Introduction

1.1 Convex Functions

Definition 1.1 (SEE [119, DEFINITION 1.1])   (a) *Let I be an interval in* $\mathbb{R}$. *Then* $\Phi :$
I $\to \mathbb{R}$ *is said to be convex if for all* $x, y \in I$ *and all* $\alpha \in [0,1]$,

$$\Phi(\alpha x + (1-\alpha)y) \leq \alpha\Phi(x) + (1-\alpha)\Phi(y) \tag{1.1}$$

holds. If (1.1) *is strict for all* $x \neq y$ *and* $\alpha \in (0,1)$, *then* Φ *is said to be strictly convex.*

(b) *If the inequality in* (1.1) *is reversed, then* Φ *is said to be concave. If it is strict for all* $x \neq y$ *and* $\alpha \in (0,1)$, *then* Φ *is said to be strictly concave.*

There are several equivalent ways to define convex functions, sometimes it is better to define convex function in one way than the other.

Remark 1.1 (SEE [119, REMARKS 1.2])   (a) For $x, y \in I$, $p, q \geq 0$, $p + q > 0$, (1.1) is equivalent to

$$\Phi\left(\frac{px + qy}{p + q}\right) \leq \frac{p\Phi(x) + q\Phi(y)}{p + q}.$$

(b) Let x_1, x_2, x_3 be three points in I such that $x_1 < x_2 < x_3$. Then (1.1) is equivalent to

$$\begin{vmatrix} x_1 & \Phi(x_1) & 1 \\ x_2 & \Phi(x_2) & 1 \\ x_3 & \Phi(x_3) & 1 \end{vmatrix} = (x_3 - x_2)\Phi(x_1) + (x_1 - x_3)\Phi(x_2) + (x_2 - x_1)\Phi(x_3) \geq 0,$$

which is further equivalent to

$$\Phi(x_2) \leq \frac{x_2 - x_3}{x_1 - x_3}\Phi(x_1) + \frac{x_1 - x_2}{x_1 - x_3}\Phi(x_3). \tag{1.2}$$

More symmetrically and without the condition of monotonicity on x_1, x_2, x_3, we can write

$$\frac{\Phi(x_1)}{(x_1 - x_2)(x_1 - x_3)} + \frac{\Phi(x_2)}{(x_2 - x_3)(x_2 - x_1)} + \frac{\Phi(x_3)}{(x_3 - x_1)(x_3 - x_2)} \geq 0.$$

(c) Φ is both convex and concave if and only if

$$\Phi(x) = \lambda x + c$$

for some $\lambda, c \in \mathbb{R}$.

(d) Another way of writing (1.2) is instructive:

$$\frac{\Phi(x_1) - \Phi(x_2)}{x_1 - x_2} \leq \frac{\Phi(x_2) - \Phi(x_3)}{x_2 - x_3}, \tag{1.3}$$

where $x_1 < x_3$ and $x_1, x_3 \neq x_2$. Hence the following result is valid:

A function Φ is convex on I if and only if for every point $c \in I$, the function $(\Phi(x) - \Phi(c))/(x - c)$ is increasing on I $(x \neq c)$.

(e) By using (1.3), we can easily prove the following result:

If Φ is a convex function on I and if $x_1 \leq y_1$, $x_2 \leq y_2$, $x_1 \neq x_2$, $y_1 \neq y_2$, then the following inequality is valid:

$$\frac{\Phi(x_2) - \Phi(x_1)}{x_2 - x_1} \leq \frac{\Phi(y_2) - \Phi(y_1)}{y_2 - y_1}.$$

The following two theorems concern derivatives of convex functions.

