of absolute value we have

$$g(\tau) - g(t) - g^\Delta(t)(\tau - t) \geq -\frac{\epsilon}{2}|t - \tau|.$$

Thus for any τ such that $|t - \tau| < \delta$,

$$\begin{aligned}
|f(\tau) - f(r)| &\leq |f(\tau) - f(t)| + |f(t) - f(r)| \\
&\leq [|f^\Delta(t)| + \frac{\epsilon}{2}]|t - \tau| + |f(t) - f(r)| \\
&\leq [g^\Delta(t) + \frac{\epsilon}{2}]|t - \tau| + g(t) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\
&= g^\Delta(t)(\tau - t) + \frac{\epsilon}{2}(\tau - t) + g(t) - g(r) + \epsilon(t - r) + \epsilon \sum_{t_n < t} 2^{-n} \\
&\leq g(\tau) - g(t) + \frac{\epsilon}{2}|t - \tau| + \frac{\epsilon}{2}(\tau - t) + g(t) - g(r) + \epsilon(t - r) \\
&\quad + \epsilon \sum_{t_n < t} 2^{-n} \\
&= g(\tau) - g(r) + \epsilon \sum_{t_n < t} 2^{-n} + \epsilon(\tau - r).
\end{aligned}$$

It follows that $S(\tau)$ holds for all τ with $|t - \tau| < \delta$ when $t \in D$.

Next, we want to consider the case when $t \notin D$. If $t \notin D$ then $t = t_m$ for some $m \in \mathbb{N}$.

Again we know that f and g are pre-differentiable, thus they are both continuous and there exists a $\delta > 0$ such that when $|t - \tau| < \delta$ we have

$$|f(\tau) - f(t)| \leq \frac{\epsilon}{2} 2^{-m}$$

and

$$|g(\tau) - g(t)| \leq \frac{\epsilon}{2} 2^{-m}.$$

Thus

$$g(\tau) - g(t) \geq -\frac{\epsilon}{2} 2^{-m},$$

or

$$g(t) \leq g(\tau) + \frac{\epsilon}{2} 2^{-m}.$$

Therefore for $\tau > t$ and $\tau - t < \delta$ we have

$$\begin{aligned}
|f(\tau) - f(r)| &\leq |f(\tau) - f(t)| + |f(t) - f(r)| \\
&\leq \frac{\epsilon}{2}2^{-m} + g(t) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\
&\leq \frac{\epsilon}{2}2^{-m} + g(\tau) + \frac{\epsilon}{2}2^{-m} - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\
&= \epsilon 2^{-m} + g(\tau) - g(r) + \epsilon \left(t - r + \sum_{t_n < t} 2^{-n} \right) \\
&\leq g(\tau) - g(r) + \epsilon \left(\tau - r + \sum_{t_n < \tau} 2^{-n} \right).
\end{aligned}$$

Hence $S(\tau)$ holds for τ with $0 < \tau - t < \delta$ with $t \notin D$.

(iv) Finally let t be a left dense point and suppose $S(\tau)$ holds for all $\tau < t$. Then

$$\begin{aligned}
\lim_{\tau \rightarrow t^-} |f(\tau) - f(r)| &\leq \lim_{\tau \rightarrow t^-} \left\{ g(\tau) - g(r) + \epsilon(t - r + \sum_{t_n < \tau} 2^{-n}) \right\} \\
&\leq \lim_{\tau \rightarrow t^-} \left\{ g(\tau) - g(r) + \epsilon(t - r + \sum_{t_n < t} 2^{-n}) \right\}.
\end{aligned}$$

Thus $S(t)$ is true because both f and g are continuous at t .

Therefore by Theorem 1.6, the Induction Principle, given f and g are real valued, defined on $\mathbb{T}$ and pre-differentiable. The assumption

$$|f^\Delta(t)| \leq g^\Delta(t) \text{ for all } t \in D$$

implies

$$|f(s) - f(r)| \leq g(s) - g(r) \text{ for all } r, s \in \mathbb{T}, r \leq s.$$

□

Corollary 1.21. *Suppose f is pre-differentiable with D . If U is a compact interval with endpoints $r, s \in \mathbb{T}$ then*

$$|f(s) - f(r)| \leq \left\{ \sup_{t \in U^\kappa \cap D} |f^\Delta(t)| \right\} |s - r|.$$

Proof. Take f to be pre-differentiable with D and let $r, s \in \mathbb{T}$ with $r \leq s$. Define

$$g := \left\{ \sup_{\tau \in [r, s]^\kappa \cap D} |f^\Delta(\tau)| \right\} (t - r) \text{ for } t \in \mathbb{T}.$$

Then

$$g^\Delta(t) = \sup_{\tau \in [r, s]^\kappa \cap D} |f^\Delta(\tau)| \geq |f^\Delta(t)| \text{ for all } t \in D \cap [r, s]^\kappa.$$

By the Mean Value Theorem 1.20,

$$g(t) - g(r) \geq |f(t) - f(r)|$$

for all $t \in [r, s]$, therefore

$$|f(s) - f(r)| \leq g(s) - g(r) = g(s) = \left\{ \sup_{t \in U^\kappa \cap D} |f^\Delta(t)| \right\} |s - r|.$$

□

Corollary 1.22. *Suppose f is pre-differentiable with D . If $f^\Delta(t) = 0$ for all $t \in D$ then f is a constant function.*

Proof. By Corollary 1.21

$$|f(s) - f(r)| \leq 0|s - r| = 0$$

which implies

$$f(s) - f(r) = 0.$$

Thus

$$f(s) = f(r)$$

for all $s, t \in \mathbb{T}$. Therefore f is a constant function. □

Theorem 1.23. *Suppose $f_n : \mathbb{T} \rightarrow \mathbb{R}$ is pre-differentiable with D for each $n \in \mathbb{N}$. Assume for each $t \in \mathbb{T}^\kappa$ there exists a compact interval neighborhood $U(t)$ such that the sequence*

$$\{f_n^\Delta\}_{n \in \mathbb{N}} \text{ converges uniformly on } U(t) \cap D.$$

- (i) If $\{f_n\}_{n \in \mathbb{N}}$ converges at some $t_0 \in U(t)$ for some $t \in \mathbb{T}^\kappa$, then it converges uniformly on $U(t)$.
- (ii) If $\{f_n\}_{n \in \mathbb{N}}$ converges at some $t_0 \in \mathbb{T}$, then it converges uniformly on $U(t)$ for all $t \in \mathbb{T}^\kappa$.
- (iii) When it exists, the limit mapping $f = \lim_{n \rightarrow \infty} f_n$ is pre-differentiable with D and we have

$$f^\Delta(t) = \lim_{n \rightarrow \infty} f_n(t) \text{ for all } t \in D.$$

Proof. (i) Let $t \in \mathbb{T}^\kappa$. Then there exists $N \in \mathbb{N}$ such that

$$\sup_{s \in U(t) \cap D} |(f_n - f_m)^\Delta(s)| \text{ is finite for all } m, n \geq N.$$

Suppose $\{f_n(t_0)\}_{n \in \mathbb{N}}$ converges to some $t_0 \in U(t)$. Take $m, n \geq N$ and $r \in U(t)$. Then

$$\begin{aligned} |f_n(r) - f_m(r)| &= |f_n(r) - f_m(r) - [f_n(t_0) - f_m(t_0)] + f_n(t_0) - f_m(t_0)| \\ &\leq |f_n(t_0) - f_m(t_0)| + |(f_n - f_m)(r) - (f_n - f_m)(t_0)| \\ &\leq |f_n(t_0) - f_m(t_0)| + \left\{ \sup_{s \in U(t) \cap D} |(f_n - f_m)^\Delta(s)| \right\} |t_0 - r| \end{aligned}$$

by Corollary 1.21. Thus $\{f_n\}_{n \in \mathbb{N}}$ converges uniformly on $U(t)$. Therefore the sequence $\{f_n\}_{n \in \mathbb{N}}$ is a locally uniformly convergent sequence.

- (ii) Assume $\{f_n(t_0)\}_{n \in \mathbb{N}}$ converges for some $t_0 \in \mathbb{T}$. Now we use the Induction Principle, Theorem 1.6, to show that

$$S(t): \quad \{f_n\}_{n \in \mathbb{N}} \text{ converges}$$

is true for all $t \in [t_0, \infty)$.

- (a) $S(t_0)$ holds by the assumption.

(b) Suppose that t is right scattered and that $S(t)$ is true. Then

$$f_n(\sigma(t)) = f_n(t) + \mu(t)f_n^\Delta(t)$$

converges by our assumption and thus $S(\sigma(t))$ is also true.

(c) Suppose t is right dense and that $S(t)$ holds. Then by (i) $\{f_n\}_{n \in \mathbb{N}}$ converges uniformly on $U(t)$. Thus $S(r)$ holds for all $r \in U(t) \cap (t, \infty)$.

(d) Now, suppose that t is left dense and assume $S(t)$ is true for all $t_0 \leq r < t$. Since $U(t) \cap [t_0, \infty) \neq \emptyset$ then by (a) $\{f_n\}_{n \in \mathbb{N}}$ converges uniformly on $U(t)$. Thus $S(t)$ is true.

Therefore $S(t)$ holds for all $t \in [t_0, \infty)$. Similarly by the dual version of the Induction Principle, Theorem 1.7, $S(t)$ is true for all $t \in (-\infty, t_0)$. Part (a) has already been shown and parts (c) and (d) follow from (i). One should note for (b) if we have a left scattered t then

$$f_n(\rho(t)) = f_n(t) - \mu(\rho(t))f_n^\Delta(\rho(t)).$$

(iii) Let $t \in D$ without loss of generality assume $\sigma(t) \in U(t)$. Given $\epsilon > 0$, by the proof of (i) we know there exists a $N \in \mathbb{N}$ such that

$$|(f_n - f_m)(r) - (f_n - f_m)(\sigma(t))| \leq \left\{ \sup_{s \in U(t) \cap D} |f_n - f_m)^\Delta(s)| \right\} |\sigma(t) - r|$$

for all $r \in U(t)$ and all $m, n \geq N$. Since we know $\{f_n^\Delta\}_{n \in \mathbb{N}}$ converges uniformly on $U(t) \cap D$ there exists $\tilde{N} \geq N$ such that

$$\sup_{s \in U(t) \cap D} |(f_n - f_m)^\Delta(s)| \leq \frac{\epsilon}{3} \text{ for all } m, n \geq \tilde{N}.$$

Thus,

$$|(f_n - f_m)(r) - (f_n - f_m)(\sigma(t))| \leq \frac{\epsilon}{3} |\sigma(t) - r|$$

for all $r \in U(t)$ and all $m, n \geq \tilde{N}$. Let $m \rightarrow \infty$ then

$$|(f_n - f)(r) - f_n - f)(\sigma(t))| \leq \frac{\epsilon}{3} |\sigma(t) - r|$$

for all $r \in U(t)$ and all $n \geq \tilde{N}$. Let

$$g = \lim_{n \rightarrow \infty} f_n^\Delta.$$

Then there exists $M \geq \tilde{N}$ such that

$$|f_M^\Delta(\sigma(t)) - g(\sigma(t))| \leq \frac{\epsilon}{3}.$$

We know f_M is differentiable at t , so there exists a neighborhood W of t such that

$$|f_M(\sigma(t)) - f_M(r) - f_M^\Delta(t)(\sigma(t) - r)| \leq \frac{\epsilon}{3} |\sigma(t) - r| \text{ for all } r \in W.$$

Now, for all $r \in U(t) \cap W$

$$\begin{aligned} & |f(\sigma(t)) - f(r) - g(t)|\sigma - r| \\ & \leq |(f_M - f)(\sigma(t)) - f_m - f)(r)| + |[f_M^\Delta(t) - g(t)](\sigma(t) - r)| \\ & \quad |f_M(\sigma(t)) - f_M(r) - f_M^\Delta(t)(\sigma(t) - r)| \\ & \leq \frac{\epsilon}{3} |\sigma(t) - r| + \frac{\epsilon}{3} |\sigma(t) - r| + \frac{\epsilon}{3} |\sigma(t) - r| \\ & = \epsilon |\sigma(t) - r|. \end{aligned}$$

Thus f is differentiable at t with $f^\Delta(t) = g(t)$. □

Definition 1.24 Assume $f : \mathbb{T} \rightarrow \mathbb{R}$ is a regulated function. Any function F such that

$$F^\Delta(t) = f(t) \text{ for all } t \in D$$

is called the *pre-antiderivative* of f and D is the region of differentiation.

Theorem 1.25 (Existence and Uniqueness of Pre-Antiderivatives). *Let $t_0 \in \mathbb{T}$, $x_0 \in \mathbb{R}$, and a regulated map $f : \mathbb{T}^\kappa \rightarrow \mathbb{R}$ be given. Then there exists exactly one pre-differentiable function (F, D) satisfying*

$$F^\Delta(t) = f(t) \quad \text{for all } t \in D, \quad F(t_0) = x_0.$$

Proof. Let $n \in \mathbb{N}$. By Theorem 1.6, the Induction Principle, we can show

$$S(t) : \begin{cases} \text{There exists a pre-differentiable } (F_{nt}, D_{nt}), F_{nt} : [t_0, t] \rightarrow \mathbb{R} \text{ with} \\ F_{nt}(t_0) = x_0 \text{ and } |F_{nt}^\Delta(s) - f(s)| \leq \frac{1}{n} \text{ for } s \in D_{nt} \end{cases}$$

for all $t \in [t_0, \infty)$.

Case 1: Let $D_{nt_0} = \emptyset$ and $F_{nt_0}(t_0) = x_0$. Then by (i) of the Induction Principle $S(t_0)$ is true.

Case 2: Let t is right-scattered and $S(t)$ be true. Define

$$D_{n\sigma(t)} = D_{nt} \cup \{t\}$$

and for $F_{n\sigma(t)}$ on $[t_0, \sigma(t)]$ define

$$F_{n\sigma(t)}(s) = \begin{cases} F_{nt}(s) & \text{if } s \in D_{nt} \\ F_{nt}(t) + \mu(t)f(t) & \text{if } s = \sigma(t) \end{cases}.$$

Then $F_{n\sigma(t)}(t_0) = F_{nt} = x_0$,

$$|F_{n\sigma(t)}^\Delta(s) - f(s)| = |F_{nt}^\Delta(s) - f(s)| \leq \frac{1}{n} \text{ for } s \in D_{nt}$$

and

$$\begin{aligned} |F_{n\sigma(t)}^\Delta(t) - f(t)| &= \left| \frac{F_{n\sigma(t)}(\sigma(t)) - F_{n\sigma(t)}(t)}{\mu(t)} - f(t) \right| \\ &= \left| \frac{F_{nt} + \mu(t)f(t) - F_{n\sigma(t)}(t)}{\mu(t)} - f(t) \right| \\ &= \left| \frac{\mu(t)f(t)}{\mu(t)} - f(t) \right| \\ &= 0 \\ &\leq \frac{1}{n}. \end{aligned}$$

Therefore $S(\sigma(t))$ holds.

Case 3: Suppose t is right dense and $S(t)$ is true. We know

$$f(t^+) = \lim_{s \rightarrow t, s > t} f(s)$$

exists because t is right dense and f is regulated. Hence there exists a δ such that

$$|f(s) - f(t^+)| \leq \frac{1}{n} \text{ for all } s \in (t, t + \delta]. \quad (1.3)$$

Now, let $r \in (t, t + \delta]$ and define

$$D_{nr} = [D_{nt} \setminus \{t\}] \cup (t, r]^\kappa.$$

Define F_{nr} on $[t_0, r]$ by

$$F_{nr}(s) = \begin{cases} F_{nt}(s) & \text{if } s \in [t_0, t] \\ F_{nt}(t) + f(t^+)(s - t) & \text{if } s \in (t, r] \end{cases}.$$

F_{nr} is continuous at t and thus continuous on $[t_0, r]$. Also note that F_{nr} is differentiable on $(t, r]^\kappa$ with

$$F_{nr}^\Delta \equiv f(t^+) \quad s \in (t, r].$$

Thus F_{nr} is pre-differentiable on $[t_0, t)$. So, F_{nr} is pre-differentiable with D_{nr} because t is right dense. Now, by $S(t)$ and (1.3) we have

$$|F_{nr}^\Delta - f(s)| \leq \frac{1}{n} \text{ for all } s \in D_{nr}.$$

Thus $S(r)$ is true for all $r \in (t, t + \delta]$.

