



Scholars' Mine

[Doctoral Dissertations](#)

[Student Theses and Dissertations](#)

Fall 2011

Probability theory on time scales and applications to finance and inequalities

Thomas Matthews

Follow this and additional works at: https://scholarsmine.mst.edu/doctoral_dissertations

 Part of the [Mathematics Commons](#)

Department: Mathematics and Statistics

Recommended Citation

Matthews, Thomas, "Probability theory on time scales and applications to finance and inequalities" (2011). *Doctoral Dissertations*. 2241.

https://scholarsmine.mst.edu/doctoral_dissertations/2241

This thesis is brought to you by Scholars' Mine, a service of the Missouri S&T Library and Learning Resources. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

PROBABILITY THEORY ON TIME SCALES AND
APPLICATIONS TO FINANCE AND INEQUALITIES

by

THOMAS MATTHEWS

A DISSERTATION

Presented to the Faculty of the Graduate School of the
MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

2011

Approved by

Dr. Martin Bohner, Advisor

Dr. Elvan Akın-Bohner

Dr. Ilene H. Morgan

Dr. Xuerong Wen

Dr. Gregory Gelles

Copyright 2011
Thomas Matthews
All Rights Reserved

ABSTRACT

In this dissertation, the recently discovered concept of time scales is applied to probability theory, thus unifying discrete, continuous and many other cases. A short introduction to the theory of time scales is provided.

Following this preliminary overview, the moment generating function is derived using a Laplace transformation on time scales. Various unifications of statements and new theorems in statistics are shown.

Next, distributions on time scales are defined and their properties are studied. Most of the derived formulas and statements correspond exactly to those from discrete and continuous calculus and extend the applicability to many other cases. Some theorems differ from the ones found in the literature, but improve and simplify their handling.

Finally, applications to finance, economics and inequalities of Ostrowski and Grüss type are presented. Throughout this paper, our results are compared to their well known counterparts in discrete and continuous analysis and many examples are given.

ACKNOWLEDGMENT

First and foremost, I would like to express my deepest gratitude to my advisor Dr. Martin Bohner for his support and guidance to develop this thesis. Over the last years, he put big effort into explaining complex mathematical and economic topics related to my research work. I would like to thank Dr. Bohner for his enthusiasm exposing me to the concept of time scales. Moreover, I want to thank Dr. Elvan Akin-Bohner, Dr. Ilene Morgan, Dr. Xuerong Wen and Dr. Gregory Gelles for their support and help during my academic stay at Missouri University of Science and Technology. They were always willing to share their expertise regarding my research and extended my knowledge in various lectures, where I had the privilege to have them as professors. I would like to express my humble gratitude to Dr. Adnan Tuna for his impact on several projects. On a special note, I would like to thank the University of Ulm and Middle East Technical University in Ankara for making it possible to spend an extended research period on their campuses. Particular thanks and gratitude go to my parents Gisela and Wolfgang Matthews for their continuous support from Germany. Last, but not least, I would like to thank all German exchange students, especially Julius Heim, Dr. Benedikt Wundt, Karl Ulrich, Mathias Göggel, Christian Keller, Matthias Noller, Sabrina Streipert and Jens-Uwe Reitingner, my office- and roommates and all American and international friends I had the honor to meet during the last years.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
ACKNOWLEDGMENT	iv
LIST OF ILLUSTRATIONS	vii
LIST OF TABLES	viii
 SECTION	
1. INTRODUCTION	1
2. TIME SCALES ESSENTIALS	3
2.1. BASIC DEFINITIONS.....	3
2.2. DIFFERENTIATION AND INTEGRATION.....	6
2.3. EXPONENTIAL FUNCTION	12
2.4. LOGARITHM	15
2.5. DYNAMIC INEQUALITIES	16
2.6. SHIFT, CONVOLUTION AND LAPLACE TRANSFORM	17
3. MOMENT GENERATING FUNCTION	20
4. CUMULANT GENERATING FUNCTION	27
5. DISTRIBUTIONS	33
5.1. UNIFORM DISTRIBUTION.....	33
5.2. EXPONENTIAL DISTRIBUTION	39
5.3. GAMMA DISTRIBUTION.....	45
6. CUMULATIVE DISTRIBUTION FUNCTION	58
7. FURTHER PROPERTIES AND ENTROPY	70
7.1. THE SUM OF INDEPENDENT RANDOM VARIABLES	70
7.2. THE MEMORYLESS PROPERTY.....	73
7.3. ENTROPY.....	78
8. ECONOMIC APPLICATIONS	90
8.1. INTEREST RATES AND NET PRESENT VALUE	90
8.2. HAZARD RATES	98
8.3. PRICING OF CREDIT DEFAULT SWAPS	102
8.4. CREDIT SPREADS	108
9. OSTROWSKI INEQUALITIES INVOLVING EXPECTED VALUES	113
10. DIAMOND-ALPHA GRÜSS TYPE INEQUALITIES	130

10.1. INTRODUCTION	130
10.2. GENERAL DEFINITIONS	131
10.3. THE WEIGHTED DIAMOND-ALPHA GRÜSS INEQUALITY	133
10.4. THE CASE WHEN BOTH MAPPINGS ARE LIPSCHITZIAN	136
10.5. THE CASE WHEN F IS M - G -LIPSCHITZIAN	139
11. WEIGHTED OSTROWSKI–GRÜSS INEQUALITIES	143
11.1. INTRODUCTION	143
11.2. OSTROWSKI–GRÜSS INEQUALITIES IN L^∞ -NORM	144
11.3. OSTROWSKI–GRÜSS INEQUALITIES IN L^P -NORM	149
12. OSTROWSKI AND GRÜSS TYPE INEQUALITIES	152
12.1. INTRODUCTION	152
12.2. AUXILIARY RESULTS AND ASSUMPTIONS	154
12.3. PACHPATTE’S FIRST THEOREM ON TIME SCALES	156
12.4. PACHPATTE’S SECOND THEOREM ON TIME SCALES	160
12.5. PACHPATTE’S THIRD THEOREM ON TIME SCALES	165
BIBLIOGRAPHY	169
VITA	174

LIST OF ILLUSTRATIONS

Figure	Page
5.1 Uniform distribution	38
5.2 Exponential distribution	45
5.3 Gamma distribution (continuous case)	50
5.4 Gamma distribution (discrete case)	50
8.1 Premium payments of a credit default swap	104

LIST OF TABLES

Table	Page
2.1 Classification of Points	4
2.2 Examples of Time Scales	13

1. INTRODUCTION

Two of the main objectives in mathematics are simplification and unification. In 1988, Stefan Hilger introduced the concept of time scales in his dissertation “Ein Maßkettenkalkül mit Anwendung auf Zentrumsmannigfaltigkeiten” [41]. In order to provide the reader with the necessary definitions and theorems of time scales, this thesis summarizes important concepts in the second section, following the books of Martin Bohner and Allan Peterson “Dynamic equations on time scales” [22] and “Advances in dynamic equations on time scales” [23].

The main goal of this thesis is to establish the basics of probability theory on time scales and apply those results to finance, economics and inequalities. Thus, well known definitions, properties and theorems for discrete, continuous and many other cases will be unified. With the help of the Laplace transformation, the moment and cumulant generating functions are derived, setting the stage for concepts like expected value, variance, independence and entropy.

Furthermore, uniform, exponential and gamma distributions are defined and studied within the theory of time scales. The results are put into context with existing discrete and continuous distributions. The time scales exponential distribution corresponds to the known exponential distribution in the continuous and to the geometric distribution in the discrete setting. The time scales gamma distribution has the continuous Erlang distribution and the discrete negative binomial distribution as its counterparts. Due to the new definition of moments on time scales, some properties of distributions do not coincide with the ones found for discrete cases in the literature, but have the advantage of stating an easier and more approachable result for all time scales.

Next, the basic concepts of interest rates and net present value are presented. Financial applications also include hazard rates and the pricing of credit default swaps. The latter has become of great interest during the last decades, due to the increased risk of default for countries on the one hand and companies on the other. Throughout, the theory is derived on time scales and several examples are provided.

In 1938, Alexander Ostrowski first proved a formula to estimate the absolute deviation of a differentiable function from its integral mean. The so-called Ostrowski inequality

holds and is shown in [55]

$$\left| f(t) - \frac{1}{b-a} \int_a^b f(s) ds \right| \leq \sup_{a < t < b} |f'(t)|(b-a) \left[\frac{(t - \frac{a+b}{2})^2}{(b-a)^2} + \frac{1}{4} \right].$$

The time scales equivalent was shown by Martin Bohner and Thomas Matthews [19]. In 1935, Gerhard Grüss introduced an inequality describing lower and upper bounds for the difference of a the integral of the product of two functions from the product of the integrals. As shown in [40], we have

$$\left| \frac{1}{b-a} \int_a^b f(x)g(x) dx - \frac{1}{(b-a)^2} \int_a^b f(x) dx \int_a^b g(x) dx \right| \leq \frac{1}{4}(M_1 - m_1)(M_2 - m_2),$$

where

$$m_1 \leq f(s) \leq M_1, \quad m_2 \leq g(s) \leq M_2.$$

In [18], Martin Bohner and Thomas Matthews derived the time scales version of that inequality.

The last sections of this thesis discuss Ostrowski type inequalities on time scales using the previously derived properties from probability theory. Multiple inequalities involving expected values are presented. In addition, joint work with Martin Bohner and Adnan Tuna is given, exploring the concept of diamond-alpha theory on time scales for Grüss inequalities. Furthermore, collaborative work with Martin Bohner and Adnan Tuna, discussing Ostrowski–Grüss like and Ostrowski inequalities for two functions, is provided. Finally, research together with Martin Bohner and Elvan Akın–Bohner on Ostrowski–Grüss like and Ostrowski inequalities for three functions is presented.

2. TIME SCALES ESSENTIALS

In 1988, Stefan Hilger introduced the theory of time scales in his Ph.D. thesis “Ein Maßkettenkalkül mit Anwendung auf Zentrums Mannigfaltigkeiten”. Since then many authors like M. Bohner, A. Peterson, R. Agarwal and G. Guseinov have extended this theory and various results can be found in recent books and papers. Time scales unify the fields of discrete and continuous analysis and extend them to numerous other cases. So one could say that unification and extension are the two main features of the time scales calculus. Moreover, the time scales theory has an incredible potential for applications in economics, finance, physics and biology.

This section is meant to be an introduction to time scales. Summarizing the books of Bohner and Peterson [22, 23], basic definitions, properties and theorems will be presented, which are needed throughout this paper. Various examples to different time scales will be provided.

2.1. BASIC DEFINITIONS

Definition 1. *A time scale is an arbitrary nonempty closed subset of the real numbers.*

The most important examples of time scales are \mathbb{R} (continuous case) and \mathbb{Z} (discrete case). Other examples of time scales considered in more detail in this thesis, are

$$q^{\mathbb{N}_0} := \{q^k \mid k \in \mathbb{N}_0\}, \quad \text{where } q > 1$$

(the quantum calculus case) and

$$h\mathbb{Z} := \{hk \mid k \in \mathbb{N}_0\}, \quad \text{where } h > 0.$$

On the other hand, \mathbb{Q} , $\mathbb{R} \setminus \mathbb{Q}$ and $(0, 1)$ are not time scales since those domains fail to be a closed subset of the real numbers.

Definition 2. *If \mathbb{T} is a time scale, then we define the forward jump operator $\sigma : \mathbb{T} \rightarrow \mathbb{T}$ by*

$$\sigma(t) := \inf \{s \in \mathbb{T} \mid s > t\} \quad \text{for all } t \in \mathbb{T},$$

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

$$\rho(t) := \sup \{s \in \mathbb{T} \mid s < t\} \quad \text{for all } t \in \mathbb{T},$$

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

$$\mu(t) := \sigma(t) - t \quad \text{for all } t \in \mathbb{T}.$$

Furthermore, for a function $f : \mathbb{T} \rightarrow \mathbb{R}$, we define

$$f^\sigma(t) = f(\sigma(t)) \quad \text{for all } t \in \mathbb{T}$$

and

$$f^\rho(t) = f(\rho(t)) \quad \text{for all } t \in \mathbb{T}.$$

In this definition we use $\inf \emptyset = \sup \mathbb{T}$ (i.e., $\sigma(t) = t$ if t is the maximum of \mathbb{T}) and $\sup \emptyset = \inf \mathbb{T}$ (i.e., $\rho(t) = t$ if t is the minimum of \mathbb{T}). Moreover, this definition allows us to characterize every point in a time scale as displayed in Table 2.1.

Table 2.1. Classification of Points

t right-scattered	$t < \sigma(t)$
t right-dense	$t = \sigma(t)$
t left-scattered	$\rho(t) < t$
t left-dense	$\rho(t) = t$
t isolated	$\rho(t) < t < \sigma(t)$
t dense	$\rho(t) = t = \sigma(t)$

The following example applies these first definitions to the important time scales mentioned after Definition 1.

Example 1. *Let us consider these four cases.*

(i) *For $\mathbb{T} = \mathbb{R}$, we have*

$$\sigma(t) = \inf \{s \in \mathbb{R} \mid s > t\} = t,$$

$$\rho(t) = \sup \{s \in \mathbb{R} \mid s < t\} = t,$$

$$\mu(t) = t - t = 0,$$

all $t \in \mathbb{R}$ are dense.

(ii) *For $\mathbb{T} = \mathbb{Z}$, we have*

$$\sigma(t) = \inf \{s \in \mathbb{Z} \mid s > t\} = t + 1,$$

$$\rho(t) = \sup \{s \in \mathbb{Z} \mid s < t\} = t - 1,$$

$$\mu(t) = (t + 1) - t = 1,$$

all $t \in \mathbb{Z}$ are isolated.

(iii) *For $\mathbb{T} = q^{\mathbb{N}_0}$, we have*

$$\sigma(t) = \inf \{s \in q^{\mathbb{N}_0} \mid s > t\} = qt,$$

$$\rho(t) = \sup \{s \in q^{\mathbb{N}_0} \mid s < t\} = \frac{t}{q} \quad \text{for } t > 1,$$

$$\mu(t) = qt - t = (q - 1)t,$$

all $t \in q^{\mathbb{N}_0}$ are isolated.

(iv) *For $\mathbb{T} = h\mathbb{Z}$, we have*

$$\sigma(t) = \inf \{s \in h\mathbb{Z} \mid s > t\} = t + h,$$

$$\rho(t) = \sup \{s \in h\mathbb{Z} \mid s < t\} = t - h,$$

$$\mu(t) = (t + h) - t = h,$$

all $t \in h\mathbb{Z}$ are isolated.

2.2. DIFFERENTIATION AND INTEGRATION

Definition 3. The set \mathbb{T}^κ for a time scale \mathbb{T} is defined as follows: If t has a left-scattered maximum m , then $\mathbb{T}^\kappa = \mathbb{T} \setminus \{m\}$ and otherwise $\mathbb{T}^\kappa = \mathbb{T}$. In summary:

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

Definition 4. Let $f : \mathbb{T} \rightarrow \mathbb{R}$ and $t \in \mathbb{T}^\kappa$. Then we define $f^\Delta(t)$ to be the number (if it exists) such that for all $\varepsilon > 0$, there exists $U = (t - \delta, t + \delta) \cap \mathbb{T}$ for some $\delta > 0$ such that

$$|[f(\sigma(t)) - f(s)] - f^\Delta(t) [\sigma(t) - s]| \leq \varepsilon |\sigma(t) - s| \quad \text{for all } s \in U.$$

We call $f^\Delta(t)$ the delta (or Hilger) derivative of f at t . Furthermore, f is called delta differentiable on \mathbb{T}^κ if $f^\Delta(t)$ exists for all $t \in \mathbb{T}^\kappa$, and f^Δ is called the delta derivative of f .

Theorem 1. Let $f : \mathbb{T} \rightarrow \mathbb{R}$ be a function and $t \in \mathbb{T}^\kappa$.

- (i) If f is differentiable at t , then f is continuous at t .
- (ii) If f is continuous at t and t is right-scattered, then f is differentiable at t with

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

- (iii) If t is right-dense, then f is differentiable at t iff the limit

$$\lim_{s \rightarrow t} \frac{f(t) - f(s)}{t - s}$$

exists as a finite number, and in this case

$$f^\Delta(t) = \lim_{s \rightarrow t} \frac{f(t) - f(s)}{t - s}.$$

- (iv) If f is differentiable at t , then

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

Proof. See [22, Theorem 1.16]. □

Theorem 2. Let $f, g : \mathbb{T} \rightarrow \mathbb{R}$ be differentiable at $t \in \mathbb{T}^\kappa$.

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

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

(ii) For any constant c , $cf : \mathbb{T} \rightarrow \mathbb{R}$ is differentiable at t with

$$(cf)^\Delta(t) = cf^\Delta(t).$$

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

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

(iv) 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. See [22, Theorem 1.20]. □

Theorem 3. Let c be a constant, $m \in \mathbb{N}$ and $f(t) = (t - c)^m$. Then

$$f^\Delta(t) = \sum_{\nu=0}^{m-1} (\sigma(t) - c)^\nu (t - c)^{m-1-\nu}.$$

Proof. See [22, Theorem 1.24]. □

Therefore for $f(t) = t$, the derivative is $f^\Delta(t) = 1$ and for $f(t) = t^2$, we get the derivative $f^\Delta(t) = \sigma(t) + t$.

Example 2. Let f be differentiable.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$f^\Delta(t) = f'(t).$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$f^\Delta(t) = \Delta f(t),$$

where the backward difference operator Δ is defined as usual by $\Delta f(t) = f(t+1) - f(t)$.

(iii) If $\mathbb{T} = q^{\mathbb{N}_0}$, then

$$f^\Delta(t) = \frac{f(qt) - f(t)}{(q-1)t}.$$

(iv) If $\mathbb{T} = h\mathbb{Z}$, then

$$f^\Delta(t) = \frac{f(t+h) - f(t)}{h}.$$

Definition 5. A function $f : \mathbb{T} \rightarrow \mathbb{R}$ is called regulated if its right-sided limits exist (finite) at all right-dense points in \mathbb{T} and its left-sided limits exist (finite) at all left-dense points in \mathbb{T} .

Definition 6. A function $f : \mathbb{T} \rightarrow \mathbb{R}$ is called rd-continuous (denoted by $f \in C_{\text{rd}}$) if it is continuous at right-dense points of \mathbb{T} and its left-sided limits exist (finite) at left-dense points of \mathbb{T} .

Theorem 4 (Existence of Antiderivatives). Let f be rd-continuous and $t_0 \in \mathbb{T}$. Then f has an antiderivative (denoted by F) defined by

$$F(t) = \int_{t_0}^t f(\tau) \Delta \tau \quad \text{for } t \in \mathbb{T}. \quad (1)$$

Proof. See [22, Theorem 1.74]. □

Therefore for rd-continuous functions f , we have

$$\int_a^b f(\tau) \Delta \tau = F(b) - F(a), \quad (2)$$

where $F^\Delta = f$.

Theorem 5. Let f be rd-continuous and $t \in \mathbb{T}^\kappa$. Then

$$\int_t^{\sigma(t)} f(\tau) \Delta \tau = \mu(t) f(t). \quad (3)$$

Proof. See [22, Theorem 1.75]. □

Example 3. *Integration for the most important time scales is done as follows. Let $a, b \in \mathbb{T}$ and $f \in C_{rd}$.*

(i) *If $\mathbb{T} = \mathbb{R}$ and if f is Riemann integrable, then*

$$\int_a^b f(t)\Delta t = \int_a^b f(t)dt.$$

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$\int_a^b f(t)\Delta t = \sum_{t=a}^{b-1} f(t), \quad \text{where } a < b.$$

(iii) *If $\mathbb{T} = q^{\mathbb{N}_0}$, then*

$$\int_{q^m}^{q^n} f(t)\Delta t = \sum_{i=m}^{n-1} \mu(q^i) f(q^i) = (q-1) \sum_{i=m}^{n-1} q^i f(q^i), \quad \text{where } m < n.$$

(iv) *If $\mathbb{T} = h\mathbb{Z}$, then*

$$\int_a^b f(t)\Delta t = \sum_{t=\frac{a}{h}}^{\frac{b}{h}-1} f(th)h, \quad \text{where } a < b.$$

(v) *If $[a, b]$ consists only of isolated points, then*

$$\int_a^b f(t)\Delta t = \sum_{t \in [a, b]} \mu(t) f(t), \quad \text{where } a < b.$$

Proof. See [22, Theorem 1.79]. □

Theorem 6. *Let f, g be rd-continuous, $a, b, c \in \mathbb{T}$ and $\alpha \in \mathbb{R}$. Then*

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

$$(ii) \int_a^b [\alpha f(t)]\Delta t = \alpha \int_a^b f(t)\Delta t,$$

$$(iii) \int_a^b f(t)\Delta t = - \int_b^a f(t)\Delta t,$$

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

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

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

$$(vii) \int_a^a f(t)\Delta t = 0,$$

(viii) if $|f(t)| \leq g(t)$, then

$$\left| \int_a^b f(t)\Delta t \right| \leq \int_a^b g(t)\Delta t,$$

(ix) if $f(t) \geq 0$ for all $a \leq t < b$, then $\int_a^b f(t)\Delta t \geq 0$.

Proof. See [22, Theorem 1.77]. □

Definition 7. Let $t \in \mathbb{C}$ and $k \in \mathbb{Z}$. We define the factorial function $t^{\underline{k}}$ as follows.

(i) If $k \in \mathbb{N}$, then

$$t^{\underline{k}} = t(t-1) \cdots (t-k+1).$$

(ii) If $k = 0$, then

$$t^{\underline{0}} = 1.$$

(iii) If $-k \in \mathbb{N}$, then

$$t^{\underline{k}} = \frac{1}{(t+1)(t+2) \cdots (t-k)},$$

for $t \neq -1, -2, \dots, k$.

For all $t, k \in \mathbb{C}$, we have

$$t^{\underline{k}} = \frac{\Gamma(t+1)}{\Gamma(t-k+1)}.$$

Since the antiderivative of t is not necessarily $\frac{t^2}{2}$, we try the approach of defining the following functions.

Definition 8. Let $g_k, h_k : \mathbb{T}^2 \rightarrow \mathbb{R}$, $k \in \mathbb{N}_0$ be defined by

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

and then recursively by

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

and

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

Therefore, we have

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

and similarly

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

Moreover, we can state the properties

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

and

$$g_2(t, s) = \int_s^t (\sigma(\tau) - s) \Delta\tau, \quad h_2(t, s) = \int_s^t (\tau - s) \Delta\tau.$$

Example 4. *Let us consider the following examples.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$g_2(t, s) = h_2(t, s) = \int_s^t (\tau - s) d\tau = \frac{(t - s)^2}{2}.$$

More generally, we have

$$g_k(t, s) = h_k(t, s) = \frac{(t - s)^k}{k!}.$$

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$h_2(t, s) = \int_s^t (\tau - s) \Delta\tau = \frac{(\tau - s)(\tau - s - 1)}{2} \Big|_s^t = \binom{t - s}{2}$$

since

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

More generally, we have

$$h_k(t, s) = \binom{t - s}{k}$$

and

$$g_k(t, s) = \frac{(t - s + k - 1)^{\underline{k}}}{k!}.$$

(iii) If $\mathbb{T} = q^{\mathbb{N}_0}$, then by [22, Example 1.104]

$$h_k(t, s) = \prod_{\nu=0}^{k-1} \frac{t - q^\nu s}{\sum_{\mu=0}^{\nu} q^\mu}$$

and thus

$$h_2(t, s) = \frac{(t - s)(t - qs)}{1 + q}.$$

Table 2.2 gives a short summary of the considered topics.

2.3. EXPONENTIAL FUNCTION

Definition 9. For $h > 0$, we define the Hilger complex numbers as

$$\mathbb{C}_h := \left\{ z \in \mathbb{C} : z \neq -\frac{1}{h} \right\}$$

and the strip

$$\mathbb{Z}_h := \left\{ z \in \mathbb{C} : -\frac{\pi}{h} < \text{Im}(z) \leq \frac{\pi}{h} \right\},$$

Table 2.2. Examples of Time Scales

Time Scale \mathbb{T}	\mathbb{R}	\mathbb{Z}	$q^{\mathbb{N}_0}$
Forward jump operator $\sigma(t)$	t	$t + 1$	qt
Backward jump operator $\rho(t)$	t	$t - 1$	$\frac{t}{q}$
Graininess $\mu(t)$	0	1	$(q - 1)t$
$h_2(t, s)$	$\frac{(t-s)^2}{2}$	$\binom{t-s}{2}$	$\frac{(t-s)(t-qs)}{1+q}$

where $\mathbb{Z}_h := \mathbb{C}$ for $h = 0$.

Definition 10. For $h > 0$, we define the cylinder transformation $\xi_h : \mathbb{C}_h \rightarrow \mathbb{Z}_h$ by

$$\xi_h(z) = \frac{1}{h} \text{Log}(1 + zh),$$

where Log is the principal logarithm function.

Definition 11. We define a function $p : \mathbb{T} \rightarrow \mathbb{R}$ to be regressive provided that for all $t \in \mathbb{T}^\kappa$

$$1 + \mu(t)p(t) \neq 0.$$

The set of regressive and rd-continuous functions $f : \mathbb{T} \rightarrow \mathbb{R}$ will be denoted by

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

Moreover, we define the set \mathcal{R}^+ of all positively regressive elements of \mathcal{R} by

$$\mathcal{R}^+ = \{p \in \mathcal{R} : 1 + \mu(t)p(t) > 0 \quad \forall t \in \mathbb{T}^\kappa\}.$$

Definition 12. Let $p \in \mathcal{R}$ and ξ_h the cylinder transformation. Then we define the exponential function by

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

Definition 13. If $p \in \mathcal{R}$, then the first order linear dynamic equation

$$y^\Delta = p(t)y \tag{4}$$

is called regressive.

Theorem 7. Suppose (4) is regressive and fix $t_0 \in \mathbb{T}$. Then $e_p(\cdot, t_0)$ is the solution of the initial value problem

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

on \mathbb{T} .

Proof. See [22, Theorem 2.33]. □

Definition 14. Let $p, q \in \mathcal{R}$. We define the circle plus addition \oplus by

$$(p \oplus q)(t) := p(t) + q(t) + \mu(t)p(t)q(t) \quad \text{for all } t \in \mathbb{T}^\kappa,$$

the circle minus subtraction \ominus by

$$(p \ominus q)(t) := (p \oplus (\ominus q))(t) \quad \text{for all } t \in \mathbb{T}^\kappa,$$

and

$$(\ominus p)(t) := -\frac{p(t)}{1 + \mu(t)p(t)} \quad \text{for all } t \in \mathbb{T}^\kappa.$$

Theorem 8. Let $p, q \in \mathcal{R}$ and $t, s, r \in \mathbb{T}$. Then

- (i) $e_0(t, s) \equiv 1$ and $e_p(t, t) \equiv 1$;
- (ii) $e_p(\sigma(t), s) = (1 + \mu(t)p(t))e_p(t, s)$;
- (iii) $\frac{1}{e_p(t, s)} = e_{\ominus p}(t, s)$;
- (iv) $e_p(t, s) = \frac{1}{e_p(s, t)} = e_{\ominus p}(s, t)$;
- (v) $e_p(t, s)e_p(s, r) = e_p(t, r)$;
- (vi) $e_p(t, s)e_q(t, s) = e_{p \oplus q}(t, s)$;

$$(vii) \frac{e_p(t,s)}{e_q(t,s)} = e_{p \ominus q}(t, s);$$

$$(viii) \left(\frac{1}{e_p(\cdot, s)} \right)^\Delta = -\frac{p(t)}{e_p^\sigma(\cdot, s)}.$$

Proof. See [22, Theorem 2.36]. □

Theorem 9. *Let $p \in \mathcal{R}$ and $t_0 \in \mathbb{T}$.*

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

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

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

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

Proof. See [22, Theorem 2.44 and 2.48]. □

Example 5. *Let $\alpha \in \mathcal{R}$ be constant and $t, t_0 \in \mathbb{T}$.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$e_\alpha(t, t_0) = e^{\alpha(t-t_0)}.$$

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$e_\alpha(t, t_0) = (1 + \alpha)^{t-t_0}.$$

(iii) *If $\mathbb{T} = q^{\mathbb{N}_0}$, then*

$$e_\alpha(t, t_0) = \prod_{s \in [t_0, t)} (1 + (q-1)\alpha s).$$

(iv) *If $\mathbb{T} = h\mathbb{Z}$, then*

$$e_\alpha(t, t_0) = (1 + \alpha h)^{\frac{t-t_0}{h}}.$$

2.4. LOGARITHM

To define an inverse function of the previously introduced exponential function on time scales that still has most of the desired properties from the well known logarithm

in the literature is difficult. In 2005, Martin Bohner suggested two different approaches to the topic, see [14]. Later, Billy Jackson defined the logarithm on time scales in the following way, compare [43].

Definition 15. *Let $g : \mathbb{T} \rightarrow \mathbb{R}$ be a differentiable, nonvanishing function. Then the logarithm on time scales is defined as*

$$\log_{\mathbb{T}} g(t) = \frac{g^{\Delta}(t)}{g(t)}. \quad (5)$$

Remark 1. *Note that*

$$\log_{\mathbb{T}} e_p(t, s) = \frac{(e_p(t, s))^{\Delta}}{e_p(t, s)} = \frac{p(t)e_p(t, s)}{e_p(t, s)} = p(t),$$

and therefore the logarithm is a left inverse of the exponential function.

2.5. DYNAMIC INEQUALITIES

Next, we consider some basic inequalities, which will be useful in the upcoming sections.

Theorem 10 (Hölder's Inequality). *Let $a, b \in \mathbb{T}$ and $f, g : [a, b] \rightarrow \mathbb{R}$ be rd-continuous. Then*

$$\int_a^b |f(t)g(t)| \Delta t \leq \left\{ \int_a^b |f(t)|^p \Delta t \right\}^{\frac{1}{p}} \left\{ \int_a^b |g(t)|^q \Delta t \right\}^{\frac{1}{q}}, \quad (6)$$

where $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. See [22, Theorem 6.13]. □

Theorem 11 (Jensen's Inequality). *Let $a, b \in \mathbb{T}$ and $c, d \in \mathbb{R}$. If $g : [a, b] \rightarrow (c, d)$ is rd-continuous and $F : (c, d) \rightarrow \mathbb{R}$ is continuous and convex, then*

$$F \left(\frac{\int_a^b g(t) \Delta t}{b - a} \right) \leq \frac{\int_a^b F(g(t)) \Delta t}{b - a}. \quad (7)$$

Proof. See [22, Theorem 6.17]. □

2.6. SHIFT, CONVOLUTION AND LAPLACE TRANSFORM

The following definitions and examples can be found in the paper [17] by Bohner and Guseinov.

Definition 16. *Let \mathbb{T} be a time scale, $\sup\mathbb{T} = \infty$, $t_0, t, s \in \mathbb{T}$ and $t_0 \leq s \leq t$. For a given $f : [t_0, \infty) \rightarrow \mathbb{C}$, we call the solution of the shifting problem*

$$u^{\Delta t}(t, \sigma(s)) = -u^{\Delta s}(t, s), \quad u(t, t_0) = f(t)$$

the shift of f . We denote the shift by \hat{f} .

Example 6. *We compute the shift for the following time scales.*

(i) *If $\mathbb{T} = \mathbb{R}$, then the shifting problem*

$$\frac{\partial u(t, s)}{\partial t} = -\frac{\partial u(t, s)}{\partial s}, \quad u(t, t_0) = f(t)$$

has the unique solution $u(t, s) = f(t - s + t_0)$.

(ii) *If $\mathbb{T} = \mathbb{Z}$, then the shifting problem*

$$u(t + 1, s + 1) - u(t, s + 1) = -u(t, s + 1) + u(t, s), \quad u(t, t_0) = f(t)$$

has the unique solution $u(t, s) = f(t - s + t_0)$.

(iii) *If $\mathbb{T} = h\mathbb{Z}$, then the shifting problem*

$$\frac{u(t + h, s + h) - u(t, s + h)}{h} = \frac{-u(t, s + h) + u(t, s)}{h}, \quad u(t, t_0) = f(t)$$

has the unique solution $u(t, s) = f(t - s + t_0)$.

(iv) *If $\mathbb{T} = q^{\mathbb{N}_0}$, then the shift of $f : \mathbb{T} \rightarrow \mathbb{R}$ is given by [17]*

$$\hat{f}(q^k t, t) = \sum_{\nu=0}^k \begin{bmatrix} k \\ \nu \end{bmatrix} t^\nu (1-t)_q^{k-\nu} f(q^\nu),$$

where we use the following notation

$$[\alpha] = \frac{q^\alpha - 1}{q - 1};$$

$$[n]! = \prod_{k=1}^n [k];$$

$$\begin{bmatrix} m \\ n \end{bmatrix} = \frac{[m]!}{[n]! [m-n]!};$$

$$(t-s)_q^n = \prod_{k=0}^{n-1} (t - q^k s),$$

where $\alpha \in \mathbb{R}$, $m, n \in \mathbb{N}_0$ and $t, s \in \mathbb{T}$.

Definition 17. Let $f, g : \mathbb{T} \rightarrow \mathbb{R}$ be two functions. Their convolution $f * g$ is defined by

$$(f * g)(t) = \int_{t_0}^t \widehat{f}(t, \sigma(s))g(s)\Delta s, \quad t \in \mathbb{T},$$

where \widehat{f} is the shift of f .

The convolution on time scales has the following properties.

Theorem 12. The shift of a convolution is given by

$$(\widehat{f * g})(t, s) = \int_s^t \widehat{f}(t, \sigma(u))\widehat{g}(u, s)\Delta u.$$

Proof. See [17, Theorem 2.6]. □

Theorem 13. The convolution is associative, that is

$$(f * g) * h = f * (g * h).$$

Proof. See [17, Theorem 2.7]. □

Theorem 14. If f is delta differentiable, then

$$(f * g)^\Delta = f^\Delta * g + f(t_0)g.$$

If g is delta differentiable, then

$$(f * g)^\Delta = f * g^\Delta + fg(t_0).$$

Moreover,

$$\int_{t_0}^t \widehat{f}(t, \sigma(s)) \Delta s = \int_{t_0}^t f(s) \Delta s.$$

Proof. See [17, Theorem 2.8, Corollary 2.9]. \square

We now introduce the notion of the Laplace transform.

Definition 18. Assume $f : \mathbb{T}_0 \rightarrow \mathbb{R}$ is regulated. Then the Laplace transform of f is defined by

$$\mathcal{L}\{f\}(z) = \int_0^\infty e_{\ominus z}(\sigma(t), 0) f(t) \Delta t$$

for $z \in \mathcal{D}\{f\}$, where $\mathcal{D}\{f\}$ consists of all complex numbers z for which the improper integral exists.

The following property is given in [22, Example 3.103].

Lemma 1. Assume $f : \mathbb{T}_0 \rightarrow \mathbb{R}$ is regulated. Then

$$\mathcal{L}\{gf\}(z) = -\frac{d}{dz} \mathcal{L}\{f\}(z) \quad \text{for } z \in \mathcal{D}(f),$$

where

$$g(t) = \int_0^{\sigma(t)} \frac{1}{1 + \mu(\tau)z} \Delta \tau.$$

Theorem 15. If $f, g : \mathbb{T} \rightarrow \mathbb{R}$ are locally Δ -integrable function on \mathbb{T} , then

$$\mathcal{L}\{f * g\}(z) = \mathcal{L}\{f\}(z) \cdot \mathcal{L}\{g\}(z),$$

where $z \in \mathcal{D}(f) \cap \mathcal{D}(g)$. We will refer to this property as the convolution theorem.

Proof. See [17, Theorem 2.7]. \square

3. MOMENT GENERATING FUNCTION

Let a moment generating function be defined, with the help of the Laplace transformation [39, p.181], as follows

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx,$$

where $f(x)$ is a continuous probability density function. This function allows us to compute the moments by

$$\mathbb{E}(X^n) = M_X^{(n)}(0) = \frac{d^n M_X}{dt^n}(0),$$

compare [11, Section 2.5].

The following definition introduces the modified Laplace transform on time scales (see Definition 18). We assume that the time scale \mathbb{T}_0 is such that $0 \in \mathbb{T}_0$ and $\sup \mathbb{T}_0 = \infty$. Moreover, note the slight change from Definition 18 as we replace $\sigma(t)$ with t and $\ominus z$ with z in the exponential function. The main reason for this change is to derive a more approachable formula for the moments and especially for the expected value.

$$\tilde{\mathcal{L}}\{f\}(z) = \int_0^{\infty} e_z(t, 0) f(t) \Delta t.$$

Throughout this section, we assume that z is positively regressive and that interchanging the order of differentiation and integration does not cause any problems.

Lemma 2. *Assume $f : \mathbb{T}_0 \rightarrow \mathbb{R}$ is regulated. Then*

$$\frac{d}{dz} \tilde{\mathcal{L}}\{f\}(z) = \int_0^{\infty} e_z(t, 0) f(t) g_t(z) \Delta t,$$

where

$$g_t(z) = \int_0^t \frac{1}{1 + \mu(\tau)z} \Delta \tau.$$

Proof. Using [22, Example 2.42], we have

$$\frac{d}{dz} e_z(t, 0) = e_z(t, 0) \int_0^t \frac{1}{1 + \mu(\tau)z} \Delta \tau.$$

As interchanging the order of differentiation and integration is possible, we get

$$\begin{aligned}
\frac{d}{dz} \tilde{\mathcal{L}}\{f\}(z) &= \frac{d}{dz} \int_0^\infty e_z(t,0) f(t) \Delta t \\
&= \int_0^\infty \frac{d}{dz} e_z(t,0) f(t) \Delta t \\
&= \int_0^\infty e_z(t,0) f(t) \int_0^t \frac{1}{1 + \mu(\tau)z} \Delta \tau \Delta t \\
&= \int_0^\infty e_z(t,0) f(t) g_t(z) \Delta t.
\end{aligned}$$

This completes the proof. □

Example 7. *First moment on time scales. Computation yields*

$$\begin{aligned}
\frac{d}{dz} \tilde{\mathcal{L}}\{f\}(0) &= \int_0^\infty f(t) \left(\int_0^t 1 \Delta \tau \right) \Delta t \\
&= \int_0^\infty t f(t) \Delta t \\
&= \int_0^\infty h_1(t,0) f(t) \Delta t.
\end{aligned}$$

Example 8. *Second moment on time scales. Using the usual product rule and Lemma 2, we get*

$$\begin{aligned}
\frac{d^2}{dz^2} \tilde{\mathcal{L}}\{f\}(z) &= \int_0^\infty e_z(t,0) f(t) g'_t(z) \Delta t \\
&\quad + \int_0^\infty e_z(t,0) f(t) g_t^2(z) \Delta t \\
&= \int_0^\infty e_z(t,0) f(t) [g_t^2(z) + g'_t(z)] \Delta t.
\end{aligned}$$

As

$$g'_t(z) = - \int_0^t \frac{\mu(\tau)}{(1 + \mu(\tau)z)^2} \Delta \tau,$$

we have

$$\begin{aligned}
\frac{d^2}{dz^2} \tilde{\mathcal{L}}\{f\}(0) &= \int_0^\infty f(t) \left(\int_0^t 1 \Delta \tau \right)^2 \Delta t - \int_0^\infty f(t) \int_0^t \mu(\tau) \Delta \tau \Delta t \\
&= \int_0^\infty f(t) \left[t^2 - \int_0^t \mu(\tau) \Delta \tau \right] \Delta t \\
&= \int_0^\infty 2h_2(t,0) f(t) \Delta t.
\end{aligned}$$

The last equality holds as

$$\begin{aligned}
\left(t^2 - \int_0^t \mu(\tau) \Delta\tau\right)^\Delta &= t + \sigma(t) - \mu(t) \\
&= t + \sigma(t) - \sigma(t) + t \\
&= 2t.
\end{aligned}$$

Moreover,

$$(2h_2(t, 0))^\Delta = 2h_1(t, 0) = 2t,$$

and for $t = 0$, both expressions are 0. Equivalently, this fact could have been shown by the string of equations

$$\begin{aligned}
t^2 - \int_0^t \mu(\tau) \Delta\tau &= \int_0^t t \Delta\tau - \int_0^t \sigma(\tau) \Delta\tau + \int_0^t \tau \Delta\tau \\
&= \int_t^0 (\sigma(\tau) - t) \Delta\tau + h_2(t, 0) \\
&= g_2(0, t) + h_2(t, 0) \\
&= 2h_2(t, 0).
\end{aligned}$$

More generally for an arbitrary function $H(t, z)$, denoting by $H'(t, z)$ the derivative of H with respect to z , we have

$$\begin{aligned}
\frac{d}{dz} \tilde{\mathcal{L}}\{Hf\}(z) &= \frac{d}{dz} \int_0^\infty e_z(t, 0) f(t) H(t, z) \Delta t \\
&= \int_0^\infty e_z(t, 0) f(t) H'(t, z) \Delta t \\
&\quad + \int_0^\infty e_z(t, 0) f(t) g_t(z) H(t, z) \Delta t \\
&= \int_0^\infty e_z(t, 0) f(t) [H'(t, z) + H(t, z) g_t(z)] \Delta t \\
&= \tilde{\mathcal{L}}\{f(H' + Hg_t)\}(z).
\end{aligned}$$

Theorem 16. Let $g_t(z) = \int_0^t \frac{1}{1+z\mu(\tau)} \Delta\tau$ and define H_k recursively by

$$H_0(t, z) = 1,$$

$$H_{k+1}(t, z) = H'_k(t, z) + H_k(t, z)g_t(z), \quad k \in \mathbb{N}_0.$$

Then for $a_k = k + 1$, we have

$$H_{k+1}(t, z) = a_k \int_0^t \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau.$$

Proof. We will prove this theorem by induction. From the derivation of the first moment, we know that $H_1(t, z) = g_t(z)$, $a_0 = 1$ and $H_0 = 1$. Thus the claim holds for $k = 0$. Now, assuming that the claim holds for $k \in \mathbb{N}_0$, we get

$$\begin{aligned} H_{k+2}(t, z) &= H'_{k+1}(t, z) + H_{k+1}(t, z)g_t(z) \\ &= \frac{d}{dz} a_k \int_0^t \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau + a_k \int_0^t \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau g_t(z) \\ &= a_k \int_0^t \frac{H'_k(\tau, z)(1 + z\mu(\tau)) - \mu(\tau)H_k(\tau, z)}{(1 + z\mu(\tau))^2} \Delta\tau \\ &\quad + a_k \int_0^t \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau g_t(z) \\ &= a_k \int_0^t \frac{H'_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau - a_k \int_0^t \frac{\mu(\tau)H_k(\tau, z)}{(1 + z\mu(\tau))^2} \Delta\tau \\ &\quad + a_k \int_0^t \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau g_t(z) \\ &= : H(t). \end{aligned}$$

Now, we have $H(0) = 0$ and

$$\begin{aligned} H^\Delta(t) &= \frac{a_k H'_k(t, z)}{1 + z\mu(t)} - \frac{a_k \mu(t) H_k(t, z)}{(1 + z\mu(t))^2} \\ &\quad + \frac{a_k H_k(t, z)}{1 + z\mu(t)} g_t(z) + \frac{a_k}{1 + z\mu(t)} \int_0^{\sigma(t)} \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau \\ &= \frac{a_k}{1 + z\mu(t)} [H'_k(t, z) + H_k(t, z)g_t(z)] \\ &\quad + \frac{a_k}{1 + z\mu(t)} \left[- \int_t^{\sigma(t)} \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau + \int_0^{\sigma(t)} \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau \right] \\ &= \frac{a_k}{1 + z\mu(t)} H_{k+1}(t, z) + \frac{a_k}{1 + z\mu(t)} \int_0^t \frac{H_k(\tau, z)}{1 + z\mu(\tau)} \Delta\tau \\ &= \frac{a_k}{1 + z\mu(t)} H_{k+1}(t, z) + \frac{1}{1 + z\mu(t)} H_{k+1}(t, z) \\ &= \frac{a_k + 1}{1 + z\mu(t)} H_{k+1}(t, z) \\ &= \frac{a_{k+1}}{1 + z\mu(t)} H_{k+1}(t, z) \end{aligned}$$

as $a_k + 1 = k + 1 + 1 = k + 2 = a_{k+1}$. Thus

$$H_{k+2}(t, z) = a_{k+1} \int_0^t \frac{H_{k+1}(\tau, z)}{1 + z\mu(\tau)} \Delta\tau,$$

which completes the proof. \square

Theorem 17. *Let $H_k(t, z)$ be defined as in Theorem 16. Then*

$$H_k(t, 0) = k!h_k(t, 0).$$

Proof. Again, we will prove this theorem by induction. As $H_0(t, 0) = 1$ and $h_0(t, 0) = 1$, the claim holds for $k = 0$. Now, assuming that the claim holds for $k \in \mathbb{N}_0$, we have

$$\begin{aligned} H_{k+1}(t, 0) &= (k+1) \int_0^t \frac{H_k(\tau, 0)}{1 + 0\mu(\tau)} \Delta\tau \\ &= (k+1) \int_0^t k!h_k(\tau, 0) \Delta\tau \\ &= (k+1)!h_{k+1}(t, 0), \end{aligned}$$

and the proof is complete. \square

Some interesting properties regarding the functions h_k and g_k result from here.