Theorem 1.1 (SEE [119, THEOREM 1.3]) *Let I be an interval in $\mathbb{R}$ and $\Phi : I \to \mathbb{R}$ be convex. Then*

(i) *Φ is Lipschitz on any closed interval in I;*

(ii) *Φ'_+ and Φ'_- exist and are increasing in I, and $\Phi'_- \leq \Phi'_+$ (if Φ is strictly convex, then these derivatives are strictly increasing); and*

(iii) *Φ' exists, except possibly on a countable set, and on the complement of which it is continuous.*

Remark 1.2 (SEE [119, THEOREM 1.4]) In Theorem 1.1, if Φ'' exists on I, then Φ is convex if and only if $\Phi''(x) \geq 0$. If $\Phi''(x) > 0$, then Φ is strictly convex.

Theorem 1.2 (SEE [119, THEOREM 1.6]) *Let I be an open interval in $\mathbb{R}$.*

(a) *$\Phi : I \to \mathbb{R}$ is convex if and only if there is at least one line of support for Φ at each $x_0 \in I$, i.e.,*

$$\Phi(x) \geq \Phi(x_0) + \lambda(x - x_0) \quad \text{for all} \quad x \in (a, b),$$

where λ depends on x_0 and is given by $\lambda = \Phi'(x_0)$ when Φ' exists, and $\lambda \in [\Phi'_-(x_0), \Phi'_+(x_0)]$ when $\Phi'_-(x_0) \neq \Phi'_+(x_0)$.

(b) *$\Phi : I \to \mathbb{R}$ is convex if the function $\Phi(x) - \Phi(x_0) - \lambda(x - x_0)$ (the difference between the function and its support) is decreasing for $x < x_0$ and increasing for $x > x_0$.*

When dealing with functions with different degree of smoothness, divided differences are found to be very useful.

Definition 1.2 (SEE [119, PAGE 14]) *Let Φ be a real-valued function defined on $[a, b] \subset \mathbb{R}$. The kth-order divided difference of Φ at distinct points $x_0, \ldots, x_k$ in $[a, b]$ is defined recursively by*

$$[x_i; \Phi] = \Phi(x_i), \quad i \in \{0, 1, \ldots, k\}$$

and

$$[x_0, \ldots, x_k; \Phi] = \frac{[x_1, \ldots, x_k; \Phi] - [x_0, \ldots, x_{k-1}; \Phi]}{x_k - x_0}.$$

Remark 1.3 In Definition 1.2, the value $[x_0, \ldots, x_k; \Phi]$ is independent of the order of the points $x_0, \ldots, x_k$. This definition may be extended to include the case in which some or all of the points coincide by assuming that $x_0 \leq \ldots \leq x_k$ and letting

$$\underset{(j+1\,times)}{[x, \ldots, x; \Phi]} = \Phi^{(j)}(x)/j!,$$

provided that $\Phi^{(j)}$ exists.

Definition 1.3 (SEE [119, PAGE 15]) *A real-valued function Φ defined on $[a, b] \subset \mathbb{R}$ is said to be n-convex, $n \geq 0$, on $[a, b]$ if and only if for all choices of $(n+1)$ distinct points in $[a, b]$,*

$$[x_0, \ldots, x_n; \Phi] \geq 0.$$

Remark 1.4 A function $\Phi : I \to \mathbb{R}$ is convex if and only if for every choice of three mutually different points $x_0, x_1, x_2 \in I$, $[x_0, x_1, x_2; \Phi] \geq 0$ holds.

The definition of a convex function has a very natural generalization to real-valued functions defined on $\mathbb{R}^n$. Here we merely require that the domain U of Φ be convex, i.e., $\alpha \mathbf{x} + (1 - \alpha)\mathbf{y} \in U$ whenever $\mathbf{x}, \mathbf{y} \in U$ and $\alpha \in [0, 1]$.

Definition 1.4 *Let U be a convex set in $\mathbb{R}^n$. Then $\Phi : U \to \mathbb{R}$ is said to be convex if for all $\mathbf{x}, \mathbf{y} \in U$ and all $\alpha \in [0, 1]$, we have*

$$\Phi(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha \Phi(\mathbf{x}) + (1 - \alpha)\Phi(\mathbf{y}). \tag{1.4}$$

J-Convex Function

In the theory of convex functions, the most known case is that of *J*-convex functions, which deals with the arithmetic mean.