Case 4: Finally, suppose t is left dense and $S(r)$ is true for all $r < t$. Since f is regulated we know

$$f(t^-) = \lim_{s \rightarrow t, s < t} f(s)$$

exists. Hence there exists a $\delta > 0$ such that

$$|f(s) - f(t^-)| \leq \frac{1}{n} \text{ for all } s \in [t - \delta, t). \quad (1.4)$$

Now, fix $r \in [t - \delta, t)$ and define

$$D_{nt} = \begin{cases} D_{nr} \cup (r, t) & \text{if } r \text{ is right dense} \\ D_{nr} \cup [r, t) & \text{if } r \text{ is right scattered} \end{cases}$$

and F_{nt} on $[t_0, r]$ by

$$F_{nt}(s) = \begin{cases} F_{nr}(s) & \text{if } s \in [t_0, r] \\ F_{nr}(t) + f(t^-)(s - t) & \text{if } s \in (r, t] \end{cases}.$$

Similar to the previous cases, F_{nt} is continuous with respect to r and thus it is continuous on $[t_0, t]$. We should note F_{nt} might not be differentiable at r if r is also right dense. However we do know

$$F_{nt}^\Delta \equiv f(t^-) \text{ for all } s \in (r, t].$$

Hence F_{nt} is pre-differentiable with D_{nt} . $S(t)$ and (1.4) imply

$$|F_{nt}^\Delta - f(s)| \leq \frac{1}{n} \text{ for all } s \in D_{nt}.$$

Thus $S(t)$ holds.

Thus by the Induction Principle, Theorem 1.6, $S(t)$ is true for all $t \geq t_0$ and $t \in \mathbb{T}$. Similarly $S(t)$ is valid for $t \leq t_0$. Then F_n is pre-differentiable with D_n , $F_n(t_0) = x_0$, and

$$|F_n^\Delta - f(t)| \leq \frac{1}{n} \text{ for all } t \in D_n.$$

Let

$$F = \lim_{n \rightarrow \infty} F_n \quad \text{and} \quad D = \bigcap_{n \in \mathbb{N}} D_n.$$

Then $F(t_0) = x_0$, F is pre-differentiable on D and

$$F^\Delta(t) = \lim_{n \rightarrow \infty} F_n^\Delta(t) = f(t) \text{ for all } t \in D$$

by Theorem 1.23. □

Definition 1.26 A function $F : \mathbb{T} \rightarrow \mathbb{R}$ is called the *antiderivative* of $f : \mathbb{T} \rightarrow \mathbb{R}$ provided

$$F^\Delta(t) = f(t) \text{ holds for all } t \in \mathbb{T}^\kappa.$$

Theorem 1.27 (Existence of Antiderivatives). *Every right dense continuous function has an antiderivative.*

Proof. Suppose $f \in C_{rd}$. Then by Theorem 1.12 (iii), f is regulated. By Theorem 1.25 we know there exists a function F and a region of differentiation D such that

$$F^\Delta(t) = f(t) \text{ for all } t \in D.$$

We want to show that $F^\Delta(t) = f(t)$ for all $t \in \mathbb{T}^\kappa$. Let $t \in \mathbb{T}^\kappa \setminus D$. Then t is right dense by Definition 1.3. Since f is right dense continuous at t it is continuous at t . Given $\epsilon > 0$ there exists a $\delta > 0$ such that when $|t - s| < \delta$ then

$$|f(s) - f(t)| \leq \epsilon.$$

Define for $\tau \in \mathbb{T}$

$$h(\tau) := F(\tau) - f(t)(\tau - t_0).$$

Therefore, h is pre-differentiable on D with

$$h^\Delta(\tau) = F^\Delta(\tau) - f(t) = f(\tau) - f(t) \text{ for all } \tau \in D.$$

Hence for $s \in D$ and $|t - s| < \delta$ we have

$$|h^\Delta(s)| = |f(s) - f(t)| \leq \epsilon.$$

Thus

$$\sup_{s \in D \cap U} |h^\Delta(s)| \leq \epsilon.$$

Now, by Corollary 1.21, for any r such that $|t - r| < \delta$,

$$\begin{aligned}
 |F(t) - F(r) - f(t)(t - r)| &= |h(t) + f(t)(t - t_0) - [h(r) - f(t)(r - t_0)] - f(t)(t - r)| \\
 &= |h(t) - h(r)| \\
 &\leq \left\{ \sup_{s \in D \cap U} |h^\Delta(s)| \right\} |t - r| \\
 &\leq \epsilon |t - r|.
 \end{aligned}$$

Thus F is differentiable at t with $F^\Delta(t) = f(t)$. □

With the above theorem we can introduce indefinite integrals.

Definition 1.28 Assume $f : \mathbb{T} \rightarrow \mathbb{R}$ is a regulated function. Assume F is the pre-antiderivative of f . We define the *indefinite integral* of a regulated function f by

$$\int f(t) \Delta t = F(t) + C$$

where C is an arbitrary constant and F is the pre-antiderivative of f .

The theorem below gives a useful identity that we will need later.

Theorem 1.29. *If $f \in C_{rd}$ and $t \in \mathbb{T}$, then*

$$\int_t^{\sigma(t)} f(\tau) \Delta \tau = \mu(t) f(t).$$

Proof. We know there exists antiderivative F of f by Theorem 1.27, such that

$$\begin{aligned}
 \int_t^{\sigma(t)} f(\tau) \Delta \tau &= F(\sigma(t)) - F(t) \\
 &= F(t) + \mu(t) F^\Delta(t) - F(t) \\
 &= \mu(t) F^\Delta(t) \\
 &= \mu(t) f(t),
 \end{aligned}$$

where in the second line we use the identity from Theorem 1.15. □

We will need the following theorem to prove integration by parts for time scales.

Theorem 1.30. *If $a, b \in \mathbb{T}$ and $f, g \in C_{rd}$ then*

$$\int_a^b [f(t) + g(t)] \Delta t = \int_a^b f(t) \Delta t + \int_a^b g(t) \Delta t.$$

Proof. By Theorem 1.16 (i), $F + G$ is an antiderivative of $f + g$, therefore

$$\begin{aligned} \int_a^b (f + g)(t) \Delta t &= (F + G)(b) - (F + G)(a) \\ &= F(b) - F(a) + G(b) - G(a) \\ &= \int_a^b f(t) \Delta t + \int_a^b g(t) \Delta t. \end{aligned}$$

□

Theorem 1.31. *If $a, b, c \in \mathbb{T}$ and $f, g \in C_{rd}$ then*

$$\int_a^b f(t) \Delta t = \int_a^c f(t) \Delta t + \int_c^b f(t) \Delta t$$

Proof.

$$\begin{aligned} \int_a^b g(t) \Delta t &= F(b) - F(a) \\ &= F(c) - F(a) + F(b) - F(c) \\ &= \int_a^c f(t) \Delta t + \int_c^b f(t) \Delta t. \end{aligned}$$

□

Next we have integration by part formulas for time scales we will need later for our error estimates.

Theorem 1.32 (Integration by Parts). *Let $a, b \in \mathbb{T}$ and $f, g \in C_{rd}$ then*

$$\int_a^b f(\sigma(t))g^\Delta(t) \Delta t = (fg)(b) - (fg)(a) - \int_a^b f^\Delta(t)g(t) \Delta t$$

or

$$\int_a^b f(t)g^\Delta(t) = (fg)(b) - (fg)(a) - \int_a^b f^\Delta(t)g(\sigma(t))\Delta t.$$

Proof. First we should note that fg is an antiderivative of $f^\sigma g^\Delta + f^\Delta g$, thus

$$\int_a^b (f^\sigma g^\Delta + f^\Delta g)(t)\Delta t = (fg)(b) - (fg)(a).$$

Now by Theorem 1.30

$$\int_a^b (f^\sigma g^\Delta + f^\Delta g)(t)\Delta t = \int_a^b f(\sigma(t))g^\Delta(t) + \int_a^b f^\Delta(t)g(t).$$

Therefore

$$\int_a^b f(\sigma(t))g^\Delta(t) + \int_a^b f^\Delta(t)g(t) = (fg)(b) - (fg)(a)$$

and we get

$$\int_a^b f(\sigma(t))g^\Delta(t)\Delta t = (fg)(b) - (fg)(a) - \int_a^b f^\Delta(t)g(t)\Delta t$$

or

$$\int_a^b f(t)g^\Delta(t) = (fg)(b) - (fg)(a) - \int_a^b f^\Delta(t)g(\sigma(t))\Delta t.$$

□

We need the notion of a *partition* and the *Riemann Δ -integral* to develop a mean value theorem for integrals.

Definition 1.33 Let $\mathbb{T}$ be a time scale, $a < b$ be points in $\mathbb{T}$, and $[a, b]$ the closed and bounded interval is also in $\mathbb{T}$. A *partition* of $[a, b]$ is any finite ordered subset

$$P = \{t_0, t_1, \dots, t_n : t_i \in \mathbb{T} \text{ for } 0 \leq i \leq n\} \text{ where } a = t_0 < t_1 < \dots < t_n = b.$$

A particular partition depends on the number n . We denote the set of all partitions of $[a, b]$ by $\mathcal{P} = \mathcal{P}(a, b)$. $\mathcal{P}_\delta = \mathcal{P}_\delta(a, b)$ denotes the set of all $P \in \mathcal{P}$ such that for every $\delta > 0$ there

exists a partition $P \in \mathcal{P}(a, b)$ given by $a = t_0 < t_1 < \dots < t_n = b$ such that for each $i \in \{1, 2, \dots, n\}$ either

$$t_i - t_{i-1} \leq \delta$$

or

$$t_i - t_{i-1} > \delta \text{ and } \rho(t_i) = t_{i-1}.$$

Definition 1.34 Let f be a bounded function on $[a, b]$ and let $P \in \mathcal{P}(a, b)$ be given by $a = t_0 < t_1 < \dots < t_n = b$. In each interval $[t_{i-1}, t_i)$ with $1 \leq i \leq n$, choose an arbitrary point $\xi_i \in \mathbb{T}$ and form the sum

$$S = \sum_{i=1}^n f(\xi_i)(t_i - t_{i-1}).$$

S is called the *Riemann Δ -sum* of f corresponding to $P \in \mathcal{P}$. We say f is the *Riemann Δ -integral* from a to b if there exists a unique number I such that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|S - I| < \epsilon$$

for every Riemann Δ -sum S of f corresponding to $P \in \mathcal{P}_\delta(a, b)$ and independent of the choice $\xi_i \in [t_{i-1}, t_i)$ for $1 < i < n$. The number I is called the *Riemann Δ -integral* of f from a to b .

A function is considered *Δ -integrable* when the Riemann integral exists. The following few theorems are needed to derive a mean value theorem for integrals.

Theorem 1.35. *Let f and g be Δ -integrable functions on $[a, b]$, and let $c \in \mathbb{R}$. Then*

$$(i) \quad cf \text{ is } \Delta\text{-integrable and } \int_a^b cf(t)\Delta t = c \int_a^b f(t)\Delta t,$$

$$(ii) \quad f + g \text{ is } \Delta\text{-integrable and } \int_a^b c(f + g)(t)\Delta t = \int_a^b f(t)\Delta t + \int_a^b g(t)\Delta t.$$

Proof. (i) If $c = 0$, then the statement is obvious. Let $c \neq 0$ and take any $\epsilon > 0$. By Definition 1.3 there exists a $\delta > 0$ such that

$$\left| S_f - \int_a^b f(t) \Delta t \right| < \frac{\epsilon}{|c|}$$

for every Riemann Δ -sum S_f of f for $P \in \mathcal{P}_\delta(a, b)$. Let S_{cf} denote the Riemann Δ -sum for cf associated with $P = \{a = t_0 < t_1 < \dots < t_n = b\} \in \mathcal{P}_\delta(a, b)$. Since

$$S_{cf} = \sum_{i=1}^n (cf)(\xi_i)(t_i - t_{i-1}) = c \sum_{i=1}^n f(\xi_i)(t_i - t_{i-1}) = cS_f$$

we have

$$\begin{aligned} \left| S_{cf} - c \int_a^b f(t) \Delta t \right| &= \left| cS_f - c \int_a^b f(t) \Delta t \right| \\ &= |c| \left| S_f - \int_a^b f(t) \Delta t \right| \\ &< |c| \frac{\epsilon}{|c|} \\ &= \epsilon. \end{aligned}$$

Thus cf is Δ -integrable with $\int_a^b cf(t) \Delta t = c \int_a^b f(t) \Delta t$.

(ii) Let $\epsilon > 0$ be given. Since f and g are Δ -integrable on $[a, b]$ there exists $\delta > 0$ such that for any $P \in \mathcal{P}(a, b)$,

$$\left| S_f - \int_a^b f(t) \Delta t \right| \leq \frac{\epsilon}{2} \quad \text{and} \quad \left| S_g - \int_a^b g(t) \Delta t \right| \leq \frac{\epsilon}{2}$$

where S_f and S_g are Riemann Δ -sums for f and g respectively, associated with P . We also know

$$\begin{aligned} S_{f+g} &= \sum_{i=1}^n (f+g)(\xi_i)(t_i - t_{i-1}) \\ &= \sum_{i=1}^n (f)(\xi_i)(t_i - t_{i-1}) + \sum_{i=1}^n (g)(\xi_i)(t_i - t_{i-1}) \\ &= S_f + S_g. \end{aligned}$$

Thus

$$\begin{aligned}
 \left| S_{f+g} - \left[\int_a^b f(t) \Delta t + \int_a^b g(t) \Delta t \right] \right| &\leq \left| S_f - \int_a^b f(t) \Delta t \right| + \left| S_g - \int_a^b g(t) \Delta t \right| \\
 &= \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
 &= \epsilon.
 \end{aligned}$$

Therefore $f + g$ is Δ -integrable with $\int_a^b c(f + g)(t) \Delta t = \int_a^b f(t) \Delta t + \int_a^b g(t) \Delta t$. $\square$

The following theorem is need to prove Theorem 1.37. We will not prove it in this paper but the proof can be found in Bohner and Peterson [2].

Theorem 1.36. *Let f be Δ -integrable on $[a, b]$ and let M and m be the supremum and infimum on $[a, b]$, respectively. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a function defined on $[m, M]$ satisfying a Lipschitz condition. Then the composite function $h = \phi \circ f$ is Δ -integrable on $[a, b]$.*

Theorem 1.37. *The product of two Δ -integrable functions is again Δ -integrable.*

Proof. First we should note that if f is Δ -integrable then f^2 is Δ -integrable. Since $f^2 = \phi(f(t))$ with $\phi(t) = x^2$ and ϕ satisfies a Lipschitz condition on any interval $[m, M]$,

$$\begin{aligned}
 |\phi(x) - \phi(y)| &= |x - y||x + y| \\
 &= L|x - y|
 \end{aligned}$$

where $L = \max\{2|m|, 2|M|\}$. Thus by Theorem 1.36 f^2 is Δ -integrable on $[a, b]$. Now by using the identity

$$fg = \frac{(f + g)^2 - (f - g)^2}{4}.$$

Theorem 1.35 implies that the product of any two Δ -integrable functions is also Δ -integrable. $\square$

Theorem 1.38. *if f and g are Δ -integrable on $[a, b]$ and if $f(t) < g(t)$ for all $t \in [a, b]$ then*

$$\int_a^b f(t) \Delta t \leq \int_a^b g(t) \Delta t. \tag{1.5}$$

Proof. By Theorem 1.35, $h = g - f$ is Δ -integrable on $[a, b]$. $h(t) \geq 0$ on $[a, b]$ which implies $L(h, P) \geq 0$ for all $P \in \mathcal{P}(a, b)$. Now Theorem 1.35 implies

$$\int_a^b h(t) \Delta t = \int_a^b f(t) \Delta t - \int_a^b g(t) \Delta t \geq 0$$

and (1.5) follows directly. $\square$

We need the following useful mean value theorems for integrals to bound our error terms in the next chapter.