Corollary 1. *We have*

$$h_2(t, 0) + g_2(t, 0) = h_1^2(t, 0).$$

Proof. With the $H_k(t, z)$ as defined in Theorem 16, we have

$$H_1(t, z) = H_0'(t, z) + H_0(t, z)g_t(z) = g_t(z)$$

and

$$\begin{aligned} H_2(t, z) &= H_1'(t, z) + H_1(t, z)g_t(z) \\ &= g_t'(z) + g_t^2(z) \\ &= - \int_0^t \frac{\mu(\tau)}{(1 + z\mu(\tau))^2} \Delta\tau + \left(\int_0^t \frac{1}{1 + z\mu(\tau)} \Delta\tau \right)^2. \end{aligned}$$

Therefore

$$\begin{aligned}
H_2(t, 0) &= -\int_0^t \mu(\tau) \Delta\tau + \left(\int_0^t 1 \Delta\tau\right)^2 \\
&= -\int_0^t (\sigma(\tau) - \tau) \Delta\tau + h_1^2(t, 0) \\
&= -\int_0^t (\sigma(\tau) - 0) \Delta\tau + \int_0^t (\tau - 0) \Delta\tau + h_1^2(t, 0) \\
&= -g_2(t, 0) + h_2(t, 0) + h_1^2(t, 0).
\end{aligned}$$

By Theorem 17, we also have

$$H_2(t, 0) = 2h_2(t, 0)$$

and thus

$$h_2(t, 0) = h_1^2(t, 0) - g_2(t, 0),$$

which completes the proof of the corollary. \square

Remark 2. *With the help of the previous results, we can define the moments on time scales by*

$$\mathbb{E}_{\mathbb{T}}(X^k) := \int_0^\infty k! h_k(t, 0) f(t) \Delta t.$$

Example 9. *Let $\mathbb{T} = \mathbb{R}$. Then*

$$\begin{aligned}
\mathbb{E}_{\mathbb{R}}(X^k) &= \int_0^\infty k! h_k(t, 0) f(t) \Delta t \\
&= \int_0^\infty k! \frac{t^k}{k!} f(t) dt \\
&= \int_0^\infty t^k f(t) dt,
\end{aligned}$$

which corresponds to the continuous definition found for example in [11, 2.3.9] for continuous distributions with positive support.

Example 10. *Let $\mathbb{T} = \mathbb{Z}$. Then*

$$\mathbb{E}_{\mathbb{Z}}(X^k) = \int_0^\infty k! h_k(t, 0) f(t) \Delta t$$

$$\begin{aligned} &= \sum_{t=0}^{\infty} k! \binom{t}{k} f(t) \\ &= \sum_{t=0}^{\infty} t(t-1)\cdots(t-k+1)f(t), \end{aligned}$$

which is slightly different from the discrete definition found in [11, 2.2.14], applied for discrete distributions with positive support

$$\mathbb{E}(X^k) = \sum_{t=0}^{\infty} t^k f(t).$$

4. CUMULANT GENERATING FUNCTION

Definition 19. f is called a time scales probability density function if

1. $f(t) \geq 0 \quad \forall t \in \mathbb{T}_0$,
2. $\int_0^\infty f(t)\Delta t = 1$.

Definition 20. Assume $f : \mathbb{T}_0 \rightarrow \mathbb{R}$ is a regulated time scales probability density function and let the moment generating function be defined as

$$M(z) = \int_0^\infty e_z(t, 0)f(t)\Delta t.$$

Then the cumulant generating function $C(z)$ is the logarithm of $M(z)$

$$C(z) = \log M(z).$$

Definition 21. Let f be a time scales probability density function of the random variable X .

1. The expected value $\mathbb{E}_{\mathbb{T}}(X)$ is defined as

$$\mathbb{E}_{\mathbb{T}}(X) := \frac{dC}{dz}(0).$$

2. The variance $\text{Var}_{\mathbb{T}}(X)$ is defined as

$$\text{Var}_{\mathbb{T}}(X) := \frac{d^2C}{dz^2}(0).$$

Remark 3. The expected value and the variance correspond therefore with the first and second cumulant. Moreover, the expected value matches the definition of the first moment of a random variable on time scales.

Computation yields

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X) &= \frac{dC}{dz}(0) = \frac{1}{M(0)} \frac{dM}{dz}(0) \\ &= \frac{1}{\int_0^\infty f(t)\Delta t} \int_0^\infty h_1(t, 0)f(t)\Delta t \end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty h_1(t, 0)f(t)\Delta t \\
&= \int_0^\infty tf(t)\Delta t
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{V}\text{ar}_{\mathbb{T}}(X) &= \frac{d^2C}{dz^2}(0) = \frac{M''(0)}{M(0)} - \frac{(M'(0))^2}{M^2(0)} \\
&= \int_0^\infty 2h_2(t, 0)f(t)\Delta t - \left(\int_0^\infty h_1(t, 0)f(t)\Delta t \right)^2.
\end{aligned}$$

Remark 4. *Note that, as in the usual definition, we have*

$$\mathbb{V}\text{ar}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

see [39, p.51], and therefore on time scales

$$\mathbb{V}\text{ar}_{\mathbb{T}}(X) = \mathbb{E}_{\mathbb{T}}(X^2) - (\mathbb{E}_{\mathbb{T}}(X))^2.$$

Example 11. *For the continuous and the discrete cases, we have the following.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$\mathbb{E}_{\mathbb{R}}(X) = \int_0^\infty tf(t)dt$$

and

$$\mathbb{V}\text{ar}_{\mathbb{R}}(X) = \int_0^\infty t^2f(t)dt - \left(\int_0^\infty tf(t)dt \right)^2.$$

Note that, as previously mentioned, the expected value corresponds to the known definition, and due to the matching definition of the second moment, also the variance coincides.

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$\mathbb{E}_{\mathbb{Z}}(X) = \sum_{t=0}^{\infty} tf(t)$$

and

$$\text{Var}_{\mathbb{Z}}(X) = \sum_{t=0}^{\infty} t(t-1)f(t) - \left(\sum_{t=0}^{\infty} tf(t) \right)^2.$$

Note that the expected value is the same as the one for usual discrete distributions, but due to the different definition of the second moment, the variance is slightly different.

Theorem 18. Assume $f : \mathbb{T}_0 \rightarrow \mathbb{R}$ is a regulated time scales probability density function, then

$$\text{Var}_{\mathbb{T}}(X) \geq - \int_0^{\infty} f(t) \int_0^t \mu(\tau) \Delta\tau \Delta t.$$

Moreover, if $X = c$ for $c \in \mathbb{R}$, $c \geq 0$, then the variance is minimized with value

$$\text{Var}_{\mathbb{T}}(X) = - \int_0^{\infty} f(t) \int_0^t \mu(\tau) \Delta\tau \Delta t = 2h_2(c, 0) - h_1^2(c, 0).$$

Proof. We have

$$\begin{aligned} \text{Var}_{\mathbb{T}}(X) &= \int_0^{\infty} 2h_2(t, 0)f(t)\Delta t - \left(\int_0^{\infty} h_1(t, 0)f(t)\Delta t \right)^2 \\ &= \int_0^{\infty} 2h_2(t, 0)f(t)\Delta t - 2\mathbb{E}_{\mathbb{T}}(X) \int_0^{\infty} h_1(t, 0)f(t)\Delta t \\ &\quad + (\mathbb{E}_{\mathbb{T}}(X))^2 \int_0^{\infty} f(t)\Delta t \\ &= \int_0^{\infty} [2h_2(t, 0) - 2\mathbb{E}_{\mathbb{T}}(X)h_1(t, 0) + (\mathbb{E}_{\mathbb{T}}(X))^2] f(t)\Delta t \\ &= \int_0^{\infty} [2h_2(t, 0) - h_1^2(t, 0) + (h_1(t, 0) - \mathbb{E}_{\mathbb{T}}(X))^2] f(t)\Delta t. \end{aligned}$$

As

$$f(t) \geq 0$$

and

$$(h_1(t, 0) - \mathbb{E}_{\mathbb{T}}(X))^2 \geq 0,$$

and with Corollary 1

$$\begin{aligned}
2h_2(t, 0) - h_1^2(t, 0) &= h_2(t, 0) - g_2(t, 0) \\
&= \int_0^t \tau \Delta\tau - \int_0^t \sigma(\tau) \Delta\tau \\
&= - \int_0^t \mu(\tau) \Delta\tau,
\end{aligned}$$

we get

$$\mathbb{V}\text{ar}_{\mathbb{T}}(X) \geq - \int_0^\infty f(t) \int_0^t \mu(\tau) \Delta\tau \Delta t.$$

If $X = c$, then the entire density is distributed at c and

$$\mathbb{E}_{\mathbb{T}}(X) = \int_0^\infty h_1(t, 0) f(t) \Delta t = h_1(c, 0) \int_0^\infty f(t) \Delta t = h_1(c, 0) = c,$$

and similarly

$$\int_0^\infty h_2(t, 0) f(t) \Delta t = h_2(c, 0).$$

Therefore, we have

$$\mathbb{V}\text{ar}_{\mathbb{T}}(X) = 2h_2(c, 0) - h_1^2(c, 0)$$

and

$$\int_0^\infty (h_1(t, 0) - \mathbb{E}_{\mathbb{T}}(X))^2 f(t) \Delta t = \int_0^\infty (h_1(t, 0) - c)^2 f(t) \Delta t = 0.$$

This results in

$$\begin{aligned}
\mathbb{V}\text{ar}_{\mathbb{T}}(X) &= \int_0^\infty [2h_2(t, 0) - h_1^2(t, 0) + (h_1(t, 0) - \mathbb{E}_{\mathbb{T}}(X))^2] f(t) \Delta t \\
&= \int_0^\infty (2h_2(t, 0) - h_1^2(t, 0)) f(t) \Delta t \\
&= - \int_0^\infty f(t) \int_0^t \mu(\tau) \Delta\tau \Delta t
\end{aligned}$$

and concludes the proof of the theorem. \square

Remark 5. Note that, in comparison to the classical definition of variance, the new quantity $\mathbb{V}\text{ar}_{\mathbb{T}}$ is not necessarily greater or equal to 0. Only in the continuous case is that property necessarily achieved. Nevertheless, we get a lower bound for the variance. This new definition will yield great advantages in the computation of variance and moments for the upcoming new time scales distributions.

Remark 6. We have $\mathbb{E}_{\mathbb{T}}(X) = \mathbb{E}(X)$. The relationship of $\mathbb{V}\text{ar}_{\mathbb{T}}(X)$ with the classical variance $\mathbb{V}\text{ar}(X)$ and expectation $\mathbb{E}(X)$ is

$$\begin{aligned}\mathbb{V}\text{ar}_{\mathbb{T}}(X) &= \int_0^{\infty} [2h_2(t, 0) - h_1^2(t, 0) + (h_1(t, 0) - \mathbb{E}_{\mathbb{T}}(X))^2] f(t) \Delta t \\ &= \int_0^{\infty} [2h_2(t, 0) - h_1^2(t, 0) + (t - \mathbb{E}_{\mathbb{T}}(X))^2] f(t) \Delta t \\ &= \mathbb{V}\text{ar}(X) + \int_0^{\infty} [2h_2(t, 0) - h_1^2(t, 0)] f(t) \Delta t \\ &= \mathbb{V}\text{ar}(X) + \mathbb{E}(2H(X)),\end{aligned}$$

where $H(X)$ has the time scales probability density function $f(t) = h_2(t, 0) - \frac{t^2}{2}$.

Example 12. We apply the previous properties to different time scales.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\mathbb{V}\text{ar}_{\mathbb{R}}(X) \geq - \int_0^{\infty} f(t) \int_0^t \mu(\tau) \Delta \tau \Delta t = - \int_0^{\infty} f(t) \int_0^t 0 d\tau dt = 0,$$

and if $X = c$, then

$$\mathbb{V}\text{ar}_{\mathbb{R}}(X) = 2h_2(c, 0) - h_1^2(c, 0) = 2\frac{c^2}{2} - c^2 = 0.$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$\begin{aligned}\mathbb{V}\text{ar}_{\mathbb{Z}}(X) &\geq - \int_0^{\infty} f(t) \int_0^t \mu(\tau) \Delta \tau \Delta t \\ &= - \sum_{t=0}^{\infty} f(t) \sum_{j=0}^{t-1} 1 \\ &= - \sum_{t=0}^{\infty} t f(t) \\ &= -\mathbb{E}_{\mathbb{Z}}(X),\end{aligned}$$

and if $X = c$, then

$$\mathbb{V}\text{ar}_{\mathbb{T}}(X) = 2h_2(c, 0) - h_1^2(c, 0) = 2\frac{c(c-1)}{2} - c^2 = -c.$$

(iii) If $\mathbb{T} = h\mathbb{Z}$, then

$$\begin{aligned} \mathbb{V}\text{ar}_{h\mathbb{Z}}(X) &\geq -\int_0^\infty f(t) \int_0^t \mu(\tau) \Delta\tau \Delta t \\ &= -\int_0^\infty f(t) \int_0^t h \Delta\tau \Delta t \\ &= -h \int_0^\infty t f(t) \Delta t \\ &= -h\mathbb{E}_{h\mathbb{Z}}(X), \end{aligned}$$

and if $X = c$, then

$$\mathbb{V}\text{ar}_{h\mathbb{Z}}(X) = -h\mathbb{E}_{h\mathbb{Z}}(X) = -hc.$$

5. DISTRIBUTIONS

5.1. UNIFORM DISTRIBUTION

Definition 22. Let $a, b \in \mathbb{T}_0$ and $a \leq t \leq b$. Then we define the time scales probability density function of the uniform distribution by

$$f(t) = \begin{cases} \frac{1}{\sigma(b)-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise.} \end{cases}$$

Remark 7. Clearly we have $f(t) \geq 0$ and

$$\int_0^\infty f(t)\Delta t = \frac{1}{\sigma(b)-a} \int_a^{\sigma(b)} \Delta t = 1,$$

and therefore $f(t)$ is a well-defined time scales probability density function.

Theorem 19. Let $a, b \in \mathbb{T}_0$, $a \leq t \leq b$ and the time scales probability density function of X be

$$f(t) = \begin{cases} \frac{1}{\sigma(b)-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\mathbb{E}_{\mathbb{T}}(X) = \frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a \tag{8}$$

and

$$\text{Var}_{\mathbb{T}}(X) = 2 \frac{h_3(\sigma(b), 0) - h_3(a, 0)}{\sigma(b) - a} - \left(\frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a \right)^2 \tag{9}$$

and

$$\mathbb{E}_{\mathbb{T}}(X^k) = k! \frac{h_{k+1}(\sigma(b), 0) - h_{k+1}(a, 0)}{\sigma(b) - a}. \tag{10}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}_{\mathbb{T}}(X) &= \int_0^{\infty} h_1(t, 0) f(t) \Delta t \\
&= \frac{1}{\sigma(b) - a} \int_a^{\sigma(b)} t \Delta t \\
&= \frac{1}{\sigma(b) - a} \int_a^{\sigma(b)} (t - a) \Delta t + \frac{1}{\sigma(b) - a} \int_a^{\sigma(b)} a \Delta t \\
&= \frac{h_2(\sigma(b), a)}{\sigma(b) - a} + \frac{a}{\sigma(b) - a} (\sigma(b) - a) \\
&= \frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a,
\end{aligned}$$

and (8) is shown. Moreover, we have

$$\begin{aligned}
\text{Var}_{\mathbb{T}}(X) &= \int_0^{\infty} 2h_2(t, 0) f(t) \Delta t - \left(\int_0^{\infty} h_1(t, 0) f(t) \Delta t \right)^2 \\
&= \frac{2}{\sigma(b) - a} \int_a^{\sigma(b)} h_2(t, 0) \Delta t - \left(\frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a \right)^2 \\
&= \frac{2}{\sigma(b) - a} \left(\int_0^{\sigma(b)} h_2(t, 0) \Delta t - \int_0^a h_2(t, 0) \Delta t \right) - \left(\frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a \right)^2 \\
&= 2 \frac{h_3(\sigma(b), 0) - h_3(a, 0)}{\sigma(b) - a} - \left(\frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a \right)^2,
\end{aligned}$$

which is the desired equation (9). Finally

$$\begin{aligned}
\mathbb{E}_{\mathbb{T}}(X^k) &= \int_0^{\infty} k! h_k(t, 0) f(t) \Delta t \\
&= \frac{k!}{\sigma(b) - a} \int_a^{\sigma(b)} h_k(t, 0) \Delta t \\
&= \frac{k!}{\sigma(b) - a} \left(\int_0^{\sigma(b)} h_k(t, 0) \Delta t - \int_0^a h_k(t, 0) \Delta t \right) \\
&= k! \frac{h_{k+1}(\sigma(b), 0) - h_{k+1}(a, 0)}{\sigma(b) - a},
\end{aligned}$$

completing the proof of (10). □

Example 13 (Continuous case). Let $\mathbb{T} = \mathbb{R}$. Then the probability density function is

$$f(t) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise.} \end{cases}$$

Hence

$$\begin{aligned}
 \mathbb{E}_{\mathbb{R}}(X) &= \frac{h_2(\sigma(b), a)}{\sigma(b) - a} + a \\
 &= \frac{1}{b-a} \frac{(b-a)^2}{2} + a \\
 &= \frac{b-a}{2} + a = \frac{a+b}{2}.
 \end{aligned}$$

This is the expected value for the continuous uniform distribution that can be found in [11, p.110]. Moreover,

$$\begin{aligned}
 \text{Var}_{\mathbb{R}}(X) &= \frac{2}{b-a} \frac{b^3 - a^3}{3!} - \left(\frac{a+b}{2} \right)^2 \\
 &= \frac{2}{b-a} \frac{b^3 - a^3}{6} - \frac{a^2 + ab + b^2}{4} \\
 &= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + ab + b^2}{4} \\
 &= \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12}.
 \end{aligned}$$

This is the variance for the continuous uniform distribution, which can be found in the literature for example in [11, p.110].

Example 14 (Discrete case). Let $\mathbb{T} = \mathbb{Z}$. Then the probability density function is

$$f(t) = \begin{cases} \frac{1}{b+1-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise,} \end{cases}$$

which corresponds to $\frac{1}{n}$, where n represents the number of points having density greater than zero. Furthermore,

$$\begin{aligned}
 \mathbb{E}_{\mathbb{Z}}(X) &= \frac{1}{b+1-a} \frac{(b+1-a)(b-a)}{2} + a \\
 &= \frac{b-a}{2} + a = \frac{a+b}{2}.
 \end{aligned}$$

This is exactly the expected value for the discrete uniform distribution with $n = b+1-a$, which can be found in [11, p.108]. Moreover,

$$\text{Var}_{\mathbb{Z}}(X) = \frac{2}{b+1-a} \frac{(b+1)b(b-1) - a(a-1)(a-2)}{6} - \left(\frac{a+b}{2} \right)^2$$

$$\begin{aligned}
&= \frac{b^3 - a^3 + 3a^2b - 3ab^2 + 9a^2 - 3b^2 - 6ab - 8a - 4b}{12(b+1-a)} \\
&= \frac{(b+1-a)^2 - 1}{12} - \frac{a+b}{2}.
\end{aligned}$$

This is not exactly the variance for the discrete uniform distribution from the literature, $\frac{n^2-1}{12}$, compare [11, p.108], due to the slightly changed definition of the variance. As expected, we get $\text{Var}(X) - \mathbb{E}(X)$, as in the probability density function of $2H(X)$, we have

$$f(t) = 2h_2(t, 0) - 2\frac{t^2}{2} = t(t-1) - t^2 = -t.$$

Now, the moment generating function will be derived for the uniform distribution on time scales.

Theorem 20. Let $a, b \in \mathbb{T}_0$, $a \leq t \leq b$ and the time scales probability density function of X be

$$f(t) = \begin{cases} \frac{1}{\sigma(b)-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$M_X(z) = \frac{1}{z} \frac{e_z(\sigma(b), 0) - e_z(a, 0)}{\sigma(b) - a}. \quad (11)$$

Proof.

$$\begin{aligned}
M_X(z) &= \int_0^\infty f(t)e_z(t, 0)\Delta t \\
&= \int_a^{\sigma(b)} \frac{1}{\sigma(b) - a} e_z(t, 0)\Delta t \\
&= \frac{1}{z} \frac{1}{\sigma(b) - a} \int_a^{\sigma(b)} z e_z(t, 0)\Delta t \\
&= \frac{1}{z} \frac{1}{\sigma(b) - a} e_z(t, 0) \Big|_a^{\sigma(b)} \\
&= \frac{1}{z} \frac{e_z(\sigma(b), 0) - e_z(a, 0)}{\sigma(b) - a}
\end{aligned}$$

and (11) is shown. □

Example 15. We apply (11) to different time scales.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$M_X(z) = \frac{1}{z} \frac{e^{bz} - e^{az}}{b - a}.$$

Note that this is exactly the formula of the moment generating function of the continuous uniform distribution, see [48, Example 10.1].

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$M_X(z) = \frac{1}{z} \frac{(1+z)^{b+1} - (1+z)^a}{b+1-a}.$$

Due to the different definition of the Laplace transform and exponential function, the moment generating function differs from the discrete case, but still follows the same structure. In the literature, we find

$$M_X(z) = \frac{e^{(b+1)z} - e^{az}}{n(e^z - 1)},$$

where n equals the number of points with density different from 0, i.e., $n = b+1-a$, compare [13, p.72].

(iii) If $\mathbb{T} = h\mathbb{Z}$, then

$$M_X(z) = \frac{1}{z} \frac{(1+hz)^{\frac{b+h}{h}} - (1+hz)^{\frac{a}{h}}}{b-a}.$$

The diagrams in Figure 5.1 represent the time scales probability density function (pdf) and the time scales cumulative distribution function (cdf) for an uniformly distributed random variable, with support $[1, 4]$. The diagrams include continuous, discrete, $h\mathbb{Z}$ (with $h = \frac{1}{2}$) and $q^{\mathbb{N}_0}$ (with $q = 1.1$) time scales cases. A formal definition of the cumulative density function will be presented in Section 6. Note that the cumulative distribution function is equal to one at $\sigma(4)$, due to the slightly different definition.

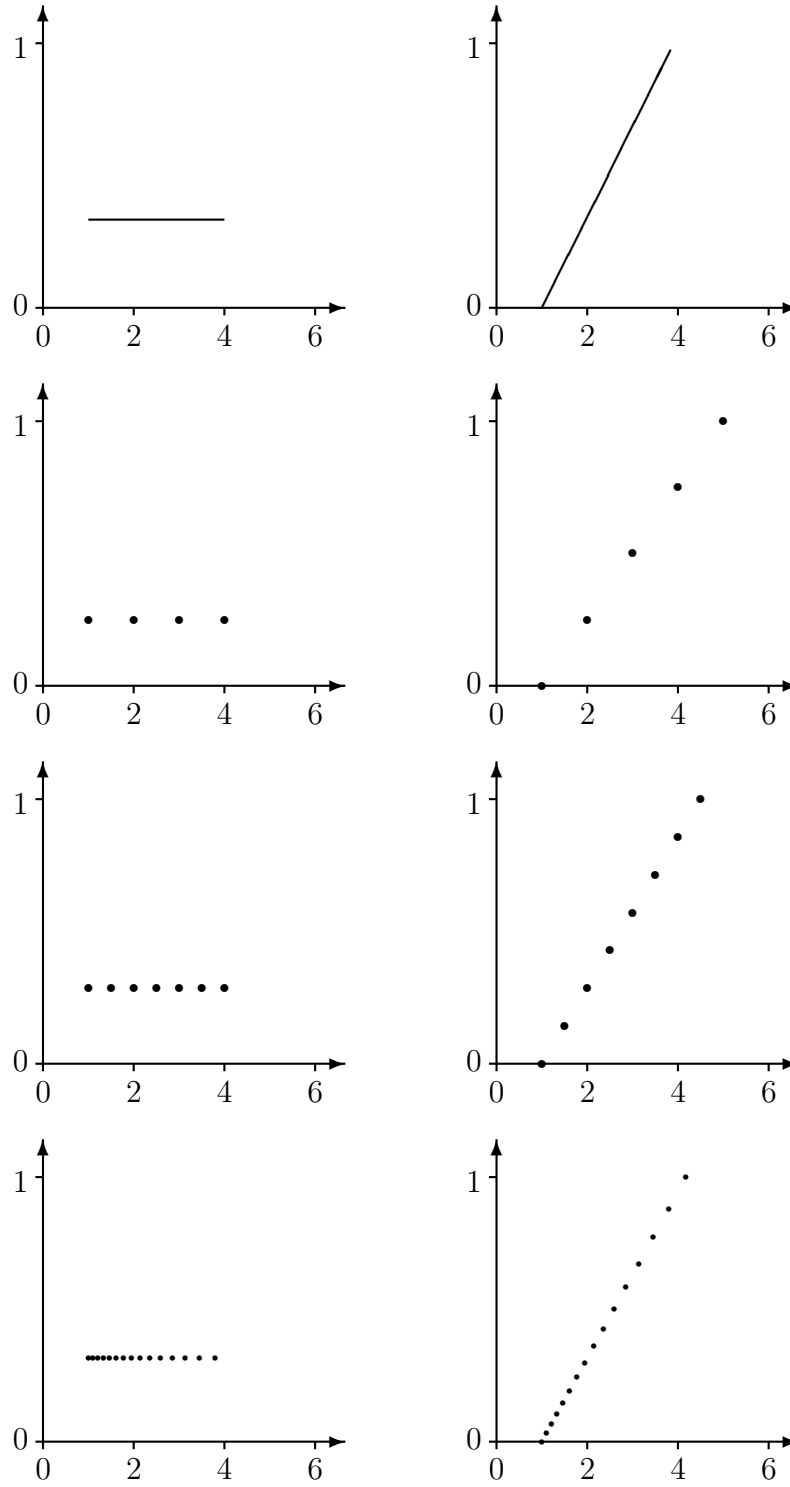


Figure 5.1. Uniform distribution

5.2. EXPONENTIAL DISTRIBUTION

Definition 23. Let $\lambda > 0$, $\ominus\lambda$ be positively regressive and $t \in \mathbb{T}_0$. Then we define the time scales probability density function of the exponential distribution by

$$f(t) = \begin{cases} -(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Remark 8. Note that

$$-(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0) = \frac{\lambda}{1 + \mu(t)\lambda} e_{\ominus\lambda}(t, 0) = \lambda e_{\ominus\lambda}(\sigma(t), 0) = \frac{\lambda}{e_{\lambda}(\sigma(t), 0)},$$

and therefore the time scales probability density function is equivalent to

$$f(t) = \begin{cases} \frac{\lambda}{e_{\lambda}(\sigma(t), 0)}, & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Remark 9. Clearly, we have $f(t) \geq 0$ as $\lambda > 0$ and $e_{\lambda}(\sigma(t), 0) > 0$ due to the fact that $\ominus\lambda$ is positively regressive [22, Theorem 2.44 (i)] and

$$\begin{aligned} \int_0^{\infty} f(t)\Delta t &= \int_0^{\infty} -(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0)\Delta t \\ &= - \int_0^{\infty} e_{\ominus\lambda}^{\Delta}(\cdot, 0)(t)\Delta t \\ &= e_{\ominus\lambda}(t, 0)|_0^{\infty} = 1, \end{aligned}$$

and therefore f is a well-defined time scales probability density function. The last equality holds, as for nonnegative, rd-continuous λ , we have

$$e_{\lambda}(t, t_0) \geq 1 + \int_{t_0}^t \lambda \Delta u = 1 + \lambda(t - t_0), \quad t \geq t_0,$$

compare [15, Remark 2], which goes to infinity for $t \rightarrow \infty$.

Theorem 21. Let $\lambda > 0$ and $\ominus\lambda$ be positively regressive, $t \in \mathbb{T}_0$, and the time scales probability density function of X be

$$f(t) = \begin{cases} -(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then

$$\mathbb{E}_{\mathbb{T}}(X) = \frac{1}{\lambda} \quad (12)$$

and

$$\mathbb{V}\text{ar}_{\mathbb{T}}(X) = \frac{1}{\lambda^2}. \quad (13)$$

Proof. Integration by parts yields

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X) &= \int_0^{\infty} tf(t)\Delta t \\ &= -\int_0^{\infty} t(\ominus\lambda)(t)e_{\ominus\lambda}(t,0)\Delta t \\ &= -te_{\ominus\lambda}(t,0)|_0^{\infty} + \int_0^{\infty} e_{\ominus\lambda}(\sigma(t),0)\Delta t \\ &= 0 + \frac{1}{\lambda} \int_0^{\infty} \frac{\lambda}{1+\mu(t)\lambda} e_{\ominus\lambda}(t,0)\Delta t \\ &= -\frac{1}{\lambda} \int_0^{\infty} (\ominus\lambda)(t)e_{\ominus\lambda}(t,0)\Delta t \\ &= \frac{1}{\lambda} \int_0^{\infty} f(t)\Delta t \\ &= \frac{1}{\lambda}, \end{aligned}$$

and (12) is shown. Moreover, we have

$$\begin{aligned} \mathbb{V}\text{ar}_{\mathbb{T}}(X) &= \int_0^{\infty} 2h_2(t,0)f(t)\Delta t - \left(\int_0^{\infty} h_1(t,0)f(t)\Delta t \right)^2 \\ &= -2 \int_0^{\infty} h_2(t,0)(\ominus\lambda)(t)e_{\ominus\lambda}(t,0)\Delta t - \frac{1}{\lambda^2} \\ &= -2h_2(t,0)e_{\ominus\lambda}(t,0)|_0^{\infty} + 2 \int_0^{\infty} h_1(t,0)e_{\ominus\lambda}(\sigma(t),0)\Delta t - \frac{1}{\lambda^2} \\ &= -\frac{2}{\lambda} \int_0^{\infty} t(\ominus\lambda)e_{\ominus\lambda}(t,0)\Delta t - \frac{1}{\lambda^2} \\ &= \frac{2}{\lambda} \mathbb{E}_{\mathbb{T}}(X) - \frac{1}{\lambda^2} \\ &= \frac{2}{\lambda} \frac{1}{\lambda} - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}, \end{aligned}$$

which is the desired equation (13). □

Theorem 22. Let $\lambda > 0$, $t \in \mathbb{T}_0$ and the time scales probability density function of X be

$$f(t) = \begin{cases} -(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then the k -th moment of X is given by

$$\mathbb{E}_{\mathbb{T}}(X^k) = \frac{k!}{\lambda^k}. \quad (14)$$

Proof. This can be shown by induction. Note that $\mathbb{E}_{\mathbb{T}}(X) = \frac{1}{\lambda}$ by Theorem 21, so the statement holds for $k = 1$. Now, assume that the statement holds for $k - 1$. Then we have

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X^k) &= \int_0^{\infty} k!h_k(t, 0)f(t)\Delta t \\ &= -\int_0^{\infty} k!h_k(t, 0)(\ominus\lambda)e_{\ominus\lambda}(t, 0)\Delta t \\ &= -k!h_k(t, 0)e_{\ominus\lambda}(t, 0)|_0^{\infty} + \int_0^{\infty} k!h_{k-1}(t, 0)e_{\ominus\lambda}(\sigma(t), 0)\Delta t \\ &= -\frac{1}{\lambda} \int_0^{\infty} k!h_{k-1}(t, 0)(\ominus\lambda)e_{\ominus\lambda}(t, 0)\Delta t \\ &= \frac{k}{\lambda} \mathbb{E}_{\mathbb{T}}(X^{k-1}) \\ &= \frac{k(k-1)!}{\lambda \lambda^{k-1}} \\ &= \frac{k!}{\lambda^k}. \end{aligned}$$

This completes the proof of (14). □

Example 16 (Continuous case). Let $\mathbb{T} = \mathbb{R}$. Then $f(t) = \lambda e^{-\lambda t}$. This is the exact definition of the continuous exponential distribution. Also expected value ($\frac{1}{\lambda}$), variance ($\frac{1}{\lambda^2}$) and k -th moment ($\frac{k!}{\lambda^k}$) match with results found in the literature. Moreover, considering $\ominus\lambda$ as positively regressive, gives $1 + \mu(t)(\ominus\lambda) > 0$ and therefore $1 > 0$, which is true for all $\lambda \in \mathbb{R}$. So with the initial restriction of $\lambda > 0$, we get the exact definition from the continuous case, see [64, Section 5.2].

Example 17 (Discrete case). Let $\mathbb{T} = \mathbb{Z}$. Then

$$f(t) = \frac{\lambda}{(1 + \lambda)^{t+1}} = \frac{\lambda}{1 + \lambda} \left(1 - \frac{\lambda}{1 + \lambda}\right)^t.$$

If we let $p = \frac{\lambda}{1 + \lambda}$, then

$$f(t) = p(1 - p)^t.$$

This is the exact definition of the geometric distribution, which is the discrete equivalent of the continuous exponential function. p represents here the probability that a success occurs. Therefore $f(t)$ represents the probability of a success after t failures. Considering that $\ominus\lambda$ is positively regressive yields $1 + \mu(t)(\ominus\lambda) > 0$ and therefore $1 - \frac{\lambda}{1 + \lambda} > 0$ or $p < 1$. So with the initial restriction of $\lambda > 0$ follows also $p > 0$ and we get $p \in (0, 1)$, the property from the geometric distribution. Moreover, we have $\lambda = \frac{p}{1 - p}$ and

$$\mathbb{E}_{\mathbb{Z}}(X) = \frac{1}{\lambda} = \frac{1 - p}{p},$$

which matches exactly the property of the geometric distribution. We also have

$$\text{Var}_{\mathbb{Z}}(X) = \frac{1}{\lambda^2} = \frac{(1 - p)^2}{p^2},$$

which slightly differs from the variance of the geometric distribution $\frac{1 - p}{p^2}$. Moreover,

$$\mathbb{E}_{\mathbb{Z}}(X^k) = \frac{k!}{\lambda^k} = \frac{k!(1 - p)^k}{p^k}.$$

For the known corresponding discrete version of the geometric distribution, see [25, 39] applied to $f(t) = p(1 - p)^t$.

Example 18. Let $\mathbb{T} = h\mathbb{Z}$. Then

$$f(t) = \frac{\lambda}{(1 + \lambda h)^{\frac{t+h}{h}}} = \frac{\lambda}{1 + \lambda h} \left(\frac{1}{1 + \lambda h}\right)^{\frac{t}{h}} = \frac{1}{h} \left[\frac{\lambda h}{1 + \lambda h} \left(1 - \frac{\lambda h}{1 + \lambda h}\right)^{\frac{t}{h}} \right].$$

This is a new discrete distribution. Considering that $\ominus\lambda$ is positive regressive yields $1 - \frac{\lambda h}{1 + \lambda h} > 0$ and therefore $\frac{\lambda h}{1 + \lambda h} < 1$, which is true for all λ .

Now, a closed formula for the moment generating function for exponentially distributed random variables will be derived for all time scales.

Theorem 23. *Let $\lambda > 0$ be constant, $z \ominus \lambda < 0$, $t \in \mathbb{T}_0$ and the time scales probability density function of X be*

$$f(t) = \begin{cases} -(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then

$$M_X(z) = \frac{\lambda}{\lambda - z}. \quad (15)$$

Proof. First, note $\lambda > 0$ implies

$$\begin{aligned} (z \ominus \lambda)(t) &= (z \oplus (\ominus\lambda))(t) \\ &= z - \frac{\lambda}{1 + \mu(t)\lambda} - \frac{\mu(t)\lambda z}{1 + \mu(t)\lambda} \\ &= \frac{z + \mu(t)\lambda z - \lambda - \mu(t)\lambda z}{1 + \mu(t)\lambda} \\ &= \frac{z - \lambda}{1 + \mu(t)\lambda} \end{aligned}$$

and

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

Using this identity, we have

$$\begin{aligned} M_X(z) &= \int_0^\infty f(t)e_z(t, 0)\Delta t \\ &= - \int_0^\infty (\ominus\lambda)(t)e_{\ominus\lambda}(t, 0)e_z(t, 0)\Delta t \\ &= - \int_0^\infty (\ominus\lambda)(t)e_{z \ominus \lambda}(t, 0)\Delta t \\ &= - \int_0^\infty \left(\frac{\ominus\lambda}{z \ominus \lambda} (z \ominus \lambda) \right) (t)e_{z \ominus \lambda}(t, 0)\Delta t \\ &= - \frac{\lambda}{\lambda - z} \int_0^\infty (z \ominus \lambda)(t)e_{z \ominus \lambda}(t, 0)\Delta t \end{aligned}$$

$$\begin{aligned}
&= -\frac{\lambda}{\lambda - z} e_{z \ominus \lambda}(t, 0) \Big|_0^\infty \\
&= \frac{\lambda}{\lambda - z},
\end{aligned}$$

which is the desired property (15). □

Example 19. *Computation for the continuous and discrete cases yields:*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$M_X(z) = \frac{\lambda}{\lambda - z}.$$

This corresponds to the known result for the exponential function, see [64, p.66].

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$\begin{aligned}
M_X(z) &= \frac{\lambda}{\lambda - z} \\
&= \frac{\frac{p}{1-p}}{\frac{p}{1-p} - z} \\
&= \frac{p}{p - z + pz} \\
&= \frac{p}{1 - (1-p)(1+z)}.
\end{aligned}$$

Note that this corresponds to the structure of the moment-generating function of the geometric distribution $\frac{p}{1-(1-p)e^z}$, compare [25], again for $f(t) = p(1-p)^t$.

(iii) *If $\mathbb{T} = h\mathbb{Z}$, then*

$$M_X(z) = \frac{\lambda}{\lambda - z}.$$

The following three diagrams in 5.2 represent the time scales probability density function (pdf) and the time scales cumulative density function (cdf) for an exponentially distributed random variable with $\lambda = \frac{1}{2}$. The diagrams include continuous, discrete and $h\mathbb{Z}$ (with $h = \frac{1}{2}$) time scales cases.

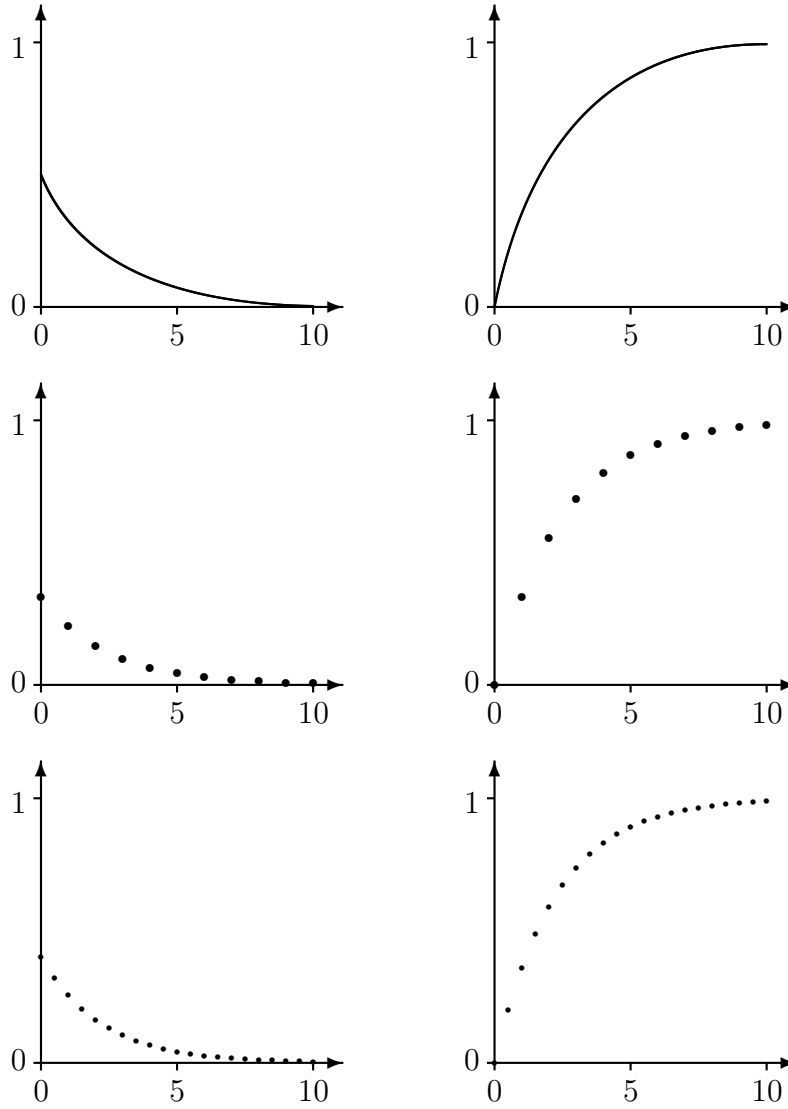


Figure 5.2. Exponential distribution

5.3. GAMMA DISTRIBUTION

Definition 24. Let $\lambda \in \mathbb{R}$, $\lambda > 0$ and define

$$\Lambda_0(t, t_0) = 0, \quad \Lambda_1(t, t_0) = 1$$

and then recursively

$$\Lambda_{k+1}(t, t_0) = - \int_{t_0}^t (\ominus \lambda)(\tau) \Lambda_k(\sigma(\tau), t_0) \Delta \tau \quad \text{for } k \in \mathbb{N}.$$

Remark 10. Note that

$$\Lambda_{k+1}^\Delta(t, t_0) = -(\ominus \lambda)(t) \Lambda_k(\sigma(t), t_0) \quad \text{for } k \in \mathbb{N}.$$

Definition 25. Let $\lambda > 0$ and $k \in \mathbb{N}$ and $t \in \mathbb{T}_0$. Then we define the time scales probability density function of the gamma distribution by

$$f_k(t) = \begin{cases} \frac{\lambda}{e_\lambda(\sigma(t), 0)} \Lambda_k(\sigma(t), 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases} \quad (16)$$

Moreover, we define

$$F_k(t) := - \left(\sum_{\nu=1}^k \Lambda_\nu(t, 0) \right) e_{\ominus \lambda}(t, 0).$$

If X is a random variable, which is gamma distributed, we write $X \sim \text{Gam}(k, \lambda)$.