Definition 1.5 (SEE [119, DEFINITION 1.8]) *Let $I \subset \mathbb{R}$ be an interval. A function $\Phi : I \to \mathbb{R}$ is called convex in the Jensen sense (or J-convex) on I if for all $x, y \in I$, the inequality*

$$\Phi\left(\frac{x+y}{2}\right) \leq \frac{\Phi(x) + \Phi(y)}{2} \tag{1.5}$$

holds. A J-convex function Φ is said to be strictly J-convex if for all pairs of points (x, y), $x \neq y$, strict inequality holds in (1.5).

Remark 1.5 (SEE [119, THEOREM 1.10]) (i) It can be easily seen that a convex function is *J*-convex. For continuous functions, *J*-convex functions are equivalent to convex functions.

(ii) The inequality (1.5) can easily be extended to the convex combination of finitely many points and next to random variables associated to arbitrary probability spaces. These extensions are known as the discrete Jensen inequality and integral Jensen inequality, respectively.

Log-Convex Function

An important sub-class of convex functions is that of log-convex functions.

Definition 1.6 (SEE [119, DEFINITION 1.15]) *A function $\Phi : I \to \mathbb{R}$, I an interval in $\mathbb{R}$, is said to be log-convex, or multiplicative convex if $\log \Phi$ is convex, or equivalently if for all $x, y \in I$ and all $\alpha \in [0, 1]$,*

$$\Phi(\alpha x + (1 - \alpha)y) \leq \Phi(x)^{\alpha} \Phi(y)^{1-\alpha}. \tag{1.6}$$

It is said to be log-concave if the inequality in (1.6) *is reversed.*

Remark 1.6 (a) If we take $\alpha = 1/2$, then (1.6) becomes

$$\Phi\left(\frac{x+y}{2}\right)^2 \leq \Phi(x)\Phi(y),$$

and the function Φ is said to be log-convex in the Jensen sense. If the function Φ is log-convex in the Jensen sense and is continuous, then Φ is also log-convex.

(b) If $x_1, x_2, x_3 \in I$ such that $x_1 < x_2 < x_3$, then (1.6) is equivalent to

$$[\Phi(x_2)]^{(x_3-x_1)} \leq [\Phi(x_1)]^{(x_3-x_2)}[\Phi(x_3)]^{(x_2-x_1)}.$$

Furthermore, if $x_1, x_2, y_1, y_2 \in I$ such that $x_1 \leq y_1, x_2 \leq y_2, x_1 \neq x_2, y_1 \neq y_2$, then

$$\left(\frac{\Phi(x_2)}{\Phi(x_1)}\right)^{\frac{1}{x_2-x_1}} \leq \left(\frac{\Phi(y_2)}{\Phi(y_1)}\right)^{\frac{1}{y_2-y_1}}.$$

(c) $\Phi : I \rightarrow \mathbb{R}$ is log-convex in the Jensen sense if and only if

$$\alpha^2 \Phi(x) + 2\alpha\beta\Phi\left(\frac{x+y}{2}\right) + \beta^2\Phi(y) \geq 0$$

holds for all $\alpha, \beta \in \mathbb{R}$ and $x, y \in I$.

1.2 Exponential and *n*-Exponential Convexity

Exponentially convex functions were introduced by S. N. Bernstein [31] over eighty years ago and later D. V. Widder [132]. The notion of *n*-exponential convexity was introduced by J. Pečarić and J. Perić in [115] (see also [89, 78, 88]). Now we quote some definitions and results about exponential and *n*-exponential convexity.