Theorem 1.39. *Let f and g be bounded and integrable functions on $[a, b]$, and let g be nonnegative (or nonpositive) on $[a, b]$. Let us set*

$$m = \inf \{f(t) : t \in [a, b]\} \text{ and } M = \sup \{f(t) : t \in [a, b]\}.$$

Then there exists a real number Λ satisfying the inequalities $m \leq \Lambda \leq M$ such that

$$\int_a^b f(t)g(t) \Delta t = \Lambda \int_a^b g(t) \Delta t. \quad (1.6)$$

Proof. We are given

$$m \leq f(t) \leq M. \quad (1.7)$$

Suppose $g(t)$ is nonnegative, $g(t) \geq 0$. Multiplying (1.7) by $g(t)$,

$$mg(t) \leq f(t)g(t) \leq Mg(t) \text{ for all } t \in [a, b].$$

By Theorem 1.35 and Theorem 1.37 mg , Mg and fg are all integrable from a to b . Thus by Theorem 1.38

$$m \int_a^b g(t) \Delta t \leq \int_a^b f(t)g(t) \Delta t \leq M \int_a^b g(t) \Delta t. \quad (1.8)$$

If $\int_a^b g(t) \Delta t = 0$, then (1.8) implies $\int_a^b f(t)g(t) \Delta t = 0$ and thus the inequality (1.6) is clear.

If $\int_a^b g(t) \Delta t > 0$ then (1.8) implies

$$m \leq \frac{\int_a^b f(t)g(t) \Delta t}{\int_a^b g(t) \Delta t} \leq M.$$

Then

$$\Lambda = \frac{\int_a^b f(t)g(t)\Delta t}{\int_a^b g(t)\Delta t}$$

satisfies $m \leq \Lambda \leq M$. □

Corollary 1.40. *Let f be an integrable function on $[a, b]$ and let m and M be the infimum and supremum, respectively, of f on $[a, b]$. Then there exists a number Λ between m and M such that*

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

Proof. Let $g(t) = 1$ then Theorem 1.39 implies the above result. □

1.4 POLYNOMIALS AND TAYLOR'S FORMULA

We know 0 is the derivative of 1 and 1 is the derivative of t . We would like a closed formula for the anti-derivative of t for an arbitrary time scale but it is not possible. For instance on $\mathbb{R}$, t is the derivative of $\frac{t^2}{2}$ but for an arbitrary time scale

$$\begin{aligned} \left(\frac{t^2}{2}\right)^\Delta &= \left(\frac{1}{2} \cdot t \cdot t\right)^\Delta \\ &= \frac{1}{2} (t \cdot t^\Delta + t^\Delta \cdot \sigma(t)) \\ &= \frac{t + \sigma(t)}{2}. \end{aligned}$$

Moreover, although $\sigma(t)$ is continuous, $\frac{t + \sigma(t)}{2}$ is not necessarily differentiable, as we will see in the following example.

Example 1.41 Let

$$\mathbb{T} = \left\{ t_n = \left(\frac{1}{2}\right)^{2^n} : n \in \mathbb{N}_0 \right\} \cup \{0, -1\}.$$

Then

$$\sigma(t_n) = t_{n-1} \rightarrow 0 = \sigma(0), \quad n \rightarrow \infty,$$

therefore $\lim_{s \rightarrow 0} \sigma(s) = \sigma(0)$ and thus σ is continuous at 0.

However,

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{\sigma(\sigma(0)) - \sigma(s)}{\sigma(0) - s} &= \lim_{s \rightarrow 0} \frac{\sigma(s)}{s} \\ &= \lim_{s \rightarrow 0} \frac{\sqrt{s}}{s} \\ &= \lim_{s \rightarrow 0} \frac{1}{\sqrt{s}} \\ &= \infty, \end{aligned}$$

so σ is not differentiable at 0 despite being continuous at 0. We should note that this is true regardless of whether 0 is left dense or left scattered.

We want to derive a function that will act like a polynomial in the sense that its Δ -derivative yields the polynomial of one less in degree all the way down the the polynomial of zeroth degree being 1. In time scale calculus, we do this by starting with 1 and recursively define functions by integration.

Definition 1.42 A *time scale polynomial* is defined as

$g_k, h_k : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}$ for $k \in \mathbb{N}_0$ as follows:

$$g_0(t, s) = h_0(t, s) \equiv 1 \text{ for all } s, t \in \mathbb{T}, \quad (1.9)$$

$$g_{k+1}(t, s) = \int_s^t g_k(\sigma(\tau), s) \Delta\tau \text{ for all } s, t \in \mathbb{T}, \quad (1.10)$$

$$h_{k+1}(t, s) = \int_s^t h_k(\tau, s) \Delta\tau \text{ for all } s, t \in \mathbb{T}. \quad (1.11)$$

By Theorem 1.27 and Theorem 1.12 we know the g_k and h_k are well defined. Next we will develop some useful equations in manipulating polynomials.

Theorem 1.43. Let $s \in \mathbb{T}$ and let $h_k^\Delta(t, s)$ denote the Δ -derivative of $h_k(t, s)$ for fixed s . Then the following hold:

$$(i) \quad g_2(t, s) = h_2(s, t) \text{ for all } t \in \mathbb{T}.$$

$$(ii) \quad h_k^\Delta(t, s) = h_{k-1}(t, s) \text{ for all } t \in \mathbb{T}^\kappa, k \in \mathbb{N}.$$

(iii) $g_k^\Delta(t, s) = g_{k-1}(\sigma(t), s)$ for all $t \in \mathbb{T}^\kappa$, $k \in \mathbb{N}$.

(iv) $g_1(t, s) = h_1(t, s) = t - s$ for all $t \in \mathbb{T}$.

(v) $g_n(t, t) = h_n(t, t) = 0$ for all $t \in \mathbb{T}$ and $n > 0$.

Proof. (i) From the definition of $g_2(t, s)$ and (iv),

$$\begin{aligned}
 g_2(t, s) &= \int_s^t (\sigma(\tau) - s) \Delta\tau \\
 &= \int_s^t (\sigma(\tau) + \tau) \Delta\tau - \int_s^t \tau \Delta\tau - \int_s^t s \Delta\tau \\
 &= \int_s^t (\tau^2)^\Delta \Delta\tau + \int_t^s \tau \Delta\tau - s(t - s) \\
 &= \int_s^t \tau \Delta\tau + t^2 - s^2 - s(t - s) \\
 &= \int_s^t (\tau - t) \Delta\tau \\
 &= h_2(s, t).
 \end{aligned}$$

(ii) For a fixed s let $h_k^\Delta(t, s)$ denote the derivative of $h_k(t, s)$ with respect to t then

$$h_k^\Delta(t, s) = h_{k-1}(t, s) \text{ for all } t \in \mathbb{T}^\kappa, k \in \mathbb{N}.$$

(iii) Similar to (ii), fix s let $g_k^\Delta(t, s)$ denote the derivative of $g_k(t, s)$ with respect to t then

$$g_k^\Delta(t, s) = g_{k-1}(\sigma(t), s) \text{ for all } t \in \mathbb{T}^\kappa, k \in \mathbb{N}.$$

(iv) Let $t \in \mathbb{T}$. By Definition 1.4 we have

$$\begin{aligned}
 g_1(t, s) &= \int_s^t g_0(\sigma(\tau), s) \Delta\tau \\
 &= \int_s^t \Delta\tau \\
 &= t - s.
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 h_1(t, s) &= \int_s^t h_k(\tau, s) \Delta\tau \\
 &= \int_s^t \Delta\tau \\
 &= t - s.
 \end{aligned}$$

(v) By Definition 1.4 we have

$$\begin{aligned}
 g_n(t, t) &= \int_t^t g_{n-1}(\sigma(\tau), t) \Delta\tau \\
 &= 0.
 \end{aligned}$$

Similarly

$$\begin{aligned}
 h_n(t, t) &= \int_t^t h_{n-1}(\sigma(\tau), t) \Delta\tau \\
 &= 0.
 \end{aligned}$$

□

Remark 1.44. In (iv) and (v) of Theorem 1.43 if we have a scattered time scale where t and s are neighboring points, say $t = \sigma(s)$, then (iv) yields

$$\begin{aligned}
 h_1(\sigma(s), s) &= \int_s^{\sigma(s)} h_k(\tau, s) \Delta\tau \\
 &= \int_s^{\sigma(s)} \Delta\tau \\
 &= \sigma(s) - s \\
 &= \mu(s).
 \end{aligned}$$

To prove Taylor's Theorem for an arbitrary time scale we need the following results.

Lemma 1.45. Let $n \in \mathbb{N}$. If f is n times differentiable and p_k , $0 \leq k \leq n - 1$, are differentiable at some $t \in \mathbb{T}$ with

$$p_{k+1}^\Delta(t) = p_k^\sigma(t) \text{ for all } k \text{ such that } 0 \leq k \leq n - 2, \quad (1.12)$$

then at t we have

$$\left[\sum_{k=0}^{n-1} (-1)^k f^{\Delta^k} p_k \right]^{\Delta} = (-1)^{n-1} f^{\Delta^n} p_{n-1}^{\sigma} + f p_0^{\Delta}.$$

Proof. By using the Δ -differentiation rules we developed in Theorem 1.16 and (1.12) then

$$\begin{aligned} & \left[\sum_{k=0}^{n-1} (-1)^k f^{\Delta^k} p_k \right]^{\Delta} \\ &= \sum_{k=0}^{n-1} (-1)^k \left[f^{\Delta^k} p_k \right]^{\Delta} \\ &= \sum_{k=0}^{n-1} (-1)^k \left[f^{\Delta^{k+1}} p_k^{\sigma} + f^{\Delta^k} p_k^{\Delta} \right] \\ &= \sum_{k=0}^{n-2} (-1)^k f^{\Delta^{k+1}} p_k^{\sigma} + (-1)^{n-1} f^{\Delta^n} p_{n-1}^{\sigma} + \sum_{k=1}^{n-1} (-1)^k f^{\Delta^k} p_k^{\Delta} + f p_0^{\Delta} \\ &= \sum_{k=0}^{n-2} (-1)^k f^{\Delta^{k+1}} p_k^{\sigma} + (-1)^{n-1} f^{\Delta^n} p_{n-1}^{\sigma} - \sum_{k=0}^{n-2} (-1)^k f^{\Delta^{k+1}} p_{k+1}^{\Delta} + f p_0^{\Delta} \\ &= (-1)^{n-1} f^{\Delta^n} p_{n-1}^{\sigma} + f p_0^{\Delta} \end{aligned}$$

holds at t . □

Lemma 1.46. *The functions g_k defined in (1.10) and (1.11) satisfy for all $t \in \mathbb{T}$*

$$g_n(\rho^k(t), t) = 0 \text{ for all } n \in \mathbb{N} \text{ for all } k \text{ such that } 0 \leq k \leq n-1.$$

Proof. We will prove this lemma by induction. Let $k = 0$ then

$$g_n(\rho^0(t), t) = g_n(t, t) = 0.$$

Now, we want to show

$$g_{n-1}(\rho^k(t), t) = g_n(\rho^k(t), t) = 0 \text{ for } 0 \leq k < n$$

which implies

$$g_n(\rho^{k+1}(t), t) = 0.$$

There are two cases: when $\rho^k(t)$ is left dense and when $\rho^k(t)$ is left scattered.

Case 1: If $\rho^k(t)$ is left dense, then $\rho^{k+1}(t) = \rho^k(t)$ and

$$g_n(\rho^{k+1}(t), t) = g_n(\rho^k(t), t) = 0.$$

Case 2: If $\rho^k(t)$ is left scattered, then $\sigma(\rho^{k+1}(t)) = \rho^k(t)$. Now by applying Theorem 1.15 (i) and (1.11) we have

$$\begin{aligned} g_n(\rho^{k+1}(t), t) &= g_n(\sigma(\rho^{k+1}(t)) - \mu(\rho^{k+1}(t))g_n^\Delta(\rho^{k+1}(t), t) \\ &= g_n(\rho^k(t), t) - \mu(\rho^{k+1}(t))g_{n-1}(\sigma(\rho^{k+1}(t)), t) \\ &= g_n(\rho^k(t), t) - \mu(\rho^{k+1}(t))g_{n-1}(\rho^k(t), t) \\ &= 0. \end{aligned}$$

Thus by induction we have $g_n(\rho^k(t), t) = 0$ for all $n \in \mathbb{N}$ for all k such that $0 \leq k \leq n-1$. $\square$

Lemma 1.47. *Let $n \in \mathbb{N}$, $t \in \mathbb{T}$ and suppose f is $n-1$ times differentiable at $\rho^{n-1}(t)$. Then we have*

$$\sum_{k=0}^{n-1} (-1)^k f^{\Delta^k}(\rho^{n-1}(t)) g_k(\rho^{n-1}(t), t) = f(t), \quad (1.13)$$

where the function g_k are defined by (1.10) and (1.11).

Proof. We will also prove this lemma by induction. Let $n = 0$ and then

$$(-1)^0 f^{\Delta^0}(\rho^0(t)) g_0(\rho^0(t), t) = f(t) g_0(t, t) = f(t).$$

Thus (1.13) holds for $n = 0$.

Suppose that (1.13) holds for some $m \in \mathbb{N}$. Now there are two cases; when $\rho^{m-1}(t)$ is left dense and when $\rho^{m-1}(t)$ is left scattered.

Case 1: Assume $\rho^{m-1}(t)$ is left dense. Then $\rho^m(t) = \rho^{m-1}(t)$ and

$$\begin{aligned}
& \sum_{k=0}^m (-1)^k f^{\Delta^k}(\rho^m(t)) g_k(\rho^m(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k f^{\Delta^k}(\rho^m(t)) g_k(\rho^m(t), t) \\
&\quad + (-1)^m f^{\Delta^m}(\rho^m(t)) g_m(\rho^m(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k f^{\Delta^k}(\rho^{m-1}(t)) g_k(\rho^{m-1}(t), t) \\
&\quad + (-1)^m f^{\Delta^m}(\rho^{m-1}(t)) g_m(\rho^{m-1}(t), t) \\
&= f(t) + (-1)^m f^{\Delta^m}(\rho^{m-1}(t)) g_m(\rho^{m-1}(t), t) \quad \text{by Lemma 1.46} \\
&= f(t).
\end{aligned}$$

Thus for left dense points (1.13) holds for all $n \in \mathbb{N}$.