Lemma 3. Let f_k and F_k be as in Definition 25. Then

$$F_k(\sigma(t)) = -\frac{1}{\lambda} \sum_{\nu=1}^k f_\nu(t).$$

Proof. Using the definitions, we have

$$\begin{aligned} F_k(\sigma(t)) &= - \left(\sum_{\nu=1}^k \Lambda_\nu(\sigma(t), 0) \right) e_{\ominus \lambda}(\sigma(t), 0) \\ &= -\frac{1}{\lambda} \frac{\lambda}{e_\lambda(\sigma(t), 0)} \sum_{\nu=1}^k \Lambda_\nu(\sigma(t), 0) \\ &= -\frac{1}{\lambda} \sum_{\nu=1}^k f_\nu(t), \end{aligned}$$

which completes the proof. □

Lemma 4. *Let f_k and F_k be as in Definition 25. Then*

$$F_k^\Delta = f_k \quad \text{on} \quad [0, \infty), \quad k \in \mathbb{N}.$$

Proof. Using the time scales product rule, we have

$$\begin{aligned} F_k^\Delta(t) &= \left(- \left(\sum_{\nu=1}^k \Lambda_\nu(\cdot, 0) \right) e_{\ominus\lambda}(\cdot, 0) \right)^\Delta(t) \\ &= (\ominus\lambda)(t) e_{\ominus\lambda}(t, 0) \sum_{\nu=1}^k \Lambda_{\nu-1}(\sigma(t), 0) - (\ominus\lambda)(t) e_{\ominus\lambda}(t, 0) \sum_{\nu=1}^k \Lambda_\nu(\sigma(t), 0) \\ &= (\ominus\lambda)(t) e_{\ominus\lambda}(t, 0) \left(\sum_{\nu=0}^{k-1} \Lambda_\nu(\sigma(t), 0) - \sum_{\nu=1}^k \Lambda_\nu(\sigma(t), 0) \right) \\ &= (\ominus\lambda)(t) e_{\ominus\lambda}(t, 0) (\Lambda_0(\sigma(t), 0) - \Lambda_k(\sigma(t), 0)) \\ &= -(\ominus\lambda)(t) e_{\ominus\lambda}(t, 0) \Lambda_k(\sigma(t), 0) \\ &= \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \Lambda_k(\sigma(t), 0) \\ &= f_k(t), \end{aligned}$$

which completes the proof. □

Remark 11. *The function f_k , defined in Definition 25 is a valid time scales probability density function. Note that, as $\ominus\lambda$ is positively regressive, we have*

$$\frac{\lambda}{e_{\lambda}(\sigma(t), 0)} = -\ominus\lambda e_{\ominus\lambda}(t, 0) > 0$$

and inductively $\Lambda_k(\sigma(t), 0) > 0$. For $k = 1$, we have $\Lambda_1(\sigma(t), 0) = 1 > 0$ and moreover,

$$\begin{aligned} \Lambda_{k+1}(\sigma(t), 0) &= - \int_0^{\sigma(t)} (\ominus\lambda)(\tau) \Lambda_k(\sigma(\tau), 0) \Delta\tau \\ &= \int_0^{\sigma(t)} \frac{\lambda}{1 + \mu(\tau)\lambda} \Lambda_k(\sigma(\tau), 0) \Delta\tau \\ &> 0. \end{aligned}$$

Therefore $f_k(t) > 0$ for $k \in \mathbb{N}$. Finally, using Lemma 4, we have

$$\begin{aligned} \int_0^\infty f_k(t) \Delta t &= \int_0^\infty F_k^\Delta(t) \Delta t \\ &= \int_0^\infty \left[- \left(\sum_{\nu=1}^k \Lambda_\nu(\cdot, 0) \right) e_{\ominus\lambda}(\cdot, 0) \right]^\Delta(t) \Delta t \end{aligned}$$

$$\begin{aligned}
&= - \left(\sum_{\nu=1}^k \Lambda_{\nu}(t, 0) \right) e_{\ominus\lambda}(t, 0) \Big|_0^{\infty} \\
&= - \lim_{t \rightarrow \infty} \left(\sum_{\nu=1}^k \Lambda_{\nu}(t, 0) \right) e_{\ominus\lambda}(t, 0) + \Lambda_1(0, 0) \\
&= 1.
\end{aligned}$$

Now we apply the definition of the gamma distribution to the first three cases.

Example 20. *If $k = 1$, then*

$$f_1(t) = \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \Lambda_1(\sigma(t), 0) = \frac{\lambda}{e_{\lambda}(\sigma(t), 0)}.$$

Note that this is the exact time scales probability density function of the exponential distribution. If $k = 2$, then

$$f_2(t) = \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \Lambda_2(\sigma(t), 0) = - \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \int_0^{\sigma(t)} (\ominus\lambda)(\tau) \Delta\tau.$$

If $k = 3$, then

$$\begin{aligned}
f_3(t) &= \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \Lambda_3(\sigma(t), 0) \\
&= - \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \int_0^{\sigma(t)} (\ominus\lambda)(\tau) \Lambda_2(\sigma(\tau), 0) \Delta\tau \\
&= \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \int_0^{\sigma(t)} (\ominus\lambda)(\tau) \int_0^{\sigma(\tau)} (\ominus\lambda)(s) \Lambda_1(\sigma(s), 0) \Delta s \Delta\tau \\
&= \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \int_0^{\sigma(t)} (\ominus\lambda)(\tau) \int_0^{\sigma(\tau)} (\ominus\lambda)(s) \Delta s \Delta\tau.
\end{aligned}$$

Lemma 5. *Let the graininess $\mu(t)$ be constant μ . Then*

$$\Lambda_{k+1}(\sigma(t), 0) = \left(\frac{\lambda}{1 + \mu\lambda} \right)^k g_k(\sigma(t), 0).$$

Proof. First note that

$$\begin{aligned}
\Lambda_{k+1}(\sigma(t), 0) &= - \int_0^{\sigma(t)} (\ominus\lambda)(\tau) \Lambda_k(\sigma(\tau), 0) \Delta\tau \\
&= \int_0^{\sigma(t)} \frac{\lambda}{1 + \mu\lambda} \Lambda_k(\sigma(\tau), 0) \Delta\tau.
\end{aligned}$$

This can be shown by induction. If $k = 1$, then

$$\Lambda_2(\sigma(t), 0) = \frac{\lambda}{1 + \mu\lambda} \int_0^{\sigma(t)} 1 \Delta\tau = \frac{\lambda}{1 + \mu\lambda} \sigma(t) = \frac{\lambda}{1 + \mu\lambda} g_1(\sigma(t), 0).$$

Therefore

$$\begin{aligned} \Lambda_{k+1}(\sigma(t), 0) &= \frac{\lambda}{1 + \mu\lambda} \int_0^{\sigma(t)} \Lambda_k(\sigma(\tau), 0) \Delta\tau \\ &= \left(\frac{\lambda}{1 + \mu\lambda} \right)^k \int_0^{\sigma(t)} g_{k-1}(\sigma(\tau), 0) \Delta\tau \\ &= \left(\frac{\lambda}{1 + \mu\lambda} \right)^k g_k(\sigma(t), 0) \end{aligned}$$

and the proof is complete. \square

Example 21. *Let us consider the following time scales.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$\begin{aligned} f_k(t) &= \lambda e^{-\lambda t} \left(\frac{\lambda}{1 + 0\lambda} \right)^{k-1} \frac{t^{k-1}}{(k-1)!} \\ &= \frac{\lambda^k e^{-\lambda t} t^{k-1}}{\Gamma(k)}. \end{aligned}$$

This corresponds exactly to the definition of the density function of the continuous gamma distribution for $k \in \mathbb{N}$. This is also known as the Erlang distribution, see [38, p.15]. In Figure 5.3, the probability density function and the cumulative density function of a random variable, which is gamma distributed with $\lambda = \frac{1}{2}$ for different values of k , is presented.

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$\begin{aligned} f_k(t) &= \frac{\lambda}{(1 + \lambda)^{t+1}} \left(\frac{\lambda}{1 + \lambda} \right)^{k-1} \frac{(t + 1 - 0 + (k-1) - 1)^{k-1}}{(k-1)!} \\ &= \frac{1}{(1 + \lambda)^t} \left(\frac{\lambda}{1 + \lambda} \right)^k \frac{(t + k - 1)^{k-1}}{(k-1)!} \\ &= (1 - p)^t p^k \frac{\Gamma(t + k)}{\Gamma(k)t!}. \end{aligned}$$

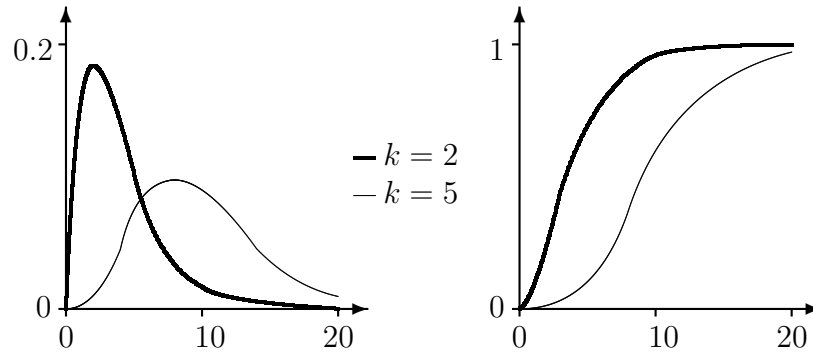


Figure 5.3. Gamma distribution (continuous case)

Note that this is equivalent to the definition of the density function of the negative binomial distribution, where t represents the number of failures until getting k successes, see [25, p.95]. The diagrams in Figure 5.4 represent the probability density function and the cumulative density function of a random variable, which is gamma distributed with $\lambda = \frac{1}{2}$ for different values of k .

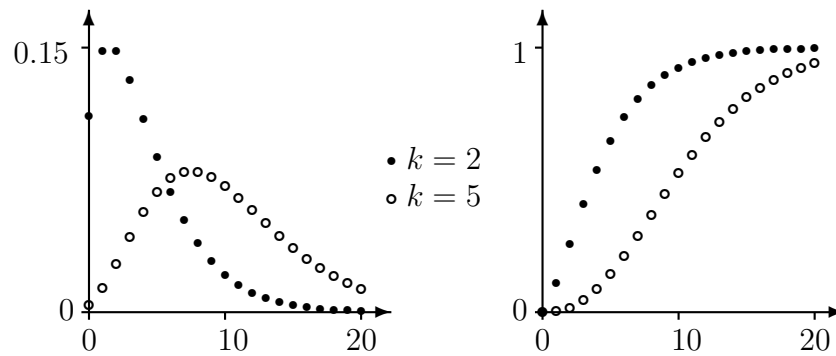


Figure 5.4. Gamma distribution (discrete case)

Theorem 24. Let $\lambda > 0$, $t \in \mathbb{T}_0$, $k \in \mathbb{N}$ and the time scales probability density function of X be

$$f_k(t) = \begin{cases} \frac{\lambda}{e_\lambda(\sigma(t), 0)} \Lambda_k(\sigma(t), 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then

$$\mathbb{E}_{\mathbb{T}}(X) = \frac{k}{\lambda} \tag{17}$$

and

$$\mathbb{V}\text{ar}_{\mathbb{T}}(X) = \frac{k}{\lambda^2}. \tag{18}$$

Proof. Using Lemma 3 and Lemma 4, we have

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X) &= \int_0^\infty t f_k(t) \Delta t = \int_0^\infty t F^D e_{\lambda} t a_k(t) \Delta t \\ &= t F_k(t) \Big|_0^\infty - \int_0^\infty F_k(\sigma(t)) \Delta t \\ &= \frac{1}{\lambda} \int_0^\infty \sum_{\nu=1}^k f_k(t) \Delta t \\ &= \frac{k}{\lambda}. \end{aligned}$$

This completes the proof of (17). Furthermore,

$$\begin{aligned} \mathbb{V}\text{ar}_{\mathbb{T}}(X) &= \int_0^\infty 2h_2(t, 0) f_k(t) \Delta t - (\mathbb{E}_{\mathbb{T}}(X))^2 \\ &= - \int_0^\infty 2h_2(t, 0) \left[\sum_{\nu=1}^k \Lambda_\nu(\cdot, 0) e_{\ominus\lambda}(\cdot, 0) \right]^\Delta (t) \Delta t - \left(\frac{k}{\lambda} \right)^2 \\ &= -2h_2(t, 0) \sum_{\nu=1}^k \Lambda_\nu(t, 0) e_{\ominus\lambda}(t, 0) \Big|_0^\infty \\ &\quad + \int_0^\infty 2h_1(t, 0) \sum_{\nu=1}^k \Lambda_\nu(\sigma(t), 0) e_{\ominus\lambda}(\sigma(t), 0) \Delta t - \left(\frac{k}{\lambda} \right)^2 \\ &= - \int_0^\infty 2t F_k(\sigma(t)) \Delta t - \left(\frac{k}{\lambda} \right)^2 \end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty \frac{2}{\lambda} t \sum_{\nu=1}^k f_\nu(t) \Delta t - \left(\frac{k}{\lambda}\right)^2 \\
&= \frac{2}{\lambda} \sum_{\nu=1}^k \mathbb{E}_{\mathbb{T}}(X_\nu) - \left(\frac{k}{\lambda}\right)^2 \\
&= \frac{2}{\lambda} \sum_{\nu=1}^k \frac{\nu}{\lambda} - \left(\frac{k}{\lambda}\right)^2 \\
&= \frac{2k(k+1)}{2\lambda^2} - \left(\frac{k}{\lambda}\right)^2 \\
&= \frac{k}{\lambda^2},
\end{aligned}$$

showing (18). □

Example 22. *We apply the expected value and variance result of Theorem 24 to the continuous and discrete time scales case.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$\mathbb{E}_{\mathbb{R}}(X) = \frac{k}{\lambda}$$

and

$$\text{Var}_{\mathbb{R}}(X) = \frac{k}{\lambda^2}.$$

Those are the same results that can be found in the literature for the continuous gamma distribution, compare [48, p.196].

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$\mathbb{E}_{\mathbb{Z}}(X) = \frac{k}{\lambda} = \frac{k}{\frac{p}{1-p}} = \frac{k(1-p)}{p}$$

and

$$\begin{aligned}
\text{Var}_{\mathbb{Z}}(X) &= \frac{k}{\lambda^2} \\
&= \frac{k}{\left(\frac{p}{1-p}\right)^2} \\
&= \frac{k(1-p)^2}{p^2}.
\end{aligned}$$

Note that the expectation matches exactly the case of the negative binomial distribution, whereas the variance is slightly different from $\frac{k(1-p)}{p^2}$, see [25, p.96].

Theorem 25. Let $\lambda > 0$, $t \in \mathbb{T}_0$, $k, p \in \mathbb{N}$ and the time scales probability density function of X be

$$f_k(t) = \begin{cases} \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \Lambda_k(\sigma(t), 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then we have the recursive formula for the p -th moment

$$\mathbb{E}_{\mathbb{T}}(X_k^p) = \frac{p}{\lambda} \sum_{\nu=1}^k \mathbb{E}_{\mathbb{T}}(X_{\nu}^{p-1}). \quad (19)$$

Proof. We have

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X_k^p) &= \int_0^{\infty} p! h_p(t, 0) f_k(t) \Delta t \\ &= - \int_0^{\infty} p! h_p(t, 0) \left[\sum_{\nu=1}^k \Lambda_{\nu}(\cdot, 0) e_{\ominus \lambda}(\cdot, 0) \right]^{\Delta} (t) \Delta t \\ &= \int_0^{\infty} p! h_{p-1}(t, 0) \sum_{\nu=1}^k \Lambda_{\nu}(\sigma(t), 0) e_{\ominus \lambda}(\sigma(t), 0) \Delta t \\ &= - \int_0^{\infty} p! h_{p-1}(t, 0) F_k(\sigma(t)) \Delta t \\ &= \int_0^{\infty} \frac{p!}{\lambda} h_{p-1}(t, 0) \sum_{\nu=1}^k f_{\nu}(t) \Delta t \\ &= \frac{p}{\lambda} \sum_{\nu=1}^k \mathbb{E}_{\mathbb{T}}(X_{\nu}^{p-1}), \end{aligned}$$

which completes the proof. □

Similarly as in Definition 24, we define $\tilde{\Lambda}(t, 0)$.

Definition 26. Let $\lambda \in \mathbb{R}$, $\lambda > 0$ and define

$$\tilde{\Lambda}_1(t, 0) = 1$$

and then recursively

$$\tilde{\Lambda}_{k+1}(t, 0) = - \int_0^t (z \ominus \lambda)(\tau) \tilde{\Lambda}_k(\sigma(\tau), 0) \Delta\tau \quad \text{for } k \in \mathbb{N}.$$

Remark 12. Note that

$$\tilde{\Lambda}_{k+1}^\Delta(t, 0) = -(z \ominus \lambda)(t) \tilde{\Lambda}_k(\sigma(t), 0) \quad \text{for } k \in \mathbb{N}.$$

Lemma 6. The relation between $\Lambda_k(t, 0)$ and $\tilde{\Lambda}_k(t, 0)$ is

$$\tilde{\Lambda}_k(t, 0) = \left(\frac{\lambda - z}{\lambda} \right)^{k-1} \Lambda_k(t, 0). \quad (20)$$

Proof. This is shown by induction. Clearly the claim is true for $k = 1$. Now assume the claim holds for $k - 1$. Then

$$\begin{aligned} \tilde{\Lambda}_k(t, 0) &= - \int_0^t (z \ominus \lambda)(\tau) \tilde{\Lambda}_{k-1}(\sigma(\tau), 0) \Delta\tau \\ &= - \int_0^t (z \ominus \lambda)(\tau) \left(\frac{\lambda - z}{\lambda} \right)^{k-2} \Lambda_{k-1}(\sigma(\tau), 0) \Delta\tau \\ &= - \left(\frac{\lambda - z}{\lambda} \right)^{k-1} \int_0^t (\ominus\lambda)(\tau) \Lambda_{k-1}(\sigma(\tau), 0) \Delta\tau \\ &= \left(\frac{\lambda - z}{\lambda} \right)^{k-1} \Lambda_k(t, 0), \end{aligned}$$

completing the proof. □

Lemma 7. Let $\Lambda_k(t, 0)$ and $\tilde{\Lambda}_k(t, 0)$ as previously defined.

$$\left[- \left(\frac{\lambda}{\lambda - z} \right)^k e_{z \ominus \lambda}(\cdot, 0) \sum_{\nu=1}^k \tilde{\Lambda}_\nu(\cdot, 0) \right]^\Delta(t) = -(\ominus\lambda)(t) e_{z \ominus \lambda}(t, 0) \Lambda_k(\sigma(t), 0). \quad (21)$$

Proof.

$$\begin{aligned} &\left[- \left(\frac{\lambda}{\lambda - z} \right)^k e_{z \ominus \lambda}(\cdot, 0) \sum_{\nu=1}^k \tilde{\Lambda}_\nu(\cdot, 0) \right]^\Delta(t) \\ &= - \left(\frac{\lambda}{\lambda - z} \right)^k (z \ominus \lambda)(t) e_{z \ominus \lambda}(t, 0) \sum_{\nu=1}^k \tilde{\Lambda}_\nu(\sigma(t), 0) \\ &\quad + \left(\frac{\lambda}{\lambda - z} \right)^k (z \ominus \lambda)(t) e_{z \ominus \lambda}(t, 0) \sum_{\nu=1}^k \tilde{\Lambda}_{\nu-1}(\sigma(t), 0) \end{aligned}$$

$$\begin{aligned}
&= - \left(\frac{\lambda}{\lambda - z} \right)^{k-1} (\ominus \lambda)(t) e_{z \ominus \lambda}(t, 0) \tilde{\Lambda}_k(\sigma(t), 0) \\
&= - \ominus \lambda(t) e_{z \ominus \lambda}(t, 0) \Lambda_k(\sigma(t), 0),
\end{aligned}$$

which completes the proof. \square

Now, a closed formula for the moment generating function for gamma distributed random variables will be derived for all time scales.

Theorem 26. *Let $\lambda > 0$ be constant, $z \ominus \lambda < 0$, and $t \in \mathbb{T}_0$ and the time scales probability density function of X be*

$$f_k(t) = \begin{cases} \frac{\lambda}{e_{\lambda}(\sigma(t), 0)} \Lambda_k(\sigma(t), 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then

$$M_X(z) = \left(\frac{\lambda}{\lambda - z} \right)^k. \tag{22}$$

Proof. Using the previous lemmas, we have

$$\begin{aligned}
M_X(z) &= \int_0^{\infty} e_z(t, 0) f(t) \Delta t \\
&= \int_0^{\infty} e_z(t, 0) (-\ominus \lambda(t)) e_{z \ominus \lambda}(t, 0) \Lambda_k(\sigma(t), 0) \Delta t \\
&= \int_0^{\infty} (-\ominus \lambda(t)) e_{z \ominus \lambda(t)}(t, 0) \Lambda_k(\sigma(t), 0) \Delta t \\
&= \left(\frac{\lambda}{\lambda - z} \right)^k \int_0^{\infty} \left[-e_{z \ominus \lambda}(\cdot, 0) \sum_{\nu=1}^k \tilde{\Lambda}_k(\cdot, 0) \right]^{\Delta} (t) \Delta t \\
&= \left(\frac{\lambda}{\lambda - z} \right)^k \tilde{\Lambda}_1(0, 0) \\
&= \left(\frac{\lambda}{\lambda - z} \right)^k.
\end{aligned}$$

The proof is complete. \square

Now, we apply the previous results to examine the chi-squared distribution with even degrees of freedom.

Definition 27. Let $\lambda > 0$, $\frac{\nu}{2} \in \mathbb{N}$ and $t \in \mathbb{T}_0$. Then, we define the time scales probability density function of the chi squared distribution by

$$f_\nu(t) = \begin{cases} \frac{1}{2e^{\frac{1}{2}(\sigma(t),0)}} \Lambda_{\frac{\nu}{2}}(\sigma(t), 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0, \end{cases} \quad (23)$$

where Λ is defined as before.

Remark 13. Note, that if X is a chi-squared distributed random variable, then

$$X \sim \text{Gam}\left(\frac{\nu}{2}, \frac{1}{2}\right),$$

where ν is an even positive integer, representing the degrees of freedom.

Theorem 27. Let X be a chi squared distributed random variable. Then

$$f_\nu(t) = \begin{cases} \frac{1}{2e^{\frac{1}{2}(\sigma(t),0)}} \Lambda_{\frac{\nu}{2}}(\sigma(t), 0), & t \geq 0 \\ 0, & t < 0 \end{cases}$$

is a proper time scales probability density function and

$$\mathbb{E}_{\mathbb{T}}(X) = \nu$$

and

$$\text{Var}_{\mathbb{T}}(X) = 2\nu.$$

Proof. As $X \sim \text{Gam}\left(\frac{\nu}{2}, \frac{1}{2}\right)$, we have $k = \frac{\nu}{2}$ and $\lambda = \frac{1}{2}$ in Theorem 24. Therefore f_ν is well defined as time scales probability density function and

$$\mathbb{E}_{\mathbb{T}}(X) = \frac{k}{\lambda} = \frac{\frac{\nu}{2}}{\frac{1}{2}} = \nu$$

and

$$\text{Var}_{\mathbb{T}}(X) = \frac{k}{\lambda^2} = \frac{\frac{\nu}{2}}{\left(\frac{1}{2}\right)^2} = 2\nu,$$

which completes the proof. □

Example 23. We apply the definition of the chi squared distribution to the continuous and discrete time scales case.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\begin{aligned} f_\nu(t) &= \frac{1}{2e_{\frac{1}{2}}(\sigma(t), 0)} \left(\frac{\frac{1}{2}}{1 + \mu_{\frac{1}{2}}} \right)^{\frac{\nu}{2}-1} g_{\frac{\nu}{2}-1}(\sigma(t), 0) \\ &= \frac{1}{2e^{\frac{1}{2}t}} \left(\frac{1}{2} \right)^{\frac{\nu}{2}-1} \frac{t^{\frac{\nu}{2}-1}}{(\frac{\nu}{2}-1)!} \\ &= \frac{e^{-\frac{1}{2}t} t^{\frac{\nu}{2}-1}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})}. \end{aligned}$$

Moreover,

$$\mathbb{E}_{\mathbb{R}}(X) = \nu \quad \text{and} \quad \text{Var}_{\mathbb{R}}(X) = 2\nu.$$

This is the same result that can be found in the literature for the chi squared distribution, see [48, p.196].

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$\begin{aligned} f_\nu(t) &= \frac{1}{2e_{\frac{1}{2}}(\sigma(t), 0)} \left(\frac{\frac{1}{2}}{1 + \mu_{\frac{1}{2}}} \right)^{\frac{\nu}{2}-1} g_{\frac{\nu}{2}-1}(\sigma(t), 0) \\ &= \frac{1}{2 \left(1 + \frac{1}{2}\right)^{t+1}} \left(\frac{\frac{1}{2}}{1 + \frac{1}{2}} \right)^{\frac{\nu}{2}-1} g_{\frac{\nu}{2}-1}(t+1, 0) \\ &= \left(\frac{2}{3} \right)^t \left(\frac{1}{3} \right)^{\frac{\nu}{2}} \frac{\Gamma(t + \frac{\nu}{2})}{\Gamma(\frac{\nu}{2})t!}. \end{aligned}$$

Moreover,

$$\mathbb{E}_{\mathbb{Z}}(X) = \nu \quad \text{and} \quad \text{Var}_{\mathbb{Z}}(X) = 2\nu.$$

6. CUMULATIVE DISTRIBUTION FUNCTION

Definition 28. Let f be a time scales probability density function. Then we define

$$F(x) := \int_0^x f(t)\Delta t \quad \text{for all } x \in \mathbb{T} \quad (24)$$

to be the cumulative distribution function (cdf).

Definition 29. Let f be a time scales probability density function. Then we define

$$p(X < x) := F(x) = \int_0^x f(t)\Delta t \quad \text{for all } x \in \mathbb{T} \quad (25)$$

to be the probability that X is less than a given value x .

Remark 14. Note that this definition differs slightly from the one found in [25, p.35], as we have $p(X < x)$ instead of $p(X \leq x)$. This is due to the fact that the upper bound of the integral has density 0 in the time scales theory.

Remark 15. Note that we also have

$$F(x) = p(X < x) = \int_0^x f(x)\Delta x = 1 - \int_x^\infty f(x)\Delta x = 1 - p(X \geq x).$$

Definition 30. Let $A, B \subset \mathbb{T}$. Then, we define the conditional probability by

$$p(A|B) := \frac{p(A \cap B)}{p(B)}. \quad (26)$$

Lemma 8. We have

$$\int_0^\infty \int_{\sigma(y)}^\infty f(x)\Delta x \Delta y = \int_0^\infty \int_0^x f(x)\Delta y \Delta x. \quad (27)$$

Proof. Let us define the two functions F and G by

$$F(z) = \int_0^z \int_{\sigma(y)}^z f(x)\Delta x \Delta y$$

and

$$G(z) = \int_0^z \int_0^x f(x)\Delta y \Delta x.$$

Clearly $F(0) = G(0) = 0$ and

$$G^\Delta(z) = \int_0^z f(z)\Delta y = zf(z).$$

Letting $f(z, y) = \int_{\sigma(y)}^z f(x)\Delta x$ and therefore $F(z) = \int_0^z f(z, y)\Delta y$ in [22, Theorem 1.117 (i)] and applying [22, Theorem 1.75], we get

$$\begin{aligned} F^\Delta(z) &= \int_0^z f^{\Delta z}(z, y)\Delta y + f(\sigma(z), z) \\ &= \int_0^z f(z)\Delta y + \int_{\sigma(z)}^{\sigma(z)} f(x)\Delta x \\ &= zf(z). \end{aligned}$$

Therefore $F(z) = G(z)$ for all $z \in \mathbb{T}_0$. This completes the proof of (27). \square

Theorem 28. *Let f be a time scales probability density function and F_X be the corresponding cumulative distribution function. Then we have*

$$\mathbb{E}_{\mathbb{T}}(X) = \int_0^\infty (1 - F_X(\sigma(y))) \Delta y. \quad (28)$$

Proof. Using the fact that f is a time scales probability density function and Lemma 8, we get

$$\begin{aligned} \int_0^\infty (1 - F_X(\sigma(y))) \Delta y &= \int_0^\infty \left(1 - \int_0^{\sigma(y)} f(x)\Delta x \right) \Delta y \\ &= \int_0^\infty \int_{\sigma(y)}^\infty f(x)\Delta x \Delta y \\ &= \int_0^\infty \int_0^x f(x)\Delta y \Delta x \\ &= \int_0^\infty f(x) \int_0^x 1\Delta y \Delta x \\ &= \int_0^\infty xf(x)\Delta x \\ &= \mathbb{E}_{\mathbb{T}}(X), \end{aligned}$$

which completes the proof. \square

The following theorem considers some basic properties of the cumulative distribution function.

Theorem 29. *Let f be a time scales probability density function and F_X be the corresponding cumulative distribution function. Then the following properties hold.*

- (i) $F_X(0) = 0$.
- (ii) $F_X(\infty) = 1$.
- (iii) *If $x, h(x) \in \mathbb{T}$ and $x \leq h(x)$, then*

$$F_X(x) \leq F_X(h(x)). \quad (29)$$

Proof. To see (i) and (ii), note that

$$F_X(0) = \int_0^0 f(t)\Delta t = 0$$

and

$$F_X(\infty) = \lim_{x \rightarrow \infty} F_X(x) = \lim_{x \rightarrow \infty} \int_0^x f(t)\Delta t = \int_0^\infty f(t)\Delta t = 1.$$

To see (29), note that $f(t) \geq 0$ and therefore

$$F_X(x) = \int_0^x f(t)\Delta t \leq \int_0^{h(x)} f(t)\Delta t = F_X(h(x)).$$

The proof is complete. □

Remark 16. *Moreover, it holds, that*

$$p(a \leq X \leq b) = \int_a^{\sigma(b)} f(t)\Delta t = F_X(\sigma(b)) - F_X(a).$$

Similarly, we get

$$p(a \leq X < b) = \int_a^b f(t)\Delta t = F_X(b) - F_X(a),$$

$$p(a < X < b) = \int_{\sigma(a)}^b f(t)\Delta t = F_X(b) - F_X(\sigma(a))$$

and

$$p(a < X \leq b) = \int_{\sigma(a)}^{\sigma(b)} f(t)\Delta t = F_X(\sigma(b)) - F_X(\sigma(a)).$$

Note that in the original probability theory $p(a \leq X \leq b)$ was given as

$$p(a \leq X \leq b) = F_X(b) - \lim_{h \searrow 0} F_X(a - h),$$

where

$$\lim_{h \searrow 0} F_X(a - h) = F_X(a) - p(X = a).$$

Due to the slightly changed definition on time scales, we need to deal with $\sigma(a)$ and $\sigma(b)$.

We apply the definition of the cumulative distribution function to the uniform and the exponential distributions.

Theorem 30. Let $a, b \in \mathbb{T}$, $a \leq t \leq b$ and the time scales probability density function of X be

$$f(t) = \begin{cases} \frac{1}{\sigma(b)-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$F(x) = \frac{x - a}{\sigma(b) - a}. \quad (30)$$

Proof. Using the definition, we have

$$F(x) = \int_0^x f(t) \Delta t = \int_a^x \frac{1}{\sigma(b) - a} \Delta t = \frac{x - a}{\sigma(b) - a},$$

which shows (30). □

Theorem 31. Let $\lambda > 0$, $t \in \mathbb{T}_0$ and the time scales probability density function of X be

$$f(t) = \begin{cases} -\ominus \lambda e_{\ominus \lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then

$$F(x) = 1 - e_{\ominus \lambda}(x, 0). \quad (31)$$

Proof. Using the definition, we have

$$F(x) = \int_0^x f(t)\Delta t = \int_0^x -\ominus \lambda e_{\ominus\lambda}(t, 0)\Delta t = -e_{\ominus\lambda}(t, 0)|_0^x = 1 - e_{\ominus\lambda}(x, 0),$$

which shows (31). \square

Definition 31. Let X and Y be two random variables. We say $f_{X,Y}(x, y)$ is a joint time scales probability density function if

1. $f_{X,Y}(x, y) \geq 0 \quad \forall x, y \in \mathbb{T}$,
2. $\int_0^\infty \int_0^\infty f_{X,Y}(x, y)\Delta y\Delta x = 1$.

Moreover, we define the joint time scales cumulative distribution function by

$$F_{X,Y}(x, y) := p(X < x, Y < y) = \int_0^x \int_0^y f_{X,Y}(s, t)\Delta t\Delta s.$$

The following theorem considers some basic properties of the joint cumulative distribution function.

Theorem 32. Let $f_{X,Y}$ be a joint time scales probability density function and F_X be the corresponding cumulative distribution function. Then the following properties hold.

- (i) $F_{X,Y}(x, 0) = F_{X,Y}(0, y) = 0$,
- (ii) $F_{X,Y}(\infty, \infty) = 1$,
- (iii) $F_{X,Y}(x, \infty) = F_X(x)$ and $F_{X,Y}(\infty, y) = F_Y(y)$,
- (iv) If $x, y, h_1(x), h_2(y) \in \mathbb{T}$, $x \leq h_1(x)$ and $y \leq h_2(y)$, then

$$F_{X,Y}(x, y) \leq F_{X,Y}(h_1(x), h_2(y)).$$

Proof. We have

$$F_{X,Y}(x, 0) = \int_0^x \int_0^0 f_{X,Y}(s, t)\Delta t\Delta s = \int_0^x 0\Delta s = 0$$

and

$$F_{X,Y}(0, y) = \int_0^0 \int_0^y f_{X,Y}(s, t)\Delta t\Delta s = 0,$$

which shows (i),

$$F_{X,Y}(\infty, \infty) = \int_0^\infty \int_0^\infty f_{X,Y}(s, t) \Delta t \Delta s = 1,$$

which shows (ii),

$$F_{X,Y}(x, \infty) = p(X < x, Y < \infty) = p(X < x) = F_X(x)$$

and similarly

$$F_{X,Y}(\infty, y) = p(X < \infty, Y < y) = p(Y < y) = F_Y(y),$$

which shows (iii), and

$$\begin{aligned} F_{X,Y}(x, y) &= \int_0^x \int_0^y f_{X,Y}(s, t) \Delta t \Delta s \\ &\leq \int_0^x \int_0^{h_2(y)} f_{X,Y}(s, t) \Delta t \Delta s \\ &\leq \int_0^{h_1(x)} \int_0^{h_2(y)} f_{X,Y}(s, t) \Delta t \Delta s \\ &= F_{X,Y}(h_1(x), h_2(y)), \end{aligned}$$

which shows (iv). □

Theorem 33. *Let X, Y be two random variables with joint time scales probability density function $f_{X,Y}(x, y)$. Then, we have*

$$f_X(x) = \int_0^\infty f_{X,Y}(x, y) \Delta y \tag{32}$$

and

$$f_Y(y) = \int_0^\infty f_{X,Y}(x, y) \Delta x. \tag{33}$$

Proof. To show (32), we use the cumulative density function

$$\int_0^x f_X(s) \Delta s = F_X(x) = F_{X,Y}(x, \infty) = \int_0^x \int_0^\infty f_{X,Y}(s, t) \Delta t \Delta s.$$

Therefore

$$f_X(s) = \int_0^\infty f_{X,Y}(s,t)\Delta t.$$

Finally, (33) is shown in the same way. \square

Remark 17. Note that f_X and f_Y are called the marginal time scales density functions.

Now the bivariate weighted uniform distribution is considered on time scales, see [10, p.41] for the nonweighted version.

Example 24. Let X, Y be two random variables with joint time scales probability density function

$$f_{X,Y}(x,y) = \begin{cases} \frac{xy}{(h_2(b,a)+a(b-a))^2}, & \text{if } a \leq x, y < b \\ 0, & \text{otherwise.} \end{cases}$$

Clearly as $0 \leq a < b$, we have $f_{X,Y}(x,y) \geq 0$ and moreover,

$$\begin{aligned} \int_a^b t\Delta t &= \int_a^b (t-a)\Delta t + \int_a^b a\Delta t \\ &= h_2(b,a) + a(b-a). \end{aligned}$$

Therefore

$$\begin{aligned} \int_a^b \int_a^b st\Delta t\Delta s &= \int_a^b t\Delta t \int_a^b s\Delta s \\ &= (h_2(b,a) + a(b-a))^2, \end{aligned}$$

which results in

$$\begin{aligned} \int_a^b \int_a^b f_{X,Y}(s,t)\Delta t\Delta s &= \int_a^b \int_a^b \frac{st}{(h_2(b,a) + a(b-a))^2}\Delta t\Delta s \\ &= \frac{1}{(h_2(b,a) + a(b-a))^2} \int_a^b \int_a^b st\Delta t\Delta s \\ &= \frac{1}{(h_2(b,a) + a(b-a))^2} (h_2(b,a) + a(b-a))^2 \\ &= 1. \end{aligned}$$

So $f_{X,Y}(x, y)$ is a valid joint time scales probability density function. The marginal density functions are given by

$$\begin{aligned}
 f_X(x) &= \int_0^\infty f_{X,Y}(x, y) \Delta y \\
 &= \int_a^b \frac{xy}{(h_2(b, a) + a(b - a))^2} \Delta y \\
 &= \frac{x}{(h_2(b, a) + a(b - a))^2} \int_a^b y \Delta y \\
 &= \frac{x}{h_2(b, a) + a(b - a)}
 \end{aligned}$$

and

$$f_Y(y) = \frac{y}{h_2(b, a) + a(b - a)}.$$

The next example applies the previous result to different time scales.

Example 25. We compute the joint and marginal time scales density functions.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\begin{aligned}
 f_{X,Y}(x, y) &= \frac{xy}{\left(\frac{(b-a)^2}{2} + a(b-a)\right)^2} \\
 &= \frac{xy}{\left((b-a)\left(\frac{b-a}{2} + a\right)\right)^2} \\
 &= \frac{xy}{\left((b-a)\left(\frac{b+a}{2}\right)\right)^2} \\
 &= \frac{4xy}{(b^2 - a^2)^2},
 \end{aligned}$$

$$f_X(x) = \frac{2x}{b^2 - a^2} \quad \text{and} \quad f_Y(y) = \frac{2y}{b^2 - a^2}.$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$\begin{aligned}
 f_{X,Y}(x, y) &= \frac{xy}{\left(\binom{b-a}{2} + a(b-a)\right)^2} \\
 &= \frac{xy}{\left(\frac{(b-a)!}{2!(b-a-2)!} + a(b-a)\right)^2} \\
 &= \frac{4xy}{\left((b-a)(b-a-1) + a(b-a)\right)^2} \\
 &= \frac{4xy}{\left((b-a)(b+a-1)\right)^2},
 \end{aligned}$$

$$f_X(x) = \frac{2x}{(b-a)(b+a-1)} \quad \text{and} \quad f_Y(y) = \frac{2y}{(b-a)(b+a-1)}.$$

(iii) If $\mathbb{T} = q^{\mathbb{N}_0}$, $a = q^m$ and $b = q^n$, then

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{xy}{\left(\frac{(q^n - q^m)(q^n - qq^m)}{1+q} + q^m(q^n - q^m)\right)^2} \\ &= \frac{xy}{\left((q^n - q^m)\left(\frac{q^n - q^{m+1}}{1+q} + q^m\right)\right)^2} \\ &= \frac{(1+q)^2 xy}{\left((q^n - q^m)(q^n + q^m)\right)^2} \\ &= \frac{(1+q)^2 xy}{(q^{2n} - q^{2m})^2}, \end{aligned}$$

$$f_X(x) = \frac{(1+q)x}{q^{2n} - q^{2m}} \quad \text{and} \quad f_Y(y) = \frac{(1+q)y}{q^{2n} - q^{2m}}.$$

Definition 32. We say that two random variables X and Y are independent if

$$F_{X,Y}(x,y) = F_X(x)F_Y(y),$$

or equivalently

$$p(X < x, Y < y) = p(X < x)p(Y < y).$$

Theorem 34. Let X and Y be two linearly independent random variables with marginal probability density function f_X and f_Y , respectively, and joint probability density function $f_{X,Y}$. Then we have

$$f_{X,Y}(x,y) = f_X(x)f_Y(y). \tag{34}$$

Proof. We have

$$\begin{aligned} \int_0^x \int_0^y f_{X,Y}(s,t) \Delta t \Delta s &= F_{X,Y}(x,y) = F_X(x)F_Y(y) \\ &= \int_0^x f_X(s) \Delta s \int_0^y f_Y(t) \Delta t \\ &= \int_0^x \int_0^y f_X(s)f_Y(t) \Delta t \Delta s, \end{aligned}$$

and therefore (34) is shown. □

Definition 33. Let X and Y be two random variables with marginal probability density function f_X and f_Y , respectively, and joint probability density function $f_{X,Y}$. Then the time scales conditional probability density functions are defined as

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \quad \text{where } f_Y(y) > 0$$

and

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad \text{where } f_X(x) > 0.$$

Example 26. Considering the bivariate weighted uniform distribution from Example 24, we have

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{xy}{(h_2(b,a)+a(b-a))^2}}{\frac{y}{h_2(b,a)+a(b-a)}} = \frac{x}{h_2(b,a) + a(b-a)}$$

and similarly

$$f_{Y|X=x}(y) = \frac{y}{h_2(b,a) + a(b-a)}.$$

Note that here the conditional density functions are equal to the marginal density functions. Moreover, X and Y are linearly independent.

Theorem 35 (Markov inequality). Let $a \in \mathbb{T}_0$. Then

$$p(X \geq a) \leq \frac{\mathbb{E}_{\mathbb{T}}(X)}{a}. \quad (35)$$

Proof. We have

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X) &= \int_0^{\infty} tf(t)\Delta t \\ &= \int_0^a tf(t)\Delta t + \int_a^{\infty} tf(t)\Delta t \\ &\geq \int_a^{\infty} tf(t)\Delta t \\ &\geq a \int_a^{\infty} f(t)\Delta t \\ &= ap(X \geq a) \end{aligned}$$

This completes the proof of (35). □

Remark 18. *As we consider random variables with positive support, this version of the Markov inequality matches exactly the one given in the literature, see [39, p.311].*

Theorem 36 (Tschebycheff inequality). *Let $\varepsilon > 0$. Then*

$$\frac{\text{Var}_{\mathbb{T}}(X) - \mathbb{E}_{\mathbb{T}}(2H(X))}{\varepsilon^2} \geq p((X - \mathbb{E}_{\mathbb{T}}(X))^2 \geq \varepsilon^2), \quad (36)$$

where the density function of $H(X)$ is $h_2(t, 0) - \frac{t^2}{2}$.

Proof. To prove (36), just note that

$$\text{Var}_{\mathbb{T}}(X) - \mathbb{E}_{\mathbb{T}}(2H(X)) = \text{Var}(X)$$

and

$$\mathbb{E}_{\mathbb{T}}(X) = \mathbb{E}(X),$$

where $\mathbb{E}(X)$ and $\text{Var}(X)$ represent the usual expected value and variance. Then applying the usual Tschebycheff inequality [64, p.77]

$$\frac{\text{Var}(X)}{\varepsilon^2} \geq p(|X - \mathbb{E}(X)| \geq \varepsilon)$$

yields the desired result. □

Example 27. *For different time scales, we derive the following inequalities.*

(i) *If $\mathbb{T} = \mathbb{R}$, then note that the density function of $H(X)$ is $\frac{t^2}{2} - \frac{t^2}{2} = 0$, and therefore*

$$\begin{aligned} \frac{\text{Var}_{\mathbb{R}}(X)}{\varepsilon^2} &\geq p((X - \mathbb{E}_{\mathbb{R}}(X))^2 \geq \varepsilon^2) \\ &= p(|X - \mathbb{E}_{\mathbb{R}}(X)| \geq \varepsilon). \end{aligned}$$

This matches exactly the Tschebycheff inequality found in [64, p.77].

(ii) *If $\mathbb{T} = \mathbb{Z}$, then note that the density function of $H(X)$ is $\frac{t(t-1)}{2} - \frac{t^2}{2} = -\frac{t}{2}$, and therefore*

$$\begin{aligned} \frac{\text{Var}_{\mathbb{Z}}(X) + \mathbb{E}_{\mathbb{Z}}(X)}{\varepsilon^2} &\geq p((X - \mathbb{E}_{\mathbb{Z}}(X))^2 \geq \varepsilon^2) \\ &= p(|X - \mathbb{E}_{\mathbb{Z}}(X)| \geq \varepsilon). \end{aligned}$$

Note that, due to the new definition of the variance, this is a slightly different Tschebycheff inequality for the discrete case, compared to the one, that is found in the literature.

7. FURTHER PROPERTIES AND ENTROPY

7.1. THE SUM OF INDEPENDENT RANDOM VARIABLES

Let X, Y be two independent random variables. Then the density function of $Z = X + Y$ can be expressed by the convolution of the density function of X with the density function of Y . In the discrete case, we have

$$p_Z(z) = (p_X * p_Y)(z) = \sum_{k=-\infty}^{\infty} p_X(k)p_Y(z - k).$$

Equivalently, for the continuous case, we have

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x)dx.$$

Proofs of these theorems can be found in any standard probability theory book, see [39]. For the time scales equivalent, we will take a different approach. First, we will examine some convolution properties, which will lead to a joint time scales probability density function representing the sum of two independent random variables for certain time scales.