Definition 1.7 *A function* $\Phi : I \rightarrow \mathbb{R}$ $(I \subseteq \mathbb{R})$ *is n-exponentially convex in the Jensen sense on I, if*

$$\sum_{i,j=1}^{n} \xi_i \xi_j \Phi\left(\frac{x_i + x_j}{2}\right) \geq 0$$

holds for all choices $\xi_i \in \mathbb{R}$ *and* $x_i \in I$, $i \in \{1, \ldots, n\}$. *A function* $\Phi : I \rightarrow \mathbb{R}$ *is n-exponentially convex if it is n-exponentially convex in the Jensen sense and continuous on I.*

Remark 1.7 It is clear from the definition that 1-exponentially convex functions in the Jensen sense are in fact nonnegative functions. Also, *n*-exponentially convex functions in the Jensen sense are *k*-exponentially convex in the Jensen sense for every $k \in \mathbb{N}$, $k \leq n$.

By definition of positive semi-definite matrices and some basic linear algebra, we have the following proposition.

Proposition 1.1 *If* Φ *is an n-exponentially convex function in the Jensen sense, then the matrix* $\left[\Phi\left(\frac{x_i + x_j}{2}\right)\right]_{i,j=1}^{k}$ *is positive semi-definite for all* $k \in \mathbb{N}$, $k \leq n$. *Particularly,* $\det\left[\Phi\left(\frac{x_i + x_j}{2}\right)\right]_{i,j=1}^{k} \geq 0$ *for all* $k \in \mathbb{N}$, $k \leq n$.

Definition 1.8 *A function* $\Phi : I \rightarrow \mathbb{R}$ *is exponentially convex in the Jensen sense on I, if it is n-exponentially convex in the Jensen sense for all* $n \in \mathbb{N}$. *A function* $\Phi : I \rightarrow \mathbb{R}$ *is exponentially convex if it is exponentially convex in the Jensen sense and continuous.*

Proposition 1.2 *[See [19, Proposition 1]] Let* $\Phi : (a, b) \rightarrow \mathbb{R}$. *The following are equivalent:*

(i) Φ *is exponentially convex.*

(ii) Φ *is continuous and*

$$\sum_{i,j=1}^{n} v_i v_j \Phi\left(\frac{x_i+x_j}{2}\right) \geq 0$$

for all $n \in \mathbb{N}$, $v_i \in \mathbb{R}$, and $x_i + x_j \in (a,b)$, $1 \leq i, j \leq n$.

(iii) Φ *is continuous and*

$$\det\left[\Phi\left(\frac{x_i+x_j}{2}\right)\right]_{i,j=1}^{m} \geq 0, \quad 1 \leq m \leq n$$

for all $n \in \mathbb{N}$ and for every $x_i \in (a,b)$, $i \in \{1,\dots,n\}$.

Remark 1.8 Some examples of exponentially convex functions are:

(i) $\Phi : I \to \mathbb{R}$ defined by $\Phi(x) = ce^{kx}$, where $c \geq 0$ and $k \in \mathbb{R}$.

(ii) $\Phi : \mathbb{R}^+ \to \mathbb{R}$ defined by $\Phi(x) = x^{-k}$, where $k > 0$.

(iii) $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ defined by $\Phi(x) = e^{-k\sqrt{x}}$, where $k > 0$.

Remark 1.9 From Remark 1.6 (c) it follows that a positive function is log-convex in the Jensen sense if and only if it is 2-exponentially convex in the Jensen sense. Also, using basic convexity theory, it follows that a positive function is log-convex if and only if it is 2-exponentially convex.

1.3 Superquadratic Functions

The concept of superquadratic functions in one variable, as a generalization of the class of convex functions, was recently introduced by S. Abramovich, G. Jameson and G. Sinnamon in [6] and [5]. More examples and properties of superquadratic functions can be found in [1, 25, 26, 24] and its references.