Case 2: Assume $k \in \mathbb{N}$ and $\rho^{m-1}(t)$ is left scattered then $\sigma(\rho^m(t)) = \rho^{m-1}(t)$. Thus by applying

Theorem 1.15 and (1.11) we have

$$\begin{aligned}
g_k(\rho^{m-q}(t), t) &= g_k(\sigma(\rho^m(t)), t) \\
&= g_k(\rho^m(t), t) + \mu(\rho^m(t)) g_k^\Delta(\rho^m(t), t) \\
&= g_k(\rho^m(t), t) + \mu(\rho^m(t)) g_{k-1}(\sigma(\rho^m(t)), t) \\
&= g_k(\rho^m(t), t) + \mu(\rho^m(t)) g_{k-1}(\rho^{m-1}(t), t).
\end{aligned}$$

By applying Lemma 1.46 to the above we have

$$\begin{aligned}
& \sum_{k=0}^m (-1)^k f^{\Delta^k}(\rho^m(t)) g_k(\rho^m(t), t) \\
&= f(\rho^m(t)) + \sum_{k=1}^m (-1)^k f^{\Delta^k}(\rho^m(t)) g_k(\rho^m(t), t) \\
&= f(\rho^m(t)) \\
&\quad + \sum_{k=1}^m (-1)^k f^{\Delta^k}(\rho^m(t)) [g_k(\rho^{m-1}(t), t) - \mu(\rho^m(t)) g_{k-1}(\rho^{m-1}(t), t)] \\
&= f(\rho^m(t)) + \sum_{k=1}^{m-1} (-1)^k f^{\Delta^k}(\rho^m(t)) g_k(\rho^{m-1}(t), t) \\
&\quad + \sum_{k=1}^m (-1)^{k-1} \mu(\rho^m(t)) f^{\Delta^k}(\rho^m(t)) g_{k-1}(\rho^{m-1}(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k f^{\Delta^k}(\rho^m(t)) g_k(\rho^{m-1}(t), t) \\
&\quad + \sum_{k=0}^{m-1} (-1)^k \mu(\rho^m(t)) f^{\Delta^{k+1}}(\rho^m(t)) g_k(\rho^{m-1}(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k \left[f^{\Delta^k}(\rho^m(t)) + \mu(\rho^m(t)) (f^{\Delta^k})^\Delta(\rho^m(t)) \right] g_k(\rho^{m-1}(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k (f^{\Delta^k})^\sigma(\rho^m(t)) g_k(\rho^{m-1}(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k (f^{\Delta^k})(\sigma(\rho^m(t))) g_k(\rho^{m-1}(t), t) \\
&= \sum_{k=0}^{m-1} (-1)^k (f^{\Delta^k})(\rho^{m-1}(t)) g_k(\rho^{m-1}(t), t) \\
&= f(t).
\end{aligned}$$

Thus for left scattered points (1.13) holds for all $n \in \mathbb{N}$.

Therefore we have proved Lemma 1.47 by the principle of mathematical induction. $\square$

Now we are ready for our first version of Taylor's Formula.

Theorem 1.48 (Taylor's Formula). *Let $n \in \mathbb{N}$. Suppose f is n times differentiable on $\mathbb{T}^{k^n}$. Let $\alpha \in \mathbb{T}^{k^{n-1}}$, $t \in \mathbb{T}$, and define the functions g_k by (1.10) and (1.11). Then we have*

$$f(t) = \sum_{k=0}^{n-1} (-1)^k g_k(\alpha, t) f^{\Delta^k}(\alpha) + \int_{\alpha}^{\rho^{n-1}(t)} (-1)^{n-1} g_{n-1}(\sigma(\tau), t) f^{\Delta^n}(\tau) \Delta\tau.$$

Proof. By Lemma 1.45 we have

$$\left[\sum_{k=0}^{n-1} (-1)^k g_k(\cdot, t) f^{\Delta^k} \right]^{\Delta}(\tau) = (-1)^{n-1} g_{n-1}(\sigma(\tau), t) f^{\Delta^n}(\tau)$$

for all $\tau \in \mathbb{T}^{k^n}$. We know $\alpha, \rho^{n-1}(t) \in \mathbb{T}^{k^{n-1}}$, so we can integrate from α to ρ^{n-1} . Thus

$$\begin{aligned} & \int_{\alpha}^{\rho^{n-1}(t)} (-1)^{n-1} g_{n-1}(\sigma(\tau), t) f^{\Delta^n}(\tau) \Delta\tau \\ &= \int_{\alpha}^{\rho^{n-1}(t)} \left[\sum_{k=0}^{n-1} (-1)^k g_k(\cdot, t) f^{\Delta^k} \right]^{\Delta}(\tau) \Delta\tau \\ &= \sum_{k=0}^{n-1} (-1)^k g_k(\rho^{n-1}(t), t) f^{\Delta^k}(\rho^{n-1}(t)) - \sum_{k=0}^{n-1} (-1)^k g_k(\alpha, t) f^{\Delta^k}(\alpha) \\ &= f(t) - \sum_{k=0}^{n-1} (-1)^k g_k(\alpha, t) f^{\Delta^k}(\alpha) \end{aligned}$$

by Lemma 1.47. □

Now, we can use our first version of Taylor's formula, Theorem 1.48 to generate another Taylor's Formula for h_k instead of g_k . First we need to derive another relation between h_k and g_k .

Theorem 1.49. *The functions g_k and h_k defined by (1.10), (1.11) and (1.11) satisfy*

$$h_n(t, s) = (-1)^n g_n(s, t) \text{ for all } t \in \mathbb{T} \text{ and } s \in \mathbb{T}^{k^n}.$$

Proof. Let $t \in \mathbb{T}$ and $s \in \mathbb{T}^{k^n}$. Take

$$\alpha = s \quad \text{and} \quad f = h_n(\cdot, s).$$

By Theorem 1.48 we have

$$f^\Delta = h_{n-1}(\cdot, s),$$

and by (1.11) we have

$$f^{\Delta^k}(s) = h_{s-k}(\cdot, s) \text{ for all } k \text{ such that } 0 \leq k \leq n.$$

Thus

$$f^{\Delta^k}(s) = h_{n-k}(s, s) = \text{ for all } k \text{ such that } 0 \leq k \leq n-1,$$

$f^{\Delta^n} = h_0(s, s) = 1$ and $f^{\Delta^{n-1}}(\tau) \equiv 0$. Now by applying our first version of Taylor's Theorem, Theorem 1.48, we have

$$\begin{aligned} h_n(s, t) &= f(t) \\ &= \sum_{k=0}^n (-1)^k g_k(\alpha, t) f^{\Delta^k}(\alpha) \int_{\alpha}^{\rho^n(t)} (-1)^n g_n(\sigma(\tau, t)) f^{\Delta^{n+1}}(\tau) \Delta\tau \\ &= \sum_{k=0}^n (-1)^k g_k(s, t) f^{\Delta^k}(s) \int_s^{\rho^n(t)} (-1)^n g_n(\sigma(\tau, t)) f^{\Delta^{n+1}}(\tau) \Delta\tau \\ &= (-1)^n g_n(s, t) f^{\Delta^n}(s) \\ &= (-1)^n g_n(s, t). \end{aligned}$$

□

Now we are ready for our second version of Taylor's Theorem.

Theorem 1.50 (Taylor's Formula). *Let $n \in \mathbb{N}$. Suppose f is n times differentiable on $\mathbb{T}^{k^n}$. Let $\alpha \in \mathbb{T}^{k^{n-1}}$, $t \in \mathbb{T}$ and define the functions h_k by (1.10) and (1.11). Then we have*

$$f(t) = \sum_{k=0}^{n-1} h_k(t, \alpha) f^{\Delta^k}(\alpha) + \int_{\alpha}^{\rho^{n-1}(t)} h_{n-1}(t, \sigma(\tau)) f^{\Delta^n}(\tau) \Delta\tau.$$

Proof. Let $h_0(r, s) \equiv 1$ and $h_{k+1}(r, s) = \int_s^r h_k(\tau, s) \Delta\tau$ for $k \in \mathbb{N}_0$ then by Theorem 1.48 we have

$$f(t) = \sum_{k=0}^{n-1} (-1)^k g_k(\alpha, t) f^{\Delta^k}(\alpha) + \int_{\alpha}^{\rho^{n-1}(t)} (-1)^{n-1} g_{n-1}(\sigma(\tau), t) f^{\Delta^n}(\tau) \Delta\tau.$$

By Theorem 1.49 we can substitute $h_n(t, s) = (-1)^n g_n(s, t)$, thus

$$f(t) = \sum_{k=0}^{n-1} h_k(t, \alpha) f^{\Delta^k}(\alpha) + \int_{\alpha}^{\rho^{n-1}(t)} h_{n-1}(t, \sigma(\tau)) f^{\Delta^n}(\tau) \Delta\tau.$$

□

1.5 THE EXPONENTIAL FUNCTION

Before we can discuss the exponential function we need to define a new class of functions.

Definition 1.51 We say that a function $f : \mathbb{T} \rightarrow \mathbb{R}$ is regressive if

$$1 + \mu(t)f(t) \neq 0 \text{ for all } t \in \mathbb{T}.$$

We denote the set of regressive and right dense continuous function $\mathcal{R} = \mathcal{R}(\mathbb{T}) = \mathcal{R}(\mathbb{T}, \mathbb{R})$.

Definition 1.52 For $h > 0$ define the function

$$\xi_h(x) = \frac{1}{h} \text{Log}(1 + xh),$$

for any real number $x \neq -\frac{1}{h}$, where Log is the principal logarithm function. In other words,

$$\xi_h(x) = \frac{1}{h} \begin{cases} \log(1 + xh) & \text{for } x > -\frac{1}{h} \\ \log|1 + xh| + i\pi & \text{for } x < -\frac{1}{h} \end{cases}.$$

When $h = 0$ we define $\xi_0(x) = x$.

Now we are ready to define the exponential function.

Definition 1.53 If $f : \mathbb{T} \rightarrow \mathbb{R}$ is regressive and right dense continuous, then we define the *exponential function* by

$$e_f(t, s) = \exp \left(\int_s^t \xi_{\mu(\tau)}(f(\tau)) \Delta\tau \right)$$

for $t, s \in \mathbb{T}$.

We should remark that time scale exponential function, $x_0 e_p(t, t_0)$, is the unique solution to the dynamic initial value problem

$$x^\Delta = p(t)x, \quad x(t_0) = x_0$$

where $t_0 \in \mathbb{T}$ and $p \in \mathcal{R}$.

Theorem 1.54. *If $x^\Delta = p(t)x$ is regressive then the unique solution to*

$$x^\Delta = p(t)x, \quad x(t_0) = 1 \tag{1.14}$$

is given by $e_p(\cdot, t_0)$.

Proof. Assume x is a second solution to the initial value problem (1.14). By Definition 1.5, $e_p(t, s) \neq 0$ for all $s, t \in \mathbb{T}$. Thus Theorem 1.16 (v) yields

$$\begin{aligned} \left(\frac{x}{e_p(\cdot, t_0)} \right)^\Delta &= \frac{x^\Delta e_p(t, t_0) x(t) e_p^\Delta(t, t_0)}{e_p(t, t_0) e_p(\sigma(t), t_0)} \\ &= \frac{p(t)x(t)e_p(t, t_0) - x(t)p(t)e_p(t, t_0)}{e_p(t, t_0)e_p(\sigma(t), t_0)} \\ &= 0. \end{aligned}$$

Now by Corollary 1.22

$$\frac{x(t)}{e_p(t, t_0)} \equiv \frac{x(t_0)}{e_p(t_0, t_0)} = 1.$$

Thus $x = e_p(\cdot, t_0)$. □

Next we will present some useful properties of the exponential function.

Theorem 1.55. *If $p, q \in \mathcal{R}$ and $t, s, r \in \mathbb{T}$, then*

$$(i) \quad e_0 \equiv 1 \text{ and } e_p(t, t) \equiv 1.$$

$$(ii) \quad e_p(\sigma(t), s) = (1 + \mu(t)p(t))e_p(t, s).$$

$$(iii) \quad e_p(t, s)e_p(s, r) = e_p(t, r).$$

Proof. (i) An obvious solution to the initial value problem $x^\Delta(t) = 0$, $x(s) = 1$ is $x(t) \equiv 1$.

By Theorem 1.54 the initial value problem only has one solution $e_0(t, s)$. Therefore $e_0(t, s) \equiv x(t) \equiv 1$.

(ii) By Theorem 1.15 (i)

$$\begin{aligned} e_p(\sigma(t), s) &= e_p(t, s) + \mu(t)e_p^\Delta(t, s) \\ &= e_p(t, s) + \mu(t)p(t)e_p(t, s) \\ &= (1 + \mu(t)p(t))e_p(t, s). \end{aligned}$$

(iii) Let $p \in \mathcal{R}$ and $r, s, t \in \mathbb{T}$. Then by Definition 1.5 and by Theorem 1.31

$$\begin{aligned} e_p(t, r)e_p(r, s) &= \exp\left(\int_r^t \xi_{\mu(t)}(p(\tau))\Delta\tau\right) \exp\left(\int_s^r \xi_{\mu(t)}(p(\tau))\Delta\tau\right) \\ &= \exp\left(\int_r^t \xi_{\mu(t)}(p(\tau))\Delta\tau + \int_s^r \xi_{\mu(t)}(p(\tau))\Delta\tau\right) \\ &= \exp\left(\int_s^t \xi_{\mu(t)}(p(\tau))\Delta\tau\right) \\ &= e_p(t, s). \end{aligned}$$

□

The construction of the exponential function yields $e_p(t, t_0) = e^{p(t-t_0)}$ on $\mathbb{R}$. Similarly $e_p(t, t_0) = (1 + p)^{t-t_0}$ on $\mathbb{Z}$ as we will see in the next example.

Example 1.56 Let $\mathbb{T} = \mathbb{Z}$ and $p \in \mathcal{R}$. Then

$$e_p(t, 0) = (1 + p)^t$$

for all $t \in \mathbb{T}$. First we should note that the above equation satisfies

$$y(0) = (1 + p)^0 = 1.$$

and

$$\begin{aligned}
 y^\Delta(t) &= \frac{y(\sigma(t)) - y(t)}{\mu(t)} \\
 &= \frac{y(t+1) - y(t)}{1} \\
 &= (1+p)^{t+1} - (1+p)^t \\
 &= (1+p-1)(1+p)^t \\
 &= p(1+p)^t \\
 &= py(t)
 \end{aligned}$$

for all $t \in \mathbb{T}$.

To use for later approximations we want to develop the exponential function for the time scale $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$.

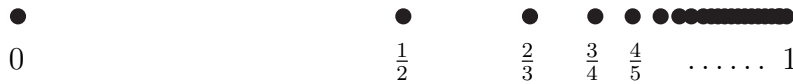


Figure 1.5: $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$

Example 1.57 Let $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$, and $p \in \mathcal{R}$. We will need the following identities to develop this exponential

$$t = 1 - \frac{1}{n}, \quad \sigma(t) = \frac{1}{2-t}, \quad \text{and} \quad n = \frac{1}{1-t} - 1. \quad (1.15)$$

We should note that σ holds for all $t \in \mathbb{T}$ even 1 but the formula for n does not hold for $t = 1$. We want

$$\frac{y(\sigma(t)) - y(t)}{\mu(t)} = py(t)$$

for $n \neq 1$. Therefore

$$y(\sigma(t)) = (\mu(t)p + 1)y(t).$$

Now by using our identities in (1.15) we have

$$y\left(\frac{1}{2-t}\right) = \left[\left(\frac{1}{2-t} - t\right)p + 1\right] y(t).$$

By working out the first few iterations we can see a pattern:

$$\begin{aligned} y\left(\frac{1}{2}\right) &= \left(\frac{1}{2}p + 1\right) y(0) \\ y\left(\frac{2}{3}\right) &= \left(\frac{1}{6}p + 1\right) y\left(\frac{1}{2}\right) \\ &= \left(\frac{1}{6}p + 1\right) \left(\frac{1}{2}p + 1\right) y(0) \\ y\left(\frac{3}{4}\right) &= \left(\frac{1}{12}p + 1\right) y\left(\frac{2}{3}\right) \\ &= \left(\frac{1}{12}p + 1\right) \left(\frac{1}{6}p + 1\right) \left(\frac{1}{2}p + 1\right) y(0) \\ &\vdots \\ y(t) &= y(0) \prod_{i=1}^{\frac{1}{1-t}-1} \left(\frac{p}{i(i+1)} + 1\right). \end{aligned}$$

Thus the exponential on our time scale $\mathbb{T}$ is

$$e_p(t, t_0) = y(t_0) \prod_{i=1}^{\frac{1}{1-t}-1} \left(\frac{p}{i(i+1)} + 1\right). \quad (1.16)$$

where $\frac{1}{1-t} - 1$ is taken to be ∞ when $t = 1$.