Lemma 9. *Let f be a time scales probability density function and $F(t) = \int_{-\infty}^t f(u)\Delta u$. Moreover, let \widehat{f} be the shift of f and $\lim_{t \rightarrow -\infty} \widehat{f}(t, \sigma(s)) = 0$. Then*

$$\widehat{F}(t, s) = \int_{-\infty}^t \widehat{f}(u, s)\Delta u. \tag{37}$$

Proof. First, define

$$G(t, s) = \int_{-\infty}^t \widehat{f}(u, s)\Delta u.$$

Now

$$G(t, t_0) = \int_{-\infty}^t \widehat{f}(u, t_0)\Delta u = \int_{-\infty}^t f(u)\Delta u = F(t),$$

$$G^{\Delta t}(t, s) = \widehat{f}(t, s)$$

and

$$\begin{aligned}
G^{\Delta s}(t, s) &= \int_{-\infty}^t \widehat{f}^{\Delta s}(u, s) \Delta u \\
&= - \int_{-\infty}^t \widehat{f}^{\Delta t}(u, \sigma(s)) \Delta u \\
&= -\widehat{f}(t, \sigma(s)) + \lim_{t \rightarrow -\infty} \widehat{f}(t, \sigma(s)) \\
&= -\widehat{f}(t, \sigma(s)) \\
&= -G^{\Delta t}(t, \sigma(s)).
\end{aligned}$$

Therefore

$$G(t, s) = \widehat{F}(t, s),$$

proving (37). □

Theorem 37. *Let f_X, f_Y be time scales probability density functions and F_Z the joint time scales probability density function, where $Z = Z(X, Y)$. Moreover, let*

$$F_Z(z) = (F_X * f_Y)(z).$$

Then the time scales probability density function of Z is

$$f_Z = f_X * f_Y. \tag{38}$$

Proof. We have

$$\begin{aligned}
\int_{-\infty}^z (f_X * f_Y)(u) \Delta u &= \int_{-\infty}^z \int_{-\infty}^{\infty} \widehat{f}_X(u, \sigma(s)) f_Y(s) \Delta s \Delta u \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^z \widehat{f}_X(u, \sigma(s)) f_Y(s) \Delta u \Delta s \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^z f_Y(s) \left(\widehat{f}_X(u, \sigma(s)) \Delta u \right) \Delta s \\
&= \int_{-\infty}^{\infty} f_Y(s) \widehat{F}_X(z, \sigma(s)) \Delta s \\
&= (F_X * f_Y)(z) \\
&= F_Z(z),
\end{aligned}$$

completing the proof of (38). □

Example 28. For different time scales, we derive the following joint distributions, if $F_Z(z) = (F_X * f_Y)(z)$.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\begin{aligned}
 F_Z(z) &= (F_X * f_Y)(z) \\
 &= \int_{-\infty}^{\infty} F_X(z-s) f_Y(s) ds \\
 &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-s} f_X(t) dt \right) f_Y(s) ds \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-s} f_X(t) f_Y(s) dt ds \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-s} f_{X,Y}(t, s) dt ds \\
 &= \iint_{\{(t,s)|s+t \leq z\}} f_{X,Y}(t, s) dt ds \\
 &= p(X + Y \leq z).
 \end{aligned}$$

So we can conclude that the unknown random variable Z is $Z = X + Y$.

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$\begin{aligned}
 F_Z(z) &= (F_X * f_Y)(z) \\
 &= \sum_{s=-\infty}^{\infty} F_X(z-s-1) f_Y(s) \\
 &= \sum_{s=-\infty}^{\infty} \left(\sum_{t=-\infty}^{z-s-2} f_X(t) \right) f_Y(s) \\
 &= \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{z-s-2} f_{X,Y}(t, s) \\
 &= \sum_{\{(t,s)|s+t+1 < z\}} f_{X,Y}(t, s) \\
 &= p(X + Y + 1 < z).
 \end{aligned}$$

So we can conclude that the unknown random variable Z is $Z = X + Y + 1$.

(iii) If $\mathbb{T} = h\mathbb{Z}$, then

$$F_Z(z) = (F_X * f_Y)(z)$$

$$\begin{aligned}
&= \sum_{s=\mathbb{R} \cap \mathbb{T}} F_X(z-s-h) f_Y(s) h \\
&= \sum_{s=\mathbb{R} \cap \mathbb{T}} \left(\sum_{t \in [-\infty, z-s-h] \cap \mathbb{T}} f_X(t) h \right) f_Y(s) h \\
&= \sum_{s=\mathbb{R} \cap \mathbb{T}} \sum_{t \in [-\infty, z-s-h] \cap \mathbb{T}} f_{X,Y}(t, s) h^2 \\
&= \sum_{\{(t,s) \in \mathbb{T}^2 \mid s+t+h < z\}} f_{X,Y}(t, s) \\
&= p(X + Y + h < z).
\end{aligned}$$

So we can conclude that the unknown random variable Z is $Z = X + Y + h$.

7.2. THE MEMORYLESS PROPERTY

Definition 34. Let X be a time scales random variable, $t, \tau \in \mathbb{T}_0$ and $\tau \geq t$. Then we say that X satisfies the memoryless property if

$$p(X \geq \tau | X \geq t) = p(\widehat{X \geq \cdot})(\tau, t). \quad (39)$$

Remark 19. Note that, because of

$$p(X \geq \tau | X \geq t) = \frac{p(X \geq \tau \cap X \geq t)}{p(X \geq t)} = \frac{p(X \geq \tau)}{p(X \geq t)},$$

(39) is equivalent to

$$p(X \geq \tau) = p(X \geq t) p(\widehat{X \geq \cdot})(\tau, t).$$

Example 29. For the continuous and the discrete cases, we have the following representations of the memoryless property (39).

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\begin{aligned}
p(X \geq \tau) &= p(X \geq t) p(\widehat{X \geq \cdot})(\tau, t) \\
&= p(X \geq t) p(X \geq \tau - t).
\end{aligned}$$

If we let $s \in \mathbb{R}_0^+$ and $\tau = t + s$, then

$$p(X \geq t + s) = p(X \geq t) p(X \geq t + s - t)$$

$$= p(X \geq t)p(X \geq s).$$

This is also equivalent to

$$p(X \geq t + s | X \geq t) = p(X \geq s),$$

which is the well-known memoryless property found for example in [64].

(ii) If $\mathbb{T} = \mathbb{Z}$, then we get similarly

$$\begin{aligned} p(X \geq \tau) &= p(X \geq t) \widehat{p(X \geq \cdot)}(\tau, t) \\ &= p(X \geq t)p(X \geq \tau - t). \end{aligned}$$

If we let $s \in \mathbb{N}_0$ and $\tau = t + s$, then

$$\begin{aligned} p(X \geq t + s) &= p(X \geq t)p(X \geq t + s - t) \\ &= p(X \geq t)p(X \geq s). \end{aligned}$$

This is also equivalent to

$$p(X \geq t + s | X \geq t) = p(X \geq s),$$

which is the well-known memoryless property found for example in [64].

Theorem 38. Let $h : \mathbb{T}_0 \rightarrow \mathbb{T}_0$, $h(t) > t$ and the time scales probability density function of X be defined by

$$f(t) = \begin{cases} -\ominus \lambda e_{\ominus \lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Then

$$p(X \geq h(t) | X \geq t) = e_{\ominus \lambda}(h(t), t). \quad (40)$$

Proof. First note that

$$p(X < t) = \int_0^t -(\ominus \lambda)(\tau) e_{\ominus \lambda}(\tau, 0) \Delta \tau$$

$$\begin{aligned}
&= -e_{\ominus\lambda}(\tau, 0)|_0^t \\
&= 1 - e_{\ominus\lambda}(t, 0).
\end{aligned}$$

Therefore, we have

$$p(X \geq t) = e_{\ominus\lambda}(t, 0).$$

Now, we have

$$\begin{aligned}
p(X \geq h(t)|X \geq t) &= \frac{p(X \geq h(t) \cap X \geq t)}{p(X \geq t)} \\
&= \frac{p(X \geq h(t))}{p(X \geq t)} \\
&= \frac{e_{\ominus\lambda}(h(t), 0)}{e_{\ominus\lambda}(t, 0)} \\
&= e_{\ominus\lambda}(h(t), t),
\end{aligned}$$

which is the desired memoryless property for the exponential function (40). \square

Remark 20. Note that, with the help of [17, Example 2.3], Theorem 38 matches the memoryless property from (39), as for constant $\ominus\lambda$, i.e., constant graininess

$$\begin{aligned}
\frac{p(X \geq \tau)}{p(X \geq t)} &= p(\widehat{X \geq \cdot})(\tau, t) \\
&= e_{\ominus\lambda}(\widehat{\cdot}, 0)(\tau, t) \\
&= e_{\ominus\lambda}(\tau, t).
\end{aligned}$$

Replacing τ with $h(t)$ gives the desired property.

Example 30 (Continuous case). Let $\mathbb{T} = \mathbb{R}$, $t \in \mathbb{R}$ and $h(t) = t + s$, where $s \in \mathbb{R}^+$. Then

$$p(X \geq h(t)|X \geq t) = p(X \geq t + s|X \geq t) = e_{\ominus\lambda}(t + s, t) = e^{-\lambda(t+s-t)} = e^{-\lambda s}.$$

This property is independent of t and matches exactly with the memoryless property for the exponential distribution, found in [64, p.284].

Example 31 (Discrete case). Let $\mathbb{T} = \mathbb{Z}$, $t \in \mathbb{Z}$ and $h(t) = t + k$, where $k \in \mathbb{N}$. Then

$$p(X \geq h(t)|X \geq t) = p(X \geq t + k|X \geq t) = e_{\ominus\lambda}(t + k, t)$$

$$= \frac{1}{(1+\lambda)^{t+k-t}} = \frac{1}{(1+\lambda)^k} = \left(1 - \frac{\lambda}{1+\lambda}\right)^k = (1-p)^k.$$

This property is independent of t and matches exactly with the memoryless property for the geometric distribution, compare [35, p.48].

Example 32. Let $\mathbb{T} = h\mathbb{Z}$, $h > 0$, $t \in h\mathbb{Z}$ and $h(t) = t + hk$, where $k \in \mathbb{N}$. Then

$$\begin{aligned} p(X \geq h(t)|X \geq t) &= p(X \geq t + hk|X \geq t) = e_{\ominus\lambda}(t + hk, t) \\ &= \frac{1}{(1+\lambda h)^{\frac{t+hk-t}{h}}} = \frac{1}{(1+\lambda h)^k} \\ &= \left(1 - \frac{h\lambda}{1+h\lambda}\right)^k. \end{aligned}$$

This property is independent of t and represents a memoryless property for $\mathbb{T} = h\mathbb{Z}$.

Example 33 (Quantum calculus case). Let $\mathbb{T} = q^{\mathbb{N}_0}$, $q > 1$ and $h(t) = q^k t$, where $k \in \mathbb{N}$. Then

$$\begin{aligned} p(X \geq h(t)|X \geq t) &= p(X \geq q^k t|X \geq t) \\ &= e_{\ominus\lambda}(q^k t, t) \\ &= \frac{1}{e_{\lambda}(q^k t, t)} \\ &= \frac{1}{\prod_{s \in [t, q^k t)} (1 + \lambda s (q-1))} \\ &= \frac{1}{[1 + \lambda t (q-1)] [1 + \lambda q t (q-1)] \cdots [1 + \lambda q^{k-1} t (q-1)]}. \end{aligned}$$

This property is not independent of t , and therefore does not match the known definition of the memoryless property.

Remark 21. The memoryless property in the old sense, i.e., that $p(X \geq h(t)|X \geq t)$ is independent of t , holds for all time scales with constant graininess μ . In order to get the result for all time scales, we can redefine the memoryless property to be

$$p(X \geq h(t)|X \geq t) = e_{\ominus\lambda}(h(t), t).$$

A further property of the exponential distribution with constant graininess is computing the probability that $X_1 \leq X_2$ for two exponentially distributed random variables X_1 and X_2 .

Theorem 39. Let the time scales probability density function of X_i be defined by

$$f_{X_i}(t) = \begin{cases} -\ominus \lambda_i e_{\ominus \lambda_i}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0 \end{cases}$$

for $i = 1, 2$ and constant graininess $\mu(t) = \mu$. Then

$$p(X_1 \leq X_2) = \frac{\lambda_1(1 + \mu\lambda_2)}{\lambda_1 + \lambda_2 + \mu\lambda_1\lambda_2}. \quad (41)$$

Proof. First note, that

$$\begin{aligned} (\ominus \lambda_1) \oplus (\ominus \lambda_2) &= \frac{-\lambda_1}{1 + \mu\lambda_1} \oplus \frac{-\lambda_2}{1 + \mu\lambda_2} \\ &= \frac{-\lambda_1}{1 + \mu\lambda_1} + \frac{-\lambda_2}{1 + \mu\lambda_2} + \mu \frac{\lambda_1\lambda_2}{(1 + \mu\lambda_1)(1 + \mu\lambda_2)} \\ &= -\frac{\lambda_1 + \lambda_2 + \mu\lambda_1\lambda_2}{(1 + \mu\lambda_1)(1 + \mu\lambda_2)}, \end{aligned}$$

and therefore by using conditional probability

$$\begin{aligned} p(X_1 \leq X_2) &= \int_0^\infty p(X_1 \leq X_2 | X_1 = t) f_{X_1}(t) \Delta t \\ &= \int_0^\infty p(t \leq X_2) f_{X_1}(t) \Delta t \\ &= \int_0^\infty (1 - p(X_2 < t)) (-\ominus \lambda_1 e_{\ominus \lambda_1}(t, 0)) \Delta t \\ &= \int_0^\infty e_{\ominus \lambda_2}(t, 0) (-\ominus \lambda_1 e_{\ominus \lambda_1}(t, 0)) \Delta t \\ &= \int_0^\infty \frac{\lambda_1}{1 + \mu\lambda_1} e_{(\ominus \lambda_1) \oplus (\ominus \lambda_2)}(t, 0) \Delta t \\ &= -\frac{\lambda_1}{1 + \mu\lambda_1} \frac{\lambda_1 + \lambda_2 + \mu\lambda_1\lambda_2}{(1 + \mu\lambda_1)(1 + \mu\lambda_2)} \\ &\quad \times \int_0^\infty ((\ominus \lambda_1) \oplus (\ominus \lambda_2)) e_{(\ominus \lambda_1) \oplus (\ominus \lambda_2)}(t, 0) \Delta t \\ &= \frac{\lambda_1(1 + \mu\lambda_2)}{\lambda_1 + \lambda_2 + \mu\lambda_1\lambda_2} \int_0^\infty (e_{(\ominus \lambda_1) \oplus (\ominus \lambda_2)}(t, 0))^\Delta \Delta t \\ &= \frac{\lambda_1(1 + \mu\lambda_2)}{\lambda_1 + \lambda_2 + \mu\lambda_1\lambda_2}, \end{aligned}$$

which completes the proof. □

Example 34. For different time scales, we have the following.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$p(X_1 \leq X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

which can be found in [64, p.292] with $<$ instead of \leq . As $p(X_1 = X_2) = 0$, we get the same result.

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$p(X_1 \leq X_2) = \frac{\lambda_1(1 + \lambda_2)}{\lambda_1 + \lambda_2 + \lambda_1\lambda_2}.$$

(iii) If $\mathbb{T} = h\mathbb{Z}$, then

$$p(X_1 \leq X_2) = \frac{\lambda_1(1 + h\lambda_2)}{\lambda_1 + \lambda_2 + h\lambda_1\lambda_2}.$$

7.3. ENTROPY

In information theory, the quantity *entropy* describes the uncertainty of a random variable. The word itself refers to the Shannon entropy (“A Mathematical Theory of Communication” [68]), which measures the average information one is missing if the exact value of the random variable is unknown. As in information theory most of the computation is done in bits, it is not surprising that we find a base 2 for the logarithm in the definition of the entropy. In other papers and books, see [27], the natural logarithm is used, which will be done in this thesis as well.

Definition 35. Let f be a time scales probability density function of the random variable X . Then we define the entropy as follows

$$H(X) := - \int_0^\infty f(t) \log f(t) \Delta t. \quad (42)$$

We assume here that the integral exists.

Example 35. For different time scales, we derive the following entropies.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$H(X) = - \int_0^\infty f(t) \log f(t) dt.$$

This matches the definition of the differential entropy for a random variable with positive support found in [27, p.243].

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$H(X) = - \sum_{t=0}^{\infty} f(t) \log f(t).$$

This is the entropy of a discrete random variable, found in [27, p.14]

(iii) If $\mathbb{T} = q^{\mathbb{N}_0}$, then

$$H(X) = -(q-1) \sum_{t=0}^{\infty} q^t f(q^t) \log f(q^t).$$

Note that the time scales integral starts at 1 as this is the first possible point having a density greater than zero in the quantum calculus case.

(iv) If $\mathbb{T} = h\mathbb{Z}$, then

$$H(X) = - \sum_{t=0}^{\infty} hf(ht) \log f(ht).$$

This is now slightly different from the previously defined entropy of a discrete random variable. In time scales theory, we have an additional factor h in the definition.

Next, the entropy for different distributions will be considered. We start with the uniform distribution.

Theorem 40. Let $a, b \in \mathbb{T}_0$, $a \leq t \leq b$ and the time scales probability density function of X be

$$f(t) = \begin{cases} \frac{1}{\sigma(b)-a}, & \text{if } a \leq t \leq b \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$H(X) = \log(\sigma(b) - a). \tag{43}$$

Proof. Computation yields

$$\begin{aligned}
H(X) &= - \int_0^\infty f(t) \log f(t) \Delta t \\
&= - \int_a^{\sigma(b)} \frac{1}{\sigma(b) - a} \log \frac{1}{\sigma(b) - a} \Delta t \\
&= - \frac{1}{\sigma(b) - a} \log \frac{1}{\sigma(b) - a} \int_a^{\sigma(b)} \Delta t \\
&= \frac{1}{\sigma(b) - a} \log(\sigma(b) - a)(\sigma(b) - a) \\
&= \log(\sigma(b) - a),
\end{aligned}$$

which shows (43). □

Remark 22. *This corresponds to the entropies for $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$. In the first case we have $\log(b - a)$, see [27, p.244], and in the second $\log n$, see [28, p.440], where n is the number of points where the density is larger than zero. However, for other isolated time scales this result is different from the entropy of a discrete random variable, as it depends now on the length of the interval, instead of the number of points with positive density.*

Now, the exponential distribution will be considered.

Theorem 41. *Let $\lambda > 0$ be constant, $t \in \mathbb{T}_0$ and the time scales probability density function of X be*

$$f(t) = \begin{cases} -\ominus \lambda e_{\ominus\lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Moreover, assume that $\mu(t) \neq 0$. Then

$$H(X) = -\log \lambda + \int_0^\infty e_{\ominus\lambda}(t, 0) \frac{\log(1 + \mu(t)\lambda)}{\mu(t)} \Delta t. \tag{44}$$

If the graininess $\mu(t)$ is constant $\mu \neq 0$, then we have

$$H(X) = -\log \lambda + \frac{(1 + \mu\lambda) \log(1 + \mu\lambda)}{\mu\lambda}. \tag{45}$$

Proof. Note that

$$\begin{aligned}
(\log e_\lambda(\cdot, 0))^\Delta(t) &= \frac{\log e_\lambda(\sigma(t), 0) - \log e_\lambda(t, 0)}{\mu(t)} \\
&= \frac{\log\left(\frac{e_\lambda(\sigma(t), 0)}{e_\lambda(t, 0)}\right)}{\mu(t)} \\
&= \frac{\log\left((1 + \mu(t)\lambda)\frac{e_\lambda(t, 0)}{e_\lambda(t, 0)}\right)}{\mu(t)} \\
&= \frac{\log(1 + \mu(t)\lambda)}{\mu(t)}.
\end{aligned}$$

Using this and the integration by parts formula, we have

$$\begin{aligned}
H(X) &= -\int_0^\infty f(t) \log f(t) \Delta t \\
&= \int_0^\infty (\ominus\lambda)(t) e_{\ominus\lambda}(t, 0) \log(-(\ominus\lambda)(t) e_{\ominus\lambda}(t, 0)) \Delta t \\
&= -\int_0^\infty \frac{\lambda}{e_\lambda(\sigma(t), 0)} \log\left(\frac{\lambda}{e_\lambda(\sigma(t), 0)}\right) \Delta t \\
&= -\int_0^\infty \frac{\lambda}{e_\lambda(\sigma(t), 0)} \log \lambda \Delta t + \int_0^\infty \frac{\lambda}{e_\lambda(\sigma(t), 0)} \log e_\lambda(\sigma(t), 0) \Delta t \\
&= -\log \lambda + \int_0^\infty [-e_{\ominus\lambda}(\cdot, 0)]^\Delta(t) \log e_\lambda(\sigma(t), 0) \Delta t \\
&= -\log \lambda + \left[-e_{\ominus\lambda}(t, 0) \log e_\lambda(t, 0) \Big|_0^\infty + \int_0^\infty e_{\ominus\lambda}(t, 0) (\log e_\lambda(\cdot, 0))^\Delta(t) \Delta t \right] \\
&= -\log \lambda + \int_0^\infty e_{\ominus\lambda}(t, 0) (\log e_\lambda(\cdot, 0))^\Delta(t) \Delta t \\
&= -\log \lambda + \int_0^\infty e_{\ominus\lambda}(t, 0) \frac{\log(1 + \mu(t)\lambda)}{\mu(t)} \Delta t.
\end{aligned}$$

This proves equation (44). If $\mu \neq 0$ is constant, then

$$\begin{aligned}
H(X) &= -\log \lambda + \int_0^\infty e_{\ominus\lambda}(t, 0) \frac{\log(1 + \mu\lambda)}{\mu} \Delta t \\
&= -\log \lambda + \frac{\log(1 + \mu\lambda)}{\mu} \int_0^\infty \frac{-\ominus\lambda}{-\ominus\lambda} e_{\ominus\lambda}(t, 0) \Delta t \\
&= -\log \lambda + \frac{\log(1 + \mu\lambda)}{-(\ominus\lambda)\mu} \\
&= -\log \lambda + \frac{\log(1 + \mu\lambda)}{\frac{\lambda}{1 + \mu\lambda}\mu} \\
&= -\log \lambda + \frac{(1 + \mu\lambda) \log(1 + \mu\lambda)}{\mu\lambda},
\end{aligned}$$

which completes the proof of (45) and therefore is the conclusion of the proof of Theorem 41. \square

Example 36. *For different time scales with constant graininess function, we have the following.*

(i) *If $\mathbb{T} = \mathbb{R}$, then using the definition, we have*

$$\begin{aligned}
 H(X) &= - \int_0^{\infty} f(t) \log f(t) dt \\
 &= - \int_0^{\infty} \lambda e^{-\lambda t} \log(\lambda e^{-\lambda t}) dt \\
 &= - \int_0^{\infty} \lambda \log(\lambda e^{-\lambda t}) dt + \int_0^{\infty} \lambda^2 t e^{-\lambda t} dt \\
 &= \log \lambda e^{-\lambda t} \Big|_0^{\infty} - \lambda t e^{-\lambda t} \Big|_0^{\infty} + \int_0^{\infty} \lambda e^{-\lambda t} dt \\
 &= -\log \lambda - e^{-\lambda t} \Big|_0^{\infty} \\
 &= -\log \lambda + 1.
 \end{aligned}$$

Also note, that this matches the statement of Theorem 41, as we consider the limit as μ approaches 0

$$\begin{aligned}
 H(X) &= \lim_{\mu \rightarrow 0} -\log \lambda + \frac{(1 + \mu\lambda) \log(1 + \mu\lambda)}{\mu\lambda} \\
 &= -\log \lambda + \lim_{\mu \rightarrow 0} \frac{\frac{(1 + \mu\lambda)\lambda}{1 + \mu\lambda} + \lambda \log(1 + \mu\lambda)}{\lambda} \\
 &= -\log \lambda + 1.
 \end{aligned}$$

This matches the definition of the differential entropy for a random variable with exponential distribution, see [45, p.61].

(ii) *If $\mathbb{T} = \mathbb{Z}$, then $\mu = 1$ and*

$$\begin{aligned}
 H(X) &= -\log \lambda + \frac{(1 + \lambda) \log(1 + \lambda)}{\lambda} \\
 &= \frac{-\lambda \log \lambda + (1 + \lambda) \log(1 + \lambda)}{\lambda} \\
 &= \frac{-\frac{p}{1-p} \log \frac{p}{1-p} + (1 + \frac{p}{1-p}) \log(1 + \frac{p}{1-p})}{\frac{p}{1-p}} \\
 &= \frac{-p \log \frac{p}{1-p} + \log \frac{1}{1-p}}{p}
 \end{aligned}$$

$$\begin{aligned}
&= \frac{-p \log p + p \log(1-p) - \log(1-p)}{p} \\
&= \frac{-p \log p - (1-p) \log(1-p)}{p}.
\end{aligned}$$

This is the entropy of a discrete random variable with geometric distribution, see [44, p.210].

(iii) If $\mathbb{T} = h\mathbb{Z}$, then $\mu = h$ and

$$\begin{aligned}
H(X) &= -\log \lambda + \frac{(1+h\lambda) \log(1+h\lambda)}{h\lambda} \\
&= \frac{-h\lambda \log \lambda + (1+h\lambda) \log(1+h\lambda)}{h\lambda}.
\end{aligned}$$

Note that, similarly to the discrete case, if $p = \frac{h\lambda}{1+h\lambda}$ or equivalently $\lambda = \frac{p}{h-hp}$, we have

$$\begin{aligned}
H(X) &= \frac{-h \frac{p}{h-hp} \log \frac{p}{h-hp} + (1 + h \frac{p}{h-hp}) \log(1 + h \frac{p}{h-hp})}{h \frac{p}{h-hp}} \\
&= \frac{-\frac{p}{1-p} \log \frac{p}{h-hp} + (1 + \frac{p}{1-p}) \log(1 + h \frac{p}{h-hp})}{\frac{p}{1-p}} \\
&= \frac{-p \log \frac{p}{h-hp} + \log \frac{h-hp+p}{h-hp}}{p} \\
&= \frac{-p \log p - (1-p) \log(h-hp) + \log(h-hp+p)}{p}.
\end{aligned}$$

The next definitions and theorems unify some properties of entropy from [27], using time scales theory.

Definition 36. Let X and Y be two random variables with density function f_X and f_Y , respectively, and joint density distribution $f_{X,Y}$. Then we define the joint entropy by

$$H(X, Y) := - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_{X,Y}(x, y) \Delta x \Delta y \quad (46)$$

and the conditional entropy by

$$H(X|Y) := - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_{X|Y=y}(x) \Delta x \Delta y. \quad (47)$$

Theorem 42. Let X and Y be two random variables with density function f_X and f_Y , respectively, and joint density distribution $f_{X,Y}$. Then

$$H(X|Y) = H(X, Y) - H(Y). \quad (48)$$

If X and Y are independent, then

$$H(X, Y) = H(X) + H(Y). \quad (49)$$

Proof. Using Definition 36, we have

$$\begin{aligned} H(X|Y) &= - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_{X|Y=y}(x) \Delta x \Delta y \\ &= - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log \frac{f_{X,Y}(x, y)}{f_Y(y)} \Delta x \Delta y \\ &= - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_{X,Y}(x, y) \Delta x \Delta y \\ &\quad + \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_Y(y) \Delta x \Delta y \\ &= H(X, Y) + \int_0^\infty f_Y(y) \log f_Y(y) \Delta y \\ &= H(X, Y) - H(Y), \end{aligned}$$

and if X and Y are independent, then

$$\begin{aligned} H(X, Y) &= - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_{X,Y}(x, y) \Delta x \Delta y \\ &= - \int_0^\infty \int_0^\infty f_X(x) f_Y(y) \log(f_X(x) f_Y(y)) \Delta x \Delta y \\ &= - \int_0^\infty f_Y(y) \int_0^\infty f_X(x) \log f_X(x) \Delta x \Delta y \\ &\quad - \int_0^\infty f_X(x) \int_0^\infty f_Y(y) \log f_Y(y) \Delta y \Delta x \\ &= H(Y) \int_0^\infty f_X(x) \Delta x + H(X) \int_0^\infty f_Y(y) \Delta y \\ &= H(Y) + H(X), \end{aligned}$$

which completes the proof of (48) and (49). \square

Definition 37. Let f and g be two time scales probability density functions. Then we define relative entropy by

$$D(f||g) := \int_0^\infty f(t) \log \frac{f(t)}{g(t)} \Delta t,$$

where we assume that we have $0 \log \frac{0}{g} = 0$ for $f = 0$.

Definition 38. Let X and Y be two random variables with density function f_X and f_Y , respectively, and joint density distribution $f_{X,Y}$. Then we define the mutual information by

$$I(X, Y) := \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log \frac{f_{X,Y}(x, y)}{f_X(x)f_Y(y)} \Delta x \Delta y. \quad (50)$$

The previous definitions match exactly the definitions for relative entropy and mutual information given in [27].

Theorem 43. Let X and Y be two random variables with density function f_X and f_Y , respectively, and joint density distribution $f_{X,Y}$. Then the following relationship between mutual information and entropy holds:

$$I(X, Y) = H(X) + H(Y) - H(X, Y). \quad (51)$$

Proof. Using the properties of the logarithm and marginal densities, we have

$$\begin{aligned} I(X, Y) &= \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log \frac{f_{X,Y}(x, y)}{f_X(x)f_Y(y)} \Delta x \Delta y \\ &= \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_{X,Y}(x, y) \Delta x \Delta y \\ &\quad - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_X(x) \Delta x \Delta y \\ &\quad - \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log f_Y(y) \Delta x \Delta y \\ &= -H(X, Y) - \int_0^\infty f_X(x) \log f_X(x) \Delta x - \int_0^\infty f_Y(y) \log f_Y(y) \Delta y \\ &= -H(X, Y) + H(X) + H(Y), \end{aligned}$$

and (51) is shown. □

Theorem 44. *Let X and Y be two random variables with density function f_X and f_Y , respectively, and joint density distribution $f_{X,Y}$. Then, we have*

$$I(X, Y) = I(Y, X) \quad (52)$$

$$= H(Y) - H(Y|X) \quad (53)$$

$$= H(X) - H(X|Y) \quad (54)$$

$$= H(X) + H(Y) - H(X, Y). \quad (55)$$

Proof. To see (52), note that $f_{X,Y} = f_{Y,X}$ and therefore

$$\begin{aligned} I(X, Y) &= \int_0^\infty \int_0^\infty f_{X,Y}(x, y) \log \frac{f_{X,Y}(x, y)}{f_X(x)f_Y(y)} \Delta x \Delta y \\ &= \int_0^\infty \int_0^\infty f_{Y,X}(y, x) \log \frac{f_{Y,X}(y, x)}{f_Y(y)f_X(x)} \Delta x \Delta y \\ &= I(Y, X). \end{aligned}$$

Equation (55) follows from Theorem 43. Finally as

$$H(X, Y) = H(X|Y) + H(Y),$$

we have

$$\begin{aligned} I(X, Y) &= H(X) + H(Y) - H(X, Y) \\ &= H(X) + H(Y) - (H(X|Y) + H(Y)) \\ &= H(X) - H(X|Y), \end{aligned}$$

and 54 follows. Similarly, we can show (53). \square

Next, we consider a time scales probability theory version of the Jensen inequality.

Theorem 45 (Jensen's inequality). *Let f be a time scales probability density function. Let $g : \mathbb{T} \rightarrow \mathbb{R}$ be rd-continuous and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be continuous and convex. Then*

$$\varphi \left(\int_0^\infty f(t)g(t)\Delta t \right) \leq \int_0^\infty f(t)\varphi(g(t))\Delta t. \quad (56)$$

Proof. We define $x_0 \in \mathbb{R}$ as follows

$$x_0 := \int_0^\infty f(t)g(t)\Delta t.$$

As φ is a continuous and convex function, the existence of subderivatives is guaranteed, compare [25, p.190]. Therefore there exist a and b such that

$$ax + b \leq \varphi(x) \quad \forall x \in \mathbb{R}$$

and

$$ax_0 + b = \varphi(x_0).$$

If we let $x = g(t)$, then

$$ag(t) + b \leq \varphi(g(t)),$$

and we have

$$\begin{aligned} \int_0^\infty f(t)\varphi(g(t))\Delta t &\geq \int_0^\infty f(t) [ag(t) + b] \Delta t \\ &= a \int_0^\infty f(t)g(t)\Delta t + b \int_0^\infty f(t)\Delta t \\ &= ax_0 + b \\ &= \varphi(x_0) \\ &= \varphi\left(\int_0^\infty f(t)g(t)\Delta t\right), \end{aligned}$$

proving (56). □

Example 37. For the continuous and the discrete time scales cases, we get well known results.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\varphi\left(\int_0^\infty f(t)g(t)dt\right) \leq \int_0^\infty f(t)\varphi(g(t))dt.$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$\varphi \left(\sum_{k=0}^{\infty} f(k)g(k) \right) \leq \sum_{k=0}^{\infty} f(k)\varphi(g(k)).$$

Note that these properties in the original probability theory translate to

$$\varphi(E(g(t))) \leq E(\varphi(g(t))),$$

see [25, p.190].

Next, we use the Jensen inequality to prove the information inequality.

Theorem 46 (Information inequality). *Let f and g be two time scales probability density functions. Then*

$$D(f||g) \geq 0, \tag{57}$$

and

$$D(f||g) = 0 \tag{58}$$

if and only if $f = g$.

Proof. Let $A := \{t \in \mathbb{T} | f(t) > 0\}$. Using that the logarithm is a strictly concave function and applying the Jensen inequality, we have

$$\begin{aligned} -D(f||g) &= - \int_0^{\infty} f(t) \log \frac{f(t)}{g(t)} \Delta t \\ &= \int_0^{\infty} f(t) \log \frac{g(t)}{f(t)} \Delta t \\ &= \int_A f(t) \log \frac{g(t)}{f(t)} \Delta t \\ &\leq \log \left(\int_A f(t) \frac{g(t)}{f(t)} \Delta t \right) \\ &= \log \left(\int_A g(t) \Delta t \right) \\ &\leq \log \left(\int_0^{\infty} g(t) \Delta t \right) \\ &= \log 1 = 0, \end{aligned}$$

and therefore, (57) holds. To achieve equality in (57), we need to have equality for both \leq in the previous proof. Equality in the first estimation holds if and only if $\frac{g(t)}{f(t)}$ is constant, say $\frac{g(t)}{f(t)} = c$. Therefore $g(t) = cf(t)$ and

$$\int_A g(t)\Delta t = c \int_A f(t)\Delta t = c.$$

Equality in the second estimation holds if and only if

$$\int_A g(t)\Delta t = \int_0^\infty g(t)\Delta t = 1.$$

Now it follows that $c = 1$ and therefore $f(t) = g(t)$. The proof of (58) is complete. \square

8. ECONOMIC APPLICATIONS

8.1. INTEREST RATES AND NET PRESENT VALUE

Definition 39. We define the nominal rate of interest to be the interest that is paid before accounting for effects of inflation or compounding. This interest rate is usually given in the basic time unit (per annum). We denote the nominal rate of interest by r . The interest rate that accounts for compounding is called effective rate of interest, denoted by r_E , and the one accounting for inflation is called real rate of interest, denoted by r_R . If i is the rate of inflation, we have

$$1 + r_R = \frac{1 + r}{1 + i}.$$

Note, that this is considered after exactly one year without compounding. A common approximation for the real interest rate is

$$r_R = r - i,$$

see [42, 24].

Definition 40. Let $K(t_0)$ be the initial wealth at time $t = t_0$ and let there be no inflation. Then we define the wealth at time t , denoted by $K(t)$ to be

$$K(t) := K(t_0)e_r(t, t_0),$$

and the effective interest rate over the time interval from t_0 to t can be computed by

$$r_E := e_r(t, t_0) - 1.$$

Example 38. Let $K(t_0)$ be the initial wealth at $t = t_0$.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$K(t) = K(t_0)e^{r(t-t_0)}$$

and

$$r_E = e^{r(t-t_0)} - 1.$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$K(t) = K(t_0)(1+r)^{t-t_0}$$

and

$$r_E = (1+r)^{t-t_0} - 1.$$

(iii) If $\mathbb{T} = q^{\mathbb{N}_0}$, then

$$K(t) = K(t_0) \prod_{s \in [t_0, t)} (1 + (q-1)rs)$$

and

$$r_E = \prod_{s \in [t_0, t)} (1 + (q-1)rs) - 1.$$

(iv) If $\mathbb{T} = \frac{1}{m}\mathbb{Z}$, then

$$K(t) = K(t_0) \left(1 + \frac{r}{m}\right)^{m(t-t_0)}$$

and

$$r_E = \left(1 + \frac{r}{m}\right)^{m(t-t_0)} - 1$$

This corresponds to periodic compounding with frequency m .

For time scales with constant graininess, similar results can be found in [42, 24].

Example 39. *Let $t_0 = 1$, $t = 2$, $K(t_0) = 100$ and $r = 0.05$. Then, we get the following wealths for different time scales.*

(i) If $\mathbb{T} = \mathbb{R}$, then

$$K(2) = 100e^{0.05(2-1)} \approx 105.1271.$$

Here the compounding is done continuously. The effective rate of interest over this one year time period is 5.1271%.

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$K(2) = 100(1 + 0.05)^{2-1} = 105.$$

Here the compounding is done once a year. The effective rate of interest over this one year time period is 5% and therefore equal to the nominal rate of interest.

(iii) If $\mathbb{T} = \frac{1}{12}\mathbb{Z}$, then

$$K(2) = 100 \left(1 + \frac{1}{12}0.05 \right)^{\frac{2-1}{\frac{1}{12}}} \approx 105.1162.$$

Here the compounding is done monthly. The effective rate of interest over this one year time period is 5.1162%.

(iv) If $\mathbb{T} = \frac{1}{365}\mathbb{Z}$, then

$$K(2) = 100 \left(1 + \frac{1}{365}0.05 \right)^{\frac{2-1}{\frac{1}{365}}} \approx 105.1267.$$

Here the compounding is done daily. The effective rate of interest over this one year time period is 5.1267%.

Definition 41. Let $K(t)$ be the wealth at time t and let there be no inflation. Then we define the present value at time t_0 , denoted by $PV(t_0)$, to be

$$PV(t_0) := K(t)e_{\ominus r}(t, t_0).$$

Example 40. Let $K(t)$ be the wealth at t and $t_0 \in \mathbb{T}$.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$PV(t_0) = \frac{K(t)}{e^{r(t-t_0)}}.$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$PV(t_0) = \frac{K(t)}{(1+r)^{t-t_0}}.$$

(iii) If $\mathbb{T} = q^{\mathbb{N}_0}$, then

$$PV(t_0) = \frac{K(t)}{\prod_{s \in [t_0, t)} (1 + (q-1)rs)}.$$

(iv) If $\mathbb{T} = h\mathbb{Z}$, then

$$PV(t_0) = \frac{K(t)}{(1+hr)^{\frac{t-t_0}{h}}}.$$

For time scales with constant graininess, equivalent results can be found in [42, 24] for the present value.

Definition 42. Assume that $\mu(t) \neq 0$ for all $t \in \mathbb{T}$ (\star). Let $K(t)$ be a fixed stream of payments over a given time period \mathbb{T} , called annuity. Then we define the accumulated value of this annuity at time t , denoted by $V(t)$, by

$$V(t) := \int_{t_0}^t \frac{K(\tau)e_r(t, \tau)}{\mu(\tau)} \Delta\tau. \quad (59)$$

Remark 23. Note, that $\mathbb{T} = \mathbb{R}$ or any continuous part of our time scale does not make sense in the economic setting. This would mean that you constantly receive or make payments. Therefore the accumulated value of the annuity would be infinity. This could be fixed by letting the stream of payments equal zero for any continuous part of the time scale. So we consider time scales where $\mu(t) \neq 0$ for all $t \in \mathbb{T}$. Consequently, without loss of generality we only consider isolated points. Therefore, we can rewrite $V(t)$ as

$$\begin{aligned} V(t) &= \int_{t_0}^t \frac{K(\tau)e_r(t, \tau)}{\mu(\tau)} \Delta\tau \\ &= \sum_{\tau \in [t_0, t)} \frac{\mu(\tau)K(\tau)e_r(t, \tau)}{\mu(\tau)} \\ &= \sum_{\tau \in [t_0, t)} K(\tau)e_r(t, \tau). \end{aligned}$$

Example 41. Let $K(t)$ be a fixed stream of payments, r the nominal rate of interest and $t_0 \in \mathbb{T}$.

(i) If $\mathbb{T} = \mathbb{Z}$, then

$$\begin{aligned} V(t) &= \sum_{\tau=t_0}^{t-1} K(\tau)(1+r)^{t-\tau} \\ &= K(t_0)(1+r)^{t-t_0} + K(t_0+1)(1+r)^{t-t_0-1} + \dots + K(t-1)(1+r). \end{aligned}$$

(ii) If $\mathbb{T} = h\mathbb{Z}$, then

$$\begin{aligned} V(t) &= \sum_{\tau \in [t_0, t)} K(\tau)(1+hr)^{\frac{t-\tau}{h}} \\ &= K(t_0)(1+hr)^{\frac{t-t_0}{h}} + K(t_0+h)(1+hr)^{\frac{t-t_0-h}{h}} \\ &\quad + \dots + K(t-h)(1+hr)^{\frac{t-t+h}{h}} \\ &= K(t_0)(1+hr)^{\frac{t-t_0}{h}} + K(t_0+h)(1+hr)^{\frac{t-t_0}{h}-1} \\ &\quad + \dots + K(t-h)(1+hr). \end{aligned}$$

Example 42. Let us consider the following payment structure. At times $t = 0$ and $t = 1$, we receive the amounts of 100 dollars each. From year 2 on, we receive quarterly payments of 30 Dollars. The interest rate for the entire time is 5%. How much is the money worth at $t = 3$? *Solution:*

$$\begin{aligned} V(3) &= \int_0^3 \frac{K(\tau)e_{0.05}(3, \tau)}{\mu(\tau)} \Delta\tau \\ &= \sum_{\tau \in [t_0, t)} K(\tau)e_{0.05}(3, \tau) \\ &= \sum_{\tau \in [0, 2)} 100(1+0.05)^{3-\tau} + \sum_{\tau \in [2, 3)} 30 \left(1 + \frac{1}{4}0.05\right)^{\frac{3-\tau}{4}} \\ &= \sum_{\tau=0}^1 100(1+0.05)^{3-\tau} + \sum_{\tau=\frac{2}{4}}^{\frac{3}{4}-1} 30 \left(1 + \frac{1}{4}0.05\right)^{\frac{3-\tau}{4}} \\ &= \sum_{\tau=0}^1 100(1+0.05)^{3-\tau} + \sum_{\tau=8}^{11} 30(1+0.0125)^{12-\tau} \\ &\approx \$349.81. \end{aligned}$$

Theorem 47. Assume (\star) . Let the stream of payments K , the interest rate r and the graininess μ be all constant. Then we have

$$V(t) = \frac{K}{\mu}(1 + \mu r) \frac{e_r(t, t_0) - 1}{r}. \quad (60)$$

Proof. First, note that as r and μ are constant, so is $\ominus r = -\frac{r}{1+\mu r}$. Moreover, we have

$$\begin{aligned} \int_{t_0}^t e_r(t, \tau) \Delta \tau &= \int_{t_0}^t e_{\ominus r}(\tau, t) \Delta \tau \\ &= \frac{1}{\ominus r} \int_{t_0}^t (\ominus r) e_{\ominus r}(\tau, t) \Delta \tau \\ &= -\frac{1 + \mu r}{r} e_{\ominus r}(\tau, t) \Big|_{t_0}^t \\ &= -\frac{1 + \mu r}{r} (1 - e_{\ominus r}(t_0, t)) \\ &= (1 + \mu r) \frac{e_r(t, t_0) - 1}{r}. \end{aligned}$$

Therefore

$$\begin{aligned} V(t) &= \int_{t_0}^t \frac{K(\tau) e_r(t, \tau)}{\mu(\tau)} \Delta \tau \\ &= \frac{K}{\mu} \int_{t_0}^t e_r(t, \tau) \Delta \tau \\ &= \frac{K}{\mu} (1 + \mu r) \frac{e_r(t, t_0) - 1}{r}, \end{aligned}$$

which is equation (60). □

Example 43. Using (60), we get the following.

(i) If $\mathbb{T} = \mathbb{Z}$, then

$$V(t) = K(1 + r) \frac{(1 + r)^{t-t_0} - 1}{r}.$$

This property can be found in [42, 24] for annually paying annuities.

(ii) If $\mathbb{T} = h\mathbb{Z}$, then

$$V(t) = K(1 + hr) \frac{(1 + hr)^{\frac{t-t_0}{h}} - 1}{hr}.$$

This property can be found in [42, 24] for periodically paying annuities, where hr represents the interest rate in the period, i.e., daily, monthly or quarterly.

The accumulated value of an annuity defined in (59) considers a fixed stream of payments, independent of the time scale in the sense that the payable amount at time t is fixed from the beginning. That amount does not depend on the gaps between the points in time when an amount is due. If this gap is also included in our consideration, we can derive a *time scale adjusted value*, denoted by $V_a(t)$. Note that this payable amount would be $\mu(t)K(t)$.