Definition 1.9 *A function $\Psi : [0,\infty) \to \mathbb{R}$ is called superquadratic if there exists a function $C : [0,\infty) \to \mathbb{R}$ such that*

$$\Psi(y) - \Psi(x) - \Psi(|y-x|) \geq C(x)(y-x) \quad \text{for all} \quad x, y \geq 0. \tag{1.7}$$

We say that Ψ is subquadratic if $-\Psi$ is superquadratic. If for all $x, y > 0$ with $x \neq y$, there is strict inequality in (1.7), then Ψ is called strictly superquadratic.

For example, the function $\Psi(x) = x^p$ is superquadratic for $p \geq 2$ and subquadratic for $p \in (0,2]$.

The following lemma shows essentially that positive superquadratic functions are also convex functions.

Lemma 1.1 *Let Ψ be a superquadratic function with $C(x)$ as in Definition 1.9. Then*

(i) $\Psi(0) \leq 0$;

(ii) *if $\Psi(0) = \Psi'(0) = 0$, then $C(x) = \Psi'(x)$ whenever Ψ is differentiable at $x > 0$;*

(iii) *if $\Psi \geq 0$, then Ψ is convex and $\Psi(0) = \Psi'(0) = 0$.*

In the following theorem, some characterizations of superquadratic functions are given analogous to the well-known characterizations of convex functions.

Theorem 1.3 (SEE [26, THEOREM 9]) *For the function $\Psi : [0,\infty) \to \mathbb{R}$, the following conditions are equivalent:*

(i) *The function Ψ is a superquadratic function, i.e., (1.7) holds.*

(ii) *For any two nonnegative n-tuples $(x_1,\ldots,x_n)$ and $(p_1,\ldots,p_n)$ such that $P_n = \sum\limits_{i=1}^{n} p_i > 0$, the inequality*

$$\Psi(\overline{x}) \leq \frac{1}{P_n}\sum_{i=1}^{n} p_i\Psi(x_i) - \frac{1}{P_n}\sum_{i=1}^{n} p_i\Psi(|x_i - \overline{x}|)$$

holds, where $\overline{x} = \dfrac{1}{P_n}\sum\limits_{i=1}^{n} p_i x_i$.

(iii) *The inequality*

$$\Psi(\lambda y_1 + (1-\lambda)y_2) \leq \lambda\Psi(y_1) + (1-\lambda)\Psi(y_2) \\ - \lambda\Psi(1-\lambda)|y_1 - y_2| - (1-\lambda)\Psi(\lambda|y_1 - y_2|)$$

holds for all $y_1, y_2 \geq 0$ and $\lambda \in [0,1]$.

(iv) *For all $x, y_1, y_2 \geq 0$, such that $y_1 < x < y_2$, we have*

$$\Psi(x) \leq \frac{y_2 - x}{y_2 - y_1}(\Psi(y_1) - \Psi(x - y_1)) + \frac{x - y_1}{y_2 - y_1}(\Psi(y_2) - \Psi(y_2 - x)),$$

i.e.,

$$\frac{\Psi(y_1) - \Psi(x) - \Psi(x - y_1)}{y_1 - x} \leq \frac{\Psi(y_2) - \Psi(x) - \Psi(y_2 - x)}{y_2 - x}.$$

In the following, for any function $\Psi \in C^1([0,\infty),\mathbb{R})$, we define an associated function $\overline{\Psi} \in C^1((0,\infty),\mathbb{R})$ by

$$\overline{\Psi}(x) = \frac{\Psi'(x)}{x} \quad \text{for all} \quad x > 0. \tag{1.8}$$

Lemma 1.2 (SEE [3, LEMMA 1]) *Let $\Psi \in C^1([0,\infty),\mathbb{R})$ such that $\Psi(0) \leq 0$. If $\overline{\Psi}$ is increasing (strictly increasing) or Ψ' is superadditive (strictly superadditive), then Ψ is superquadratic (strictly superquadratic).*

Lemma 1.3 *[See [3, Lemma 3]] Let* $\Psi \in C^2([0,\infty), \mathbb{R})$ *be such that*

$$m_1 \leq \frac{x\Psi''(x) - \Psi'(x)}{x^2} \leq M_1 \quad \text{for all} \quad x > 0.$$