We are interested in values of p for which the exponential function in (1.16) converges so we can estimate $e_p(1, t_0)$. Therefore we need the following theorems for convergence of a infinite product.

Theorem 1.58. *Let $\operatorname{Re}(z_n) > 0$ for all $n \geq 1$. Then $\prod_{n=1}^{\infty} z_n$ converges to a nonzero number if and only if the series $\sum_{n=1}^{\infty} \log z_n$ converges.*

Proof. Suppose $\prod_{n=1}^{\infty} z_n$ converges to a nonzero number. Let $p_n = \prod_{n=1}^{\infty} z_n = (z_1 \cdots z_n)$, where $z = re^{i\theta}$, and $-\pi < \theta \leq \pi$. Also let $l(p_n) = \log |p_n| + i\theta_n$ where $\theta - \pi < \theta_n \leq \theta + \pi$. Now if $s_n = \log z_n + \cdots + \log z_n$ then $\exp(s_n) = p_n$ such that $s_n = l(p_n) + 2\pi i k_n$ for some integer k_n . Now as $p_n \rightarrow z$, $s_n - s_{n-1} = \log z_n \rightarrow 0$ and $l(p_n) - l(p_{n-1}) \rightarrow 0$. Thus $k_n - k_{n-1} \rightarrow 0$

as $n \rightarrow \infty$. Since $k_n \in \mathbb{Z}$ there exists an n_0 and a k such that $k_m = k_n = k$ for $m, n \geq n_0$. Therefore $s_n \rightarrow l(z) + 2\pi ik$ and thus the series $\sum_{n=1}^{\infty} \log z_n$ converges.

Suppose the series $\sum_{n=1}^{\infty} \log z_n$ converges. Let $s_n = \sum_{n=1}^n z_n$ and $s_n \rightarrow s$ then $\exp(s_n) \rightarrow \exp(s)$. We also have $\exp(s_n) = \prod_{k=1}^n z_k$ so $\prod_{n=1}^{\infty} z_n$ is convergent to $z = \exp(s) \neq 0$. $\square$

Before our next theorem we want to consider the power series expansion for $\log(1+z)$ about $z=0$:

$$\log(1+z) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{z^n}{n} = z - \frac{z^2}{2} + \frac{z^3}{3} - \cdots.$$

The power series expansion for $\log(1+z)$ has a radius of convergence of 1. Therefore when $|z| < 1$ we have

$$\begin{aligned} \left| 1 - \frac{\log(1+z)}{z} \right| &= \left| \frac{1}{2}z - \frac{1}{3}z^2 + \cdots \right| \\ &\leq \frac{1}{2}(|z| + |z|^2 + \cdots) \\ &= \frac{|z|}{2(1-|z|)}. \end{aligned}$$

In fact if $|z| < \frac{1}{2}$ then

$$\left| 1 - \frac{\log(1+z)}{z} \right| \leq \frac{1}{2}$$

which implies

$$\frac{1}{2}|z| \leq |\log(1+z)| \leq \frac{3}{2}|z|. \quad (1.17)$$

Theorem 1.59. *Let $\operatorname{Re}(z_n) > -1$ then the series $\sum_{n=1}^{\infty} \log(z_n + 1)$ converges absolutely if and only if the series $\sum_{n=1}^{\infty} z_n$ converges absolutely.*

Proof. Suppose the series $\sum_{n=1}^{\infty} |z_n|$ converges. Then $z_n \rightarrow 0$ and eventually $|z_n| < \frac{1}{2}$. Therefore $\sum_{n=1}^{\infty} |\log(1+z_n)|$ is bounded by a convergent series in (1.17) and therefore must also converge.

Suppose $\sum_{n=1}^{\infty} \log(1+z_n)$ converges thus $\log(1+z_n) \rightarrow 0$. Therefore $1+z_n \rightarrow 1$ which implies $z_n \rightarrow 0$. So for sufficiently large n , $|z_n| < \frac{1}{2}$. By (1.17) we know $\sum_{n=1}^{\infty} |z_n|$ is bounded by a convergent sum and therefore converges. $\square$

Definition 1.60 A product $\prod_{n=1}^{\infty}(z_n + 1)$ is said to converge absolutely if the series $\sum_{n=1}^{\infty} \log(z_n)$ converges absolutely.

Now by combining the above definition and previous two theorems we can conclude the following corollary.

Corollary 1.61. *If $\operatorname{Re}(z_n) > 0$ then the product $\prod_{n=1}^{\infty}(z_n + 1)$ converges absolutely if and only if the series $\sum_{n=1}^{\infty} z_n$ converges absolutely.*

Therefore our exponential function, (1.16), for $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$ converges for any p such that $\sum_{n=1}^{\infty} |\frac{1}{n(n+1)}p|$ converges.

Example 1.62 Let $p \equiv 1$. The sum $\sum_{n=1}^{\infty} |\frac{1}{n(n+1)}|$ converges to 1. Thus by Corollary 1.61 the product $\prod_{n=1}^{\infty} (\frac{1}{n(n+1)} + 1)$ also converges absolutely. By using Maple

$$\prod_{n=1}^{\infty} \left(\frac{1}{n(n+1)} + 1 \right) \rightarrow \frac{\sin \left(\pi \left(\frac{1}{2} + \frac{1}{2}i\sqrt{3} \right) \right)}{\pi}.$$

We should note that this is not actually a complex number since it is a product of real numbers. In fact it is approximately 2.428189793.

1.6 EULER'S METHOD

Euler's Method is used to obtain a linear approximation to the initial value problem

$$\frac{\delta y}{\delta t} = f(t, y) \quad a \leq t \leq b \text{ and } y(a) = \alpha. \quad (1.18)$$

Euler's Method does not acquire a continuous solution $y(t)$ but it creates a piecewise linear approximation to $y(t)$. The approximation to y is calculated at various mesh points in the interval $[a, b]$. First we require that these mesh points are evenly distributed through the interval $[a, b]$. The even step size, h , can be guaranteed by selecting a positive integer N and

choosing the mesh points as

$$t_i = a + ih, \text{ for each } i = 0, 1, 2, \dots, N-1 \text{ where } h = \frac{(b-a)}{N}.$$

We will need Taylor's theorem to derive Euler's method.

Theorem 1.63 (Taylor's Theorem). *Suppose $f \in C^n[a, b]$, that f^{n+1} exists on $[a, b]$ and $x_0 \in [a, b]$ For every $x \in [a, b]$ there exists a number $\xi(x)$ between x_0 and x with*

$$f(x) = P_n(x) + R_n(x),$$

where

$$\begin{aligned} P_n(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^n(x_0)}{n!}(x - x_0)^n \\ &= \sum_{k=0}^n \frac{f^k(x_0)}{k!}(x - x_0)^k \end{aligned}$$

and

$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)^{n+1}.$$

The Taylor's Theorem above is just the special case of our time scale versions of Taylor's Theorem, specifically when $\mathbb{T} = \mathbb{R}$.

Suppose that $y(t)$ is the unique solution to (1.18), and has two continuous derivatives on $[a, b]$ such that for each $i = 0, 1, 2, \dots, N-1$, we have

$$y(t_{i+1}) = y(t_i) + (t_{i+1} - t_i)y'(t_i) + \frac{(t_{i+1} - t_i)^2}{2}y''(\xi_i),$$

for some number $\xi_i \in (t_i, t_{i+1})$. Since we know $h = t_{i+1} - t_i$

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(\xi_i),$$

and we also know that $y(t)$ satisfies (1.18) thus

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2}y''(\xi_i). \quad (1.19)$$

For Euler's method we approximate $y(t_i)$, for each $i = 1, 2, \dots, N$ by ignoring the remainder term. So we construct a sequence $w_i \approx y(t_i)$ such that

$$w_0 = \alpha \quad (1.20)$$

$$w_{i+1} = w_i + hf(t_i, w_i), \quad (1.21)$$

for each $i = 0, 1, 2, \dots, N - 1$.

Example 1.64 Let us use Euler's Method to approximate the solution to

$$\frac{\delta y}{\delta t} = -ty \quad -1 \leq t \leq 1 \text{ and } y(-1) = 0.6065306597 \quad (1.22)$$

with $N = 10$. Then $h = 0.2$, $w_0 = 0.6065306597$ and

$$w_{i+1} = w_i + h(t_i w_i)$$

for $i = 0, 1, 2, \dots, 9$.



Figure 1.6: Points used for Euler's Method Approximation

We know the exact solution is $y = e^{\frac{-t^2}{2}}$.

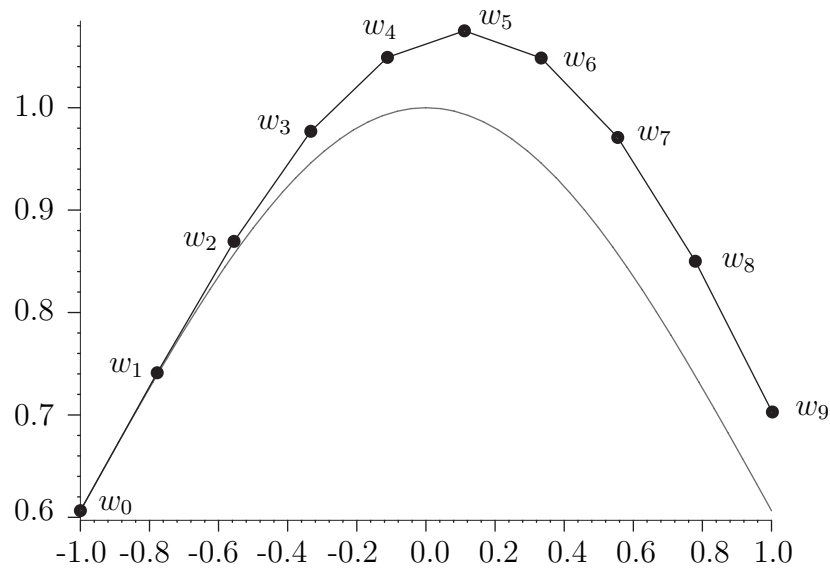


Figure 1.7: Euler's Method Approximation for $y = e^{-\frac{t^2}{2}}$

In Figure 1.7 we can see how Euler's Method decays the further out we push the approximation. Table 1.1 shows the comparison between the actual values and the approximate values generated by Euler's Method.

Table 1.1: Euler's Method Comparison for $y = e^{\frac{-t^2}{2}}$

t_i	w_i	$y_i = e^{\frac{-(t_i)^2}{2}}$	$ y_i - t_i $
-1.00000000	0.606530660	0.606530660	0.000000000
-0.77777778	0.741315251	0.738991296	0.002323954
-0.55555556	0.869443812	0.856996891	0.012446921
-0.33333333	0.976782555	0.945959469	0.030823086
-0.11111111	1.049136818	0.993846173	0.055290645
0.11111111	1.075041431	0.993846173	0.081195258
0.33333333	1.048497198	0.945959469	0.102537729
0.55555556	0.970830739	0.856996891	0.113833848
0.77777778	0.850975093	0.738991296	0.111983797
1.00000000	0.703892978	0.606530660	0.097362318

We should note that error grows as we reach the end of the interval. The error is expected to grow in no worse than a linear manner, yet the approximation decays rather quickly. This is why Euler's method is not often used in practice.

Euler's method gives us a decent approximation to $y(t_i)$ but is not exact. To derive the error bound for Euler's method we need the following lemmas and a Lipschitz condition.

Lemma 1.65. *For all $x \geq -1$ and any m , we have*

$$0 \leq (1+x)^m \leq e^{mx}.$$

Proof. By applying Taylor's Theorem, Theorem 1.63, of $f(x) = e^x$ with $x_0 = 0$ and $n = 1$ we have

$$e^x = 1 + x + \frac{1}{2}x^2e^\xi$$

where $\xi \in (0, x)$. Thus

$$0 \leq 1 + x \leq 1 + x + \frac{1}{2}x^2 e^\xi.$$

Since we know $0 \leq (1 + x)$,

$$0 \leq (1 + x)^m \leq (e^x)^m = e^{mx}.$$

□

Lemma 1.66. *If s and t are positive real numbers, $\{a_i\}_{i=0}^k$ is a sequence satisfying $a_0 \geq \frac{-t}{s}$ and*

$$a_{i+1} \leq (1 + s)a_i + t, \quad (1.23)$$

for each $i = 0, 1, 2, \dots, k$, then

$$a_{i+1} \leq e^{(i+1)s} \left(a_0 + \frac{t}{s} \right) - \frac{t}{s}.$$

Proof. Fix $i \in \mathbb{N}$ then by (1.23) we have

$$\begin{aligned} a_{i+1} &\leq (1 + s)a_i + t \\ &\leq (1 + s)[(1 + s)a_{i-1} + t] + t \\ &\leq (1 + s)\{(1 + s)[(1 + s)a_{i-2} + t] + t\} + t \\ &\vdots \\ &\leq (1 + s)^{i+1}a_0 + [1 + (1 + s) + (1 + s)^2 + \dots + (1 + s)^i] t. \end{aligned}$$

By geometric series we know

$$\begin{aligned} 1 + (1 + s) + \dots + (1 + s)^i &= \sum_{j=0}^i (1 + s)^j \\ &= \frac{1 - (1 + s)^{i+1}}{1 - (1 + s)} \\ &= \frac{1}{s} [(1 + s)^{i+1} - 1]. \end{aligned}$$

Thus,

$$a_{i+1} \leq (1 + s)^{i+1}a_0 + \left(\frac{(1 + s)^{i+1} - 1}{s} \right) t = (1 + s)^{i+1} \left(a_0 + \frac{t}{s} \right) - \frac{t}{s}.$$

Now let $x = 1 + s$. Then by Lemma 1.65 we have

$$a_{i+1} \leq e^{(i+1)s} \left(a_0 + \frac{t}{s} \right) - \frac{t}{s}.$$

□

Definition 1.67 A function $f(t, y)$ is said to satisfy a *Lipschitz condition* in the second variable y on a set $D \subset \mathbb{R}^2$ if there exists a constant $L > 0$ such that

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2|$$

whenever $(t, y_1), (t, y_2) \in D$. The constant L is called the *Lipschitz constant* for f .

Theorem 1.68. Suppose f is continuous and satisfies a Lipschitz condition with constant L on

$$D = \{(t, y) | a \leq t \leq b, -\infty < y < \infty\},$$

and there exists a constant M such that

$$|y''(t)| \leq M \text{ for all } t \in [a, b].$$

Let $y(t)$ denote the unique solution to the initial value problem

$$\frac{\delta y}{\delta t} = f(t, y) \quad a \leq t \leq b \text{ and } y(a) = \alpha,$$

and $w_0, w_1, \dots, w_N$ be the approximations generated by (1.21) and (1.21) for some $N \in \mathbb{N}$.