Lemma 10. *The time scales adjusted value of an annuity is*

$$V_a(t) = \int_{t_0}^t K(\tau)e_r(t, \tau)\Delta\tau. \quad (61)$$

Proof. Note that (59) holds now with $\mu(t)K(t)$ instead of $K(t)$. Therefore

$$V_a(t) = \int_{t_0}^t \frac{\mu(\tau)K(\tau)e_r(t, \tau)}{\mu(\tau)}\Delta\tau = \int_{t_0}^t K(\tau)e_r(t, \tau)\Delta\tau,$$

which shows (61). □

Example 44. *For simplicity, let $K(t) = K$ be a fixed, constant stream of payments, r the nominal rate of interest and $t_0 \in \mathbb{T}$.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$\begin{aligned} V_a(t) &= \int_{t_0}^t K e^{r(t-\tau)} d\tau \\ &= \frac{K}{r} e^{r(t-t_0)} - 1. \end{aligned}$$

(ii) *If $\mathbb{T} = \mathbb{Z}$, then*

$$\begin{aligned} V_a(t) &= \sum_{\tau=t_0}^{t-1} K(1+r)^{t-\tau} \\ &= K(1+r)^t \sum_{\tau=t_0}^{t-1} \left(\frac{1}{1+r}\right)^\tau \\ &= K(1+r)^t \left(\sum_{\tau=0}^{t-1} \left(\frac{1}{1+r}\right)^\tau - \sum_{\tau=0}^{t_0-1} \left(\frac{1}{1+r}\right)^\tau \right) \end{aligned}$$

$$\begin{aligned}
&= K(1+r)^t \left(\frac{1 - \left(\frac{1}{1+r}\right)^t}{1 - \frac{1}{1+r}} - \frac{1 - \left(\frac{1}{1+r}\right)^{t_0}}{1 - \frac{1}{1+r}} \right) \\
&= \frac{K(1+r)^{t+1}}{r} \left((1+r)^{-t_0} - (1+r)^{-t} \right) \\
&= \frac{K(1+r)}{r} \left((1+r)^{t-t_0} - 1 \right).
\end{aligned}$$

(iii) If $\mathbb{T} = h\mathbb{Z}$, then

$$\begin{aligned}
V_a(t) &= \sum_{\tau=\frac{t_0}{h}}^{\frac{t}{h}-1} hK(1+hr)^{\frac{t-h\tau}{h}} \\
&= hK(1+hr)^{\frac{t}{h}} \sum_{\tau=\frac{t_0}{h}}^{\frac{t}{h}-1} \left(\frac{1}{1+hr} \right)^\tau \\
&= hK(1+hr)^t \left(\sum_{\tau=0}^{\frac{t}{h}-1} \left(\frac{1}{1+hr} \right)^\tau - \sum_{\tau=0}^{\frac{t_0}{h}-1} \left(\frac{1}{1+hr} \right)^\tau \right) \\
&= hK(1+hr)^t \left(\frac{1 - \left(\frac{1}{1+hr}\right)^{\frac{t}{h}}}{1 - \frac{1}{1+hr}} - \frac{1 - \left(\frac{1}{1+hr}\right)^{\frac{t_0}{h}}}{1 - \frac{1}{1+hr}} \right) \\
&= \frac{hK(1+hr)^{\frac{t}{h}+1}}{hr} \left((1+hr)^{-\frac{t_0}{h}} - (1+hr)^{-\frac{t}{h}} \right) \\
&= \frac{K(1+hr)}{r} \left((1+hr)^{\frac{t-t_0}{h}} - 1 \right).
\end{aligned}$$

Remark 24. Note that this time scales adjusted definition of the accumulated value of an annuity also allows continuous consideration.

Similarly to the definition of an annuity in (59) and (61), we can define the *net present value* and the *time scales adjusted net present value* of a stream of payments conducted at $t \in \mathbb{T}$.

Definition 43. Assume (\star) . We define the net present value of a stream of payments $K(t)$ by

$$NPV = \int_{t_0}^t \frac{K(\tau)e_{\ominus r}(\tau, t_0)}{\mu(\tau)} \Delta\tau. \quad (62)$$

Definition 44. We define the time scales adjusted net present value of a stream of payments $K(t)$ by

$$NPV_a = \int_{t_0}^t K(\tau) e_{\ominus r}(\tau, t_0) \Delta\tau. \quad (63)$$

8.2. HAZARD RATES

In this subsection, we will be considering a component, which may be part or even the entire system. After putting this component under some sort of stress, let X be a random variable representing the life time or time to failure of the component. Examples of these components could be fuses, light bulbs or more interesting in the light of economics, credits, compare [48].

Definition 45. For the random variable X , we define the reliability R by

$$R(t) = p(X \geq t). \quad (64)$$

Note that we have the following equivalent expressions for the reliability:

$$\begin{aligned} R(t) &= p(X \geq t) \\ &= \int_t^{\infty} f(\tau) \Delta\tau \\ &= 1 - p(X < t) \\ &= 1 - F(t). \end{aligned}$$

Definition 46. For the random variable X with time scales probability density function f and cumulative distribution function F , we define the failure or hazard function a by

$$a(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)}. \quad (65)$$

Remark 25. Note that

$$a(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{1 - p(X < t)} = \frac{f(t)}{p(X \geq t)},$$

and if X is exponentially distributed, then

$$a(t) = \frac{f(t)}{p(X \geq t)} = \frac{-(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0)}{e_{\ominus\lambda}(t, 0)} = -(\ominus\lambda)(t).$$

Example 45. Applying Remark 25 to different time scales yields the following

(i) If $\mathbb{T} = \mathbb{R}$, then

$$a(t) = \frac{\lambda}{1 + 0\lambda} = \lambda,$$

which can be found in [64, p.289].

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$a(t) = \frac{\lambda}{1 + \lambda}.$$

(iii) If $\mathbb{T} = h\mathbb{Z}$, then

$$a(t) = \frac{\lambda}{1 + h\lambda}.$$

(iv) If $\mathbb{T} = q_0^{\mathbb{N}}$ and $q > 1$, then

$$a(t) = \frac{\lambda}{1 + \lambda t(q - 1)}.$$

Note that for constant graininess, we have constant hazard functions, whereas this is not necessarily true for other time scales, for example, the quantum calculus case.

Remark 26. Using the definition of conditional probability, we have

$$\begin{aligned} p(t \leq X < \sigma(t) | X \geq t) &= \frac{p(t \leq X < \sigma(t), X \geq t)}{p(X \geq t)} \\ &= \frac{p(t \leq X < \sigma(t))}{p(X \geq t)} \\ &= \frac{\int_t^{\sigma(t)} f(\tau) \Delta\tau}{1 - F(t)} \\ &= \frac{\mu(t)f(t)}{R(t)} \\ &= \mu(t)a(t), \end{aligned}$$

where $\mu(t)a(t)$ represents the proportion of items failing in $[t, \sigma(t))$ among items functioning before time t . Note that this property also holds for $\mu(t) = 0$, as left and right hand side compute to 0.

Remark 27. Let X represent a random variable describing the time of failure (default event). Then the probability of surviving up to time T given survival until time t is

$$p(X \geq T | X \geq t) = \frac{p(X \geq T, X \geq t)}{p(X \geq t)} = \frac{p(X \geq T)}{p(X \geq t)},$$

and if X is exponentially distributed, then

$$p(X \geq T | X \geq t) = \frac{p(X \geq T)}{p(X \geq t)} = \frac{e_{\ominus\lambda}(T, 0)}{e_{\ominus\lambda}(t, 0)} = e_{\ominus\lambda}(T, t).$$

Theorem 48. Let X represent a random variable describing the survival time and a the hazard rate. Then we have

$$p(X \geq t) = e_{-a}(t, 0).$$

Proof. First, consider the case, where $\mu(t) > 0$. From Remark 26, we know

$$\begin{aligned} a(t) &= \frac{p(t \leq X < \sigma(t) | X \geq t)}{\mu(t)} \\ &= \frac{p(t \leq X < \sigma(t))}{p(X \geq t)\mu(t)} \\ &= \frac{\int_0^{\sigma(t)} f(\tau)\Delta\tau - \int_0^t f(\tau)\Delta\tau}{p(X \geq t)\mu(t)} \\ &= \frac{\left(1 - \int_{\sigma(t)}^{\infty} f(\tau)\Delta\tau\right) - \left(1 - \int_t^{\infty} f(\tau)\Delta\tau\right)}{p(X \geq t)\mu(t)} \\ &= \frac{p(X \geq t) - p(X \geq \sigma(t))}{p(X \geq t)\mu(t)} \\ &= \frac{R(t) - R(\sigma(t))}{\mu(t)R(t)} \\ &= -\frac{R^\Delta(t)}{R(t)}. \end{aligned}$$

Therefore, we have

$$R^\Delta(t) = -a(t)R(t)$$

and $R(0) = 1$, which yields the unique solution

$$p(X \geq t) = R(t) = e_{-a}(t, 0).$$

If $\mu(t) = 0$, then from the definition

$$a(t) = \frac{f(t)}{R(t)} = \frac{F^\Delta(t)}{R(t)} = \frac{\lim_{s \rightarrow t} \frac{F(s) - F(t)}{s - t}}{R(t)}$$

and therefore

$$R(t)a(t) = \lim_{s \rightarrow t} \frac{F(s) - F(t)}{s - t} = -\lim_{s \rightarrow t} \frac{R(s) - R(t)}{s - t} = -R^\Delta(t).$$

As in the first case, the claim follows. \square

Remark 28. Note that this result is independent from the actual distribution of X .

Example 46. Applying Theorem 48 to the continuous and the discrete cases yields the following.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$p(X \geq t) = e_{-a}(t, 0) = e^{-\int_0^t a(\tau) d\tau}.$$

This can be found in the literature in [64, p.289].

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$p(X \geq t) = e_{-a}(t, 0) = \prod_{\tau=0}^{t-1} (1 - a(\tau)).$$

Note, that here

$$a(t) = \frac{f(t)}{p(X \geq t)} = \frac{f(t)}{\sum_{k=t}^{\infty} f(k)} \leq 1,$$

which is necessary in the discrete representation of the exponential function.

Theorem 49. Let $\mathbb{E}_{\mathbb{T}}(X)$ be finite. Then we have

$$\mathbb{E}_{\mathbb{T}}(X) = \int_0^{\infty} R(\sigma(t)) \Delta t.$$

Proof. Integration by parts yields

$$\begin{aligned} \int_0^\infty R(\sigma(t))\Delta t &= \int_0^\infty \int_{\sigma(t)}^\infty f(\tau)\Delta\tau\Delta t \\ &= t \int_t^\infty f(\tau)\Delta\tau \Big|_0^\infty + \int_0^\infty tf(t)\Delta t \\ &= \mathbb{E}_{\mathbb{T}}(X), \end{aligned}$$

which completes the proof. \square

Remark 29. *When considering economic applications, we will have changing hazard rates over time. So it would be useful to define the exponential distribution with varying λ . Let $\lambda(t) > 0$ and $t \in \mathbb{T}_0$. Then we define*

$$f(t) = \begin{cases} -(\ominus\lambda)(t)e_{\ominus\lambda}(t, 0), & \text{if } t \geq 0 \\ 0, & \text{if } t < 0. \end{cases}$$

Note that this still represents a time scales probability density function, $p(X \geq t) = e_{\ominus\lambda}(t, 0)$ and $a(t) = -(\ominus\lambda)(t)$. Calculation of the expected value will now not yield a nice closed formula.

8.3. PRICING OF CREDIT DEFAULT SWAPS

The dictionary defines default as “the absence of something needed” or “failure to do something required by duty or law” [1]. In economics and finance, default describes the state where one party fails to fulfill its obligations according to a debt contract, which was set up with at least one other party. The event of default can occur for example to countries, companies and single individuals, compare [42]. In the last 25 years, several countries in the world had to default due to different reasons. The most prominent in recent time was Argentina in 2002, when defaulting on part of its external debt, see [37]. One measurement of judging the solvency is the bond credit rating conducted by credit rating agencies like Moody’s, Standard & Poor’s and Fitch. The ratings range from AAA (S&P [71], Fitch [62]) as best rating to D describing default. In the beginning of August 2011, S&P cut the rating of the United States of America from AAA to AA+ as a reaction to the debt crises [71]. Greece in the European Union is rated with CC by S&P, as of August 2011, causing tremendous problems in refinancing on the bond

markets [71]. Usually a low credit rating increases the interest rates a country has to offer to sell bonds on the financial market.

Possible instruments in order to secure the portfolio from the event of default of particular assets are credit default swaps (CDS). A CDS is the most fundamental credit derivative. It is a contract between two parties, enabling the bond holding party to isolate the default risk of the obligator (see [67]). Based on [9], we make the following assumptions to simplify the model:

1. Default and interest rate are two independent processes.
2. Default can occur at an arbitrary $t \in \mathbb{T}$ between start and maturity of the contract. Note that these dates are not necessarily discrete, as it was assumed in [9].
3. The default payment is paid immediately upon default.

During the term of the contract, the protection buyer pays a certain amount, depending on the time scale, at time $t \in \mathbb{T}$, whereas the protection seller is obligated to make a contingent payment in the default case. The term “default” is here defined in the contract between the two parties. That might be bankruptcy or failure of the bond/asset seller to make an interest or the principal payment. Figure 8.1 explains the payment structure of a CDS for a particular time scale containing the discrete points $\{0, 1, 3, 4, 7, 8\}$. The first part shows the premium payments in case no default occurs. All payments will be conducted at the beginning of a time period. This will coincide with the premium leg of the credit default swap. The second part, describing the default leg, represents the payment structure in the case default occurs. Note that here the premium payments are paid as in the first case until the event of default. Once the event of default occurs, the insuring party has to pay $1 - \delta$ to the protection buyer, where δ represents the recovery rate of the underlying asset. Often these recovery rates are low (see [66]) so that the insurance has to compensate for almost the entire amount of the asset. However, in some cases the defaulting party might be able to pay a more substantial portion after restructuring its debts.

First we will price the premium leg. Premium payments of the amount of $A\mu(t)$ are made at any point $t \in \mathbb{T}$. The random variable X represents the time of default with hazard rate $a(t)$, $p(X \geq t)$ is the probability that the default event did not occur until time t , T is the maturity of the CDS contract and $B(0, t)$ is the discount factor for

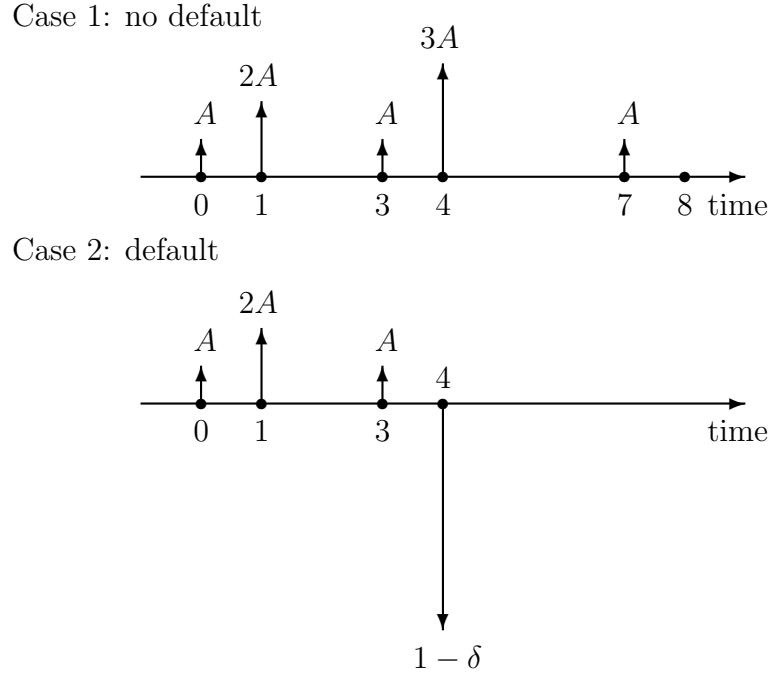


Figure 8.1. Premium payments of a credit default swap

payments done at time t . Moreover, we know, that the risk-free discount factor is

$$B(0, t) = e_{\ominus r}(t, 0),$$

where r is the risk-free interest rate. Therefore, the present value of the premium leg is

$$\begin{aligned}
 PV(PL) &= \int_0^T AB(0, t)p(X \geq t)\Delta t \\
 &= A \int_0^T e_{\ominus r}(t, 0)p(X \geq t)\Delta t \\
 &= A \int_0^T e_{\ominus r}(t, 0)e_{-a}(t, 0)\Delta t \\
 &= A \int_0^T \frac{e_{-a}(t, 0)}{e_r(t, 0)} \Delta t.
 \end{aligned}$$

If the default event is exponentially distributed and therefore $a(t) = -(\ominus\lambda)(t)$, then we have

$$PV(PL) = \int_0^T \frac{A}{e_{\lambda \oplus r}(t, 0)} \Delta t.$$

Second we will price the default leg. The probability that the event of default occurs between t and $\sigma(t)$, using Remark 26, is

$$p(t \leq X < \sigma(t)) = a(t)\mu(t)p(X \geq t) = a(t)\mu(t)e_{-a}(t, 0).$$

In case the default event occurs, the insuring party has to make a payment of $1 - \delta$. Therefore, the present value of the default leg can be computed by

$$\begin{aligned} PV(DL) &= \int_0^T (1 - \delta)B(0, t)a(t)e_{-a}(t, 0)\Delta t \\ &= (1 - \delta) \int_0^T e_{\ominus r}(t, 0)a(t)e_{-a}(t, 0)\Delta t \\ &= (1 - \delta) \int_0^T \frac{a(t)e_{-a}(t, 0)}{e_r(t, 0)}\Delta t. \end{aligned}$$

If the default event is exponentially distributed and therefore $a(t) = -(\ominus\lambda)(t)$, then we have

$$PV(DL) = (1 - \delta) \int_0^T \frac{-(\ominus\lambda)(t)}{e_{\lambda \oplus r}(t, 0)}\Delta t.$$

We can also compute the present value of the swap as

$$\begin{aligned} PV(\text{swap}) &= PV(DL) - PV(PL) \\ &= (1 - \delta) \int_0^T \frac{a(t)e_{-a}(t, 0)}{e_r(t, 0)}\Delta t - A \int_0^T \frac{e_{-a}(t, 0)}{e_r(t, 0)}\Delta t \\ &= \int_0^T ((1 - \delta)a(t) - A) \frac{e_{-a}(t, 0)}{e_r(t, 0)}\Delta t, \end{aligned}$$

and if X is exponentially distributed, then

$$PV(\text{swap}) = \int_0^T \frac{-(\ominus\lambda)(t)(1 - \delta) - A}{e_{\lambda \oplus r}(t, 0)}\Delta t.$$

Finally, if we assume the equilibrium state, where

$$PV(DL) = PV(PL),$$

then we get a closed formula to compute the premium payments A

$$A = \frac{(1 - \delta) \int_0^T \frac{a(t)e^{-a(t,0)}}{e_r(t,0)} \Delta t}{\int_0^T \frac{e^{-a(t,0)}}{e_r(t,0)} \Delta t},$$

and if X is exponentially distributed, then

$$A = \frac{(1 - \delta) \int_0^T \frac{-(\ominus\lambda)(t)}{e_{\lambda \oplus r}(t,0)} \Delta t}{\int_0^T \frac{1}{e_{\lambda \oplus r}(t,0)} \Delta t}.$$

Example 47. *We apply the previous derivation of the price of a credit default swap to the continuous and the discrete time scales. The present value for the premium leg, default leg, swap and the amount of the premium payments will be given.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$PV(PL) = A \int_0^T e^{-rt - \int_0^t a(u) du} dt,$$

$$PV(DL) = (1 - \delta) \int_0^T a(t) e^{-rt - \int_0^t a(u) du} dt,$$

$$PV(\text{swap}) = \int_0^T [(1 - \delta)a(t) - A] e^{-rt - \int_0^t a(u) du} dt,$$

$$A = \frac{(1 - \delta) \int_0^T a(t) e^{-rt - \int_0^t a(u) du} dt}{\int_0^T e^{-rt - \int_0^t a(u) du} dt}.$$

If X is exponentially distributed, then

$$PV(PL) = A \int_0^T e^{-rt - \int_0^t \lambda(u) du} dt,$$

$$PV(DL) = (1 - \delta) \int_0^T \lambda(t) e^{-rt - \int_0^t \lambda(u) du} dt,$$

$$PV(\text{swap}) = \int_0^T [(1 - \delta)\lambda(t) - A] e^{-rt - \int_0^t \lambda(u) du} dt,$$

$$A = \frac{(1 - \delta) \int_0^T \lambda(t) e^{-rt - \int_0^t \lambda(u) du} dt}{\int_0^T e^{-rt - \int_0^t \lambda(u) du} dt}.$$

If λ is constant, then

$$PV(PL) = A \int_0^T e^{-t(r+\lambda)} dt = \frac{A}{r+\lambda} (1 - e^{-T(r+\lambda)}),$$

$$PV(DL) = (1 - \delta) \int_0^T \lambda e^{-t(r+\lambda)} dt = \frac{\lambda(1 - \delta)}{r + \lambda} (1 - e^{-T(r+\lambda)}),$$

$$PV(\text{swap}) = (1 - e^{-T(r+\lambda)}) \frac{\lambda(1 - \delta) - A}{r + \lambda},$$

$$A = (1 - \delta)\lambda.$$

(ii) If $\mathbb{T} = \mathbb{Z}$, then

$$PV(PL) = A \sum_{t=0}^{T-1} \frac{\prod_{u=0}^{t-1} (1 - a(u))}{(1+r)^t},$$

$$PV(DL) = (1 - \delta) \sum_{t=0}^{T-1} \frac{a(t) \prod_{u=0}^{t-1} (1 - a(u))}{(1+r)^t},$$

$$PV(\text{swap}) = \sum_{t=0}^{T-1} ((1 - \delta)a(t) - A) \frac{\prod_{u=0}^{t-1} (1 - a(u))}{(1+r)^t},$$

$$A = \frac{(1 - \delta) \sum_{t=0}^{T-1} \frac{a(t) \prod_{u=0}^{t-1} (1 - a(u))}{(1+r)^t}}{\sum_{t=0}^{T-1} \frac{\prod_{u=0}^{t-1} (1 - a(u))}{(1+r)^t}}.$$

These results can be found similarly in [9]. If X is exponentially distributed, then

$$PV(PL) = \sum_{t=0}^{T-1} \frac{A}{(1+r)^t \prod_{u=0}^{t-1} (1 + \lambda(u))},$$

$$PV(DL) = (1 - \delta) \sum_{t=0}^{T-1} \frac{\frac{\lambda(t)}{1+\lambda(t)}}{(1+r)^t \prod_{u=0}^{t-1} (1 + \lambda(u))},$$

$$PV(\text{swap}) = \sum_{t=0}^{T-1} \frac{(1-\delta) \frac{\lambda(t)}{1+\lambda(t)} - A}{(1+r)^t \prod_{u=0}^{t-1} (1+\lambda(u))},$$

$$A = \frac{(1-\delta) \sum_{t=0}^{T-1} \frac{\frac{\lambda(t)}{1+\lambda(t)}}{(1+r)^t \prod_{u=0}^{t-1} (1+\lambda(u))}}{\sum_{t=0}^{T-1} \frac{1}{(1+r)^t \prod_{u=0}^{t-1} (1+\lambda(u))}}.$$

If λ is constant, then

$$PV(PL) = \sum_{t=0}^{T-1} \frac{A}{(1+r)^t (1+\lambda)^t} = \frac{A(1+r)(1+\lambda)}{r+\lambda+r\lambda} \left(1 - \frac{1}{(1+r)^T (1+\lambda)^T} \right),$$

$$PV(DL) = (1-\delta) \sum_{t=0}^{T-1} \frac{\frac{\lambda}{1+\lambda}}{(1+r)^t (1+\lambda)^t}$$

$$= \frac{\lambda(1-\delta)(1+r)}{r+\lambda+r\lambda} \left(1 - \frac{1}{(1+r)^T (1+\lambda)^T} \right),$$

$$PV(\text{swap}) = \left(\frac{\lambda}{1+\lambda} (1-\delta) - A \right) \frac{(1+r)(1+\lambda)}{r+\lambda+r\lambda} \left(1 - \frac{1}{(1+r)^T (1+\lambda)^T} \right),$$

$$A = \frac{(1-\delta)\lambda}{1+\lambda}.$$

8.4. CREDIT SPREADS

In this section, we will compare the prices of two bonds of the same structure, one being risk-free and the other one being risky in the sense that the probability of default is greater than zero. We already derived today's price of a risk free zero coupon bond of face value 1 and maturity time T by

$$B(0, T) = e_{\ominus r}(T, 0),$$

where r represents the risk-free interest rate. Following [9], a compensation of c has to be paid in order to sell a bond that is risky, i.e., that might default before time T . We define the new interest rate of the risky bond by $r \oplus c$ and therefore get today's price of

that asset of face value 1 and maturity time T by

$$B_{\text{risky}}(0, T) = e_{\ominus(r \oplus c)}(T, 0).$$

Let X be a random variable that describes the time of default, and therefore $p(X \geq T)$ is the probability that the default event does not happen until maturity of the underlying bond. Finally δ will again represent the recovery rate in the default case and $a(t)$ the hazard rate. So the fair price of a risky bond is therefore

$$\begin{aligned} B_{\text{risky}}(0, t) &= p(X \geq T)B(0, T) + (1 - p(X \geq T))\delta B(0, T) \\ &= e_{-a}(T, 0)e_{\ominus r}(T, 0) + (1 - e_{-a}(T, 0))\delta e_{\ominus r}(T, 0) \\ &= e_{\ominus r}(T, 0)(e_{-a}(T, 0)(1 - \delta) + \delta). \end{aligned}$$

If X is exponentially distributed, then

$$B_{\text{risky}}(0, T) = e_{\ominus r}(T, 0)(e_{\ominus \lambda}(T, 0)(1 - \delta) + \delta).$$

Now we derive the credit spread c in terms of the default probability. Note that

$$\begin{aligned} \frac{B_{\text{risky}}(0, T)}{B(0, T)} &= \frac{e_{\ominus(r \oplus c)}(T, 0)}{e_{\ominus r}(T, 0)} \\ &= \frac{e_r(T, 0)}{e_r(T, 0)e_c(T, 0)} \\ &= e_{\ominus c}(T, 0) \\ &= e_{-a}(T, 0)(1 - \delta) + \delta. \end{aligned}$$

If X is exponentially distributed, then

$$\begin{aligned} \frac{B_{\text{risky}}(0, T)}{B(0, T)} &= e_{\ominus c}(T, 0) \\ &= e_{\ominus \lambda}(T, 0)(1 - \delta) + \delta. \end{aligned}$$

Using the definition of the logarithm from (5), we get for the credit spread

$$c = \log_{\mathbb{T}} \frac{1}{e_{-a}(T, 0)(1 - \delta) + \delta},$$

and if X is exponentially distributed, then

$$c = \log_{\mathbb{T}} \frac{1}{e_{\infty\lambda}(T, 0)(1 - \delta) + \delta}.$$

Usually exact hazard rates are unknown, and therefore getting a good estimate for the default probability is achieved by using historical data, see [9]. If our time scale is for example \mathbb{Z} , then one could use the annual transition matrix and raising it to the appropriate power. That matrix is published by rating agencies like Moody's [51]. If a bond has a particular rating right now, that matrix gives all transition probabilities to have any other rating in one year. Most importantly it gives an estimate of defaulting during the one year time period. Of course going from an Aaa rating to default is almost impossible. As those probabilities are not risk neutral (see [9]), this method gives a close estimate for the computation of the credit spread but is not a valid way of pricing risky bonds.

Example 48. *Next, we apply the previously derived results to the continuous and the discrete time scales cases.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$B_{\text{risky}}(0, T) = e^{-(r+c)T},$$

$$B_{\text{risky}}(0, T) = e^{-rT} \left(e^{-\int_0^T a(t)dt} (1 - \delta) + \delta \right)$$

and

$$c = -\frac{\log \left(e^{-\int_0^T a(t)dt} (1 - \delta) + \delta \right)}{T}.$$

If X is exponentially distributed, then

$$B_{\text{risky}}(0, T) = e^{-rT} \left(e^{-\int_0^T \lambda(t)dt} (1 - \delta) + \delta \right)$$

and

$$c = -\frac{\log \left(e^{-\int_0^T \lambda(t)dt} (1 - \delta) + \delta \right)}{T}.$$

If λ is constant, then

$$B_{\text{risky}}(0, T) = e^{-rT} (e^{-\lambda T} (1 - \delta) + \delta)$$

and

$$c = -\frac{\log (e^{-\lambda T} (1 - \delta) + \delta)}{T}.$$

Similar results can be found in [9, p.180].

(ii) If $\mathbb{T} = \mathbb{Z}$, then we have the price of a risky bond given by

$$B_{\text{risky}}(0, T) = (1 + r + c + rc)^{-T} = (1 + r)^{-T} (1 + c)^{-T},$$

where we note that this definition is slightly different from the intuitive assumption of just adding the spread to the risk-free interest rate. This occurs as we deal with $r \oplus c$ in the time scales case. Moreover,

$$B_{\text{risky}}(0, T) = (1 + r)^{-T} \left(\prod_{t=0}^{T-1} (1 - a(t))(1 - \delta) + \delta \right)$$

and

$$c = \left(\prod_{t=0}^{T-1} (1 - a(t))(1 - \delta) + \delta \right)^{-\frac{1}{T}} - 1.$$

If X is exponentially distributed, then

$$B_{\text{risky}}(0, T) = (1 + r)^{-T} \left(\frac{1 - \delta}{\prod_{t=0}^{T-1} (1 + \lambda(t))} + \delta \right)$$

and

$$c = \left(\frac{1 - \delta}{\prod_{t=0}^{T-1} (1 + \lambda(t))} + \delta \right)^{-\frac{1}{T}} - 1.$$

If λ is constant, then

$$B_{\text{risky}}(0, T) = (1 + r)^{-T} ((1 + \lambda)^{-T}(1 - \delta) + \delta)$$

and

$$c = ((1 + \lambda)^{-T}(1 - \delta) + \delta)^{-\frac{1}{T}} - 1.$$

9. OSTROWSKI INEQUALITIES INVOLVING EXPECTED VALUES

This section is motivated by the paper “Some Inequalities for Probability, Expectation, and Variance of Random Variables Defined over a Finite Interval” by N.S. Barnett, S.S. Dragomir and R.P. Agarwal (see [12]). We will apply the definitions of previous sections in this thesis to theorems in [12], and, therefore, generalize the results for arbitrary time scales. Throughout, multiple examples will be provided.

Definition 47. *Throughout this section, let X be a random variable with support on the interval $[a, b]$, $a, b \in \mathbb{T}_0$, i.e., the time scales probability density function $f(t) \geq 0$ on $[a, b]$ and $f(t) = 0$ elsewhere. Moreover, we use the cumulative distribution function*

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

and we have

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

Lemma 11 (Montgomery identity). *We have*

$$\int_a^b p(t, s) f(s) \Delta s = (b - a) F(t) - \int_a^b F(\sigma(s)) \Delta s, \quad (66)$$

where

$$p(t, s) = \begin{cases} s - a, & a \leq s < t \\ s - b, & t \leq s \leq b. \end{cases}$$

Proof. Using the integration by parts formula, we have

$$\begin{aligned} \int_a^b p(t, s) f(s) \Delta s &= \int_a^t (s - a) f(s) \Delta s + \int_t^b (s - b) f(s) \Delta s \\ &= (s - a) F(s) \Big|_a^t - \int_a^t F(\sigma(s)) \Delta s \\ &\quad + (s - b) F(s) \Big|_t^b - \int_t^b F(\sigma(s)) \Delta s \\ &= (t - a) F(t) - (t - b) F(t) - \int_a^b F(\sigma(s)) \Delta s \end{aligned}$$

$$= (b-a)F(t) - \int_a^b F(\sigma(s))\Delta s,$$

which proves (66). \square

Theorem 50. *The following inequalities hold*

$$\begin{aligned} \left| F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} \right| &\leq \frac{1}{b-a} \left((2t - (a+b))F(t) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s \right) \quad (67) \\ &\leq \frac{1}{b-a} ((b-t)p(X \geq t) + (t-a)p(X < t)) \\ &\leq \frac{1}{2} + \frac{|t - \frac{a+b}{2}|}{b-a}, \end{aligned}$$

where all inequalities are sharp.

Proof. First, using the definition of the expected value and integration by parts, we have

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X) &= \int_a^b s f(s) \Delta s \\ &= sF(s)|_a^b - \int_a^b F(\sigma(s)) \Delta s \\ &= bF(b) - aF(a) - \int_a^b F(\sigma(s)) \Delta s \\ &= b - \int_a^b F(\sigma(s)) \Delta s. \end{aligned}$$

Applying (66) leads to

$$\begin{aligned} \int_a^b p(t,s)f(s)\Delta s &= (b-a)F(t) - \int_a^b F(\sigma(s))\Delta s \\ &= (b-a)F(t) + \mathbb{E}_{\mathbb{T}}(X) - b \end{aligned} \quad (68)$$

and therefore

$$F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} = \frac{1}{b-a} \int_a^b p(t,s)f(s)\Delta s.$$

Now we have

$$\begin{aligned} \left| \int_a^b p(t,s)f(s)\Delta s \right| &= \left| \int_a^t (s-a)f(s)\Delta s + \int_t^b (s-b)f(s)\Delta s \right| \\ &\leq \int_a^t |s-a||f(s)|\Delta s + \int_t^b |s-b||f(s)|\Delta s \end{aligned}$$

$$\begin{aligned}
&= \int_a^t (s-a)f(s)\Delta s + \int_t^b (b-s)f(s)\Delta s \\
&= (s-a)F(s)|_a^t - \int_a^t F(\sigma(s))\Delta s \\
&\quad + (b-a)F(s)|_t^b + \int_t^b F(\sigma(s))\Delta s \\
&= (t-a)F(t) - (b-t)F(t) - \int_a^t F(\sigma(s))\Delta s + \int_t^b F(\sigma(s))\Delta s \\
&= (2t - (a+b))F(t) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s,
\end{aligned}$$

showing the first inequality of (67). As $F(t)$ is monotonically increasing for $t \in [a, b]$, we have

$$\int_a^t F(\sigma(s))\Delta s \geq \int_a^t F(\sigma(a))\Delta s \geq \int_a^t F(a)\Delta s = 0$$

and

$$\int_t^b F(\sigma(s))\Delta s \leq \int_t^b F(\sigma(b))\Delta s = \int_t^b F(b)\Delta s = b-t.$$

Then

$$\begin{aligned}
&\frac{1}{b-a} \left((2t - (a+b))F(t) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s \right) \\
&\leq \frac{1}{b-a} ((2t - (a+b))F(t) + b-t) \\
&= \frac{1}{b-a} ((b-t)(1-F(t)) + (t-a)F(t)) \\
&= \frac{1}{b-a} ((b-t)p(X \geq t) + (t-a)p(X < t)),
\end{aligned}$$

finishing the second part of (67). Finally, we have

$$\begin{aligned}
&\frac{1}{b-a} ((b-t)p(X \geq t) + (t-a)p(X < t)) \\
&\leq \frac{1}{b-a} \max\{b-t, t-a\} (p(X \geq t) + p(X > t)) \\
&= \frac{1}{b-a} \left(\frac{b-a}{2} + \left| t - \frac{a+b}{2} \right| \right) = \frac{1}{2} + \frac{|t - \frac{a+b}{2}|}{b-a},
\end{aligned}$$

showing the last inequality of (67). To show the sharpness, let $t = t_1$, $a = t_1$, $b = t_2$ and

$$F(t) = \begin{cases} 0, & \text{if } t = t_1 \\ 1, & \text{if } t \in (t_1, t_2]. \end{cases}$$

Then, on noting that

$$\begin{aligned} \mathbb{E}_{\mathbb{T}}(X) &= \int_{t_1}^{t_2} s f(s) \Delta s \\ &= t F(t) \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} F(\sigma(s)) \Delta s \\ &= t_2 F(t_2) - t_1 F(t_1) - (t_2 - t_1) \\ &= t_1, \end{aligned}$$

we have

$$\left| F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| = \left| F(t_1) - \frac{t_2 - t_1}{t_2 - t_1} \right| = 1.$$

On the other hand

$$\frac{1}{2} + \frac{\left| t - \frac{a+b}{2} \right|}{b - a} = \frac{1}{2} + \frac{\left| t_1 - \frac{t_1+t_2}{2} \right|}{t_2 - t_1} = \frac{1}{2} + \frac{\left| \frac{t_1-t_2}{2} \right|}{t_2 - t_1} = \frac{1}{2} + \frac{1}{2} = 1,$$

which completes the proof of the sharpness of inequality (67). \square

Remark 30. Note that in the derivation of the sharpness, the function f is not rd-continuous, thus causing problems in the proof of Theorem 50 with the integration by parts formula. If t_1 is right-scattered, there is no problem, as

$$f(t_1) = F^\Delta(t_1) = \frac{F(\sigma(t_1)) - F(t_1)}{\mu(t_1)} = \frac{1 - 0}{\mu(t_1)} = \frac{1}{\mu(t_1)} < \infty.$$

For t_1 being right-dense, it would be necessary to use the Riemann–Stieltjes definition for integrals on time scales, see [52]. Here the expected value would be defined as $E_{\mathbb{T}}(X) = \int_a^b t \Delta F$, and a corresponding integration by parts formula holds, compare [52, Theorem 4.4].

Example 49. We apply Theorem 50 to different time scales

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\begin{aligned} & \left| p(X \leq t) - \frac{b - \mathbb{E}_{\mathbb{R}}(X)}{b - a} \right| \\ & \leq \frac{1}{b - a} \left((2t - (a + b))p(X \leq t) + \int_a^b \operatorname{sgn}(s - t)F(s)ds \right) \\ & \leq \frac{1}{b - a} ((b - t)p(X \geq t) + (t - a)p(X \leq t)) \leq \frac{1}{2} + \frac{|t - \frac{a+b}{2}|}{b - a}, \end{aligned}$$

which corresponds exactly to [12, Theorem 6].

(ii) If $\mathbb{T} = \mathbb{Z}$, $a = 0$, $b = n$ and $t \in \mathbb{N}_0$, then

$$\begin{aligned} & \left| p(X < t) - \frac{n - \mathbb{E}_{\mathbb{Z}}(X)}{n} \right| \\ & \leq \frac{1}{n} \left((2t - n)p(X < t) + \sum_{s=1}^n \operatorname{sgn}(s - 1 - t)p(X < s) \right) \\ & \leq \frac{1}{n} ((n - t)p(X \geq t) + tp(X < t)) \leq \frac{1}{2} + \frac{|t - \frac{n}{2}|}{n}, \end{aligned}$$

which is a new version of an Ostrowski type inequality in the discrete case.

Remark 31. Note that the left-hand side of (67) can be replaced by

$$\begin{aligned} \left| F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| &= \left| 1 - p(X \geq t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| \\ &= \left| \frac{b - a}{b - a} - p(X \geq t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| \\ &= \left| -p(X \geq t) + \frac{\mathbb{E}_{\mathbb{T}}(X) - a}{b - a} \right| \\ &= \left| p(X \geq t) - \frac{\mathbb{E}_{\mathbb{T}}(X) - a}{b - a} \right|. \end{aligned}$$

In the continuous case, we get [12, Remark 1].

Corollary 2. The following upper and lower bounds for the cumulative distribution function hold:

$$F(t) \leq \frac{1}{b - t} \int_a^b F(\sigma(s)) \left(\frac{1 + \operatorname{sgn}(s - t)}{2} \right) \Delta s \quad (69)$$

and

$$F(t) \geq \frac{1}{t-a} \int_a^b F(\sigma(s)) \left(\frac{1 - \operatorname{sgn}(s-t)}{2} \right) \Delta s. \quad (70)$$

Proof. Using the properties of the absolute value in the first inequality of (67), we get the following equivalent statements

$$\begin{aligned} F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} &\leq \frac{1}{b-a} \left((2t - (a+b))F(t) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s \right) \\ \Leftrightarrow (b-a)F(t) - (2t - (a+b))F(t) &\leq b - \mathbb{E}_{\mathbb{T}}(X) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s \\ \Leftrightarrow (2b - 2t)F(t) &\leq b - \mathbb{E}_{\mathbb{T}}(X) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s. \end{aligned}$$

Moreover,

$$b - \mathbb{E}_{\mathbb{T}}(X) = \int_a^b F(\sigma(s))\Delta s,$$

and therefore

$$\begin{aligned} F(t) &\leq \frac{1}{2(b-t)} \left(\int_a^b F(\sigma(s))\Delta s + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s \right) \\ &= \frac{1}{b-t} \int_a^b F(\sigma(s)) \left(\frac{1 + \operatorname{sgn}(s-t)}{2} \right) \Delta s, \end{aligned}$$

showing (69). Similarly, we have

$$\begin{aligned} F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} &\geq -\frac{1}{b-a} \left((2t - (a+b))F(t) + \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s \right) \\ \Leftrightarrow (2t - 2a)F(t) &\geq b - \mathbb{E}_{\mathbb{T}}(X) - \int_a^b \operatorname{sgn}(s-t)F(\sigma(s))\Delta s, \end{aligned}$$

and therefore

$$F(t) \geq \frac{1}{t-a} \int_a^b F(\sigma(s)) \left(\frac{1 - \operatorname{sgn}(s-t)}{2} \right) \Delta s, \quad (71)$$

completing the proof of (70). \square

Remark 32. Note, that in the continuous setting, this result coincides with [12, Corollary 2].

Theorem 51. *We have the inequality*

$$\left| F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| \leq \frac{M}{b - a} (h_2(t, a) + h_2(t, b)), \quad (72)$$

where

$$M = \sup_{t \in [a, b]} f(t).$$

This inequality is sharp, in the sense, that the right-hand side cannot be replaced by a smaller expression.

Proof. From (68), we know

$$\int_a^b p(t, s) f(s) \Delta s = (b - a) F(t) + \mathbb{E}_{\mathbb{T}}(X) - b$$

and if

$$p(t, s) = \begin{cases} s - a, & a \leq s < t \\ s - b, & t \leq s \leq b, \end{cases}$$

then

$$\begin{aligned} \left| \int_a^b p(t, s) f(s) \Delta s \right| &\leq M \int_a^b |p(t, s)| \Delta s \\ &= M \left(\int_a^t |s - a| \Delta s + \int_t^b |s - b| \Delta s \right) \\ &= M \left(\int_a^t (s - a) \Delta s + \int_t^b (b - s) \Delta s \right) \\ &= M \left(\int_a^t (s - a) \Delta s + \int_b^t (s - b) \Delta s \right) \\ &= M (h_2(t, a) + h_2(t, b)), \end{aligned}$$

implying (72). To show the sharpness of (72), let $F(t) = \frac{t-t_1}{t_2-t_1}$, $a = t_1$, $b = t_2$, $f(t) = \frac{1}{t_2-t_1}$ and $t = t_1$. Then

$$\begin{aligned} \left| F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| &= \left| \frac{t_1 - t_1}{t_2 - t_1} - \frac{t_2 - \frac{1}{t_2-t_1} \int_{t_1}^{t_2} s \Delta s}{t_2 - t_1} \right| \\ &= \left| -\frac{t_2 - \frac{1}{t_2-t_1} \int_{t_1}^{t_2} s \Delta s}{t_2 - t_1} \right| \end{aligned}$$

$$= \frac{1}{(t_2 - t_1)^2} \left(t_2(t_2 - t_1) - \int_{t_1}^{t_2} s \Delta s \right)$$

and

$$\begin{aligned} \frac{M}{b-a} (h_2(t, a) + h_2(t, b)) &= \frac{\frac{1}{t_2 - t_1}}{t_2 - t_1} (h_2(t_1, t_1) + h_2(t_1, t_2)) \\ &= \frac{1}{(t_2 - t_1)^2} \int_{t_2}^{t_1} (s - t_2) \Delta s \\ &= \frac{1}{(t_2 - t_1)^2} \left(t_2(t_2 - t_1) - \int_{t_1}^{t_2} s \Delta s \right), \end{aligned}$$

completing the proof of the sharpness. \square

Remark 33. Also note that we can, as previously done, replace $\left| F(t) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} \right|$ with $\left| p(X \geq t) - \frac{\mathbb{E}_{\mathbb{T}}(X) - a}{b-a} \right|$ in (72).

Remark 34. As

$$\mathbb{E}_{\mathbb{T}}(X) = b - \int_a^b F(\sigma(s)) \Delta s,$$

we have

$$\left| F(t) - \frac{1}{b-a} \int_a^b F(\sigma(s)) \Delta s \right| \leq \frac{M}{b-a} (h_2(t, a) + h_2(t, b)),$$

which corresponds exactly to the Ostrowski inequality on time scales in [19].

Example 50. We apply (72) to different time scales.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\left| F(t) - \frac{b - \mathbb{E}_{\mathbb{R}}(X)}{b-a} \right| \leq M(b-a) \left(\frac{1}{4} + \frac{(t - \frac{a+b}{2})^2}{(b-a)^2} \right),$$

where

$$M = \sup_{t \in [a, b]} f(t),$$

which corresponds exactly to [12, Theorem 7].