Let the functions ϑ_1, ϑ_2 *be defined by*

$$\vartheta_1(x) = \frac{M_1 x^3}{3} - \Psi(x), \quad \vartheta_2(x) = \Psi(x) - \frac{m_1 x^3}{3}. \tag{1.9}$$

Then $\overline{\vartheta_1}, \overline{\vartheta_2}$ *are increasing. If also* $\Psi(0) = 0$, *then* ϑ_1, ϑ_2 *are superquadratic.*

Lemma 1.4 *Let* $s > 0$ *and* $\Psi_s : [0,\infty) \to \mathbb{R}$ *be defined by*

$$\Psi_s(x) = \begin{cases} \dfrac{x^s}{s(s-2)}, & s \neq 2, \\[2mm] \dfrac{x^2}{2} \log x, & s = 2. \end{cases} \tag{1.10}$$

Then Ψ_s *is superquadratic, with the convention* $0 \log 0 := 0$.

Lemma 1.5 *Let* $s \in \mathbb{R}$ *and* $\varphi_s : [0,\infty) \to \mathbb{R}$ *be defined by*

$$\varphi_s(x) = \begin{cases} \dfrac{sxe^{sx} - e^{sx} + 1}{s^3}, & s \neq 0, \\[2mm] \dfrac{x^3}{3}, & s = 0. \end{cases} \tag{1.11}$$

Then φ_s *is superquadratic.*

1.4 Time Scales Theory

The theory of time scales was introduced by Stefan Hilger in his PhD thesis [69] in 1988 as a unification of the theory of difference equations with that of differential equations, unifying integral and differential calculus with the calculus of finite differences, extending to cases "in between", and offering a formalism for studying hybrid discrete-continuous dynamic systems. It has applications in any field that requires simultaneous modelling of discrete and continuous data. Now, we briefly introduce the time scales calculus and refer to [70, 71] and the monograph [45] for further details.

By a time scale $\mathbb{T}$ we mean any nonempty closed subset of $\mathbb{R}$. The two most popular examples of time scales are the real numbers $\mathbb{R}$ and the integers $\mathbb{Z}$. Since the time scale $\mathbb{T}$ may or may not be connected, we need the concept of jump operators.

For $t \in \mathbb{T}$, we define the forward jump operator $\sigma : \mathbb{T} \to \mathbb{T}$ by

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

and the backward jump operator by

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

In this definition, the convention is $\inf \emptyset = \sup \mathbb{T}$ and $\sup \emptyset = \inf \mathbb{T}$.

If $\sigma(t) > t$, then we say that t is right-scattered, and if $\rho(t) < t$, then we say that t is left-scattered. Points that are right-scattered and left-scattered at the same time are called isolated. Also, if $\sigma(t) = t$, then t is said to be right-dense, and if $\rho(t) = t$, then t is said to be left-dense. Points that are simultaneously right-dense and left-dense are called dense.

If $\mathbb{T}$ has a left-scattered maximum M_1, then we define $\mathbb{T}^{\kappa} = \mathbb{T} \setminus \{M_1\}$; otherwise $\mathbb{T}^{\kappa} = \mathbb{T}$. If $\mathbb{T}$ has a right-scattered minimum M_2, then we define $\mathbb{T}_{\kappa} = \mathbb{T} \setminus \{M_2\}$; otherwise $\mathbb{T}_{\kappa} = \mathbb{T}$. Finally we define $\mathbb{T}^* = \mathbb{T}^{\kappa} \cap \mathbb{T}_{\kappa}$.

The mappings $\mu, \nu : \mathbb{T} \to [0, \infty)$ defined by

$$\mu(t) = \sigma(t) - t \quad \text{and} \quad \nu(t) = t - \rho(t)$$

are called the forward and backward graininess functions, respectively.