Then for each $i = 0, 1, 2, \dots, N - 1$,

$$|y(t_i) - w_i| \leq \frac{hM}{2L} [e^{L(t_i-a)} - 1]. \quad (1.24)$$

Proof. Let $i = 0$ then the result is clearly true since

$$y(t_0) = w_0 = \alpha.$$

From (1.19) for $i = 0, 1, 2, \dots, N-1$ we have

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2}y''(\xi_i),$$

and from (1.21) we have

$$w_{i+1} = w_i + hf(t_i, w_i).$$

Thus

$$y(t_{i+1}) - w_{i+1} = y(t_i) - w_i + h[f(t_i, y(t_i)) - f(t_i, w_i)] + \frac{h^2}{2}y''(\xi_i).$$

Therefore

$$|y(t_{i+1}) - w_{i+1}| \leq |y(t_i) - w_i| + h|f(t_i, y(t_i)) - f(t_i, w_i)| + \frac{h^2}{2}|y''(\xi_i)|.$$

Since f satisfies a Lipschitz condition in the second variable with constant L and $|y''(t)| \leq M$,

$$|y(t_{i+1}) - w_{i+1}| \leq (1 + hL)|y(t_i) - w_i| + \frac{h^2M}{2}.$$

Let $s = hL$, $t = \frac{h^2M}{2}$ and $a_j = |y(t_j) - w_j|$ for each $j = 0, 1, 2, \dots, N$. Then by Lemma 1.66 we have

$$|y(t_{i+1}) - w_{i+1}| \leq e^{(i+1)hL} \left(|y_0 - w_0| + \frac{h^2M}{2hL} \right) - \frac{h^2M}{2hL}.$$

Since we know $|y(t_0) - w_0| = 0$ and $(i+1)h = t_{i+1} - t_0 = t_{i+1} - a$,

$$|y(t_{i+1}) - w_{i+1}| \leq \frac{hM}{2L}(e^{(t_{i+1}-a)L} - 1)$$

for each $i = 0, 1, \dots, N-1$. □

Example 1.69 Let us revisit our previous example, Example 1.22. The function $f(t, y) = -ty$ has a Lipschitz constant, $L = 1$ since

$$\begin{aligned} |f(t, y_1) - f(t, y_2)| &= |-ty_1 - (-ty_2)| \\ &= |-t(y_1 - y_2)| \\ &= |-t||y_1 - y_2| \\ &\leq 1|y_1 - y_2|. \end{aligned}$$

Also $|y''(t)| \leq e^{-\frac{1}{2}}$ on $-1 \leq t \leq 1$. Now we can apply Theorem 1.68 to estimate the error for Example 1.22.

Table 1.2: Error Estimate for Example 1.22

t_i	Actual Error	Estimated Error
-1.000000000	0.000000000	0.000000000
-0.777777778	0.0023239543	0.0167704965
-0.555555556	0.0124469209	0.0377143122
-0.333333333	0.0308230859	0.0638699726
-0.111111111	0.0552906447	0.0965344395
0.111111111	0.0811952577	0.1373274221
0.333333333	0.1025377290	0.1882716922
0.555555556	0.1138338480	0.2518933867
0.777777778	0.1119837965	0.3313472675
1.000000000	0.0973623182	0.4305731567

As we can see in Table 1.2 that Theorem 1.68 does a good job of bounding the error for Euler's method.

The flaw in Theorem 1.68 is that it relies on finding an upper bound for $y''(t)$. This condition can prevent us from obtaining a realistic error bound. We should note that if $\frac{\delta f}{\delta t}$ and $\frac{\delta f}{\delta y}$ both exist the chain rule for partial differentiation allows us to calculate an error bound for $y''(t)$ without knowing $y(t)$, by calculating $y''(t)$ as follows,

$$y''(t) = \frac{dy'}{dt}(t) = \frac{df}{dt}(t, y(t)) = \frac{\delta f}{\delta t}(t, y(t)) + \frac{\delta f}{\delta y}(t, y(t)) \cdot f(t, y(t)).$$

Theoretically Theorem 1.68 is ideal. Since the result relies linearly on the step size h , reducing the step size should result in better approximation of $y(t)$. Theorem 1.68 neglects round off error, as we reduce h more calculations necessary. Therefore we have greater round

off error. So, instead of calculating Euler's method by

$$\begin{aligned} w_0 &= \alpha \\ w_{i+1} &= w_i + hf(t_i, w_i), \quad \text{for each } i = 0, 1, 2, \dots, N-1, \end{aligned}$$

we should calculate Euler's method by

$$u_0 = \alpha + \delta_0 \tag{1.25}$$

$$u_{i+1} = u_i + hf(t_i, u_i) + \delta_{i+1}, \tag{1.26}$$

for each $i = 0, 1, 2, \dots, N-1$. where δ_i denotes the round off error associated with u_i . Now, we can develop a theorem corresponding to Theorem 1.68 that considers round off error.

Theorem 1.70. *Let $y(t)$ denotes the unique solution to the initial-value problem*

$$y' = f(t, y), \quad a \leq t \leq b, \quad y(a) = \alpha \tag{1.27}$$

and $u_0, u_1, \dots, u_N$ be the approximation obtained by (1.25) and (1.26). If $|\delta_i| < \delta$ for each $i = 0, 1, 2, \dots, N-1$ and the hypothesis of Theorem 1.68 holds for (1.27), then

$$|y(t_i) - u_i| \leq \frac{1}{L} \left(\frac{hM}{2} + \frac{\delta}{h} \right) [e^{L(t_i-a)} - 1] + |\delta_0|e^{L(t_i-a)}, \tag{1.28}$$

for each $i = 0, 1, 2, \dots, N-1$.

The proof of Theorem 1.70 is similar to the proof of Theorem 1.68. Theorem 1.70 shows the error bound is not linear in h , especially since

$$\lim_{h \rightarrow \infty} \left(\frac{hM}{2} + \frac{\delta}{h} \right) = \infty.$$

In fact the error is expected to be large for small values of h . Letting

$$E(h) = \left(\frac{hM}{2} + \frac{\delta}{h} \right)$$

implies that

$$E'(h) = \frac{M}{2} - \frac{\delta}{h^2}.$$

Thus

$$\begin{aligned} \text{if } h < \sqrt{\frac{2\delta}{M}}, \text{ then } E'(h) < 0 \text{ and } E(h) \text{ is decreasing and} \\ \text{if } h > \sqrt{\frac{2\delta}{M}}, \text{ then } E'(h) > 0 \text{ and } E(h) \text{ is increasing.} \end{aligned}$$

Thus the minimal value of $E(h)$ is

$$h = \sqrt{\frac{2\delta}{M}}.$$

If h is increases past this minimal value or if we decrease h beyond this minimal value, the error will start to increase. Even when δ is sufficiently small this lower bound will not affect the accuracy of Euler's method.

CHAPTER 2

NUMERICAL INTEGRATION FOR TIME SCALES

2.1 NUMERICAL INTEGRATION ON ISOLATED TIME SCALES

Now we are ready to develop Euler's Method for an isolated time scale. Throughout this chapter we will assume that $a < b$, $a \in \mathbb{T}$ and $b \in \mathbb{T}^\kappa$. We will want to restate the initial value problem (1.18) in the previous section in time scale terms. For time scales we would like to obtain a linear approximation to the initial value problem

$$y^\Delta(t) = f(t, y) \quad t \in \mathbb{T}^\kappa \text{ and } y(a) = \alpha \quad (2.1)$$

where $\mathbb{T}$ is a time scale over the interval $[a, b]$. The time scale interval is taken to be $\mathbb{T} \cap [a, b]$. We will calculate the solution $y(t)$ by using the time scale version of Taylor's Theorem, Theorem 1.50. We will consider any time scale of isolated points, $\mathbb{T}$, over the interval $[a, b]$. The time scale will be enumerated such that $\sigma(t_i) = t_{i+1}$. Suppose f is two times Δ -differentiable. Then by the time scale version of Taylor's Theorem for any $t \in \mathbb{T}^\kappa$

$$\begin{aligned} y(t) &= \sum_{k=0}^1 h_k(t, t_i) y^{\Delta^k}(t_i) + \int_{t_i}^{\rho(t)} h_1(t, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \\ &= h_0(t, t_i) y(t_i) + h_1(t, t_i) y^\Delta(t_i) + \int_{t_i}^{\rho(t)} h_1(t, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau. \end{aligned}$$

To ease our calculations we will simplify the integral separately

$$\int_{t_i}^{\rho(t)} h_1(t, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau. \quad (2.2)$$

We will need to use integration by parts for time scales, Theorem 1.32. We need $h_1(t, \tau)$ and $y^{\Delta^2}(\tau)$ to be right dense continuous. Since we are considering an isolated time scale the

functions are right dense continuous by default. Let

$$f(\sigma(t)) = h_1(t, \sigma(t)).$$

Then

$$f(\tau) = h_1(t, \tau) = t - \tau$$

and therefore

$$\begin{aligned} f^\Delta(\tau) &= \frac{(t - \sigma(\tau)) - (t - \tau)}{\sigma(\tau) - \tau} \\ &= \frac{-(\sigma(\tau) - \tau)}{\sigma(\tau) - \tau} \\ &= -1. \end{aligned}$$

Now let

$$g^\Delta(t)\Delta t = y^{\Delta^2}(t)\Delta t,$$

then

$$g(t) = y^\Delta(t).$$

Integration by parts for time scales, Theorem 1.32, reduces (2.2) to

$$= (t - \rho(t))y^\Delta(\rho(t)) - (t - t_i)y^\Delta(t_i) - \int_{t_i}^{\rho(t)} (-1)y^\Delta(\tau)\Delta\tau \quad (2.3)$$

$$= (t - \rho(t))y^\Delta(\rho(t)) - (t - t_i)y^\Delta(t_i) + \int_{t_i}^{\rho(t)} y^\Delta(\tau)\Delta\tau \quad (2.4)$$

and

$$\begin{aligned} y(t) &= h_0(t, t_i)y(t_i) + h_1(t, t_i)y^\Delta(t_i) + \int_{t_i}^{\rho(t)} h_1(t, \sigma(\tau))y^{\Delta^2}(\tau)\Delta\tau \\ &= h_0(t, t_i)y(t_i) + h_1(t, t_i)y^\Delta(t_i) + (t - \rho(t))y^\Delta(\rho(t)) - (t - t_i)y^\Delta(t_i) + \int_{t_i}^{\rho(t)} y^\Delta(\tau)\Delta\tau. \end{aligned}$$

We want to estimate $y(t_{i+1})$ and we have a scattered time scale where $t_{i+1} = \sigma(t_i)$ which implies that $\rho(t_{i+1}) = t_i$ then the above equation yields

$$\begin{aligned}
y(t_{i+1}) &= h_0(t_{i+1}, t_i)y(t_i) + h_1(t_{i+1}, t_i)y^\Delta(t_i) + (t_{i+1} - \rho(t_{i+1}))y^\Delta(\rho(t_{i+1})) \\
&\quad - (t_{i+1} - t_i)y^\Delta(t_i) + \int_{t_i}^{\rho(t_{i+1})} y^\Delta(\tau)\Delta\tau \\
&= y(t_i) + \mu(t_i)y^\Delta(t_i) + (t_{i+1} - t_i)y^\Delta(t_i) \\
&\quad - (t_{i+1} - t_i)y^\Delta(t_i) + \int_{t_i}^{t_i} y^\Delta(\tau)\Delta\tau \\
&= y(t_i) + \mu(t_i)y^\Delta(t_i).
\end{aligned}$$

Our result looks a lot like Euler's method except the error term disappears. For time scales we will obtain an approximation of the solution to the initial value problem (2.1) by

$$w_0 = y(a) = \alpha \quad (2.5)$$

$$w_{i+1} = w_i + \mu(t_i)y^\Delta(t_i, w_i) \quad (2.6)$$

for each $i = 0, 1, 2, \dots, N-1$ where $t_i \in \mathbb{T}$.

Since we have discovered that there is no error term we have the following theorem.

Theorem 2.1. *Given a time scale, $\mathbb{T}$, of isolated points enumerated such that $t_{i+1} = \sigma(t_i)$.*

Let $y(t)$ be the exact solution to the initial value problem

$$y^\Delta(t) = f(t, y) \quad t \in \mathbb{T} \text{ and } y(a) = \alpha, \quad (2.7)$$

and let $w_0, \dots, w_{N-1}$ be the approximations generated by (2.5) and (2.6) for some $N \in \mathbb{N}$.

Then for each $i = 0, 1, 2, \dots, N$

$$|y(t_i) - w_i| = 0. \quad (2.8)$$

Proof. We will prove this by induction. Given $w_0 = y(t_0)$ then

$$\begin{aligned}
w_1 &= w_0 + \mu(t_0)y^\Delta(t_0) \\
&= \mu(t_0) \frac{y(\sigma(t_0)) - y(t_0)}{\mu(t_0)} \\
&= y(t_0) + y(t_1) - y(t_0) \\
&= y(t_1).
\end{aligned}$$

Therefore our approximation gave us the exact solution at $t = t_1$ and thus

$$|y(t_1) - w_1| = |y(t_1) - y(t_1)| = 0.$$

Now, assume that $|y(t_k) - w_k| = 0$. This implies that $w_k = y(t_k)$ and therefore

$$\begin{aligned} w_{k+1} &= w_k + \mu(t_k)y^\Delta(t_k) \\ &= \mu(t_k) \frac{y(\sigma(t_k)) - y(t_k)}{\mu(t_k)} \\ &= y(t_k) + y(t_{k+1}) - y(t_k) \\ &= y(t_{k+1}). \end{aligned}$$

Again our approximation gave us the exact solution at $t = t_{k+1}$ and therefore

$$|y(t_{k+1}) - w_{k+1}| = |y(t_{k+1}) - y(t_{k+1})| = 0.$$

By the principle of mathematical induction our result holds for all $n \in \mathbb{N}$. □

Now, let us look at an example.

Example 2.2 Using our integration scheme for isolated time scales, we will create an approximation to the same curve in Example 1.6, $y = e^{\frac{-t^2}{2}}$. For our time scale approximation we are given

$$y^\Delta = \frac{e^{\frac{-\sigma(t)^2}{2}} - y}{\mu(t)} \quad t \in \mathbb{T}^\kappa \text{ and } y(-1) = 0.6065306597 \quad (2.9)$$

and will use the same set of points used in Example 1.6, $\mathbb{T} = \{t_i = 0.2i - 1 | i = 0, 1, 2, \dots, 9\}$, $w_0 = 0.6065306597$.



Figure 2.1: Points used for Isolated Time Scale Approximation

Then

$$\begin{aligned}
 w_{i+1} &= w_i + \mu(t_i)y^\Delta(t_i) \\
 &= w_i + \frac{e^{\frac{-\sigma(t_i)^2}{2}} - w_i}{\mu(t)} \\
 &= w_i + e^{\frac{-\sigma(t_i)^2}{2}} - w_i \\
 &= e^{\frac{-\sigma(t_i)^2}{2}} \\
 &= e^{\frac{-t_{i+1}^2}{2}}
 \end{aligned}$$

for $i = 0, 1, 2, \dots, 9$. Already we can see that this reduces to the exact solution as we can see in Figure 2.2. Our approximation scheme gives us the exact solution to the initial value problem for time scales. There is no linear approximation since our solution only exists on the time scale $\mathbb{T}$.