(ii) If $\mathbb{T} = \mathbb{Z}$, $a = 0$, $b = n$ and $t \in \mathbb{N}_0$, then

$$\left| p(X < t) - \frac{n - \mathbb{E}_{\mathbb{Z}}(X)}{n} \right| \leq \frac{M}{n} \left[\left(t - \frac{n+1}{2} \right)^2 + \frac{n^2 - 1}{4} \right],$$

where

$$M = \max_{0 \leq t \leq n-1} f(t).$$

Corollary 3. Let $a, b \in \mathbb{T}_0$. Then we have the inequality

$$b - Mh_2(a, b) \leq \mathbb{E}_{\mathbb{T}}(X) \leq a + Mh_2(b, a),$$

where

$$M = \sup_{t \in [a, b]} f(t).$$

Proof. First note that

$$\mathbb{E}_{\mathbb{T}}(X) = \int_a^b sf(s)\Delta s \leq \int_a^b bf(s)\Delta s = b$$

and similarly

$$\mathbb{E}_{\mathbb{T}}(X) = \int_a^b sf(s)\Delta s \geq \int_a^b af(s)\Delta s = a.$$

If we pick $t = a$ in (72), then

$$\left| F(a) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| = \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \leq \frac{M}{b - a} h_2(a, b),$$

resulting in

$$b - Mh_2(a, b) \leq \mathbb{E}_{\mathbb{T}}(X),$$

and if we pick $t = b$, then

$$\left| F(b) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| = \frac{\mathbb{E}_{\mathbb{T}}(X) - a}{b - a} \leq \frac{M}{b - a} h_2(b, a),$$

resulting in

$$\mathbb{E}_{\mathbb{T}}(X) \leq a + Mh_2(b, a).$$

The proof is complete. □

Example 51. *We apply Corollary 3 to different time scales.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$b - \frac{M}{2}(b - a)^2 \leq E_{\mathbb{R}}(X) \leq a + \frac{M}{2}(b - a)^2,$$

where

$$M = \sup_{t \in [a, b]} f(t),$$

which corresponds exactly to [12, Corollary 3].

(ii) *If $\mathbb{T} = \mathbb{Z}$, $a = 0$ and $b = n$, then*

$$n - \frac{M}{2}n(n + 1) \leq E_{\mathbb{Z}}(X) \leq \frac{M}{2}n(n - 1),$$

where

$$M = \max_{0 \leq t \leq n-1} f(t).$$

Theorem 52. *Let f be the time scales probability density function of a random variable X and F the corresponding cumulant distribution function. Then, we have the following inequality*

$$\begin{aligned} & \left| F(x) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b - a} \right| \tag{73} \\ & \leq \frac{q}{1 + q} \|f\|_p (b - a)^{\frac{1}{q}} \left[\left(\frac{x - \sigma(a)}{b - a} \right)^{\frac{1+q}{q}} \right. \\ & \quad \left. + \left(\frac{\sigma(b) - x}{b - a} \right)^{\frac{1+q}{q}} - \left(\frac{-\mu(x)}{b - a} \right)^{\frac{1+q}{q}} - \left(\frac{\mu(x)}{b - a} \right)^{\frac{1+q}{q}} \right] \\ & \leq \frac{q}{1 + q} \|f\|_p (b - a)^{\frac{1}{q}} \left(\frac{\sigma(b) - \sigma(a)}{b - a} \right)^{\frac{1+q}{q}}, \end{aligned}$$

where

$$\|f\|_p = \left(\int_a^b f^p(t) \Delta t \right)^{\frac{1}{p}}$$

and $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. First note that, if we apply [22, Theorem 1.90] for $\gamma \geq 1$, $g_1(x) = x^\gamma$, $g_2(t) = t - s$, then we get $g_2^\Delta(t) = 1$ and $(g_1 \circ g_2)(t) = (t - s)^\gamma$, and therefore

$$\begin{aligned} (g_1 \circ g_2)^\Delta(t) &= \int_0^1 g_1'(t - s + h\mu(t)) dh \\ &= \int_0^1 \gamma(t - s + h\mu(t))^{\gamma-1} dh \\ &\geq \int_0^1 \gamma(t - s)^{\gamma-1} dh \\ &= \gamma(t - s)^{\gamma-1}. \end{aligned}$$

So we have

$$[(t - s)^\gamma]^\Delta \geq \gamma(t - s)^{\gamma-1}.$$

Shifting t to $\sigma(t)$ and integrating t from a to b yields

$$\begin{aligned} \int_a^b (\sigma(t) - s)^{\gamma-1} \Delta t &\leq \frac{1}{\gamma} \int_a^b [(\sigma(t) - s)^\gamma]^\Delta \Delta t \\ &= \frac{1}{\gamma} [(\sigma(b) - s)^\gamma - (\sigma(a) - s)^\gamma]. \end{aligned}$$

Similarly for $g_1(x) = x^\gamma$, $g_2(t) = s - t$, we get $g_2^\Delta(t) = -1$ and $(g_1 \circ g_2)(t) = (s - t)^\gamma$, and therefore

$$\begin{aligned} (g_1 \circ g_2)^\Delta(t) &= - \int_0^1 g_1'(s - t - h\mu(t)) dh \\ &= - \int_0^1 \gamma(s - t - h\mu(t))^{\gamma-1} dh \\ &\geq - \int_0^1 \gamma(s - t)^{\gamma-1} dh \\ &= -\gamma(s - t)^{\gamma-1}. \end{aligned}$$

So we have

$$[(s-t)^\gamma]^\Delta \geq -\gamma(s-t)^{\gamma-1}.$$

Shifting t to $\sigma(t)$ and integrating t from a to b yields

$$\begin{aligned} -\int_a^b (s-\sigma(t))^{\gamma-1} \Delta t &\leq \frac{1}{\gamma} \int_a^b [(s-\sigma(t))^\gamma]^\Delta \Delta t \\ &= \frac{1}{\gamma} [(s-\sigma(b))^\gamma - (s-\sigma(a))^\gamma]. \end{aligned}$$

Now, note that Hölder's inequality (10) implies

$$\begin{aligned} |F(x) - F(y)| &= \left| \int_a^x f(t) \Delta t - \int_a^y f(t) \Delta t \right| \\ &= \left| \int_y^x f(t) \Delta t \right| \\ &\leq \left| \int_y^x |f(t)|^p \Delta t \right|^{\frac{1}{p}} \left| \int_y^x 1^q \Delta t \right|^{\frac{1}{q}} \\ &= \left| \int_y^x f^p(t) \Delta t \right|^{\frac{1}{p}} |x-y|^{\frac{1}{q}} \\ &\leq \left| \int_a^b f^p(t) \Delta t \right|^{\frac{1}{p}} |x-y|^{\frac{1}{q}} \\ &= \|f\|_p |x-y|^{\frac{1}{q}}. \end{aligned}$$

Shifting y to $\sigma(y)$, integrating y from a to b and dividing by $b-a$ yields

$$\begin{aligned} \left| F(x) - \frac{1}{b-a} \int_a^b F(\sigma(y)) \Delta y \right| &\leq \frac{1}{b-a} \int_a^b |F(x) - F(\sigma(y))| \Delta y \\ &\leq \frac{1}{b-a} \|f\|_p \int_a^b |x - \sigma(y)|^{\frac{1}{q}} \Delta y \\ &= \frac{1}{b-a} \|f\|_p \left(\int_a^x (x - \sigma(y))^{\frac{1}{q}} \Delta y + \int_x^b (\sigma(y) - x)^{\frac{1}{q}} \Delta y \right) \\ &= \frac{1}{b-a} \|f\|_p \left(-\int_x^a (x - \sigma(y))^{\frac{1}{q}} \Delta y + \int_x^b (\sigma(y) - x)^{\frac{1}{q}} \Delta y \right) \\ &\leq \frac{1}{b-a} \|f\|_p \frac{q}{1+q} \left((x - \sigma(a))^{\frac{q}{1+q}} - (x - \sigma(x))^{\frac{q}{1+q}} \right. \\ &\quad \left. + (\sigma(b) - a)^{\frac{q}{1+q}} - (\sigma(x) - x)^{\frac{q}{1+q}} \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}} \left[\left(\frac{x-\sigma(a)}{b-a} \right)^{\frac{1+q}{q}} \right. \\
&\quad \left. + \left(\frac{\sigma(b)-x}{b-a} \right)^{\frac{1+q}{q}} - \left(\frac{-\mu(x)}{b-a} \right)^{\frac{1+q}{q}} - \left(\frac{\mu(x)}{b-a} \right)^{\frac{1+q}{q}} \right].
\end{aligned}$$

As

$$\left| F(x) - \frac{1}{b-a} \int_a^b F(\sigma(y)) \Delta y \right| = \left| F(x) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} \right|,$$

the first part of (73) holds. As it can be shown easily, if $x, y \geq 0$ and $\gamma \geq 1$, then $x^\gamma + y^\gamma \leq (x+y)^\gamma$, we have

$$\begin{aligned}
&\frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}} \left[\left(\frac{x-\sigma(a)}{b-a} \right)^{\frac{1+q}{q}} \right. \\
&\quad \left. + \left(\frac{\sigma(b)-x}{b-a} \right)^{\frac{1+q}{q}} - \left(\frac{-\mu(x)}{b-a} \right)^{\frac{1+q}{q}} - \left(\frac{\mu(x)}{b-a} \right)^{\frac{1+q}{q}} \right] \\
&\leq \frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}} \left[\left(\frac{x-\sigma(a) + \sigma(b)-x}{b-a} \right)^{\frac{1+q}{q}} - \left(\frac{-\mu(x) + \mu(x)}{b-a} \right)^{\frac{1+q}{q}} \right] \\
&\leq \frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}} \left(\frac{\sigma(b)-\sigma(a)}{b-a} \right)^{\frac{1+q}{q}},
\end{aligned}$$

finishing the proof of the second inequality of (73). \square

Remark 35. Note that for all time scales different from $\mathbb{T} = \mathbb{R}$, we have the more accurate inequality

$$\left| F(x) - \frac{b - \mathbb{E}_{\mathbb{T}}(X)}{b-a} \right| \leq \frac{1}{b-a} \|f\|_p \left(\int_a^x (x-\sigma(y))^{\frac{1}{q}} \Delta y + \int_x^b (\sigma(y)-x)^{\frac{1}{q}} \Delta y \right).$$

Example 52. We apply (73) to different time scales.

(i) If $\mathbb{T} = \mathbb{R}$, then

$$\begin{aligned}
&\left| p(X \leq x) - \frac{b - \mathbb{E}_{\mathbb{R}}(X)}{b-a} \right| \\
&\leq \frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}} \left[\left(\frac{x-a}{b-a} \right)^{\frac{1+q}{q}} + \left(\frac{b-x}{b-a} \right)^{\frac{1+q}{q}} \right] \\
&\leq \frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}},
\end{aligned}$$

where

$$\|f\|_p = \left(\int_a^b f^p(t) dt \right)^{\frac{1}{p}},$$

which corresponds exactly to [12, Theorem 8].

(ii) If $\mathbb{T} = \mathbb{Z}$, $a = 0$, $b = n$ and $t \in \mathbb{N}_0$, then

$$\begin{aligned} & \left| p(X < x) - \frac{n - \mathbb{E}_{\mathbb{Z}}(X)}{n} \right| \\ & \leq \frac{q}{1+q} \|f\|_p n^{\frac{1}{q}} \left[\left(\frac{x-1}{n} \right)^{\frac{1+q}{q}} + \left(\frac{b+1-x}{n} \right)^{\frac{1+q}{q}} - \left(\frac{-1}{n} \right)^{\frac{1+q}{q}} - \left(\frac{1}{n} \right)^{\frac{1+q}{q}} \right] \\ & \leq \frac{q}{1+q} \|f\|_p n^{\frac{1}{q}}, \end{aligned}$$

where

$$\|f\|_p = \left(\sum_{t=0}^{n-1} f^p(t) \right)^{\frac{1}{p}}.$$

Corollary 4. Let $a, b \in \mathbb{T}_0$. Then we have the inequality

$$b - \frac{q}{q+1} \|f\|_p (\sigma(b) - \sigma(a))^{\frac{1+q}{q}} \leq \mathbb{E}_{\mathbb{T}}(X) \leq a + \frac{q}{q+1} \|f\|_p (\sigma(b) - \sigma(a))^{\frac{1+q}{q}}.$$

Proof. First note that, if we pick $t = a$ in (73), then

$$\begin{aligned} \left| F(a) - \frac{b - E_{\mathbb{T}}(X)}{b - a} \right| &= \frac{b - E_{\mathbb{T}}(X)}{b - a} \\ &\leq \frac{q}{1+q} \|f\|_p (b - a)^{\frac{1}{q}} \left(\frac{\sigma(b) - \sigma(a)}{b - a} \right)^{\frac{1+q}{q}}, \end{aligned}$$

resulting in

$$b - \frac{q}{q+1} \|f\|_p (\sigma(b) - \sigma(a))^{\frac{1+q}{q}} \leq \mathbb{E}_{\mathbb{T}}(X),$$

and if we pick $t = b$, then

$$\left| F(b) - \frac{b - E_{\mathbb{T}}(X)}{b - a} \right| = \frac{E_{\mathbb{T}}(X) - a}{b - a},$$

$$\leq \frac{q}{1+q} \|f\|_p (b-a)^{\frac{1}{q}} \left(\frac{\sigma(b) - \sigma(a)}{b-a} \right)^{\frac{1+q}{q}},$$

resulting in

$$\mathbb{E}_{\mathbb{T}}(X) \leq a + \frac{q}{q+1} \|f\|_p (\sigma(b) - \sigma(a))^{\frac{1+q}{q}},$$

which completes the proof. \square

Example 53. *We apply Corollary 4 to different time scales.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$b - \frac{q}{q+1} \|f\|_p (b-a)^{\frac{1+q}{q}} \leq \mathbb{E}_{\mathbb{R}}(X) \leq a + \frac{q}{q+1} \|f\|_p (b-a)^{\frac{1+q}{q}},$$

where

$$\|f\|_p = \left(\int_a^b f^p(t) dt \right)^{\frac{1}{p}},$$

which corresponds exactly to [12, Corollary 8].

(ii) *If $\mathbb{T} = \mathbb{Z}$, $a = 0$ and $b = n$, then*

$$n - \frac{q}{q+1} \|f\|_p n^{\frac{1+q}{q}} \leq \mathbb{E}_{\mathbb{Z}}(X) \leq \frac{q}{q+1} \|f\|_p n^{\frac{1+q}{q}},$$

where

$$\|f\|_p = \left(\sum_{t=0}^{n-1} f^p(t) \right)^{\frac{1}{p}}.$$

Theorem 53. *Let $a, b \in \mathbb{T}$ and X be defined on $[a, b)$. Then the inequalities*

$$\text{Var}_{\mathbb{T}}(X) \leq (b - \mathbb{E}_{\mathbb{T}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a) \leq \frac{(b-a)^2}{4} \quad (74)$$

and

$$(b - \mathbb{E}_{\mathbb{T}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a) - \text{Var}_{\mathbb{T}}(X) \leq \|f\| \int_a^b ((b-t)(t-a) + h_1^2(t,0) - 2h_2(t,0)) \Delta t \quad (75)$$

hold, where

$$\|f\| = \sup_{t \in [a, b]} |f(t)|.$$

Proof. First note that

$$\begin{aligned} 0 &\leq \int_a^b (b-t)(t-a)f(t)\Delta t + \int_a^b \int_0^t \mu(\tau)\Delta\tau f(t)\Delta t \\ &= \int_a^b [(b-t)(t-a) + h_1^2(t, 0) - 2h_2(t, 0)] f(t)\Delta t \\ &= \int_a^b (b - \mathbb{E}_{\mathbb{T}}(X) + \mathbb{E}_{\mathbb{T}}(X) - t)(t - \mathbb{E}_{\mathbb{T}}(X) + \mathbb{E}_{\mathbb{T}}(X) - a)f(t)\Delta t \\ &\quad - \int_a^b (2h_2(t, 0) - t^2)f(t)\Delta t \\ &= (b - \mathbb{E}_{\mathbb{T}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a) \int_a^b f(t)\Delta t + (b - \mathbb{E}_{\mathbb{T}}(X)) \int_a^b (t - \mathbb{E}_{\mathbb{T}}(X))f(t)\Delta t \\ &\quad + (\mathbb{E}_{\mathbb{T}}(X) - a) \int_a^b (\mathbb{E}_{\mathbb{T}}(X) - t)f(t)\Delta t - \int_a^b (t - \mathbb{E}_{\mathbb{T}}(X))^2 f(t)\Delta t \\ &\quad - \int_a^b (2h_2(t, 0) - t^2)f(t)\Delta t \\ &= (b - \mathbb{E}_{\mathbb{T}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a) - \text{Var}_{\mathbb{T}}(X), \end{aligned}$$

and therefore

$$\text{Var}_{\mathbb{T}}(X) \leq (b - \mathbb{E}_{\mathbb{T}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a).$$

To show the second part of (74), note that for all $\alpha, \beta \in \mathbb{R}$, we have

$$\alpha\beta \leq \frac{(\alpha + \beta)^2}{4}.$$

Setting $\alpha = b - \mathbb{E}_{\mathbb{T}}(X)$ and $\beta = \mathbb{E}_{\mathbb{T}}(X) - a$ completes the proof of (74). From the derivation of (74), we have

$$\begin{aligned} &(b - \mathbb{E}_{\mathbb{T}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a) - \text{Var}_{\mathbb{T}}(X) \\ &= \int_a^b [(b-t)(t-a) + h_1^2(t, 0) - 2h_2(t, 0)] f(t)\Delta t \\ &\leq \|f\| \int_a^b ((b-t)(t-a) + h_1^2(t, 0) - 2h_2(t, 0)) \Delta t, \end{aligned}$$

and therefore inequality (75) is shown. This completes the proof of both parts of Theorem 53. \square

Example 54. *We apply Theorem 53 to different time scales.*

(i) *If $\mathbb{T} = \mathbb{R}$, then*

$$\text{Var}_{\mathbb{R}}(X) \leq (b - \mathbb{E}_{\mathbb{R}}(X))(\mathbb{E}_{\mathbb{T}}(X) - a) \leq \frac{(b - a)^2}{4}$$

and

$$(b - \mathbb{E}_{\mathbb{R}}(X))(\mathbb{E}_{\mathbb{R}}(X) - a) - \text{Var}_{\mathbb{R}}(X) \leq \|f\| \frac{(b - a)^3}{6},$$

where

$$\|f\| = \sup_{t \in [a, b]} |f(t)|,$$

which corresponds exactly to [12, Theorem 23].

(ii) *If $\mathbb{T} = \mathbb{Z}$, $a = 0$ and $b = n$, then*

$$\text{Var}_{\mathbb{Z}}(X) \leq (n - \mathbb{E}_{\mathbb{Z}}(X))\mathbb{E}_{\mathbb{Z}}(X) \leq \frac{n^2}{4}$$

and

$$(n - \mathbb{E}_{\mathbb{Z}}(X))\mathbb{E}_{\mathbb{Z}}(X) - \text{Var}_{\mathbb{Z}}(X) \leq \|f\| \frac{n(n - 1)(n + 4)}{6},$$

where

$$\|f\| = \max_{0 \leq t \leq n-1} |f(t)|.$$

10. DIAMOND-ALPHA GRÜSS TYPE INEQUALITIES

This chapter contains material from a collaborative work with Martin Bohner and Adnan Tuna, which appeared in 2011 in the *International Journal of Dynamical Systems and Differential Equations* with the title “Diamond-alpha Grüss type inequalities on time scales”, see [20]. We study a more general version of Grüss type inequalities on time scales by using the recent theory of combined dynamic derivatives on time scales. In the case $\alpha = 1$, we obtain delta-integral Grüss type inequalities on time scales. For $\alpha = 0$, we obtain nabla-integral Grüss type inequalities. We supply numerous examples throughout.

10.1. INTRODUCTION

M. R. Sidi Ammi and D. F. M. Torres [7] have established the diamond- α Grüss inequality on time scales as follows.

Theorem 54 (see [7, Theorem 3.4]). *Let \mathbb{T} be a time scale and $a, b \in \mathbb{T}$ with $a < b$. If $f, g \in C(\mathbb{T}, \mathbb{R})$ satisfy $\varphi \leq f(x) \leq \Phi$ and $\gamma \leq g(x) \leq \Gamma$ for all $x \in [a, b] \cap \mathbb{T}$, then*

$$\left| \frac{1}{b-a} \int_a^b f(x)g(x) \diamond_{\alpha} x - \frac{1}{(b-a)^2} \int_a^b f(x) \diamond_{\alpha} x \int_a^b g(x) \diamond_{\alpha} x \right| \leq \frac{1}{4}(\Phi - \varphi)(\Gamma - \gamma).$$

S. S. Dragomir [29] gave some classical and new integral inequalities of Grüss type, for example, the following two results.

Theorem 55 (see [29, Theorem 2.1]). *Let $f, g : [a, b] \rightarrow \mathbb{R}$ be two Lipschitzian mappings with Lipschitz constants $L_1 > 0$ and $L_2 > 0$, respectively, i.e.,*

$$|f(x) - f(y)| \leq L_1 |x - y| \quad \text{and} \quad |g(x) - g(y)| \leq L_2 |x - y|$$

for all $x, y \in [a, b]$. If $p : [a, b] \rightarrow [0, \infty)$ is integrable, then

$$\begin{aligned} & \left| \int_a^b p(x) dx \int_a^b p(x) f(x) g(x) dx - \int_a^b p(x) f(x) dx \int_a^b p(x) g(x) dx \right| \\ & \leq L_1 L_2 \left[\int_a^b p(x) dx \int_a^b p(x) x^2 dx - \left(\int_a^b p(x) x dx \right)^2 \right], \end{aligned}$$

and the inequality is sharp.

Theorem 56 (see [29, Theorem 4.1]). *Let $f, g : [a, b] \rightarrow \mathbb{R}$ be two integrable mappings on $[a, b]$ such that*

$$|f(x) - f(y)| \leq M |g(x) - g(y)|$$

for all $x, y \in [a, b]$. If $p : [a, b] \rightarrow [0, \infty)$ is integrable, then

$$\left| \int_a^b p(x) dx \int_a^b p(x) f(x) g(x) dx - \int_a^b p(x) f(x) dx \int_a^b p(x) g(x) dx \right| \leq M \left[\int_a^b p(x) dx \int_a^b p(x) g^2(x) dx - \left(\int_a^b p(x) g(x) dx \right)^2 \right],$$

and the inequality is sharp.

In 2006, Q. Sheng, M. Fadag, J. Henderson, and J. M. Davis [69] studied a combined dynamic “diamond-alpha” derivative as a linear combination of Δ and ∇ dynamic derivatives on time scales. The diamond- α derivative reduces to the standard Δ derivative for $\alpha = 1$ and to the standard ∇ derivative for $\alpha = 0$. Since then, many authors have established diamond- α inequalities on time scales [6, 7, 16, 36, 56]. We refer the reader to [47, 53, 63, 70, 69] for an account of the calculus with diamond- α dynamic derivatives.

This section is organized as follows: In Section 10.2, we briefly present some general definitions and theorems connected to the time scales calculus. Next, in Sections 10.3–10.5, we generalize Theorem 54, Theorem 55, and Theorem 56, respectively, for general time scales by using the recent theory of combined dynamic derivatives on time scales. In the case $\alpha = 1$, we obtain delta-integral Grüss type inequalities on time scales, while for $\alpha = 0$, we obtain nabla-integral Grüss type inequalities. In order to illustrate the theoretical results, we supply numerous examples throughout.

10.2. GENERAL DEFINITIONS

For the general theory of calculus on time scales we refer to [3, 22, 23, 41]. We now introduce the diamond- α integral, referring the reader to [6, 7, 36, 56] for more on the associated calculus.

Definition 48. *Let $0 \leq \alpha \leq 1$ and $f \in C(\mathbb{T}, \mathbb{R})$. Then the diamond-alpha integral of f is defined by*

$$\int_a^b f(x) \diamond_{\alpha} x = \alpha \int_a^b f(x) \Delta x + (1 - \alpha) \int_a^b f(x) \nabla x, \quad \text{where } a, b \in \mathbb{T}.$$

Theorem 57. Let $0 \leq \alpha \leq 1$ and $f, g \in C(\mathbb{T}, \mathbb{R})$. If $a, b, c \in \mathbb{T}$ and $\beta \in \mathbb{R}$, then

- (i) $\int_a^b [f(x) + g(x)] \diamond_{\alpha} x = \int_a^b f(x) \diamond_{\alpha} x + \int_a^b g(x) \diamond_{\alpha} x;$
- (ii) $\int_a^b (\beta f)(x) \diamond_{\alpha} x = \beta \int_a^b f(x) \diamond_{\alpha} x;$
- (iii) $\int_a^b f(x) \diamond_{\alpha} x = - \int_b^a f(x) \diamond_{\alpha} x;$
- (iv) $\int_a^b f(x) \diamond_{\alpha} x = \int_a^c f(x) \diamond_{\alpha} x + \int_c^b f(x) \diamond_{\alpha} x;$
- (v) $\int_a^a f(x) \diamond_{\alpha} x = 0;$
- (vi) if $f(x) \geq 0$ for all $x \in [a, b]$, then $\int_a^b f(x) \diamond_{\alpha} x \geq 0;$
- (vii) if $f(x) \leq g(x)$ for all $x \in [a, b]$, then $\int_a^b f(x) \diamond_{\alpha} x \leq \int_a^b g(x) \diamond_{\alpha} x;$
- (viii) $\left| \int_a^b f(x) \diamond_{\alpha} x \right| \leq \int_a^b |f(x)| \diamond_{\alpha} x.$

Example 55. If we let $\mathbb{T} = \mathbb{R}$ in Definition 48, then we obtain

$$\int_a^b f(x) \diamond_{\alpha} x = \int_a^b f(x) dx, \quad \text{where } a, b \in \mathbb{R}.$$

Example 56. If we let $\mathbb{T} = \mathbb{Z}$ in Definition 48 and $m < n$, then we obtain

$$\int_m^n f(x) \diamond_{\alpha} x = \sum_{i=m}^{n-1} [\alpha f_i + (1 - \alpha) f_{i+1}], \quad \text{where } m, n \in \mathbb{N}_0,$$

and where we put for convenience $f_i = f(i)$.

Example 57. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ in Definition 48 and $m < n$, then we obtain

$$\int_{q^m}^{q^n} f(x) \diamond_{\alpha} x = (q - 1) \sum_{i=m}^{n-1} q^i [\alpha f(q^i) + (1 - \alpha) f(q^{i+1})], \quad \text{where } m, n \in \mathbb{N}_0.$$

Example 58. Let $t_i < t_{i+1}$ for all $i \in \mathbb{N}_0$. If we let $\mathbb{T} = \{t_i : i \in \mathbb{N}_0\}$ in Definition 48 and $m < n$, then we obtain

$$\int_{t_m}^{t_n} f(x) \diamond_{\alpha} x = \sum_{i=m}^{n-1} (t_{i+1} - t_i) [\alpha f(t_i) + (1 - \alpha) f(t_{i+1})], \quad \text{where } m, n \in \mathbb{N}_0,$$

and from here we may obtain Example 56 by letting $t_i = i$ for all $i \in \mathbb{N}_0$ and Example 57 by letting $t_i = q^i$ for all $i \in \mathbb{N}_0$.

10.3. THE WEIGHTED DIAMOND-ALPHA GRÜSS INEQUALITY

We first extend Theorem 54 to the weighted case.

Theorem 58. *Let \mathbb{T} be a time scale and $a, b \in \mathbb{T}$ with $a < b$. If $f, g \in C(\mathbb{T}, \mathbb{R})$ and $p \in C(\mathbb{T}, [0, \infty))$ satisfy $\varphi \leq f(x) \leq \Phi$ and $\gamma \leq g(x) \leq \Gamma$ for all $x \in [a, b] \cap \mathbb{T}$ and $\int_a^b p(x) \diamond_{\alpha} x > 0$, then*

$$\begin{aligned} & \left| \int_a^b p(x) \diamond_{\alpha} x \int_a^b p(x) f(x) g(x) \diamond_{\alpha} x - \int_a^b p(x) f(x) \diamond_{\alpha} x \int_a^b p(x) g(x) \diamond_{\alpha} x \right| \\ & \leq \frac{1}{4} (\Phi - \varphi) (\Gamma - \gamma) \left(\int_a^b p(x) \diamond_{\alpha} x \right)^2. \end{aligned} \quad (76)$$

Proof. We have

$$\begin{aligned} & \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) g(x) \diamond_{\alpha} x \\ & - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x \\ & = \frac{1}{2 \left(\int_a^b p(x) \diamond_{\alpha} x \right)^2} \int_a^b \int_a^b p(x) p(y) (f(x) - f(y)) (g(x) - g(y)) \diamond_{\alpha} x \diamond_{\alpha} y. \end{aligned} \quad (77)$$

Applying the two-dimensional diamond- α Cauchy–Schwarz inequality from [8, Theorem 3.5], we get

$$\begin{aligned} & \left[\frac{1}{2 \left(\int_a^b p(x) \diamond_{\alpha} x \right)^2} \int_a^b \int_a^b p(x) p(y) (f(x) - f(y)) (g(x) - g(y)) \diamond_{\alpha} x \diamond_{\alpha} y \right]^2 \\ & \leq \frac{1}{2 \left(\int_a^b p(x) \diamond_{\alpha} x \right)^2} \int_a^b \int_a^b p(x) p(y) (f(x) - f(y))^2 \diamond_{\alpha} x \diamond_{\alpha} y \\ & \quad \times \frac{1}{2 \left(\int_a^b p(x) \diamond_{\alpha} x \right)^2} \int_a^b \int_a^b p(x) p(y) (g(x) - g(y))^2 \diamond_{\alpha} x \diamond_{\alpha} y \\ & = \left\{ \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f^2(x) \diamond_{\alpha} x - \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x \right)^2 \right\} \\ & \quad \times \left\{ \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g^2(x) \diamond_{\alpha} x - \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x \right)^2 \right\}. \end{aligned} \quad (78)$$

We also have

$$\begin{aligned}
& \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f^2(x) \diamond_{\alpha} x - \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x \right)^2 \quad (79) \\
&= \left(\Phi - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x \right) \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x - \varphi \right) \\
&\quad - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) (\Phi - f(x))(f(x) - \varphi) \diamond_{\alpha} x \\
&\leq \left(\Phi - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x \right) \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x - \varphi \right).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
& \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g^2(x) \diamond_{\alpha} x - \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x \right)^2 \\
&\leq \left(\Gamma - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x \right) \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x - \gamma \right). \quad (80)
\end{aligned}$$

Using (79) and (80) in (78), (77) implies

$$\begin{aligned}
& \left| \frac{\int_a^b p(x) f(x) g(x) \diamond_{\alpha} x}{\int_a^b p(x) \diamond_{\alpha} x} - \frac{\int_a^b p(x) f(x) \diamond_{\alpha} x}{\int_a^b p(x) \diamond_{\alpha} x} \frac{\int_a^b p(x) g(x) \diamond_{\alpha} x}{\int_a^b p(x) \diamond_{\alpha} x} \right| \\
&\leq \left(\Phi - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x \right)^{\frac{1}{2}} \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) f(x) \diamond_{\alpha} x - \varphi \right)^{\frac{1}{2}} \\
&\quad \times \left(\Gamma - \frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x \right)^{\frac{1}{2}} \left(\frac{1}{\int_a^b p(x) \diamond_{\alpha} x} \int_a^b p(x) g(x) \diamond_{\alpha} x - \gamma \right)^{\frac{1}{2}}. \quad (81)
\end{aligned}$$

Applying the elementary inequality

$$4\beta\gamma \leq (\beta + \gamma)^2 \quad \text{for all } \beta, \gamma \in \mathbb{R},$$

we can state

$$4 \left(\Phi - \frac{\int_a^b p(x) f(x) \diamond_{\alpha} x}{\int_a^b p(x) \diamond_{\alpha} x} \right) \left(\frac{\int_a^b p(x) f(x) \diamond_{\alpha} x}{\int_a^b p(x) \diamond_{\alpha} x} - \varphi \right) \leq (\Phi - \varphi)^2 \quad (82)$$

and

$$4 \left(\Gamma - \frac{\int_a^b p(x)g(x)\diamond_{\alpha}x}{\int_a^b p(x)\diamond_{\alpha}x} \right) \left(\frac{\int_a^b p(x)g(x)\diamond_{\alpha}x}{\int_a^b p(x)\diamond_{\alpha}x} - \gamma \right) \leq (\Gamma - \gamma)^2. \quad (83)$$

Combining (81) with (82) and (83), we obtain (76). \square

Example 59. If we let $p(x) \equiv 1$ on \mathbb{T} in Theorem 58, then we obtain Theorem 54.

Example 60. If we let $\mathbb{T} = \mathbb{R}$ in Theorem 58, then we obtain the inequality

$$\left| \int_a^b p(x)dx \int_a^b p(x)f(x)g(x)dx - \int_a^b p(x)f(x)dx \int_a^b p(x)g(x)dx \right| \leq \frac{1}{4}(\Phi - \varphi)(\Gamma - \gamma) \left(\int_a^b p(x)dx \right)^2.$$

This result can be found in [29, Theorem 1.1], where the constant $\frac{1}{4}$ is also shown to be the best possible.

Example 61. If we let $\mathbb{T} = \mathbb{R}$ in Example 59, then we obtain the inequality

$$\left| \frac{1}{b-a} \int_a^b f(x)g(x)dx - \frac{1}{(b-a)^2} \int_a^b f(x)dx \int_a^b g(x)dx \right| \leq \frac{1}{4}(\Phi - \varphi)(\Gamma - \gamma).$$

Example 62. If we let $\mathbb{T} = \mathbb{Z}$ and $\alpha = 1$ in Theorem 58, then we obtain the inequality

$$\left| \sum_{i=m}^{n-1} p_i \sum_{i=m}^{n-1} p_i f_i g_i - \sum_{i=m}^{n-1} p_i f_i \sum_{i=m}^{n-1} p_i g_i \right| \leq \frac{1}{4}(\Phi - \varphi)(\Gamma - \gamma) \left(\sum_{i=m}^{n-1} p_i \right)^2.$$

Example 63. If we let $\mathbb{T} = \mathbb{Z}$ in Example 59, then we obtain the inequality

$$\left| \frac{1}{n-m} \sum_{i=m}^{n-1} [\alpha f_i g_i + (1-\alpha) f_{i+1} g_{i+1}] - \frac{1}{(n-m)^2} \sum_{i=m}^{n-1} [\alpha f_i + (1-\alpha) f_{i+1}] \sum_{i=m}^{n-1} [\alpha g_i + (1-\alpha) g_{i+1}] \right| \leq \frac{1}{4}(\Phi - \varphi)(\Gamma - \gamma).$$

If, additionally, $\alpha = 1$, then we obtain the inequality

$$\left| \frac{1}{n-m} \sum_{i=m}^{n-1} f_i g_i - \frac{1}{(n-m)^2} \sum_{i=m}^{n-1} f_i \sum_{i=m}^{n-1} g_i \right| \leq \frac{1}{4}(\Phi - \varphi)(\Gamma - \gamma).$$

Example 64. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ and $\alpha = 1$ in Theorem 58, then we obtain the inequality

$$\left| \sum_{i=m}^{n-1} q^i p(q^i) \sum_{i=m}^{n-1} q^i p(q^i) f(q^i) g(q^i) - \sum_{i=m}^{n-1} q^i p(q^i) f(q^i) \sum_{i=m}^{n-1} q^i p(q^i) g(q^i) \right| \leq \frac{1}{4} (\Phi - \varphi) (\Gamma - \gamma) \left(\sum_{i=m}^{n-1} q^i p(q^i) \right)^2.$$

Example 65. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ in Example 59, then we obtain the inequality

$$\left| \frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i [\alpha f(q^i) g(q^i) + (1-\alpha) f(q^{i+1}) g(q^{i+1})] - \left(\frac{q-1}{q^n - q^m} \right)^2 \sum_{i=m}^{n-1} q^i [\alpha f(q^i) + (1-\alpha) f(q^{i+1})] \sum_{i=m}^{n-1} q^i [\alpha g(q^i) + (1-\alpha) g(q^{i+1})] \right| \leq \frac{1}{4} (\Phi - \varphi) (\Gamma - \gamma).$$

If, additionally, $\alpha = 1$, then we obtain the inequality

$$\left| \frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i f(q^i) g(q^i) - \left(\frac{q-1}{q^n - q^m} \right)^2 \sum_{i=m}^{n-1} q^i f(q^i) \sum_{i=m}^{n-1} q^i g(q^i) \right| \leq \frac{1}{4} (\Phi - \varphi) (\Gamma - \gamma).$$

10.4. THE CASE WHEN BOTH MAPPINGS ARE LIPSCHITZIAN

We now extend Theorem 55 to time scales.

Theorem 59. Let \mathbb{T} be a time scale and $a, b \in \mathbb{T}$ with $a < b$. Let $f, g \in C(\mathbb{T}, \mathbb{R})$ be two Lipschitzian mappings with Lipschitz constants $L_1 > 0$ and $L_2 > 0$, respectively, i.e.,

$$|f(x) - f(y)| \leq L_1 |x - y| \quad \text{and} \quad |g(x) - g(y)| \leq L_2 |x - y| \quad (84)$$

for all $x, y \in [a, b] \cap \mathbb{T}$. If $p \in C(\mathbb{T}, [0, \infty))$, then

$$\left| \int_a^b p(x) \diamond_{\alpha} x \int_a^b p(x) f(x) g(x) \diamond_{\alpha} x - \int_a^b p(x) f(x) \diamond_{\alpha} x \int_a^b p(x) g(x) \diamond_{\alpha} x \right| \leq L_1 L_2 \left[\int_a^b p(x) \diamond_{\alpha} x \int_a^b p(x) x^2 \diamond_{\alpha} x - \left(\int_a^b p(x) x \diamond_{\alpha} x \right)^2 \right], \quad (85)$$

and the inequality is sharp in the sense, that the right hand side cannot be replaced by a smaller expression.

Proof. Using condition (84), we get

$$|(f(x) - f(y))(g(x) - g(y))| \leq L_1 L_2 (x - y)^2$$

for all $x, y \in [a, b] \cap \mathbb{T}$. Multiplying this inequality by $p(x)p(y) \geq 0$ and integrating over $[a, b] \times [a, b]$, we have

$$\begin{aligned} & \left| \int_a^b \int_a^b p(x)p(y)(f(x) - f(y))(g(x) - g(y)) \diamond_\alpha x \diamond_\alpha y \right| \\ & \leq \int_a^b \int_a^b p(x)p(y) |(f(x) - f(y))(g(x) - g(y))| \diamond_\alpha x \diamond_\alpha y \\ & \leq L_1 L_2 \int_a^b \int_a^b p(x)p(y)(x - y)^2 \diamond_\alpha x \diamond_\alpha y. \end{aligned}$$

We also have

$$\begin{aligned} & \frac{1}{2} \int_a^b \int_a^b p(x)p(y)(f(x) - f(y))(g(x) - g(y)) \diamond_\alpha x \diamond_\alpha y \\ & = \int_a^b p(x) \diamond_\alpha x \int_a^b p(x)f(x)g(x) \diamond_\alpha x - \int_a^b p(x)f(x) \diamond_\alpha x \int_a^b p(x)g(x) \diamond_\alpha x \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{2} \int_a^b \int_a^b p(x)p(y)(x - y)^2 \diamond_\alpha x \diamond_\alpha y \\ & = \int_a^b p(x) \diamond_\alpha x \int_a^b p(x)x^2 \diamond_\alpha x - \left(\int_a^b p(x)x \diamond_\alpha x \right)^2, \end{aligned}$$

which completes the proof of inequality (85). Moreover, if we choose $L_1, L_2 > 0$, $f(x) = L_1 x$ and $g(x) = L_2 x$ for $x \in \mathbb{T}$, then f and g are Lipschitzian with Lipschitz constants $L_1 > 0$ and $L_2 > 0$, respectively, and equality holds in (85) for any $p \in C(\mathbb{T}, [0, \infty))$. \square

Example 66. If we let $p(x) \equiv 1$ on \mathbb{T} in Theorem 59, then we obtain the inequality

$$\begin{aligned} & \left| \frac{1}{b-a} \int_a^b f(x)g(x) \diamond_\alpha x - \frac{1}{b-a} \int_a^b f(x) \diamond_\alpha x \frac{1}{b-a} \int_a^b g(x) \diamond_\alpha x \right| \\ & \leq L_1 L_2 \left[\frac{1}{b-a} \int_a^b x^2 \diamond_\alpha x - \left(\frac{1}{b-a} \int_a^b x \diamond_\alpha x \right)^2 \right]. \end{aligned}$$

Example 67. If we let $\mathbb{T} = \mathbb{R}$ in Theorem 59, then we obtain Theorem 55.

Example 68. If we let $\mathbb{T} = \mathbb{R}$ in Example 66, then we obtain the inequality

$$\left| \frac{1}{b-a} \int_a^b f(x)g(x)dx - \frac{1}{b-a} \int_a^b f(x)dx \frac{1}{b-a} \int_a^b g(x)dx \right| \leq L_1 L_2 \frac{(b-a)^2}{12},$$

which can be found in [29, Corollary 2.2].

Example 69. If we let $\mathbb{T} = \mathbb{Z}$ and $\alpha = 1$ in Theorem 59, then we obtain the inequality

$$\left| \sum_{i=m}^{n-1} p_i \sum_{i=m}^{n-1} p_i f_i g_i - \sum_{i=m}^{n-1} p_i f_i \sum_{i=m}^{n-1} p_i g_i \right| \leq L_1 L_2 \left[\sum_{i=m}^{n-1} p_i \sum_{i=m}^{n-1} p_i i^2 - \left(\sum_{i=m}^{n-1} p_i i \right)^2 \right].$$

Example 70. If we let $\mathbb{T} = \mathbb{Z}$ in Example 66, then we obtain the inequality

$$\begin{aligned} & \left| \frac{1}{n-m} \sum_{i=m}^{n-1} [\alpha f_i g_i + (1-\alpha) f_{i+1} g_{i+1}] \right. \\ & \left. - \frac{1}{(n-m)^2} \sum_{i=m}^{n-1} [\alpha f_i + (1-\alpha) f_{i+1}] \sum_{i=m}^{n-1} [\alpha g_i + (1-\alpha) g_{i+1}] \right| \\ & \leq L_1 L_2 \left[\frac{(n-m)^2 - 1}{12} + \alpha(1-\alpha) \right]. \end{aligned}$$

If, additionally, $\alpha = 1$, then we obtain the inequality

$$\left| \frac{1}{n-m} \sum_{i=m}^{n-1} f_i g_i - \frac{1}{(n-m)^2} \sum_{i=m}^{n-1} f_i \sum_{i=m}^{n-1} g_i \right| \leq L_1 L_2 \frac{(n-m)^2 - 1}{12}.$$

Note also that we have in the discrete case the same bound on the right-hand side than in the continuous case if and only if $\alpha = \frac{1}{2} - \frac{1}{\sqrt{6}}$ or $\alpha = \frac{1}{2} + \frac{1}{\sqrt{6}}$.

Example 71. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ and $\alpha = 1$ in Theorem 59, then we obtain the inequality

$$\begin{aligned} & \left| \sum_{i=m}^{n-1} q^i p(q^i) \sum_{i=m}^{n-1} q^i p(q^i) f(q^i) g(q^i) - \sum_{i=m}^{n-1} q^i p(q^i) f(q^i) \sum_{i=m}^{n-1} q^i p(q^i) g(q^i) \right| \\ & \leq L_1 L_2 \left[\sum_{i=m}^{n-1} q^i p(q^i) \sum_{i=m}^{n-1} q^{3i} p(q^i) - \left(\sum_{i=m}^{n-1} q^{2i} p(q^i) \right)^2 \right]. \end{aligned}$$

Example 72. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ in Example 66, then we obtain the inequality

$$\begin{aligned} & \left| \frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i [\alpha f(q^i)g(q^i) + (1-\alpha)f(q^{i+1})g(q^{i+1})] \right. \\ & \left. - \left(\frac{q-1}{q^n - q^m} \right)^2 \sum_{i=m}^{n-1} q^i [\alpha f(q^i) + (1-\alpha)f(q^{i+1})] \sum_{i=m}^{n-1} q^i [\alpha g(q^i) + (1-\alpha)g(q^{i+1})] \right| \\ & \leq L_1 L_2 \left[\frac{q^{2n} + q^n q^m + q^{2m}}{q^2 + q + 1} (\alpha + (1-\alpha)q^2) - \left(\frac{q^n + q^m}{q+1} \right)^2 (\alpha + (1-\alpha)q)^2 \right]. \end{aligned}$$

If, additionally, $\alpha = 1$, then we obtain the inequality

$$\begin{aligned} & \left| \frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i f(q^i)g(q^i) - \left(\frac{q-1}{q^n - q^m} \right)^2 \sum_{i=m}^{n-1} q^i f(q^i) \sum_{i=m}^{n-1} q^i g(q^i) \right| \\ & \leq L_1 L_2 \frac{(q^n - q^{m+1})(q^{n+1} - q^m)}{(q^2 + q + 1)(q+1)^2}. \end{aligned}$$

10.5. THE CASE WHEN F IS M - G -LIPSCHITZIAN

Theorem 60. Let \mathbb{T} be a time scale and $a, b \in \mathbb{T}$ with $a < b$. Let $f, g \in C(\mathbb{T}, \mathbb{R})$ be such that f is M - g -Lipschitzian with $M > 0$, i.e.,

$$|f(x) - f(y)| \leq M |g(x) - g(y)| \quad (86)$$

for all $x, y \in [a, b] \cap \mathbb{T}$. If $p \in C(\mathbb{T}, [0, \infty))$, then

$$\begin{aligned} & \left| \int_a^b p(x) \diamond_{\alpha} x \int_a^b p(x) f(x) g(x) \diamond_{\alpha} x - \int_a^b p(x) f(x) \diamond_{\alpha} x \int_a^b p(x) g(x) \diamond_{\alpha} x \right| \\ & \leq M \left[\int_a^b p(x) \diamond_{\alpha} x \int_a^b p(x) g^2(x) \diamond_{\alpha} x - \left(\int_a^b p(x) g(x) \diamond_{\alpha} x \right)^2 \right], \quad (87) \end{aligned}$$

and the inequality is sharp.