In the following considerations, $\mathbb{T}$ will denote a time scale, $I_{\mathbb{T}} = I \cap \mathbb{T}$ will denote a time scale interval (for any open or closed interval I in $\mathbb{R}$), and $[0, \infty)_{\mathbb{T}}$ will be used for the time scale interval $[0, \infty) \cap \mathbb{T}$.

Definition 1.10 *Assume $f : \mathbb{T} \to \mathbb{R}$ is a function and let $t \in \mathbb{T}^{\kappa}$. Then we define $f^{\Delta}(t)$ to be the number (provided it exists) with the property that given any $\varepsilon > 0$, there is a neighborhood $U_{\mathbb{T}}$ of t such that*

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

We call $f^{\Delta}(t)$ the delta derivative of f at t. We say that f is delta differentiable on $\mathbb{T}^{\kappa}$ provided $f^{\Delta}(t)$ exists for all $t \in \mathbb{T}^{\kappa}$.

Definition 1.11 *Assume $f : \mathbb{T} \to \mathbb{R}$ is a function and let $t \in \mathbb{T}_{\kappa}$. Then we define $f^{\nabla}(t)$ to be the number (provided it exists) with the property that given any $\varepsilon > 0$, there is a neighborhood $U_{\mathbb{T}}$ of t such that*

$$\left| (f(\rho(t)) - f(s)) - f^{\nabla}(t)[\rho(t) - s] \right| \leq \varepsilon |\rho(t) - s| \quad \text{for all} \quad s \in U_{\mathbb{T}}.$$

We call $f^{\nabla}(t)$ the nabla derivative of f at t. We say that f is nabla differentiable on $\mathbb{T}_{\kappa}$ provided $f^{\nabla}(t)$ exists for all $t \in \mathbb{T}_{\kappa}$.

Example 1.1 (i) If $\mathbb{T} = \mathbb{R}$, then

$$f^{\Delta}(t) = f^{\nabla}(t) = f'(t).$$

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

$$f^{\Delta}(t) = f(t+1) - f(t)$$

is the forward difference operator, while

$$f^{\nabla}(t) = f(t) - f(t-1)$$

is the backward difference operator.

(iii) Let $h > 0$. If $\mathbb{T} = h\mathbb{Z}$, then

$$f^{\Delta}(t) = \frac{f(t+h) - f(t)}{h} \quad \text{and} \quad f^{\nabla}(t) = \frac{f(t) - f(t-h)}{h}$$

are the h-derivatives.

(iv) Let $q > 1$. If $\mathbb{T} = q^{\mathbb{N}_0}$, where $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, then

$$f^{\Delta}(t) = \frac{f(qt) - f(t)}{(q-1)t} \quad \text{and} \quad f^{\nabla}(t) = \frac{q(f(t) - f(t/q))}{(q-1)t}$$

are the q-derivatives (or Jackson derivatives).

Definition 1.12 *Assume $f : \mathbb{T} \to \mathbb{R}$ is a function and let $t \in \mathbb{T}_{\kappa}^{\kappa}$. Then we define $f^{\diamond_{\alpha}}(t)$ to be the number (provided it exists) with the property that given any $\varepsilon > 0$, there is a neighborhood $U_{\mathbb{T}}$ of t such that*

$$|\alpha (f(\sigma(t)) - f(s))[\rho(t) - s] + (1-\alpha)(f(\rho(t)) - f(s))[\sigma(t) - s]$$
$$- f^{\diamond_{\alpha}}(t)[\rho(t) - s][\sigma(t) - s]| \le \varepsilon |[\rho(t) - s][\sigma(t) - s]| \quad \text{for all} \quad s \in U_{\mathbb{T}}.$$

We call $f^{\diamond_{\alpha}}(t)$ the diamond-α derivative of f at t. We say that f is diamond-α differentiable on $\mathbb{T}_{\kappa}^{\kappa}$ provided $f^{\diamond_{\alpha}}(t)$ exists for all $t \in \mathbb{T}_{\kappa}^{\kappa}$.