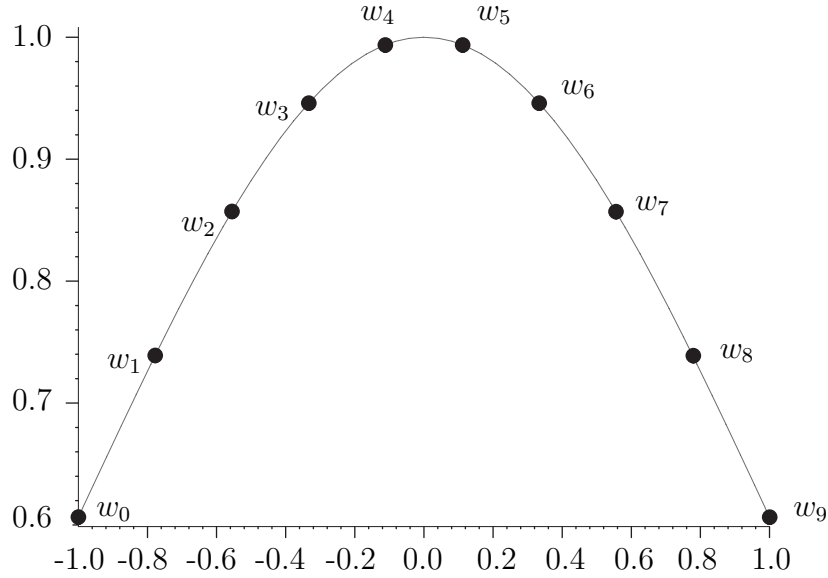


Figure 2.2: Isolated Time Scale Approximation of $y = e^{\frac{-t^2}{2}}$

Table 2.1: Isolated Time Scale Version of Euler's Method Comparison for $y = e^{-\frac{t^2}{2}}$

t_i	w_i	$y_i = e^{-\frac{(t_i)^2}{2}}$	$ y_i - w_i $
-1.000000000	0.6065306597	0.6065306597	0.000000000
-0.777777778	0.7389912963	0.7389912963	0.000000000
-0.555555556	0.8569968914	0.8569968914	0.000000000
-0.333333333	0.9459594689	0.9459594689	0.000000000
-0.111111111	0.9938461733	0.9938461733	0.000000000
0.111111111	0.9938461733	0.9938461733	0.000000000
0.333333333	0.9459594690	0.9459594690	0.000000000
0.555555556	0.8569968912	0.8569968912	0.000000000
0.777777778	0.7389912961	0.7389912961	0.000000000
1.000000000	0.6065306597	0.6065306597	0.000000000

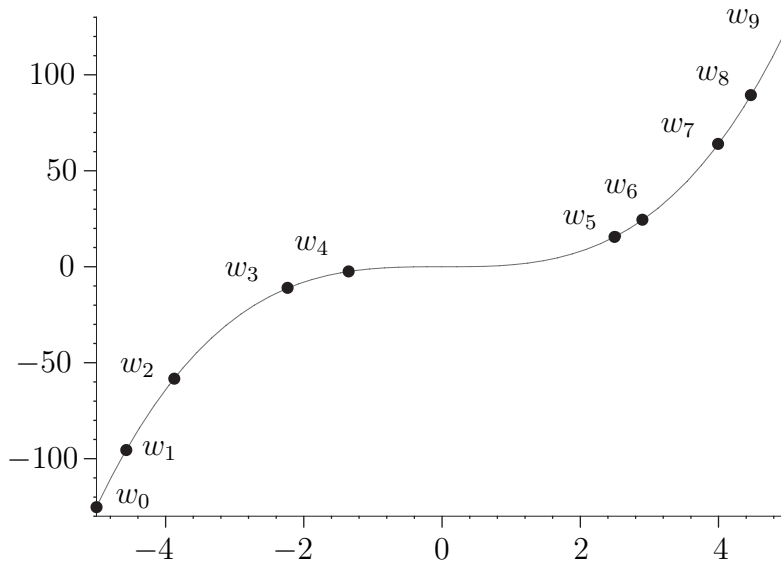
Table 2.1 shows the comparison between the actual values and the approximate values. The benefits of using a time scale approximation are obvious, but the draw back is attaining the Δ -derivative for the estimation.

The previous example is a little contrived considering $y(t)$ appeared in the Δ -derivative. Now, let us consider an example where $y(t)$ does not appear in the Δ -derivative.

Example 2.3 Let us use time scale approximation to approximate the solution to

$$y^\Delta = \sigma(t)^2 + t\sigma(t) + t^2 \quad t \in \mathbb{T}^\kappa \text{ and } y(-5) = -125 \quad (2.10)$$

given $\mathbb{T} = \{-5, 4.567872983, -3.872983346, -2.236067978, -1.347285960, 2.528597125, 2.904712595, 4.0, 4.472135955, 5\}$. We should note that in this example $y(t)$ is not apparent from $y^\Delta(t)$ and we have an unevenly spaced time scale. The exact solution for this example is $y(t) = t^3$.

Figure 2.3: Time Scale Approximation of $y = t^3$ Table 2.2: Time Scale Version of Euler's Method Comparison for $y(t) = t^3$

t_i	w_i	$y_i = t^3$	$ y_i - w_i $
-5.000000000	-125.0000000	-125.000000000	0.000000000
-4.567872983	-95.31078740	-95.31078740	0.000000000
-3.872983346	-58.09475018	-58.09475018	0.000000000
-2.236067978	-11.18033989	-11.18033990	0.000000100
-1.347285960	-2.445565794	-2.445565799	0.000000050
2.528597125	16.16735296	16.16735295	0.000000100
2.904712595	24.50809210	24.50809209	0.000000100
4.000000000	64.00000001	64.00000000	0.000000100
4.472135955	89.44271911	89.44271910	0.000000100
5.000000000	125.0000000	125.0000000	0.000000000

In Table 2.2 we can see the comparison between our actual values and are estimated values. The table suggests that our approximations are not exact. This example was calcu-

lated on Maple using 10 digits for each calculation. Since our solution $y(t)$ is not obvious from $y^\Delta(t)$, Maple had to multiply, divide, add and subtract resulting in some round off error. The error that has registered in Table 2.2 is due to the truncation and round off.

Due to possible round off error, we would not want to consider points less than some h_{min} apart. We could consider the time scale

$$\mathbb{T} = \left\{ 1 - \frac{1}{n}, n \in \mathbb{N} \right\} \cup \{1\}$$

consisting mostly of isolated points. Although it would not be practical to use this time scale with the isolated point method to find a solution to the initial value problem (2.1) since there would be infinitely many points in the time scale. Therefore we would want to use some reasonable subset of this time scale to approximate the solution.

Now we are ready to develop a numerical integration scheme for an arbitrary time scale.

2.2 NUMERICAL INTEGRATION ON ARBITRARY TIME SCALES

We wish to develop Euler's Method for an arbitrary time scale. To this end, we obtain an approximation to the same initial value problem as in the previous section

$$y^\Delta(t) = f(t, y) \quad t \in \mathbb{T}^\kappa \text{ and } y(a) = \alpha. \quad (2.11)$$

The mesh points will originate from an arbitrary time scale $\mathbb{T}$ containing time scale interval $[a, b]$. Since $\mathbb{T}$ can be any time scale, even $\mathbb{R}$, we will have to make a partition of the time scale $\mathbb{T}_p$ such that all of the points are at least some minimum step size h_{min} apart. To be specific, in $\mathbb{T}_p$ we want the end points, all of the scattered points that are at least h_{min} apart, the end points of dense intervals and any desired number of points from the dense intervals that are some h_{min} apart. After we have made this selection we want to go back through $\mathbb{T}_p$ and dispose of any points that are not at least h_{min} apart. Let us look at an example of how we will pick points.

Example 2.4 Consider $\mathbb{T} = \{0\} \cup \left\{ \frac{1}{2^n}, n \in \mathbb{N}_0 \right\} \cup \left\{ [1\frac{1}{4}, 1\frac{1}{2}], [1\frac{5}{8}, 2] \right\}$ and let $h_{min} = \frac{1}{4}$.



Figure 2.4: $\mathbb{T} = \{0\} \cup \left\{\frac{1}{2^n}, n \in \mathbb{N}_0\right\} \cup \left\{[1\frac{1}{4}, 1\frac{1}{2}], [1\frac{5}{8}, 2]\right\}$

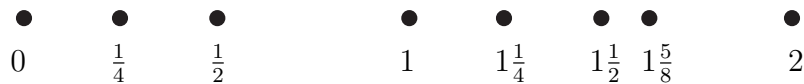
To construct $\mathbb{T}_p$, first we want the end points $\mathbb{T}_p = \{0, 2\}$.



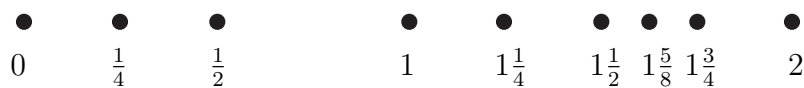
Then we want all of the scattered points at least $\frac{1}{4}$ apart thus $\mathbb{T}_p = \{0, \frac{1}{4}, \frac{1}{2}, 1, 2\}$.



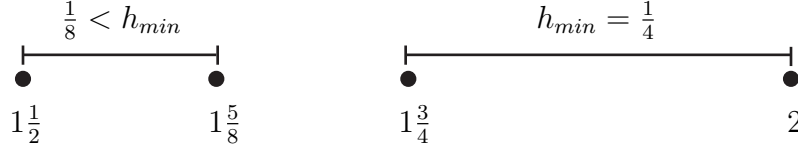
Now we need to include the end points of dense intervals thus $\mathbb{T}_p = \{0, \frac{1}{4}, \frac{1}{2}, 1, 1\frac{1}{4}, 1\frac{1}{2}, 1\frac{5}{8}, 2\}$.



Finally we can add in any points we want from the dense intervals. Let us add $1\frac{3}{4}$.



Now we have constructed $\mathbb{T}_p$, but we need to go through and check that all of the points are at least $h_{min} = \frac{1}{4}$ apart. We can see that $1\frac{5}{8}$ is not $\frac{1}{4}$ away from $1\frac{1}{2}$ and $1\frac{3}{4}$. So we will delete $1\frac{5}{8}$ from our time scale.



Thus our final $\mathbb{T}_p = \{0, \frac{1}{4}, \frac{1}{2}, 1, 1\frac{1}{4}, 1\frac{1}{2}, 2\}$.



As you can see $\mathbb{T}_p$ has the property that all of the points are at least h_{min} apart.

Again, we will calculate an approximation to the solution $y(t)$ by using the time scale version of Taylor's Theorem, Theorem 1.50. Suppose f is two times Δ -differentiable. Then for any $t \in \mathbb{T}$

$$\begin{aligned} y(t) &= \sum_{k=0}^1 h_k(t, t_i) y^{\Delta^k}(t_i) + \int_{t_i}^{\rho(t)} h_1(t, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \\ &= h_0(t, t_i) y(t_i) + h_1(t, t_i) y^{\Delta}(t_i) + \int_{t_i}^{\rho(t)} h_1(t, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau. \end{aligned}$$

Therefore

$$y(t_{i+1}) = y(t_i) + \mu_p(t_i) f(t_i, w_i) + \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau. \quad (2.12)$$

We will approximate the solution to the initial value problem (2.11) the same way we did in the isolated case. By

$$w_0 = y(a) = \alpha \quad (2.13)$$

$$w_{i+1} = w_i + \mu_p(t_i) f(t_i, w_i), \quad (2.14)$$

where $\mu_p(t_i)$ is the distance between t_i and t_{i+1} in $\mathbb{T}_p$. The integral

$$\int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \quad (2.15)$$

is our error term. We should note that $\rho(t_{i+1})$ is in the original time scale, $\mathbb{T}$, but not necessarily in secondary time scale, $\mathbb{T}_p$. Thus there are two cases to consider. The first case is $\rho(t_{i+1}) \in \mathbb{T}_p$. This is when $\rho(t_{i+1}) = t_i$ where we have a scattered point t_i that is at least h_{min} from t_{i+1} . We covered this case in the previous section.

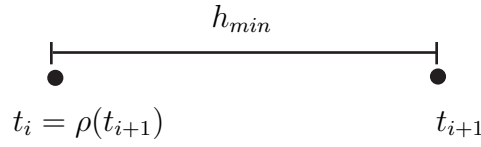


Figure 2.5: Case 1 Scattered points at least h_{min} apart

The second case is when $\rho(t_{i+1}) \notin \mathbb{T}_p$. This is the case where $\rho(t_{i+1}) \neq t_i$ which happens when t_{i+1} is a left dense point where $\rho(t_{i+1}) = t_{i+1}$ or when t_{i+1} is a left scattered point where previous point is not at least h_{min} from t_{i+1} .

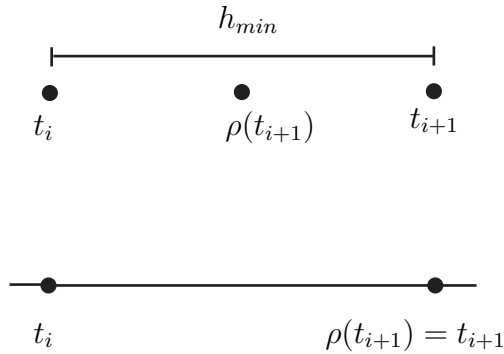


Figure 2.6: Case 2 Possibilities for $\rho(t_{i+1}) \neq t_i$

Lemma 2.5. Let $y^{\Delta^2} \in C_{rd}$ and let $\rho(t_{i+1}) \in \mathbb{T}_p$. Then

$$\int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau = 0.$$

Proof. If $\rho(t_{i+1}) \in \mathbb{T}_p$ then $\rho(t_{i+1}) = t_i$. Therefore

$$\int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau = \int_{t_i}^{t_i} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau = 0.$$

□

Lemma 2.6. Let y^{Δ^2} and y^Δ be integrable on $\mathbb{T}$, and let $|y^{\Delta^2}| \leq B$ on $\mathbb{T}$ then

$$\left| \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \right| \leq B \mu_p(t_i)^2$$

Proof.

$$\begin{aligned} \left| \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \right| &\leq \int_{t_i}^{\rho(t_{i+1})} |h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau)| \Delta\tau \\ &= \int_{t_i}^{\rho(t_{i+1})} |h_1(t_{i+1}, \sigma(\tau))| |y^{\Delta^2}(\tau)| \Delta\tau \\ &\leq B \int_{t_i}^{\rho(t_{i+1})} |h_1(t_{i+1}, \sigma(\tau))| \Delta\tau \\ &= B \int_{t_i}^{\rho(t_{i+1})} |t_{i+1} - \sigma(\tau)| \Delta\tau \\ &\leq B \int_{t_i}^{\rho(t_{i+1})} |t_{i+1} - t_i| \Delta\tau \\ &\leq B |t_{i+1} - t_i| \int_{t_i}^{\rho(t_{i+1})} \Delta\tau \\ &= B |t_{i+1} - t_i| (\rho(t_{i+1}) - t_i) \\ &\leq B |t_{i+1} - t_i| (t_{i+1} - t_i) \\ &\leq B \mu_p(t_i)^2 \end{aligned}$$

□

To estimate the error for our integration scheme for an arbitrary time scale we need to define a piecewise error function that considers both cases. Points with $\rho(t_{i+1}) \in \mathbb{T}_p$ do not contribute to the error yet they maintain the error passed to them. For points with $\rho(t_{i+1}) \notin \mathbb{T}_p$ we have the following theorem.