Proof. Using condition (86), we get

$$|(f(x) - f(y))(g(x) - g(y))| \leq M(g(x) - g(y))^2$$

for all $x, y \in [a, b] \cap \mathbb{T}$. Multiplying this inequality by $p(x)p(y) \geq 0$ and integrating over $[a, b] \times [a, b]$, we have

$$\begin{aligned} & \frac{1}{2} \left| \int_a^b \int_a^b p(x)p(y)(f(x) - f(y))(g(x) - g(y)) \diamond_{\alpha} x \diamond_{\alpha} y \right| \\ & \leq \frac{1}{2} \int_a^b \int_a^b p(x)p(y) |(f(x) - f(y))(g(x) - g(y))| \diamond_{\alpha} x \diamond_{\alpha} y \\ & \leq \frac{M}{2} \int_a^b \int_a^b p(x)p(y) ((g(x) - g(y))^2 \diamond_{\alpha} x \diamond_{\alpha} y) \\ & = M \left[\int_a^b p(x) \diamond_{\alpha} x \int_a^b p(x)g^2(x) \diamond_{\alpha} x - \left(\int_a^b p(x)g(x) \diamond_{\alpha} x \right)^2 \right], \end{aligned}$$

which completes the proof of inequality (87). Moreover, if we choose $f(x) = Mx$ with $M > 0$ and $g(x) = x$, then f is M - g -Lipschitzian and equality holds in (87) for any $p \in C(\mathbb{T}, [0, \infty))$. \square

Example 73. If we let $p(x) \equiv 1$ on \mathbb{T} in Theorem 60, then we obtain the inequality

$$\begin{aligned} & \left| \frac{1}{b-a} \int_a^b f(x)g(x) \diamond_{\alpha} x - \frac{1}{b-a} \int_a^b f(x) \diamond_{\alpha} x \frac{1}{b-a} \int_a^b g(x) \diamond_{\alpha} x \right| \\ & \leq M \left[\frac{1}{b-a} \int_a^b g^2(x) \diamond_{\alpha} x - \left(\frac{1}{b-a} \int_a^b g(x) \diamond_{\alpha} x \right)^2 \right]. \end{aligned}$$

Example 74. If we let $\mathbb{T} = \mathbb{R}$ in Theorem 60, then we obtain Theorem 56.

Example 75. If we let $\mathbb{T} = \mathbb{R}$ in Example 73, then we obtain the inequality

$$\begin{aligned} & \left| \frac{1}{b-a} \int_a^b f(x)g(x) dx - \frac{1}{b-a} \int_a^b f(x) dx \frac{1}{b-a} \int_a^b g(x) dx \right| \\ & \leq M \left[\frac{1}{b-a} \int_a^b g^2(x) dx - \left(\frac{1}{b-a} \int_a^b g(x) dx \right)^2 \right], \end{aligned}$$

which can be found in [29, Remark 4.2].

Example 76. If we let $\mathbb{T} = \mathbb{Z}$ and $\alpha = 1$ in Theorem 60, then we obtain the inequality

$$\left| \sum_{i=m}^{n-1} p_i \sum_{i=m}^{n-1} p_i f_i g_i - \sum_{i=m}^{n-1} p_i f_i \sum_{i=m}^{n-1} p_i g_i \right| \leq M \left[\sum_{i=m}^{n-1} p_i \sum_{i=m}^{n-1} p_i g_i^2 - \left(\sum_{i=m}^{n-1} p_i g_i \right)^2 \right].$$

Example 77. If we let $\mathbb{T} = \mathbb{Z}$ in Example 73, then we obtain the inequality

$$\begin{aligned} & \left| \frac{1}{n-m} \sum_{i=m}^{n-1} [\alpha f_i g_i + (1-\alpha) f_{i+1} g_{i+1}] \right. \\ & \quad \left. - \frac{1}{(n-m)^2} \sum_{i=m}^{n-1} [\alpha f_i + (1-\alpha) f_{i+1}] \sum_{i=m}^{n-1} [\alpha g_i + (1-\alpha) g_{i+1}] \right| \\ & \leq M \left[\frac{1}{n-m} \sum_{i=m}^{n-1} [\alpha g_i^2 + (1-\alpha) g_{i+1}^2] - \left(\frac{1}{n-m} \sum_{i=m}^{n-1} [\alpha g_i + (1-\alpha) g_{i+1}] \right)^2 \right]. \end{aligned}$$

If, additionally, $\alpha = 1$, then we obtain the inequality

$$\begin{aligned} & \left| \frac{1}{n-m} \sum_{i=m}^{n-1} f_i g_i - \frac{1}{(n-m)^2} \sum_{i=m}^{n-1} f_i \sum_{i=m}^{n-1} g_i \right| \\ & \leq M \left[\frac{1}{n-m} \sum_{i=m}^{n-1} g_i^2 - \left(\frac{1}{n-m} \sum_{i=m}^{n-1} g_i \right)^2 \right]. \end{aligned}$$

Example 78. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ and $\alpha = 1$ in Theorem 60, then we obtain the inequality

$$\begin{aligned} & \left| \sum_{i=m}^{n-1} q^i p(q^i) \sum_{i=m}^{n-1} q^i p(q^i) f(q^i) g(q^i) - \sum_{i=m}^{n-1} q^i p(q^i) f(q^i) \sum_{i=m}^{n-1} q^i p(q^i) g(q^i) \right| \\ & \leq M \left[\sum_{i=m}^{n-1} q^i p(q^i) \sum_{i=m}^{n-1} q^i p(q^i) g^2(q^i) - \left(\sum_{i=m}^{n-1} q^i p(q^i) g(q^i) \right)^2 \right]. \end{aligned}$$

Example 79. If we let $\mathbb{T} = q^{\mathbb{N}_0}$ in Example 73, then we obtain the inequality

$$\begin{aligned} & \left| \frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i [\alpha f(q^i) g(q^i) + (1-\alpha) f(q^{i+1}) g(q^{i+1})] \right. \\ & \quad \left. - \left(\frac{q-1}{q^n - q^m} \right)^2 \sum_{i=m}^{n-1} q^i [\alpha f(q^i) + (1-\alpha) f(q^{i+1})] \sum_{i=m}^{n-1} q^i [\alpha g(q^i) + (1-\alpha) g(q^{i+1})] \right| \\ & \leq M \left[\frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i [\alpha g^2(q^i) + (1-\alpha) g^2(q^{i+1})] \right. \\ & \quad \left. - \left(\frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i [\alpha g(q^i) + (1-\alpha) g(q^{i+1})] \right)^2 \right]. \end{aligned}$$

If, additionally, $\alpha = 1$, then we obtain the inequality

$$\left| \frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i f(q^i) g(q^i) - \left(\frac{q-1}{q^n - q^m} \right)^2 \sum_{i=m}^{n-1} q^i f(q^i) \sum_{i=m}^{n-1} q^i g(q^i) \right|$$

$$\leq M \left[\frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i g^2(q^i) - \left(\frac{q-1}{q^n - q^m} \sum_{i=m}^{n-1} q^i g(q^i) \right)^2 \right].$$

11. WEIGHTED OSTROWSKI–GRÜSS INEQUALITIES

This chapter contains also a collaborative work with Martin Bohner and Adnan Tuna and appeared in 2011 in the *African Diaspora Journal of Mathematics* with the title “Weighted Ostrowski–Grüss Inequalities on Time Scales”, see [21]. We study Ostrowski–Grüss and Ostrowski-like inequalities on time scales and thus unify and extend corresponding continuous and discrete versions from the literature. We present corresponding inequalities by using the time scales L^∞ -norm and also by using the time scales L^p -norm. Several interesting inequalities representing special cases of our general results are supplied.

11.1. INTRODUCTION

In 1938, A. Ostrowski (see [55, Formula (2)]) presented the following interesting integral inequality.

Theorem 61. *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) such that $f' \in L^\infty((a, b))$, i.e.,*

$$\|f'\|_\infty := \sup_{s \in (a, b)} |f'(s)| < \infty,$$

then for all $t \in [a, b]$, we have

$$\left| f(t) - \frac{1}{b-a} \int_a^b f(s) ds \right| \leq \left[\frac{1}{4} + \left(\frac{t - \frac{a+b}{2}}{b-a} \right)^2 \right] (b-a) \|f'\|_\infty. \quad (88)$$

In 2007, B. Pachpatte (see [60, Theorem 1 and Theorem 2]) established new generalizations of Ostrowski-type inequalities involving two functions, whose derivatives belong to L^p -spaces.

Theorem 62. *Let $p > 1$ and $q := p/(p-1)$. If $f, g : [a, b] \rightarrow \mathbb{R}$ are absolutely continuous such that $f', g' \in L^p([a, b])$, i.e.,*

$$\|f'\|_p := \left(\int_a^b |f'(s)|^p ds \right)^{\frac{1}{p}} < \infty \quad \text{and} \quad \|g'\|_p = \left(\int_a^b |g'(s)|^p ds \right)^{\frac{1}{p}} < \infty,$$

then for all $t \in [a, b]$, we have

$$\begin{aligned} \left| f(t)g(t) - \frac{1}{2(b-a)} \left[g(t) \int_a^b f(s)ds + f(t) \int_a^b g(s)ds \right] \right| \\ \leq \frac{(B(t))^{\frac{1}{q}} |g(t)| \|f'\|_p + |f(t)| \|g'\|_p}{2(b-a)} \end{aligned} \quad (89)$$

and

$$\begin{aligned} \left| f(t)g(t) - \frac{1}{b-a} \left[g(t) \int_a^b f(s)ds + f(t) \int_a^b g(s)ds \right] \right. \\ \left. + \left(\frac{1}{b-a} \int_a^b f(s)ds \right) \left(\frac{1}{b-a} \int_a^b g(s)ds \right) \right| \\ \leq \left(\frac{(B(t))^{\frac{1}{q}}}{b-a} \right)^2 \|f'\|_p \|g'\|_p, \end{aligned} \quad (90)$$

where

$$B(t) := \frac{1}{q+1} [(t-a)^{q+1} + (b-t)^{q+1}].$$

In 1988, S. Hilger [41] introduced the time scales theory to unify continuous and discrete analysis. Since then, many authors have studied certain integral inequalities on time scales, see, e.g., [2, 22, 23, 18, 19, 46, 54, 20, 65, 72]. In [19], M. Bohner and T. Matthews established the time scales version of Ostrowski's inequality, hence unifying discrete, continuous and other versions of Theorem 61.

This work is organized as follows: In Section 11.2 and Section 11.3, we obtain time scales versions of weighted Ostrowski–Grüss and Ostrowski-like inequalities using the L^∞ -norm and the L^p -norm, respectively. Our proofs utilize generalizations of so-called Montgomery inequalities, see [49, page 565] and [50, page 261].

11.2. OSTROWSKI–GRÜSS INEQUALITIES IN L^∞ -NORM

Throughout, we use the following assumption.

Assumption (H). *From now on, until the end of this paper, we assume that \mathbb{T} is a time scale and that $a, b \in \mathbb{T}$ such that $a < b$. By writing $[a, b]$, we mean $[a, b] \cap \mathbb{T}$. Moreover, $w \in C_{\text{rd}}([a, b], [0, \infty))$ is such that*

$$m(a, b) := \int_a^b w(t) \Delta t < \infty,$$

and we also define

$$p_w(t, s) := \begin{cases} \int_a^s w(\tau) \Delta \tau & \text{for } a \leq s < t \\ \int_b^s w(\tau) \Delta \tau & \text{for } t \leq s \leq b. \end{cases}$$

Theorem 63. Assume (H). If $f, g \in C_{\text{rd}}^1([a, b], \mathbb{R})$ such that $f^\Delta, g^\Delta \in L^\infty((a, b))$, i.e.,

$$\|f^\Delta\|_\infty := \sup_{s \in (a, b)} |f^\Delta(s)| < \infty \quad \text{and} \quad \|g^\Delta\|_\infty = \sup_{s \in (a, b)} |g^\Delta(s)| < \infty, \quad (91)$$

then for all $t \in [a, b]$, we have

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{2m(a, b)} \left[g(t) \int_a^b w(s) f(\sigma(s)) \Delta s + f(t) \int_a^b w(s) g(\sigma(s)) \Delta s \right] \right| \\ & \leq \left(\frac{1}{m(a, b)} \int_a^b (\sigma(s) - t) w(s) \operatorname{sgn}(s - t) \Delta s \right) \frac{|g(t)| \|f^\Delta\|_\infty + |f(t)| \|g^\Delta\|_\infty}{2} \end{aligned} \quad (92)$$

and

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{m(a, b)} \left[g(t) \int_a^b w(s) f(\sigma(s)) \Delta s + f(t) \int_a^b w(s) g(\sigma(s)) \Delta s \right] \right| \\ & + \left(\frac{1}{m(a, b)} \int_a^b w(s) f(\sigma(s)) \Delta s \right) \left(\frac{1}{m(a, b)} \int_a^b w(s) g(\sigma(s)) \Delta s \right) \\ & \leq \left(\frac{1}{m(a, b)} \int_a^b (\sigma(s) - t) w(s) \operatorname{sgn}(s - t) \Delta s \right)^2 \|f^\Delta\|_\infty \|g^\Delta\|_\infty. \end{aligned} \quad (93)$$

Proof. Using integration by parts formula twice, we have

$$\begin{aligned} \int_a^b p_w(t, s) f^\Delta(s) \Delta s &= \int_a^t \left(\int_a^s w(\tau) \Delta \tau \right) f^\Delta(s) \Delta s \\ &\quad + \int_t^b \left(\int_b^s w(\tau) \Delta \tau \right) f^\Delta(s) \Delta s \\ &= f(t) \int_a^t w(\tau) \Delta \tau - \int_a^t w(s) f(\sigma(s)) \Delta s \\ &\quad - f(t) \int_b^t w(\tau) \Delta \tau - \int_t^b w(s) f(\sigma(s)) \Delta s \\ &= m(a, b) f(t) - \int_a^b w(s) f(\sigma(s)) \Delta s \end{aligned}$$

and thus

$$f(t) - \frac{1}{m(a,b)} \int_a^b w(s)f(\sigma(s))\Delta s = \frac{1}{m(a,b)} \int_a^b p_w(t,s)f^\Delta(s)\Delta s. \quad (94)$$

Replacing f by g in (94), we obtain

$$g(t) - \frac{1}{m(a,b)} \int_a^b w(s)g(\sigma(s))\Delta s = \frac{1}{m(a,b)} \int_a^b p_w(t,s)g^\Delta(s)\Delta s. \quad (95)$$

Using a similar calculation, we find

$$\begin{aligned} \int_a^b |p_w(t,s)| \Delta s &= \int_a^t \left(\int_a^s w(\tau)\Delta\tau \right) \Delta s - \int_t^b \left(\int_b^s w(\tau)\Delta\tau \right) \Delta s \\ &= t \int_a^t w(\tau)\Delta\tau - \int_a^t w(s)\sigma(s)\Delta s + t \int_b^t w(\tau)\Delta\tau + \int_t^b w(s)\sigma(s)\Delta s \\ &= \int_a^b \sigma(s)w(s) \operatorname{sgn}(s-t)\Delta s - t \int_a^b w(s) \operatorname{sgn}(s-t)\Delta s \\ &= \int_a^b (\sigma(s) - t)w(s) \operatorname{sgn}(s-t)\Delta s. \end{aligned} \quad (96)$$

Now multiplying (94) by $g(t)$ and (95) by $f(t)$, adding the resulting identities, rewriting, and taking absolute values, we have

$$\begin{aligned} &\left| f(t)g(t) - \frac{1}{2m(a,b)} \left[g(t) \int_a^b w(s)f(\sigma(s))\Delta s + f(t) \int_a^b w(s)g(\sigma(s))\Delta s \right] \right| \\ &= \frac{1}{2m(a,b)} \left| g(t) \int_a^b p_w(t,s)f^\Delta(s)\Delta s + f(t) \int_a^b p_w(t,s)g^\Delta(s)\Delta s \right| \\ &\leq \frac{1}{2m(a,b)} \left[|g(t)| \int_a^b |p_w(t,s)| |f^\Delta(s)| \Delta s + |f(t)| \int_a^b |p_w(t,s)| |g^\Delta(s)| \Delta s \right]. \end{aligned} \quad (97)$$

Using now (91) and (96) in (97), we obtain (92).

Next, multiplying the left and right sides of (94) and (95) and taking absolute values, we get

$$\begin{aligned}
& \left| f(t)g(t) - \frac{1}{m(a,b)} \left[g(t) \int_a^b w(s)f(\sigma(s))\Delta s + f(t) \int_a^b w(s)g(\sigma(s))\Delta s \right] \right. \\
& \quad \left. + \left(\frac{1}{m(a,b)} \int_a^b w(s)f(\sigma(s))\Delta s \right) \left(\frac{1}{m(a,b)} \int_a^b w(s)g(\sigma(s))\Delta s \right) \right| \\
& = \frac{1}{m^2(a,b)} \left| \left(\int_a^b p_w(t,s)f^\Delta(s)\Delta s \right) \left(\int_a^b p_w(t,s)g^\Delta(s)\Delta s \right) \right| \\
& \leq \frac{1}{m^2(a,b)} \left(\int_a^b |p_w(t,s)| |f^\Delta(s)| \Delta s \right) \left(\int_a^b |p_w(t,s)| |g^\Delta(s)| \Delta s \right).
\end{aligned} \tag{98}$$

Using now (91) and (96) in (98), we obtain (93). \square

Corollary 5. *In addition to the assumptions of Theorem 63, let $w(t) = 1$ for all $t \in [a, b]$.*

Then for all $t \in [a, b]$, we have

$$\begin{aligned}
& \left| f(t)g(t) - \frac{1}{2(b-a)} \left[g(t) \int_a^b f(\sigma(s))\Delta s + f(t) \int_a^b g(\sigma(s))\Delta s \right] \right| \\
& \leq \frac{h_2(t,a) + g_2(b,t)}{b-a} \frac{|g(t)| \|f^\Delta\|_\infty + |f(t)| \|g^\Delta\|_\infty}{2}
\end{aligned} \tag{99}$$

and

$$\begin{aligned}
& \left| f(t)g(t) - \frac{1}{b-a} \left[g(t) \int_a^b f(\sigma(s))\Delta s + f(t) \int_a^b g(\sigma(s))\Delta s \right] \right. \\
& \quad \left. + \left(\frac{1}{b-a} \int_a^b f(\sigma(s))\Delta s \right) \left(\frac{1}{b-a} \int_a^b g(\sigma(s))\Delta s \right) \right| \\
& \leq \left(\frac{h_2(t,a) + g_2(b,t)}{b-a} \right)^2 \|f^\Delta\|_\infty \|g^\Delta\|_\infty.
\end{aligned} \tag{100}$$

Proof. We just have to use Theorem 63 and

$$\begin{aligned}
\int_a^b (\sigma(s) - t) \operatorname{sgn}(s - t) \Delta s &= - \int_a^t (\sigma(s) - t) \Delta s + \int_t^b (\sigma(s) - t) \Delta s \\
&= \int_t^a (\sigma(s) - t) \Delta s + \int_t^b (\sigma(s) - t) \Delta s \\
&= g_2(a, t) + g_2(b, t) = h_2(t, a) + g_2(b, t),
\end{aligned}$$

where we also applied [22, Theorem 1.112]. \square

Example 80. If we let $g(t) = 1$ for all $t \in [a, b]$, then (99) becomes

$$\left| f(t) - \frac{1}{b-a} \int_a^b f(\sigma(s)) \Delta s \right| \leq \frac{h_2(t, a) + g_2(b, t)}{b-a} \|f^\Delta\|_\infty, \quad (101)$$

which is the Ostrowski inequality on time scales as given in [19, Theorem 3.5]. If $\mathbb{T} = \mathbb{R}$ in (101), then we obtain (88) in Theorem 61. If $\mathbb{T} = \mathbb{Z}$, $a = 0$, and $b = n \in \mathbb{N}$ in (101), then we obtain

$$\left| f(t) - \frac{1}{n} \sum_{s=1}^n f(s) \right| \leq \frac{1}{n} \left[\frac{n^2 - 1}{4} + \left(t - \frac{n+1}{2} \right)^2 \right] \|\Delta f\|_\infty,$$

an inequality that is given by S. Dragomir in [31, Theorem 3.1].

Example 81. If we let $\mathbb{T} = \mathbb{R}$, then (99) and (100) become

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{2(b-a)} \left[g(t) \int_a^b f(s) ds + f(t) \int_a^b g(s) ds \right] \right| \\ & \leq \left[\frac{1}{4} + \left(\frac{t - \frac{a+b}{2}}{b-a} \right)^2 \right] (b-a) \frac{|g(t)| \|f'\|_\infty + |f(t)| \|g'\|_\infty}{2} \end{aligned}$$

and

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{b-a} \left[g(t) \int_a^b f(s) ds + f(t) \int_a^b g(s) ds \right] \right. \\ & \left. + \left(\frac{1}{b-a} \int_a^b f(s) ds \right) \left(\frac{1}{b-a} \int_a^b g(s) ds \right) \right| \\ & \leq \left(\left[\frac{1}{4} + \left(\frac{t - \frac{a+b}{2}}{b-a} \right)^2 \right] (b-a) \right)^2 \|f'\|_\infty \|g'\|_\infty, \end{aligned}$$

respectively.

Example 82. If we let $\mathbb{T} = \mathbb{Z}$, $a = 0$, and $b = n \in \mathbb{N}$, then (99) and (100) become

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{2n} \left[g(t) \sum_{s=1}^n f(s) + f(t) \sum_{s=1}^n g(s) \right] \right| \\ & \leq \frac{1}{n} \left[\frac{n^2 - 1}{4} + \left(t - \frac{n+1}{2} \right)^2 \right] \frac{|g(t)| \|\Delta f\|_\infty + |f(t)| \|\Delta g\|_\infty}{2} \end{aligned}$$

and

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{n} \left[g(t) \sum_{s=1}^n f(s) + f(t) \sum_{s=1}^n g(s) \right] + \left(\frac{1}{n} \sum_{s=1}^n f(s) \right) \left(\frac{1}{n} \sum_{s=1}^n g(s) \right) \right| \\ & \leq \left(\frac{1}{n} \left[\frac{n^2 - 1}{4} + \left(t - \frac{n+1}{2} \right)^2 \right] \right)^2 \|\Delta f\|_\infty \|\Delta g\|_\infty, \end{aligned}$$

respectively. This is the discrete Ostrowski–Grüss inequality, which can be found in [59, Theorem 2.1].

11.3. OSTROWSKI–GRÜSS INEQUALITIES IN L^p -NORM

Theorem 64. Assume (H). Let $p > 1$ and $q := p/(p-1)$. If $f, g \in C_{\text{rd}}^1([a, b], \mathbb{R})$ such that $f^\Delta, g^\Delta \in L^p([a, b])$, i.e.,

$$\|f^\Delta\|_p := \left(\int_a^b |f^\Delta(s)|^p \Delta s \right)^{\frac{1}{p}} < \infty \quad \text{and} \quad \|g^\Delta\|_p = \left(\int_a^b |g^\Delta(s)|^p \Delta s \right)^{\frac{1}{p}} < \infty,$$

then for all $t \in [a, b]$, we have

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{2m(a, b)} \left[g(t) \int_a^b w(s)f(\sigma(s))\Delta s + f(t) \int_a^b w(s)g(\sigma(s))\Delta s \right] \right| \\ & \leq \left\| \frac{p_w(t, \cdot)}{m(a, b)} \right\|_q \frac{|g(t)| \|f^\Delta\|_p + |f(t)| \|g^\Delta\|_p}{2} \quad (102) \end{aligned}$$

and

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{m(a, b)} \left[g(t) \int_a^b w(s)f(\sigma(s))\Delta s + f(t) \int_a^b w(s)g(\sigma(s))\Delta s \right] \right. \\ & \left. + \left(\frac{1}{m(a, b)} \int_a^b w(s)f(\sigma(s))\Delta s \right) \left(\frac{1}{m(a, b)} \int_a^b w(s)g(\sigma(s))\Delta s \right) \right| \\ & \leq \left\| \frac{p_w(t, \cdot)}{m(a, b)} \right\|_q^2 \|f^\Delta\|_p \|g^\Delta\|_p. \quad (103) \end{aligned}$$

Proof. As in the proof of Theorem 63, we obtain (97) and (98). From (97) and (98), using Hölder's inequality on time scales (see [22, Theorem 6.13]), we obtain (102) and (103), respectively. \square

Corollary 6. *In addition to the assumptions of Theorem 64, let $w(t) = 1$ for all $t \in [a, b]$. Then for all $t \in [a, b]$, we have*

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{2(b-a)} \left[g(t) \int_a^b f(\sigma(s))\Delta s + f(t) \int_a^b g(\sigma(s))\Delta s \right] \right| \\ & \leq \left(\int_a^t \left(\frac{s-a}{b-a} \right)^q \Delta s + \int_t^b \left(\frac{b-s}{b-a} \right)^q \Delta s \right)^{\frac{1}{q}} \frac{|g(t)| \|f^\Delta\|_p + |f(t)| \|g^\Delta\|_p}{2} \end{aligned} \quad (104)$$

and

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{b-a} \left[g(t) \int_a^b f(\sigma(s))\Delta s + f(t) \int_a^b g(\sigma(s))\Delta s \right] \right. \\ & \left. + \left(\frac{1}{b-a} \int_a^b f(\sigma(s))\Delta s \right) \left(\frac{1}{b-a} \int_a^b g(\sigma(s))\Delta s \right) \right| \\ & \leq \left(\int_a^t \left(\frac{s-a}{b-a} \right)^q \Delta s + \int_t^b \left(\frac{b-s}{b-a} \right)^q \Delta s \right)^{\frac{2}{q}} \|f^\Delta\|_p \|g^\Delta\|_p. \end{aligned} \quad (105)$$

Proof. We just have to use Theorem 64. □

Example 83. *If we let $g(t) = 1$ for all $t \in [a, b]$, then (104) becomes*

$$\begin{aligned} & \left| f(t) - \frac{1}{b-a} \int_a^b f(\sigma(s))\Delta s \right| \\ & \leq \left(\int_a^t \left(\frac{s-a}{b-a} \right)^q \Delta s + \int_t^b \left(\frac{b-s}{b-a} \right)^q \Delta s \right)^{\frac{1}{q}} \|f^\Delta\|_p, \end{aligned} \quad (106)$$

which is a new time scales Ostrowski inequality. If $\mathbb{T} = \mathbb{R}$ in (106), then we obtain

$$\left| f(t) - \frac{1}{b-a} \int_a^b f(s)ds \right| \leq \left(\frac{b-a}{q+1} \right)^{\frac{1}{q}} \left[\left(\frac{t-a}{b-a} \right)^{q+1} + \left(\frac{b-t}{b-a} \right)^{q+1} \right]^{\frac{1}{q}} \|f'\|_p,$$

an inequality that is given by S. Dragomir and S. Wang in [34], see also [30, Theorem 2]. If $\mathbb{T} = \mathbb{Z}$, $a = 0$, and $b = n \in \mathbb{N}$ in (106), then we obtain

$$\left| f(t) - \frac{1}{n} \sum_{s=1}^n f(s) \right| \leq \frac{1}{n} \left(\sum_{s=1}^{t-1} s^q + \sum_{s=1}^{n-t} s^q \right)^{\frac{1}{q}} \|\Delta f\|_p,$$

which turns into, e.g., when $p = q = 2$,

$$\left| f(t) - \frac{1}{n} \sum_{s=1}^n f(s) \right| \leq \frac{1}{n} \sqrt{\frac{(t-1)t(2t-1) + (n-t)(n-t+1)(2n-2t+1)}{6}} \|\Delta f\|_2.$$

Example 84. If we let $\mathbb{T} = \mathbb{R}$, then (102) and (103) become

$$\begin{aligned} \left| f(t)g(t) - \frac{1}{2m(a,b)} \left[g(t) \int_a^b w(s)f(s)ds + f(t) \int_a^b w(s)g(s)ds \right] \right| \\ \leq \left\| \frac{p_w(t, \cdot)}{m(a,b)} \right\|_q \frac{|g(t)| \|f'\|_p + |f(t)| \|g'\|_p}{2} \end{aligned}$$

and

$$\begin{aligned} \left| f(t)g(t) - \frac{1}{m(a,b)} \left[g(t) \int_a^b w(s)f(s)ds + f(t) \int_a^b w(s)g(s)ds \right] \right. \\ \left. + \left(\frac{1}{m(a,b)} \int_a^b w(s)f(s)ds \right) \left(\frac{1}{m(a,b)} \int_a^b w(s)g(s)ds \right) \right| \\ \leq \left\| \frac{p_w(t, \cdot)}{m(a,b)} \right\|_q^2 \|f'\|_p \|g'\|_p, \end{aligned}$$

respectively, and (104) and (105) become (89) and (90), respectively, in Theorem 62, and by choosing $t = (a+b)/2$, we obtain the inequalities given in [18, Remark 2].

Example 85. If we let $\mathbb{T} = \mathbb{Z}$, $a = 0$, and $b = n \in \mathbb{N}$, then (104) and (105) become

$$\begin{aligned} \left| f(t)g(t) - \frac{1}{2n} \left[g(t) \sum_{s=1}^n f(s) + f(t) \sum_{s=1}^n g(s) \right] \right| \\ \leq \frac{1}{n} \left(\sum_{s=1}^{t-1} s^q + \sum_{s=1}^{n-t} s^q \right)^{\frac{1}{q}} \frac{|g(t)| \|\Delta f\|_p + |f(t)| \|\Delta g\|_p}{2} \end{aligned}$$

and

$$\begin{aligned} \left| f(t)g(t) - \frac{1}{n} \left[g(t) \sum_{s=1}^n f(s) + f(t) \sum_{s=1}^n g(s) \right] + \left(\frac{1}{n} \sum_{s=1}^n f(s) \right) \left(\frac{1}{n} \sum_{s=1}^n g(s) \right) \right| \\ \leq \left(\frac{1}{n} \left(\sum_{s=1}^{t-1} s^q + \sum_{s=1}^{n-t} s^q \right)^{\frac{1}{q}} \right)^2 \|\Delta f\|_p \|\Delta g\|_p, \end{aligned}$$

respectively, which are new discrete Ostrowski–Grüss inequalities.

12. OSTROWSKI AND GRÜSS TYPE INEQUALITIES

This chapter contains a collaborative work with Elvan Akin-Bohner and Martin Bohner and appeared in 2011 in *Nonlinear Dynamics and Systems Theory* with the title “Time Scales Ostrowski and Grüss Type Inequalities involving Three Functions”, see [5]. We present time scales versions of Ostrowski and Grüss type inequalities containing three functions. We assume that the second derivatives of these functions are bounded. Our results are new also for the discrete case.

12.1. INTRODUCTION

Motivated by a recent paper by B. G. Pachpatte [61], our purpose is to obtain time scales versions of some Ostrowski and Grüss type inequalities including three functions, whose second derivatives are bounded. In detail, we will prove time scales analogues of the following three theorems presented in [61].

Theorem 65 (See [61, Theorem 1]). *Let $f, g, h : [a, b] \rightarrow \mathbb{R}$ be twice differentiable functions on (a, b) such that $f'', g'', h'' : (a, b) \rightarrow \mathbb{R}$ are bounded, i.e.,*

$$\|f''\|_\infty := \sup_{t \in (a, b)} |f''(t)| < \infty, \quad \|g''\|_\infty < \infty, \quad \|h''\|_\infty < \infty.$$

Moreover, let

$$A[f, g, h] := gh \int_a^b f(s) ds + fh \int_a^b g(s) ds + fg \int_a^b h(s) ds$$

and

$$B[f, g, h] := |gh| \|f''\|_\infty + |fh| \|g''\|_\infty + |fg| \|h''\|_\infty.$$

Then, for all $t \in [a, b]$, we have

$$\begin{aligned} \left| f(t)g(t)h(t) - \frac{1}{3(b-a)}A[f, g, h](t) - \frac{1}{3} \left(t - \frac{a+b}{2} \right) (fgh)'(t) \right| \\ \leq \frac{1}{6} \left\{ \left(t - \frac{a+b}{2} \right)^2 + \frac{(b-a)^2}{12} \right\} B[f, g, h](t). \end{aligned}$$

Theorem 66 (See [61, Theorem 2]). *In addition to the notation and assumptions of Theorem 65, let*

$$L[f, g, h] := gh \frac{f(a) + f(b)}{2} + fh \frac{g(a) + g(b)}{2} + fg \frac{h(a) + h(b)}{2}.$$

Then, for all $t \in [a, b]$, we have

$$\begin{aligned} & \left| f(t)g(t)h(t) - \frac{2}{3(b-a)}A[f, g, h](t) - \frac{1}{3} \left(t - \frac{a+b}{2} \right) (fgh)'(t) + \frac{1}{3}L[f, g, h](t) \right| \\ & \leq \frac{1}{3(b-a)}B[f, g, h](t) \int_a^b \left| p(t, s) \left(s - \frac{a+b}{2} \right) \right| ds, \end{aligned}$$

where $p(t, s) = s - a$ for $a \leq s < t$ and $p(t, s) = s - b$ for $t \leq s \leq b$.

Theorem 67 (See [61, Theorem 3]). *In addition to the notation and assumptions of Theorem 65, let*

$$M[f, g, h] := gh \frac{f(b) - f(a)}{b-a} + fh \frac{g(b) - g(a)}{b-a} + fg \frac{h(b) - h(a)}{b-a}.$$

Then, for all $t \in [a, b]$, we have

$$\begin{aligned} & \left| f(t)g(t)h(t) - \frac{1}{3(b-a)}A[f, g, h](t) - \frac{1}{3} \left(t - \frac{a+b}{2} \right) M[f, g, h](t) \right| \\ & \leq \frac{1}{3(b-a)^2}B[f, g, h](t) \int_a^b \int_a^b |p(t, \tau)p(\tau, s)| dsd\tau, \end{aligned}$$

where p is defined as in Theorem 66.

Our time scales versions of Theorems 65–67 will contain Theorems 65–67 as special cases when the time scale is equal to the set of all real numbers, and they will yield new discrete inequalities when the time scale is equal to the set of all integer numbers. Special cases of our results are contained in [19, 18, 20, 21, 65, 46, 54] for the general time scales case, in [57, 26, 32, 33] for the continuous case and in [58, 4] for the discrete case. One can also use our results for any other arbitrary time scale to obtain new inequalities, e.g., for the quantum case.

The set up of this chapter is as follows. Section 12.2 contains some auxiliary results as well as the assumptions and notation used in this paper. Finally, in Sections 12.3–12.5, we prove time scales analogues of Theorems 65–67. Each result is followed by several

examples and remarks. We would like to point out here that our results are new also for the discrete case.

12.2. AUXILIARY RESULTS AND ASSUMPTIONS

We start with the following auxiliary result.

Lemma 12. *The time scales monomials satisfy the following formulas:*

$$g_2(t, a) - g_2(t, b) = g_2(b, a) + (t - b)(b - a), \quad (107)$$

$$g_2(a, b) + g_2(b, a) = (b - a)^2, \quad (108)$$

$$g_3(t, a) - g_3(t, b) = g_3(b, a) + (t - b)g_2(b, a) + (b - a)g_2(t, b). \quad (109)$$

Proof. The function F defined by $F(t) := g_2(t, a) - g_2(t, b) - g_2(b, a) - (t - b)(b - a)$ satisfies $F^\Delta(t) = \sigma(t) - a - (\sigma(t) - b) - (b - a) = 0$ and $F(b) = 0$. Hence $F = 0$ and so (107) holds. Next, (108) follows by letting $t = a$ in (107). Moreover, the function G defined by $G(t) := g_3(t, a) - g_3(t, b) - g_3(b, a) - (t - b)g_2(b, a) - (b - a)g_2(t, b)$ satisfies $G^\Delta(t) = g_2(\sigma(t), a) - g_2(\sigma(t), b) - g_2(b, a) - (b - a)(\sigma(t) - b) = F(\sigma(t)) = 0$ and $G(b) = 0$. Hence $G = 0$ and so (109) holds. \square

Throughout this chapter we assume that \mathbb{T} is a time scale and that $a, b \in \mathbb{T}$ such that $a < b$. Moreover, when writing $[a, b]$, we mean the time scales interval $[a, b] \cap \mathbb{T}$. The following two Montgomery-type results are used in the proofs of our three main results.

Theorem 68. *Suppose $f \in C_{\text{rd}}^1(\mathbb{T}, \mathbb{R})$. Let $t \in [a, b]$ and $u_1, u_2 \in C_{\text{rd}}^1(\mathbb{T}, \mathbb{R})$. If*

$$u(\sigma(s)) = \begin{cases} u_1(\sigma(s)) & \text{for } a \leq s < t \\ u_2(\sigma(s)) & \text{for } t \leq s \leq b, \end{cases} \quad (110)$$

then

$$\begin{aligned} \int_a^b u(\sigma(s))f^\Delta(s)\Delta s &= (u_1(t) - u_2(t))f(t) - u_1(a)f(a) + u_2(b)f(b) \\ &\quad - \int_a^t u_1^\Delta(s)f(s)\Delta s - \int_t^b u_2^\Delta(s)f(s)\Delta s. \end{aligned} \quad (111)$$

Proof. We split the integral into two parts, each of which is evaluated by applying the integration of parts formula, i.e.,

$$\begin{aligned} \int_a^b u(\sigma(s))f^\Delta(s)\Delta s &= \int_a^t u_1(\sigma(s))f^\Delta(s)\Delta s + \int_t^b u_2(\sigma(s))f^\Delta(s)\Delta s \\ &= u_1(t)f(t) - u_1(a)f(a) - \int_a^t u_1^\Delta(s)f(s)\Delta s \\ &\quad + u_2(b)f(b) - u_2(t)f(t) - \int_t^b u_2^\Delta(s)f(s)\Delta s, \end{aligned}$$

from which (111) follows. \square

Theorem 69. *Suppose $f \in C_{\text{rd}}^2(\mathbb{T}, \mathbb{R})$. Let $t \in [a, b]$ and $u_i, v_i \in C_{\text{rd}}^1(\mathbb{T}, \mathbb{R})$ be such that $u_i^\Delta(s) = v_i(\sigma(s))$ for all $s \in [a, b]$, where $i \in \{1, 2\}$. If u satisfies (110), then*

$$\begin{aligned} \int_a^b u(\sigma(s))f^{\Delta\Delta}(s)\Delta s &= (u_1(t) - u_2(t))f^\Delta(t) - (v_1(t) - v_2(t))f(t) \\ &\quad - u_1(a)f^\Delta(a) + v_1(a)f(a) + u_2(b)f^\Delta(b) - v_2(b)f(b) \quad (112) \\ &\quad + \int_a^t v_1^\Delta(s)f(s)\Delta s + \int_t^b v_2^\Delta(s)f(s)\Delta s. \end{aligned}$$

Proof. Using (111) with f^Δ replaced by $f^{\Delta\Delta}$ and subsequently applying integration by parts twice, we obtain

$$\begin{aligned} \int_a^b u(\sigma(s))f^{\Delta\Delta}(s)\Delta s &= (u_1(t) - u_2(t))f^\Delta(t) - u_1(a)f^\Delta(a) + u_2(b)f^\Delta(b) \\ &\quad - \int_a^t u_1^\Delta(s)f^\Delta(s)\Delta s - \int_t^b u_2^\Delta(s)f^\Delta(s)\Delta s \\ &= (u_1(t) - u_2(t))f^\Delta(t) - u_1(a)f^\Delta(a) + u_2(b)f^\Delta(b) \\ &\quad - \int_a^t v_1(\sigma(s))f^\Delta(s)\Delta s - \int_t^b v_2(\sigma(s))f^\Delta(s)\Delta s \\ &= (u_1(t) - u_2(t))f^\Delta(t) - u_1(a)f^\Delta(a) + u_2(b)f^\Delta(b) \\ &\quad - \left\{ v_1(t)f(t) - v_1(a)f(a) - \int_a^t v_1^\Delta(s)f(s)\Delta s \right\} \\ &\quad - \left\{ v_2(b)f(b) - v_2(t)f(t) - \int_t^b v_2^\Delta(s)f(s)\Delta s \right\}, \end{aligned}$$

from which (112) follows. \square

Assumption (H). *For the remaining three sections of this chapter, we assume that \mathbb{T} is a time scale and that $a, b \in \mathbb{T}$ such that $a < b$. We assume that $f, g, h \in C_{\text{rd}}^2(\mathbb{T}, \mathbb{R})$ are*

such that

$$\|f^{\Delta\Delta}\|_{\infty} := \sup_{t \in (a,b)} |f^{\Delta\Delta}(t)| < \infty, \quad \|g^{\Delta\Delta}\|_{\infty} < \infty, \quad \|h^{\Delta\Delta}\|_{\infty} < \infty \quad (113)$$

and define

$$\begin{aligned} A[f, g, h] &:= gh \int_a^b f(s)\Delta s + fh \int_a^b g(s)\Delta s + fg \int_a^b h(s)\Delta s, \\ B[f, g, h] &:= |gh| \|f^{\Delta\Delta}\|_{\infty} + |fh| \|g^{\Delta\Delta}\|_{\infty} + |fg| \|h^{\Delta\Delta}\|_{\infty}, \\ C[f, g, h] &:= ghf^{\Delta} + fhg^{\Delta} + fgh^{\Delta}, \\ D[f, g, h] &:= \left(\int_a^b g(s)h(s)\Delta s \right) \left(\int_a^b f(s)\Delta s \right) \\ &\quad + \left(\int_a^b f(s)h(s)\Delta s \right) \left(\int_a^b g(s)\Delta s \right) \\ &\quad + \left(\int_a^b f(s)g(s)\Delta s \right) \left(\int_a^b h(s)\Delta s \right), \\ L[f, g, h] &:= gh \frac{g_2(b, a)f(a) + h_2(b, a)f(b)}{(b-a)^2} + fh \frac{g_2(b, a)g(a) + h_2(b, a)g(b)}{(b-a)^2} \\ &\quad + fg \frac{g_2(b, a)h(a) + h_2(b, a)h(b)}{(b-a)^2}, \\ M[f, g, h] &:= gh \frac{f(b) - f(a)}{b-a} + fh \frac{g(b) - g(a)}{b-a} + fg \frac{h(b) - h(a)}{b-a}. \end{aligned}$$

12.3. PACHPATTE'S FIRST THEOREM ON TIME SCALES

Theorem 70. *Assume (H). Then, for all $t \in [a, b]$, we have*

$$\begin{aligned} &\left| f(t)g(t)h(t) - \frac{1}{3(b-a)}A[f, g, h](t) - \frac{1}{3} \left(t - b + \frac{g_2(b, a)}{b-a} \right) C[f, g, h](t) \right| \\ &\leq \frac{1}{3} \left(h_2(b, t) + (t-b) \frac{g_2(b, a)}{b-a} + \frac{g_3(b, a)}{b-a} \right) B[f, g, h](t) \quad (114) \end{aligned}$$

and

$$\begin{aligned} &\left| \frac{1}{b-a} \int_a^b f(t)g(t)h(t)\Delta t - \frac{1}{3(b-a)^2}D[f, g, h] \right. \\ &\quad \left. - \frac{1}{3(b-a)} \int_a^b \left(t - b + \frac{g_2(b, a)}{b-a} \right) C[f, g, h](t)\Delta t \right| \\ &\leq \frac{1}{3(b-a)} \int_a^b \left(h_2(b, t) + (t-b) \frac{g_2(b, a)}{b-a} + \frac{g_3(b, a)}{b-a} \right) B[f, g, h](t)\Delta t. \quad (115) \end{aligned}$$

Proof. Fix $t \in [a, b]$ and define u by (110), where

$$u_1(s) = g_2(s, a), \quad u_2(s) = h_2(b, s).$$

With the notation as in Theorem 69, we have

$$v_1(s) = s - a, \quad v_2(s) = s - b, \quad v_1^\Delta(s) = v_2^\Delta(s) = 1$$

and $u_1(a) = v_1(a) = u_2(b) = v_2(b) = 0$. Moreover, we have

$$u_1(t) - u_2(t) \stackrel{(107)}{=} (t - b)(b - a) + g_2(b, a), \quad v_1(t) - v_2(t) = b - a.$$

By (112), we therefore obtain

$$\int_a^b u(\sigma(s))f^{\Delta\Delta}(s)\Delta s = ((t - b)(b - a) + g_2(b, a))f^\Delta(t) - (b - a)f(t) + \int_a^b f(s)\Delta s$$

and thus

$$f(t) = \frac{1}{b - a} \int_a^b f(s)\Delta s + \left(t - b + \frac{g_2(b, a)}{b - a} \right) f^\Delta(t) - \frac{1}{b - a} \int_a^b u(\sigma(s))f^{\Delta\Delta}(s)\Delta s. \quad (116)$$

Similarly, we get

$$g(t) = \frac{1}{b - a} \int_a^b g(s)\Delta s + \left(t - b + \frac{g_2(b, a)}{b - a} \right) g^\Delta(t) - \frac{1}{b - a} \int_a^b u(\sigma(s))g^{\Delta\Delta}(s)\Delta s \quad (117)$$

and

$$h(t) = \frac{1}{b - a} \int_a^b h(s)\Delta s + \left(t - b + \frac{g_2(b, a)}{b - a} \right) h^\Delta(t) - \frac{1}{b - a} \int_a^b u(\sigma(s))h^{\Delta\Delta}(s)\Delta s. \quad (118)$$

Multiplying (116), (117) and (118) by $g(t)h(t)$, $f(t)h(t)$ and $f(t)g(t)$, respectively, adding the resulting identities and dividing by three, we have

$$\begin{aligned} f(t)g(t)h(t) - \frac{1}{3(b - a)}A[f, g, h](t) - \frac{1}{3} \left(t - b + \frac{g_2(b, a)}{b - a} \right) C[f, g, h](t) \\ = -\frac{1}{3(b - a)} \int_a^b u(\sigma(s))\tilde{B}[f, g, h](t, s)\Delta s, \end{aligned} \quad (119)$$

where

$$\begin{cases} \tilde{B}[f, g, h](t, s) := g(t)h(t)f^{\Delta\Delta}(s) + f(t)h(t)g^{\Delta\Delta}(s) + f(t)g(t)h^{\Delta\Delta}(s) \\ \text{so that } \left| \tilde{B}[f, g, h](t, s) \right| \leq B[f, g, h](t). \end{cases} \quad (120)$$

By taking absolute values in (119) and using (113) and

$$\begin{aligned} \int_a^b |u(\sigma(s))| \Delta s &= \int_a^t g_2(\sigma(s), a) \Delta s + \int_t^b h_2(b, \sigma(s)) \Delta s \\ &= g_3(t, a) - g_3(t, b) \\ &\stackrel{(109)}{=} g_3(b, a) + (t - b)g_2(b, a) + (b - a)h_2(b, t), \end{aligned} \quad (121)$$

we obtain (114). Integrating (119) with respect to t from a to b , dividing by $b - a$, noting that

$$\int_a^b A[f, g, h](s) \Delta s = D[f, g, h], \quad (122)$$

taking absolute values and using (113) and (121), we obtain (115). \square

Example 86. If we let $\mathbb{T} = \mathbb{R}$ in Theorem 70, then, since $C[f, g, h] = (fgh)'$,

$$b - \frac{g_2(b, a)}{b - a} = b - \frac{(b - a)^2}{2(b - a)} = b - \frac{b - a}{2} = \frac{a + b}{2}$$

and

$$\begin{aligned} h_2(b, t) + (t - b) \frac{g_2(b, a)}{b - a} + \frac{g_3(b, a)}{b - a} &= \frac{1}{2} \left\{ (t - b)^2 + (t - b)(b - a) + \frac{(b - a)^2}{3} \right\} \\ &= \frac{1}{2} \left\{ \left(t - b + \frac{b - a}{2} \right)^2 - \frac{(b - a)^2}{4} + \frac{(b - a)^2}{3} \right\} \\ &= \frac{1}{2} \left\{ \left(t - \frac{a + b}{2} \right)^2 + \frac{(b - a)^2}{12} \right\}, \end{aligned}$$

we obtain [61, Theorem 1], in particular, Theorem 65.