Theorem 2.7. Let $\mathbb{S} = \{a\} \cup \{t_i \mid \rho(t_{i+1}) \notin \mathbb{T}_p\}$ where a is the left end point of the time scale interval. Suppose $f \in C_{rd}$ and satisfies a Lipschitz condition with constant L on

$$D = \{(t, y) \mid a \leq t \leq b, -\infty < y < \infty\},$$

and let $y(t)$ denote the unique solution to the initial value problem (2.11)

$$y^\Delta(t) = f(t, y) \quad t \in \mathbb{T}^\kappa \text{ and } y(a) = \alpha,$$

such that

$$\left| \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \right| \leq B \mu_p(t_i)^2$$

for all $t_i \in \mathbb{S}$ where $|y^\Delta(t)| \leq B$ on $\mathbb{S}$. Let $w_0, w_1, w_2, \dots, w_N$ be the approximations generated by (2.6) and (2.5) for some positive integer N where N is the number of points in $\mathbb{S}$. Then for each $i = 0, 1, 2, \dots, N$,

$$|y(t_i) - w_i| \leq \frac{B \mu_p(t_i)}{L} (e^{(i+1)\mu_p(t_i)L} - 1).$$

Proof. When $i = 0$ the result is clear since $y(t_i) = w_i = \alpha$. From (2.12) and by letting $y_{i+1} = y(t_{i+1})$ and $y_i = y(t_i)$ we have for $i = 0, 1, \dots, N-1$

$$y_{i+1} = y_i + \mu_p(t_i) f(t_i, y_i) + \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau.$$

By (2.5)

$$w_{i+1} = w_i + \mu_p(t_i) f(t_i, w_i).$$

Thus

$$|y_{i+1} - w_{i+1}| \leq |y_i - w_i| + \mu_p(t_i) |f(t_i, y_i) - f(t_i, w_i)| + \left| \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \right|.$$

Since f is Lipschitz with respect to the second variable with Lipschitz constant L ,

$$\begin{aligned} |y_{i+1} - w_{i+1}| &\leq |y_i - w_i| + \mu_p(t_i) L |y_i - w_i| + \left| \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \right| \\ &= (1 - \mu_p(t_i) L) |y_i - w_i| + \left| \int_{t_i}^{\rho(t_{i+1})} h_1(t_{i+1}, \sigma(\tau)) y^{\Delta^2}(\tau) \Delta\tau \right|. \end{aligned}$$

The integral is bounded by $B\mu_p(t_i)^2$, therefore

$$|y_{i+1} - w_{i+1}| \leq (1 - \mu_p(t_i)L) |y_i - w_i| + B\mu_p(t_i)^2.$$

By Lemma 1.66, letting $s = \mu_p(t_i)L$, $t = B\mu_p(t_i)^2$, and $a_j = |y_j - w_j|$ for $j = 0, 1, 2, \dots, N$,

$$\begin{aligned} |y_{i+1} - w_{i+1}| &\leq e^{(i+1)\mu_p(t_i)L} \left(|y_0 - w_0| + \frac{B\mu_p(t_i)^2}{\mu_p(t_i)L} \right) \\ &\quad - \frac{B\mu_p(t_i)^2}{\mu_p(t_i)L}. \end{aligned}$$

Since $|y_0 - w_0| = 0$ and $\mu_p(t_i) > 0$,

$$\begin{aligned} |y_{i+1} - w_{i+1}| &\leq e^{(i+1)\mu_p(t_i)L} \left(\frac{B\mu_p(t_i)}{L} \right) - \frac{B\mu_p(t_i)}{L} \\ &= \frac{B\mu_p(t_i)}{L} (e^{(i+1)\mu_p(t_i)L} - 1). \end{aligned}$$

□

As with the original Euler's Method, the error in the time scale version goes to zero as $\mu_p(t_i)$ tends to zero. Although with Theorem 2.7 we can not accurately estimate the error for the last point in $\mathbb{S}_p$ if it is also the last point in $\mathbb{T}_p$, since by definition $\mu_p(t_N) = 0$ and therefore the estimated error would also be zero. Now we are ready to define the error function.

Definition 2.8 We will define the error function $E : \mathbb{T}_p \rightarrow \mathbb{R}$ by

$$E(t_i) = \begin{cases} E(t_{i-1}) & \text{if } \rho(t_{i+1}) \in \mathbb{T}_p \\ \frac{B\mu_p(t_i)}{L} (e^{(i+1)\mu_p(t_i)L} - 1) & \text{if } \rho(t_{i+1}) \notin \mathbb{T}_p \end{cases}$$

where n is the number of points in $\mathbb{S}$.

Now let us consider the following example.

Example 2.9 Let us use our general time scale approximation scheme to approximate the solution to

$$y^\Delta = \sigma(t)^2 + t\sigma(t) + t^2, \quad t \in \mathbb{T}^\kappa, \text{ and } y(0) = 0$$

given $\mathbb{T} = \{[0, 1], 1\frac{1}{2}, 2, 3, [3\frac{1}{2}, 5\frac{1}{2}]\}$.



Figure 2.7: The original time scale $\mathbb{T}$

We should note that $5\frac{1}{2} \in \mathbb{T}^\kappa$ since it is left dense. We will use ten points to estimate $y(t)$. This allows us to choose three points from our dense intervals. Since $[0, 1]$ has length one and $[3\frac{1}{2}, 5\frac{1}{2}]$ has length two we will pick one point from $[0, 1]$ and two from $[3\frac{1}{2}, 5\frac{1}{2}]$. We will do this by evenly dividing up each the intervals although an even step size is not necessary for the approximation. Thus $\mathbb{T}_p = \{0, \frac{1}{2}, 1, 1\frac{1}{2}, 2, 3, 3\frac{1}{2}, 3\frac{5}{6}, 4\frac{1}{6}, 5\frac{1}{2}\}$ and $h_{min} = \frac{1}{3}$.



Figure 2.8: The partitioned time scale $\mathbb{T}_p$

In this example we will estimate $y(t)$ by

$$w_0 = 0$$

$$w_{i+1} = w_i + \mu_p(t_i)(\sigma(t_i)^2 + t_i\sigma(t_i) + t_i^2)$$

where $\mu_p(t_i)$ is generated by $\mathbb{T}_p$ and $\sigma(t_i)$ is extracted from $\mathbb{T}$. As before, we know the solution to this initial value problem is $y(t) = t^3$.

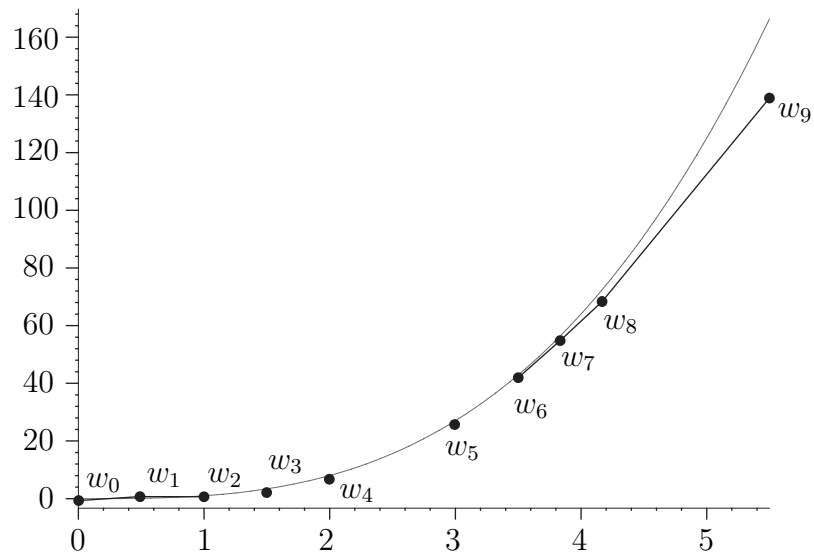


Figure 2.9: Arbitrary Time Scale Approximation of $y = t^3$

Table 2.3: Arbitrary Time Scale Version of Euler's Method Comparison for $y(t) = t^3$

t_i	w_i	$y_i = t^3$	$ y_i - w_i $
0	0.000000000	0.000000000	0.00000000
$\frac{1}{2}$	0.000000000	0.125000000	0.12500000
1	0.375000000	1.000000000	0.62500000
$1\frac{1}{2}$	2.750000000	3.375000000	0.62500000
2	7.375000000	8.000000000	0.62500000
3	26.37500000	27.00000000	0.62500000
$3\frac{1}{2}$	42.25000000	42.87500000	0.62500000
$3\frac{5}{6}$	54.50000000	56.32870370	1.82870370
$4\frac{1}{6}$	69.19444444	72.33796296	3.14351852
$5\frac{1}{2}$	138.6388889	166.3750000	27.7361111

In Table 2.3 we can see the comparison between our approximation and the actual values of $y(t)$. The table confirms the theory we developed. The table shows us that isolated points and left scattered points do not increase the cumulative error.

By Theorem 2.7 we can calculate an estimated error. The function $|y^{\Delta^2}(t)|$ is bounded by 25 on our time scale interval and $f(t, y)$ has a Lipschitz constant of 1 since

$$\begin{aligned}
 |f(t, y_1) - f(t, y_2)| &\leq |(\sigma(t_i)^2 + t_i\sigma(t_i) + t_i^2) - (\sigma(t_i)^2 + t_i\sigma(t_i) + t_i^2)| \\
 &= 0 \leq 1|y_1 - y_2|.
 \end{aligned}$$

Table 2.4: Arbitrary Time Scale Numerical Integration Error Comparison for $y(t) = t^3$

t_i	Actual Error	E(t)
$\frac{1}{2}$	0.12500000	21.47852285
1	0.62500000	43.52111338
$1\frac{1}{2}$	0.62500000	43.52111338
2	0.62500000	43.52111338
3	0.62500000	43.52111338
$3\frac{1}{2}$	0.62500000	43.52111338
$3\frac{5}{6}$	1.82870370	23.28056578
$4\frac{1}{6}$	3.14351852	26159.06648

As we can see in the Table 2.4, the error estimation does not closely bound the error. Although when $\mu(t_i)$ is small the error corrects itself. These results imply that there may be a better error approximation. Perhaps by using the time scale exponential in the error estimate rather than the real exponential since the real exponential is an upper bound for the absolute value of the time scale exponential.

Since the time scale exponential function is the only the solution to the initial value problem

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

its natural to want to use it to test our integration scheme.

Example 2.10 Recall in Example 1.57 we developed the time scale exponential for $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$,

$$e_p(t, t_0) = y(t_0) \prod_{n=1}^{\frac{1}{1-t}-1} \left(\frac{p}{n(n+1)} + 1 \right).$$

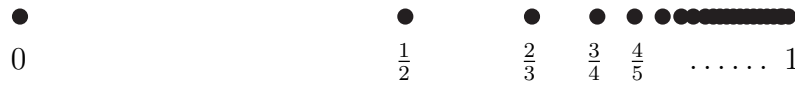


Figure 2.10: $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$

Then in Example 1.67 we let $p \equiv 1$ and saw as $n \rightarrow \infty$,

$$\prod_{n=1}^{\infty} \left(\frac{1}{n(n+1)} + 1 \right) \rightarrow \frac{\sin \left(\pi \left(\frac{1}{2} + \frac{1}{2}i\sqrt{3} \right) \right)}{\pi} \approx 2.428189793.$$

Let us use our general time scale approximation scheme to approximate a solution to

$$y^{\Delta}(t) = y(t), \quad y(0) = 1$$

on the time scale $\mathbb{T} = \{1 - \frac{1}{n}, n \in \mathbb{N}\} \cup \{1\}$. We know the solution is

$$e_1(t, 0) = \prod_{n=1}^{\frac{1}{1-t}-1} \left(\frac{1}{n(n+1)} + 1 \right).$$

Since $\mathbb{T}$ has infinitely many points as we approach 1, we will select the following partition for our approximation $\mathbb{T}_p = \{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{8}{9}, 1\}$ where $h_{min} = \frac{1}{72}$.

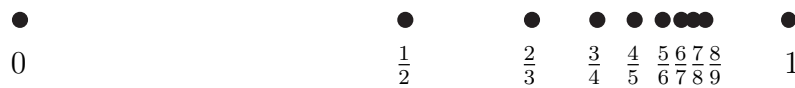


Figure 2.11: $\mathbb{T}_p = \{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{8}{9}, 1\}$

In this example we estimate $y(t)$ by

$$w_0 = 1$$

$$w_{i+1} = w_i + \mu_p(t_i)w_i$$

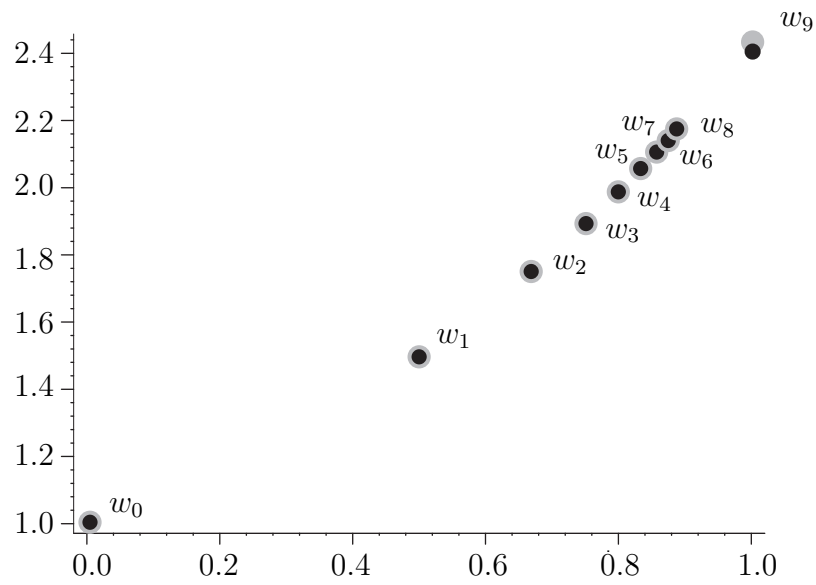


Figure 2.12: Arbitrary Time Scale Approximation of $y = e_1(t, 1)$

Table 2.5: Arbitrary Time Scale Version of Euler's Method Comparison for $y(t) = e_p(t, 1)$

t_i	w_i	$y_i = e_p(t, 1)$	$ y_i - w_i $
0	1.000000000	1.000000000	0.000000000
$\frac{1}{2}$	1.500000000	1.500000000	0.000000000
$\frac{2}{3}$	1.750000000	1.750000000	0.000000000
$\frac{3}{4}$	1.895833333	1.895833333	0.000000000
$\frac{4}{5}$	1.990625000	1.990625000	0.000000000
$\frac{5}{6}$	2.056979167	2.056979167	0.000000000
$\frac{6}{7}$	2.105954861	2.105954861	0.000000000
$\frac{7}{8}$	2.143561198	2.143561198	0.000000000
$\frac{8}{9}$	2.173332881	2.173332881	0.000000000
1	2.414814312	2.428189793	0.013375481

All of the left scattered points do not contribute to the error and the last point which is left dense does create error. Although Theorem 2.7 does not predict this error. We could estimate the error at the last point of approximation if the time scale interval $[a, b]$ is embedded $\mathbb{T}^\kappa$. Then we would have a realistic value for $\mu(t_N)$. However we do have an error approximation for the last point if it is left scattered. Since we know left scattered points maintain the error passed to them by the previous point.

2.3 CONCLUSION

In the previous chapter, we developed a numerical integration scheme for time scales. By following the development of Euler's Method, we developed a scheme that approximates the solution to

$$y^\Delta(t) = f(t, y) \quad t \in \mathbb{T}^\kappa \text{ and } y(a) = \alpha.$$

When $\mathbb{T} = \mathbb{R}$ this is the same problem posed for Euler's Method. For any time scale $\mathbb{T}$, the time scale may have to be reduced to the scattered time scale $\mathbb{T}_p$ where all the points are at

least some step size h apart to insure a finite number of approximation points. In the isolated case we found that the error term cancelled and our scheme is accurate. For the arbitrary time scale case, our scheme has the convenient property that left scattered points maintain the accuracy that was passed to them by the previous point.

Although our numerical integration scheme works really well for time scales, how practical is it in the real world? The problem is finding the Δ -derivative for a given set of data. In Example 2.9, the set up of the example seemed a little contrived since the solution $y(t)$ showed up explicitly in the Δ -derivative. This is not always the case as in Example 2.1.

For future projects one could develop a more accurate error bound than Theorem 2.7 perhaps by using the time scale exponential or by developing a theorem similar to Theorem 1.70. Also the ∇ -numerical integration scheme could be developed. The ∇ -scheme would probably be more natural for numerical analysis applications since it is based on the backwards operator. It would also allow us to estimate the error at the last point in the time scale. In addition higher order approximation schemes for time scales such as Runge-Kutta could also be developed.

BIBLIOGRAPHY

- [1] M. Bohner and A. Peterson. *Advances in Dynamic Equations on Time Scales*. Birkhäuser, Boston, 2002.
- [2] M. Bohner and A. Peterson. *Dynamic Equations on Time Scales*. Birkhäuser, Boston, 2001.
- [3] R. Burden and J. Faires. *Numerical Analysis*. 7th ed. Brooks/Cole, Pacific Grove, 2001.
- [4] J. Conway. *Functions of One Complex Variable*. Springer-Verlag, New York, 1975.
- [5] T. Gard and J. Hoffacker. Asymptotic Behavior of Natural Growth on Time Scales, *Dynamic Systems and Applications*, 12(1-2):131-148, 2003.
- [6] S. Hilger. Ein Maßkettenkalkül mit Anwendung auf Zentrumsmannigfaltigkeiten. PhD thesis, Universität Würzburg, 1988.
- [7] A. Thomas. Transforms on Time Scales. Masters thesis, University of Georgia, 2003.