Example 87. If we let $\mathbb{T} = \mathbb{Z}$ and $a = 0$, $b = n \in \mathbb{N}$ in Theorem 70, then, since

$$b - \frac{g_2(b, a)}{b - a} = b - \frac{(b - a)(b - a + 1)}{2(b - a)} = b - \frac{b - a + 1}{2} = \frac{a + b - 1}{2} = \frac{n - 1}{2}$$

and

$$\begin{aligned}
& h_2(b, t) + (t - b) \frac{g_2(b, a)}{b - a} + \frac{g_3(b, a)}{b - a} \\
&= \frac{1}{2} \left\{ (b - t)(b - t - 1) + (t - b)(b - a + 1) + \frac{(b - a + 1)(b - a + 2)}{3} \right\} \\
&= \frac{1}{2} \left\{ \left(t - b + \frac{b - a + 2}{2} \right)^2 - \frac{(b - a + 2)^2}{4} + \frac{(b - a + 1)(b - a + 2)}{3} \right\} \\
&= \frac{1}{2} \left\{ \left(t + 1 - \frac{a + b}{2} \right)^2 + \frac{(b - a + 2)(b - a - 2)}{12} \right\} \\
&= \frac{1}{2} \left\{ \left(t + 1 - \frac{n}{2} \right)^2 + \frac{n^2 - 4}{12} \right\},
\end{aligned}$$

we obtain

$$\begin{aligned}
& \left| f(t)g(t)h(t) - \frac{1}{3n}A[f, g, h](t) - \frac{1}{3} \left(t - \frac{n - 1}{2} \right) C[f, g, h](t) \right| \\
& \leq \frac{1}{6} \left\{ \left(t + 1 - \frac{n}{2} \right)^2 + \frac{n^2 - 4}{12} \right\} B[f, g, h](t)
\end{aligned}$$

and

$$\begin{aligned}
& \left| \frac{1}{n} \sum_{t=0}^{n-1} f(t)g(t)h(t) - \frac{1}{3n^2}D[f, g, h] - \frac{1}{3n} \sum_{t=0}^{n-1} \left(t - \frac{n - 1}{2} \right) C[f, g, h](t) \right| \\
& \leq \frac{1}{6n} \sum_{t=0}^{n-1} \left\{ \left(t + 1 - \frac{n}{2} \right)^2 + \frac{n^2 - 4}{12} \right\} B[f, g, h](t),
\end{aligned}$$

where

$$\begin{aligned}
A[f, g, h] &= gh \sum_{s=0}^{n-1} f(s) + fh \sum_{s=0}^{n-1} g(s) + fg \sum_{s=0}^{n-1} h(s), \\
B[f, g, h] &= |gh| \max_{1 \leq s \leq n-1} |\Delta^2 f(s)| + |fh| \max_{1 \leq s \leq n-1} |\Delta^2 g(s)| \\
& \quad + |fg| \max_{1 \leq s \leq n-1} |\Delta^2 h(s)|, \\
C[f, g, h] &= gh\Delta f + fh\Delta g + fg\Delta h, \\
D[f, g, h] &= \left(\sum_{s=0}^{n-1} g(s)h(s) \right) \left(\sum_{s=0}^{n-1} f(s) \right) + \left(\sum_{s=0}^{n-1} f(s)h(s) \right) \left(\sum_{s=0}^{n-1} g(s) \right) \\
& \quad + \left(\sum_{s=0}^{n-1} f(s)g(s) \right) \left(\sum_{s=0}^{n-1} h(s) \right).
\end{aligned}$$

Furthermore, note that these inequalities are new discrete Ostrowski–Grüss type inequalities.

Remark 36. If we let $h(t) \equiv 1$ in Theorem 70, then (114) becomes

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{2(b-a)} \left\{ g(t) \int_a^b f(s)\Delta s + f(t) \int_a^b g(s)\Delta s \right\} \right. \\ & \left. - \frac{1}{2} \left(t - b + \frac{g_2(b,a)}{b-a} \right) \{ g(t)f^\Delta(t) + f(t)g^\Delta(t) \} \right| \\ & \leq \frac{1}{2} \left(h_2(b,t) + (t-b) \frac{g_2(b,a)}{b-a} + \frac{g_3(b,a)}{b-a} \right) \{ |g(t)| \|f^{\Delta\Delta}\|_\infty + |f(t)| \|g^{\Delta\Delta}\|_\infty \} \end{aligned}$$

and (115) turns into

$$\begin{aligned} & \left| \frac{1}{b-a} \int_a^b f(t)g(t)\Delta t - \frac{1}{(b-a)^2} \left(\int_a^b f(t)\Delta t \right) \left(\int_a^b g(t)\Delta t \right) \right. \\ & \left. - \frac{1}{2(b-a)} \int_a^b \left(t - b + \frac{g_2(b,a)}{b-a} \right) \{ g(t)f^\Delta(t) + f(t)g^\Delta(t) \} \Delta t \right| \\ & \leq \frac{1}{2(b-a)} \int_a^b \left(h_2(b,t) + (t-b) \frac{g_2(b,a)}{b-a} + \frac{g_3(b,a)}{b-a} \right) \\ & \qquad \qquad \qquad \{ |g(t)| \|f^{\Delta\Delta}\|_\infty + |f(t)| \|g^{\Delta\Delta}\|_\infty \} \Delta t. \end{aligned}$$

If, moreover, we let $g(t) \equiv 1$, then (114) becomes

$$\begin{aligned} & \left| f(t) - \frac{1}{b-a} \int_a^b f(s)\Delta s - \left(t - b + \frac{g_2(b,a)}{b-a} \right) f^\Delta(t) \right| \\ & \leq \left(h_2(b,t) + (t-b) \frac{g_2(b,a)}{b-a} + \frac{g_3(b,a)}{b-a} \right) \|f^{\Delta\Delta}\|_\infty. \end{aligned}$$

From these inequalities, special cases such as discrete inequalities can be obtained.

12.4. PACHPATTE'S SECOND THEOREM ON TIME SCALES

Theorem 71. Assume (H). Then, for all $t \in [a, b]$, we have

$$\begin{aligned} & \left| f(t)g(t)h(t) - \frac{2}{3(b-a)} A[f, g, h](t) + \frac{1}{3} L[f, g, h](t) \right. \\ & \left. - \frac{1}{3} \left(t - b + \frac{g_2(b,a)}{b-a} \right) C[f, g, h](t) \right| \leq \frac{1}{3(b-a)} B[f, g, h](t) I(t) \quad (123) \end{aligned}$$

and

$$\begin{aligned} & \left| \frac{1}{b-a} \int_a^b f(t)g(t)h(t)\Delta t - \frac{2}{3(b-a)^2} D[f, g, h] + \frac{1}{3(b-a)} \int_a^b L[f, g, h](t)\Delta t \right. \\ & \left. - \frac{1}{3(b-a)} \int_a^b \left(t - b + \frac{g_2(b, a)}{b-a} \right) C[f, g, h](t)\Delta t \right| \\ & \leq \frac{1}{3(b-a)^2} \int_a^b B[f, g, h](t)I(t)\Delta t, \quad (124) \end{aligned}$$

where

$$\begin{aligned} I(t) & := \frac{1}{b-a} \int_a^t |2(b-a)g_2(\sigma(s), a) - (\sigma(s) - a)g_2(b, a)| \Delta s \\ & \quad + \frac{1}{b-a} \int_t^b |2(b-a)h_2(b, \sigma(s)) - (b - \sigma(s))h_2(b, a)| \Delta s. \end{aligned}$$

Proof. Fix $t \in [a, b]$ and define u by (110), where

$$u_1(s) = 2(b-a)g_2(s, a) - (s-a)g_2(b, a), \quad u_2(s) = 2(b-a)h_2(b, s) - (b-s)h_2(b, a).$$

With the notation as in Theorem 69, we have

$$\begin{aligned} v_1(s) & = 2(b-a)(s-a) - g_2(b, a), \quad v_2(s) = 2(b-a)(s-b) + h_2(b, a), \\ v_1^\Delta(s) & = v_2^\Delta(s) = 2(b-a) \end{aligned}$$

and $u_1(a) = u_2(b) = 0$, $v_1(a) = -g_2(b, a)$, $v_2(b) = h_2(b, a)$. Moreover, we have

$$\begin{aligned} u_1(t) - u_2(t) & = 2(b-a)(g_2(t, a) - h_2(b, t)) \\ & \quad - (t-a)g_2(b, a) + (b-t)h_2(b, a) \\ & \stackrel{(107), (108)}{=} 2(b-a)(g_2(b, a) + (t-b)(b-a)) \\ & \quad - (t-a)g_2(b, a) + (b-t)((b-a)^2 - g_2(b, a)) \\ & \stackrel{(108)}{=} (b-a)g_2(b, a) + (t-b)(b-a)^2, \\ v_1(t) - v_2(t) & = 2(b-a)^2 - g_2(b, a) - h_2(b, a) \\ & \stackrel{(108)}{=} 2(b-a)^2 - (b-a)^2 = (b-a)^2. \end{aligned}$$

By (112), we therefore obtain

$$\begin{aligned} \int_a^b u(\sigma(s))f^{\Delta\Delta}(s)\Delta s &= (b-a)(g_2(b,a) + (t-b)(b-a))f^\Delta(t) \\ &\quad - (b-a)^2f(t) - g_2(b,a)f(a) - h_2(b,a)f(b) + 2(b-a)\int_a^b f(s)\Delta s \end{aligned}$$

and thus

$$\begin{aligned} f(t) &= \frac{2}{b-a}\int_a^b f(s)\Delta s - \frac{g_2(b,a)f(a) + h_2(b,a)f(b)}{(b-a)^2} \\ &\quad + \left(t-b + \frac{g_2(b,a)}{b-a}\right)f^\Delta(t) - \frac{1}{(b-a)^2}\int_a^b u(\sigma(s))f^{\Delta\Delta}(s)\Delta s. \end{aligned} \quad (125)$$

Similarly, we get

$$\begin{aligned} g(t) &= \frac{2}{b-a}\int_a^b g(s)\Delta s - \frac{g_2(b,a)g(a) + h_2(b,a)g(b)}{(b-a)^2} \\ &\quad + \left(t-b + \frac{g_2(b,a)}{b-a}\right)g^\Delta(t) - \frac{1}{(b-a)^2}\int_a^b u(\sigma(s))g^{\Delta\Delta}(s)\Delta s \end{aligned} \quad (126)$$

and

$$\begin{aligned} h(t) &= \frac{2}{b-a}\int_a^b h(s)\Delta s - \frac{g_2(b,a)h(a) + h_2(b,a)h(b)}{(b-a)^2} \\ &\quad + \left(t-b + \frac{g_2(b,a)}{b-a}\right)h^\Delta(t) - \frac{1}{(b-a)^2}\int_a^b u(\sigma(s))h^{\Delta\Delta}(s)\Delta s. \end{aligned} \quad (127)$$

Multiplying (125), (126) and (127) by $g(t)h(t)$, $f(t)h(t)$ and $f(t)g(t)$, respectively, adding the resulting identities and dividing by three, we have

$$\begin{aligned} f(t)g(t)h(t) &- \frac{2}{3(b-a)}A[f,g,h](t) + \frac{1}{3}L[f,g,h](t) \\ &- \frac{1}{3}\left(t-b + \frac{g_2(b,a)}{b-a}\right)C[f,g,h](t) = -\frac{1}{3(b-a)^2}\int_a^b u(\sigma(s))\tilde{B}[f,g,h](t,s)\Delta s \end{aligned} \quad (128)$$

with \tilde{B} as in (120). By taking absolute values in (128) and using (113) and

$$\frac{1}{b-a}\int_a^b |u(\sigma(s))|\Delta s = I(t), \quad (129)$$

we obtain (123). Integrating (128) with respect to t from a to b , dividing by $b-a$, noting (122), taking absolute values and using (113) and (129), we obtain (124). \square

Example 88. If we let $\mathbb{T} = \mathbb{R}$ in Theorem 71, then, since $C[f, g, h] = (fgh)'$,

$$b - \frac{g_2(b, a)}{b - a} = \frac{a + b}{2}$$

and (with p as defined in Theorem 66)

$$\begin{aligned} I(t) &= \frac{1}{b - a} \int_a^t \left| (b - a)(s - a)^2 - (s - a) \frac{(b - a)^2}{2} \right| ds \\ &\quad + \frac{1}{b - a} \int_t^b \left| (b - a)(s - b)^2 - (b - s) \frac{(b - a)^2}{2} \right| ds \\ &= \int_a^t \left| (s - a) \left(s - \frac{a + b}{2} \right) \right| ds + \int_t^b \left| (s - b) \left(s - \frac{a + b}{2} \right) \right| ds \\ &= \int_a^b \left| p(t, s) \left(s - \frac{a + b}{2} \right) \right| ds, \end{aligned}$$

we obtain [61, Theorem 2], in particular, Theorem 66.

Example 89. If we let $\mathbb{T} = \mathbb{Z}$ and $a = 0$, $b = n \in \mathbb{N}$ in Theorem 71, then, since

$$b - \frac{g_2(b, a)}{b - a} = \frac{n - 1}{2}$$

and

$$\begin{aligned} I(t) &= \frac{1}{b - a} \sum_{s=a}^{t-1} \left| (b - a)(s + 1 - a)(s + 2 - a) - (s + 1 - a) \frac{(b - a)(b - a + 1)}{2} \right| \\ &\quad + \frac{1}{b - a} \sum_{s=t}^{b-1} \left| (b - a)(b - s - 1)(b - s - 2) - (b - s - 1) \frac{(b - a)(b - a - 1)}{2} \right| \\ &= \sum_{s=a}^{t-1} \left| (s + 1 - a) \left(s + 1 - \frac{a + b - 1}{2} \right) \right| \\ &\quad + \sum_{s=t}^{b-1} \left| (s + 1 - b) \left(s + 1 - \frac{a + b - 1}{2} \right) \right| \\ &= \sum_{s=0}^{t-1} \left| (s + 1) \left(s + 1 - \frac{n - 1}{2} \right) \right| + \sum_{s=t}^{n-1} \left| (s + 1 - n) \left(s + 1 - \frac{n - 1}{2} \right) \right|, \end{aligned}$$

we have

$$\left| f(t)g(t)h(t) - \frac{2}{3n}A[f, g, h](t) + \frac{1}{3}L[f, g, h](t) - \frac{1}{3} \left(t - \frac{n - 1}{2} \right) C[f, g, h](t) \right|$$

$$\leq \frac{1}{3n} B[f, g, h](t) \left\{ \sum_{s=1}^t s \left| s - \frac{n-1}{2} \right| + \sum_{s=t+1}^n (n-s) \left| s - \frac{n-1}{2} \right| \right\}$$

and

$$\begin{aligned} & \left| \frac{1}{n} \sum_{t=0}^{n-1} f(t)g(t)h(t) - \frac{2}{3n^2} D[f, g, h] \right. \\ & \left. + \frac{1}{3n} \sum_{t=0}^{n-1} L[f, g, h](t) - \frac{1}{3n} \sum_{t=0}^{n-1} \left(t - \frac{n-1}{2} \right) C[f, g, h](t) \right| \\ & \leq \frac{1}{3n^2} \sum_{t=0}^{n-1} B[f, g, h](t) \left\{ \sum_{s=1}^t s \left| s - \frac{n-1}{2} \right| + \sum_{s=t+1}^n (n-s) \left| s - \frac{n-1}{2} \right| \right\}, \end{aligned}$$

where in addition to A, B, C, D defined in Example 87,

$$\begin{aligned} L[f, g, h] &= gh \frac{(n+1)f(a) + (n-1)f(b)}{2n} + fh \frac{(n+1)g(a) + (n-1)g(b)}{2n} \\ &+ fg \frac{(n+1)h(a) + (n-1)h(b)}{2n}. \end{aligned}$$

These inequalities are new discrete Ostrowski–Grüss type inequalities.

Remark 37. If we let $h(t) \equiv 1$ in Theorem 71, then (123) becomes

$$\begin{aligned} & \left| f(t)g(t) - \frac{1}{b-a} \left\{ g(t) \int_a^b f(s) \Delta s + f(t) \int_a^b g(s) \Delta s \right\} \right. \\ & + g(t) \frac{g_2(b, a)f(a) + h_2(b, a)f(b)}{2(b-a)^2} + f(t) \frac{g_2(b, a)g(a) + h_2(b, a)g(b)}{2(b-a)^2} \\ & \left. - \frac{1}{2} \left(t - b + \frac{g_2(b, a)}{b-a} \right) \{ g(t)f^\Delta(t) + f(t)g^\Delta(t) \} \right| \\ & \leq \frac{1}{2(b-a)} \{ |g(t)| \|f^{\Delta\Delta}\|_\infty + |f(t)| \|g^{\Delta\Delta}\|_\infty \} I(t), \end{aligned}$$

(observe (108) when calculating L) and (124) turns into

$$\begin{aligned} & \left| \frac{1}{b-a} \int_a^b f(t)g(t) \Delta t - \frac{2}{(b-a)^2} \left(\int_a^b f(t) \Delta t \right) \left(\int_a^b g(t) \Delta t \right) \right. \\ & + \frac{1}{b-a} \int_a^b \left\{ g(t) \frac{g_2(b, a)f(a) + h_2(b, a)f(b)}{2(b-a)^2} + f(t) \frac{g_2(b, a)g(a) + h_2(b, a)g(b)}{2(b-a)^2} \right\} \Delta t \\ & \left. - \frac{1}{2(b-a)} \int_a^b \left(t - b + \frac{g_2(b, a)}{b-a} \right) \{ g(t)f^\Delta(t) + f(t)g^\Delta(t) \} \Delta t \right| \\ & \leq \frac{1}{2(b-a)^2} \int_a^b \{ |g(t)| \|f^{\Delta\Delta}\|_\infty + |f(t)| \|g^{\Delta\Delta}\|_\infty \} I(t) \Delta t. \end{aligned}$$

If, moreover, we let $g(t) \equiv 1$, then (123) becomes

$$\left| f(t) - \frac{2}{b-a} \int_a^b f(s) \Delta s + \frac{g_2(b,a)f(a) + h_2(b,a)f(b)}{(b-a)^2} - \left(t - b + \frac{g_2(b,a)}{b-a} \right) f^\Delta(t) \right| \leq \frac{1}{b-a} \|f^{\Delta\Delta}\|_\infty I(t).$$

From these inequalities, special cases such as discrete inequalities can be obtained.

12.5. PACHPATTE'S THIRD THEOREM ON TIME SCALES

Theorem 72. Assume (H). Then, for all $t \in [a, b]$, we have

$$\left| f(t)g(t)h(t) - \frac{1}{3(b-a)} A[f, g, h](t) - \frac{1}{3} \left(t - b + \frac{g_2(b,a)}{b-a} \right) M[f, g, h](t) \right| \leq \frac{1}{3(b-a)^2} B[f, g, h](t)H(t) \quad (130)$$

and

$$\left| \frac{1}{b-a} \int_a^b f(t)g(t)h(t) \Delta t - \frac{1}{3(b-a)^2} D[f, g, h](t) - \frac{1}{3(b-a)} \int_a^b \left(t - b + \frac{g_2(b,a)}{b-a} \right) M[f, g, h](t) \Delta t \right| \leq \frac{1}{3(b-a)^3} \int_a^b B[f, g, h](t)H(t) \Delta t, \quad (131)$$

where

$$H(t) := \int_a^b \int_a^b |p(t, \tau)p(\tau, s)| \Delta s \Delta \tau$$

and

$$p(t, s) := \begin{cases} \sigma(s) - a & \text{for } a \leq s < t \\ \sigma(s) - b & \text{for } t \leq s \leq b. \end{cases}$$

Proof. Fix $t \in [a, b]$. We use Theorem 68 three times to obtain

$$\begin{aligned} & \int_a^b \int_a^b p(t, \tau)p(\tau, s) f^{\Delta\Delta}(s) \Delta s \Delta \tau = \int_a^b p(t, \tau) \left\{ \int_a^b p(\tau, s) f^{\Delta\Delta}(s) \Delta s \right\} \Delta \tau \\ & = \int_a^b p(t, \tau) \left\{ (b-a) f^\Delta(\tau) - \int_a^b f^\Delta(s) \Delta s \right\} \Delta \tau \\ & = (b-a) \int_a^b p(t, s) f^\Delta(s) \Delta s + (f(a) - f(b)) \int_a^b p(t, s) \Delta s \end{aligned}$$

$$\begin{aligned}
&= (b-a) \left\{ (b-a)f(t) - \int_a^b f(s)\Delta s \right\} + (f(a) - f(b)) \left\{ (b-a)t - \int_a^b s\Delta s \right\} \\
&= (b-a)^2 f(t) - (b-a) \int_a^b f(s)\Delta s + (f(a) - f(b)) \int_b^a (s-t)\Delta s \\
&= (b-a)^2 f(t) - (b-a) \int_a^b f(s)\Delta s + (g_2(t, a) - h_2(b, t))(f(a) - f(b))
\end{aligned}$$

and thus (by using (107))

$$\begin{aligned}
f(t) &= \frac{1}{b-a} \int_a^b f(s)\Delta s + \left(t - b + \frac{g_2(b, a)}{b-a} \right) \frac{f(b) - f(a)}{b-a} \\
&\quad + \frac{1}{(b-a)^2} \int_a^b \int_a^b p(t, \tau)p(\tau, s)f^{\Delta\Delta}(s)\Delta s\Delta\tau. \quad (132)
\end{aligned}$$

Similarly, we get

$$\begin{aligned}
g(t) &= \frac{1}{b-a} \int_a^b g(s)\Delta s + \left(t - b + \frac{g_2(b, a)}{b-a} \right) \frac{g(b) - g(a)}{b-a} \\
&\quad + \frac{1}{(b-a)^2} \int_a^b \int_a^b p(t, \tau)p(\tau, s)g^{\Delta\Delta}(s)\Delta s\Delta\tau. \quad (133)
\end{aligned}$$

and

$$\begin{aligned}
h(t) &= \frac{1}{b-a} \int_a^b h(s)\Delta s + \left(t - b + \frac{g_2(b, a)}{b-a} \right) \frac{h(b) - h(a)}{b-a} \\
&\quad + \frac{1}{(b-a)^2} \int_a^b \int_a^b p(t, \tau)p(\tau, s)h^{\Delta\Delta}(s)\Delta s\Delta\tau. \quad (134)
\end{aligned}$$

Multiplying (132), (133) and (134) by $g(t)h(t)$, $f(t)h(t)$ and $f(t)g(t)$, respectively, adding the resulting identities and dividing by three, we have

$$\begin{aligned}
&f(t)g(t)h(t) - \frac{1}{3(b-a)} A[f, g, h](t) - \frac{1}{3} \left(t - b + \frac{g_2(b, a)}{b-a} \right) M[f, g, h](t) \\
&= \frac{1}{3(b-a)^2} \int_a^b \int_a^b p(t, \tau)p(\tau, s)\tilde{B}[f, g, h](t, s)\Delta s\Delta\tau \quad (135)
\end{aligned}$$

with \tilde{B} as in (120). By taking absolute values in (135) and using (113) and the definition of H , we obtain (130). Integrating (135) with respect to t from a to b , dividing by $b-a$, noting (122), taking absolute values and using (113) and the definition of H , we obtain (131). \square

Example 90. If we let $\mathbb{T} = \mathbb{R}$ in Theorem 72, then, by the same calculations as in Example 86, we obtain [61, Theorem 3], in particular, Theorem 67.

Example 91. If we let $\mathbb{T} = \mathbb{Z}$ and $a = 0$, $b = n \in \mathbb{N}$ in Theorem 72, then, by the same calculations as in Example 87, we obtain

$$\left| f(t)g(t)h(t) - \frac{1}{3n}A[f, g, h](t) - \frac{1}{3} \left(t - \frac{n-1}{2} \right) M[f, g, h](t) \right| \leq \frac{1}{3n^2}B[f, g, h](t)H(t)$$

and

$$\left| \frac{1}{n} \sum_{t=0}^{n-1} f(t)g(t)h(t) - \frac{1}{3n^2}D[f, g, h] - \frac{1}{3n} \sum_{t=0}^{n-1} \left(t - \frac{n-1}{2} \right) M[f, g, h](t) \right| \leq \frac{1}{3n^3} \sum_{t=0}^{n-1} B[f, g, h](t)H(t),$$

where in addition to A, B, D defined in Example 87,

$$M[f, g, h] = gh \frac{f(b) - f(a)}{b - a} + fh \frac{g(b) - g(a)}{b - a} + fg \frac{h(b) - h(a)}{b - a},$$

$$H(t) = \sum_{\tau=0}^{n-1} \sum_{s=0}^{n-1} |p(t, \tau)p(\tau, s)|,$$

$$p(t, s) = \begin{cases} s + 1 & \text{if } 0 \leq s < t \\ s + 1 - n & \text{if } t \leq s \leq n. \end{cases}$$

These inequalities are new discrete Ostrowski–Grüss type inequalities.

Remark 38. If we let $h(t) \equiv 1$ in Theorem 72, then (130) becomes

$$\left| f(t)g(t) - \frac{1}{2(b-a)} \left\{ g(t) \int_a^b f(s)\Delta s + f(t) \int_a^b g(s)\Delta s \right\} - \frac{1}{2} \left(t - b + \frac{g_2(b, a)}{(b-a)} \right) \left\{ g(t) \frac{f(b) - f(a)}{b-a} + f(t) \frac{g(b) - g(a)}{b-a} \right\} \right| \leq \frac{1}{2(b-a)^2} \{ |g(t)| \|f^{\Delta\Delta}\|_{\infty} + |f(t)| \|g^{\Delta\Delta}\|_{\infty} \} H(t)$$

and (131) turns into

$$\begin{aligned}
& \left| \frac{1}{b-a} \int_a^b f(t)g(t)\Delta t - \frac{1}{(b-a)^2} \left(\int_a^b f(t)\Delta t \right) \left(\int_a^b g(t)\Delta t \right) \right. \\
& \left. - \frac{1}{2(b-a)} \int_a^b \left(t-b + \frac{g_2(b,a)}{b-a} \right) \left\{ g(t) \frac{f(b)-f(a)}{b-a} + f(t) \frac{g(b)-g(a)}{b-a} \right\} \Delta t \right| \\
& \leq \frac{1}{2(b-a)^3} \int_a^b \{ |g(t)| \|f^{\Delta\Delta}\|_\infty + |f(t)| \|g^{\Delta\Delta}\|_\infty \} H(t)\Delta t.
\end{aligned}$$

If, moreover, we let $g(t) \equiv 1$, then (130) becomes

$$\begin{aligned}
& \left| f(t) - \frac{1}{b-a} \int_a^b f(s)\Delta s - \left(t-b + \frac{g_2(b,a)}{b-a} \right) \frac{f(b)-f(a)}{b-a} \right| \\
& \leq \frac{1}{(b-a)^2} \|f^{\Delta\Delta}\|_\infty H(t).
\end{aligned}$$

BIBLIOGRAPHY

- [1] *Webster's Third New International Dictionary*. Merriam-Webster Inc, Springfield, MA, 1986.
- [2] Ravi Agarwal, Martin Bohner, Donal O'Regan, and Allan Peterson. Dynamic equations on time scales: a survey. *J. Comput. Appl. Math.*, 141(1-2):1–26, 2002. Dynamic equations on time scales.
- [3] Ravi Agarwal, Martin Bohner, and Allan Peterson. Inequalities on time scales: a survey. *Math. Inequal. Appl.*, 4(4):535–557, 2001.
- [4] Ravi P. Agarwal. *Difference equations and inequalities*, volume 155 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1992. Theory, methods, and applications.
- [5] Elvan Akin-Bohner, Martin Bohner, and Thomas Matthews. Time Scales Ostrowski and Grüss Type Inequalities involving Three Functions. *Nonlinear Dyn. Syst. Theory*, 2011.
- [6] Moulay Rchid Sidi Ammi, Rui A. C. Ferreira, and Delfim F. M. Torres. Diamond- α Jensen's inequality on time scales. *J. Inequal. Appl.*, pages Art. ID 576876, 13, 2008.
- [7] Moulay Rchid Sidi Ammi and Delfim F. M. Torres. Combined dynamic Grüss inequalities on time scales. *arXiv:0801.1865v1*, 2010.
- [8] Moulay Rchid Sidi Ammi and Delfim F. M. Torres. Hölder's and Hardy's two dimensional diamond- α inequalities on time scales. *arXiv:0910.4115v1 [math. CA]*, 21 Oct 2009.
- [9] Angelo Arvanitis and Jon Gregory. *Credit: The Complete Guide To Pricing, Hedging And Risk Management*. Risk Books, 2001.
- [10] Jenny Baglivo. *Mathematica Laboratories for Mathematical Statistics: Emphasizing Simulation and Computer intensive Methods*. ASA-SIAM series on statistics and applied probability. Society for Industrial and Applied Mathematics, 2005.
- [11] Lee Bain and Max Engelhardt. *Introduction to probability and mathematical statistics*. Classic Series. Duxbury, 1992.
- [12] Neil Barnett, Sever Dragomir, and Ravi Agarwal. Some inequalities for probability, expectation, and variance of random variables defined over a finite interval. *Comput. Math. Appl.*, 43(10-11):1319–1357, 2002.
- [13] Purna Biswal. *Probability and Statistics*. Prentice-Hall Of India Pvt. Ltd., 2007.
- [14] Martin Bohner. The logarithm on time scales. *J. Difference Equ. Appl.*, 11(15):1305–1306, 2005.

- [15] Martin Bohner. Some oscillation criteria for first order delay dynamic equations. *Far East J. Appl. Math.*, 18(3):289–304, 2005.
- [16] Martin Bohner and Oktay Duman. Opial type inequalities for diamond-alpha derivatives and integrals on time scales. *Differ. Equ. Dyn. Syst.*, 2010, to appear.
- [17] Martin Bohner and Gusein Sh. Guseinov. The convolution on time scales. *Abstr. Appl. Anal.*, pages Art. ID 58373, 24, 2007.
- [18] Martin Bohner and Thomas Matthews. The Grüss inequality on time scales. *Commun. Math. Anal.*, 3(1):1–8 (electronic), 2007.
- [19] Martin Bohner and Thomas Matthews. Ostrowski inequalities on time scales. *JIPAM. J. Inequal. Pure Appl. Math.*, 9(1):Article 6, 8, 2008.
- [20] Martin Bohner, Thomas Matthews, and Adnan Tuna. Diamond-alpha Grüss type inequalities on time scales. *Int. J. Dyn. Syst. Differ. Equ.*, 3(1-2):234–247, 2011.
- [21] Martin Bohner, Thomas Matthews, and Adnan Tuna. Weighted Ostrowski-Grüss inequalities on time scales. *Afr. Diaspora J. Math.*, 2011.
- [22] Martin Bohner and Allan Peterson. *Dynamic equations on time scales*. Birkhäuser Boston Inc., Boston, MA, 2001. An introduction with applications.
- [23] Martin Bohner and Allan Peterson. *Advances in dynamic equations on time scales*. Birkhäuser Boston Inc., Boston, MA, 2003.
- [24] Marek Capiński and Tomasz Zastawniak. *Mathematics for Finance: An Introduction to Financial Engineering*. Springer undergraduate mathematics series. Springer, 2003.
- [25] George Casella. *Statistical inference*. Duxbury advanced series. Duxbury/Thomson Learning, 2002.
- [26] Pietro Cerone, Sever Dragomir, and John Roumeliotis. An inequality of Ostrowski type for mappings whose second derivatives belong to $L_1(a, b)$ and applications. *Honam Math. J.*, 21(1):127–137, 1999.
- [27] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.
- [28] Daryl Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes: General theory and structure*. Probability and its applications. Springer, 2007.
- [29] Sever Dragomir. Some integral inequalities of Grüss type. *Indian J. Pure Appl. Math.*, 31(4):397–415, 2000.
- [30] Sever Dragomir. A generalization of the Ostrowski integral inequality for mappings whose derivatives belong to $L_p[a, b]$ and applications in numerical integration. *J. Math. Anal. Appl.*, 255(2):605–626, 2001.

- [31] Sever Dragomir. The discrete version of Ostrowski's inequality in normed linear spaces. *JIPAM. J. Inequal. Pure Appl. Math.*, 3(1):Article 2, 16 pp. (electronic), 2002.
- [32] Sever Dragomir and Neil Barnett. An Ostrowski type inequality for mappings whose second derivatives are bounded and applications. *J. Indian Math. Soc. (N.S.)*, 66(1-4):237–245, 1999.
- [33] Sever Dragomir and Anthony Sofo. An integral inequality for twice differentiable mappings and applications. *Tamkang J. Math.*, 31(4):257–266, 2000.
- [34] Sever S. Dragomir and Song Wang. A new inequality of Ostrowski's type in L_p -norm. *Indian J. Math.*, 40(3):299–304, 1998.
- [35] Warren Ewens and Gregory Grant. *Statistical Methods in Bioinformatics: An Introduction*. Number v. 10 in Statistics for biology and health. Springer, 2005.
- [36] Rui A. C. Ferreira, Moulay Rchid Sidi Ammi, and Delfim F. M. Torres. Diamond-alpha integral inequalities on time scales. *arXiv:0805.0242v1 [math. CA]*, 2 May 2008.
- [37] International Monetary Fund. Argentina and the IMF, <http://www.imf.org>, September 2011.
- [38] Boris Gnedenko, Igor Ushakov, and James Falk. *Probabilistic reliability engineering*. Wiley-Interscience publication. Wiley, 1995.
- [39] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*. Texts from Oxford University Press. Oxford University Press, 2001.
- [40] Gerhard Grüss. Über das Maximum des absoluten Betrages von $\frac{1}{b-a} \int_a^b f(x)g(x) dx - \frac{1}{(b-a)^2} \int_a^b f(x) dx \int_a^b g(x) dx$. *Math. Z.*, 39(1):215–226, 1935.
- [41] Stefan Hilger. *Ein Maßkettenkalkül mit Anwendung auf Zentrumsmannigfaltigkeiten*. PhD thesis, Universität Würzburg, 1988.
- [42] John Hull. *Options, Futures and Other Derivatives*. Prentice Hall, 2006.
- [43] Billy Jackson. The time scale logarithm. *Appl. Math. Lett.*, 21(3):215–221, 2008.
- [44] Norman Johnson, Adrienne Kemp, and Samuel Kotz. *Univariate discrete distributions*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley, 2005.
- [45] Jagat Kapur. *Maximum-Entropy Models in Science and Engineering*. Wiley, 1989.
- [46] Wenjun Liu and Quóc-Anh Ngô. An Ostrowski-Grüss type inequality on time scales. *Comput. Math. Appl.*, 58(6):1207–1210, 2009.

- [47] Agnieszka B. Malinowska and Delfim F. M. Torres. On the diamond-alpha Riemann integral and mean value theorems on time scales. *Dynam. Systems Appl.*, 18(3-4):469–481, 2009.
- [48] Paul Meyer. *Introductory probability and statistical applications*. Addison-Wesley series in statistics. Addison-Wesley Pub. Co., 1970.
- [49] Dragoslav Mitrinović, Josip Pečarić, and Arlington Fink. *Inequalities involving functions and their integrals and derivatives*, volume 53 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers Group, Dordrecht, 1991.
- [50] Dragoslav Mitrinović, Josip Pečarić, and Arlington Fink. *Classical and new inequalities in analysis*, volume 61 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers Group, Dordrecht, 1993.
- [51] Moody's. <http://www.moody's.com>, September 2011.
- [52] Dorota Mozyrska, Ewa Pawłuszewicz, and Delfim F. M. Torres. The Riemann-Stieltjes integral on time scales. *Aust. J. Math. Anal. Appl.*, 7(1):Art. 10, 14, 2010.
- [53] Dorota Mozyrska and Delfim F. M. Torres. Diamond-alpha polynomial series on time scales. *Int. J. Math. Stat.*, 5(A09):92–101, 2009.
- [54] Quốc Anh Ngô and Wenjun Liu. A sharp Grüss type inequality on time scales and application to the sharp Ostrowski-Grüss inequality. *Commun. Math. Anal.*, 6(2):33–41, 2009.
- [55] Alexander Ostrowski. Über die Absolutabweichung einer differentiierebaren Funktion von ihrem Integralmittelwert. *Comment. Math. Helv.*, 10(1):226–227, 1937.
- [56] Umut Mutlu Özkan and Billûr Kaymakçalan. Basics of diamond- α partial dynamic calculus on time scales. *Math. Comput. Modelling*, 50(9-10):1253–1261, 2009.
- [57] Baburao Pachpatte. *Inequalities for differential and integral equations*, volume 197 of *Mathematics in Science and Engineering*. Academic Press Inc., San Diego, CA, 1998.
- [58] Baburao Pachpatte. *Inequalities for finite difference equations*, volume 247 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 2002.
- [59] Baburao Pachpatte. New discrete Ostrowski-Grüss like inequalities. *Facta Univ. Ser. Math. Inform.*, 22(1):15–20, 2007.
- [60] Baburao Pachpatte. On a new generalization of Ostrowski type inequality. *Tamkang J. Math.*, 38(4):335–339, 2007.
- [61] Baburao Pachpatte. Some new ostrowski and grüss type inequalities. *Tamkang J. Math.*, 38(2):111–120, 2007.
- [62] Fitch Ratings. <http://www.fitchratings.com>, September 2011.

- [63] James W. Rogers, Jr. and Qin Sheng. Notes on the diamond- α dynamic derivative on time scales. *J. Math. Anal. Appl.*, 326(1):228–241, 2007.
- [64] Sheldon Ross. *Introduction to Probability Models*. Introduction to Probability Models. Academic Press, 2007.
- [65] Mehmet Zeki Sarikaya. New weighted Ostrowski and Čebyšev type inequalities on time scales. *Comput. Math. Appl.*, 60(5):1510–1514, 2010.
- [66] Timo Schläfer. *Recovery Risk in Credit Default Swap Premia*. Springer.
- [67] Philipp Schönbucher. *Credit Derivatives Pricing Models: Models, Pricing And Implementation*. Wiley finance series. Wiley, 2003.
- [68] Claude Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.
- [69] Qin Sheng. Hybrid approximations via second-order crossed dynamic derivatives with the \diamond_{α} derivative. *Nonlinear Anal. Real World Appl.*, 9(2):628–640, 2008.
- [70] Qin Sheng, M. Fadag, Johnny Henderson, and John M. Davis. An exploration of combined dynamic derivatives on time scales and their applications. *Nonlinear Anal. Real World Appl.*, 7(3):395–413, 2006.
- [71] Standard&Poors. <http://www.standardandpoors.com>, September 2011.
- [72] Adnan Tuna and Durmus Daghan. Generalization of Ostrowski and Ostrowski-Grüss type inequalities on time scales. *Comput. Math. Appl.*, 60(3):803–811, 2010.

VITA

Thomas Matthews was born to Gisela and Wolfgang Matthews on September 28, 1982, in Greifswald, Germany. He received his primary and secondary education in Greifswald and graduated in June 2002 with the Abitur from the Alexander-von-Humboldt-Gymnasium in Greifswald, Germany. After finishing the Bachelor of Science degree in Mathematics with Economics in 2004 at the University of Ulm in Germany, he started his Master of Science in Mathematics with Economics (Diplom Wirtschaftsmathematik) at the same university and graduated in November 2007. As a participant in an exchange program, he came to the University of Missouri–Rolla in August 2006, where he was working towards his Master of Science in Applied Mathematics, which he completed in May 2007. Since January 2008, he is working towards his Doctor of Philosophy in Mathematics at Missouri University of Science and Technology. During the last two semesters in Ulm and the graduate studies in Rolla, he worked as a Graduate Teaching Assistant for the Faculty of Mathematics and Economics and for the Department of Mathematics and Statistics, respectively. He has published five papers, two from his Master’s Thesis [18, 19] and three from his Dissertation [5, 20, 21].