Remark 1.10 If $f : \mathbb{T} \to \mathbb{R}$ is differentiable on $\mathbb{T}$ in the sense of Δ and ∇, then f is diamond-α differentiable at $t \in \mathbb{T}_{\kappa}^{\kappa}$, and the diamond-$\alpha$ derivative is given by

$$f^{\diamond_{\alpha}}(t) = \alpha f^{\Delta}(t) + (1-\alpha) f^{\nabla}(t), \quad 0 \le \alpha \le 1.$$

Remark 1.11 From Definition 1.12, it is clear that f is diamond-α differentiable for $0 \le \alpha \le 1$ if and only if f is Δ and ∇ differentiable. It is obvious that for $\alpha = 1$, the diamond-α derivative reduces to the standard Δ derivative, and for $\alpha = 0$, the diamond-α derivative reduces to the standard ∇ derivative.

For all $t \in \mathbb{T}^{\kappa}$, we have the following properties of delta derivative.

Theorem 1.4 (SEE [45, THEOREM 1.16])   (i) *If f is delta differentiable at t, then f is continuous at t.*

(ii) *If f is continuous at t and t is right-scattered, then f is delta differentiable at t with $f^{\Delta}(t) = \frac{f(\sigma(t)) - f(t)}{\mu(t)}$.*

(iii) *If t is right-dense, then f is delta differentiable at t iff* $\lim\limits_{s \to t} \frac{f(t)-f(s)}{t-s}$ *exists as a finite number. In this case,* $f^{\Delta}(t) = \lim\limits_{s \to t} \frac{f(t)-f(s)}{t-s}$.

(iv) *If f is delta differentiable at t, then* $f(\sigma(t)) = f(t) + \mu(t) f^{\Delta}(t)$.

In the same manner, for all $t \in \mathbb{T}_{\kappa}$, we have the following properties of nabla derivative.

Theorem 1.5 (SEE [45, THEOREM 1.16])　(i) *If f is nabla differentiable at t, then f is continuous at t.*

(ii) *If f is continuous at t and t is left-scattered, then f is nabla differentiable at t with* $f^{\nabla}(t) = \frac{f(t)-f(\rho(t))}{\nu(t)}$.

(iii) *If t is left-dense, then f is nabla differentiable at t if and only if* $\lim\limits_{s \to t} \frac{f(t)-f(s)}{t-s}$ *exists as a finite number. In this case,* $f^{\nabla}(t) = \lim\limits_{s \to t} \frac{f(t)-f(s)}{t-s}$.

(iv) *If f is nabla differentiable at t, then* $f(\rho(t)) = f(t) + \nu(t) f^{\nabla}(t)$.

Definition 1.13 *A function* $f : \mathbb{T} \to \mathbb{R}$ *is called rd-continuous if it is continuous at all right-dense points in* $\mathbb{T}$ *and its left-sided limits are finite at all left-dense points in* $\mathbb{T}$. *We denote by* C_{rd} *the set of all rd-continuous functions. We say that f is rd-continuously delta differentiable (and write* $f \in C_{rd}^1$) *if* $f^{\Delta}(t)$ *exists for all* $t \in \mathbb{T}^{\kappa}$ *and* $f^{\Delta} \in C_{rd}$. *A function* $f : \mathbb{T} \to \mathbb{R}$ *is called ld-continuous if it is continuous at all left-dense points in* $\mathbb{T}$ *and its right-sided limits are finite at all right-dense points in* $\mathbb{T}$. *We denote by* C_{ld} *the set of all ld-continuous functions. We say that f is ld-continuously nabla differentiable (and write* $f \in C_{ld}^1$) *if* $f^{\nabla}(t)$ *exists for all* $t \in \mathbb{T}_{\kappa}$ *and* $f^{\nabla} \in C_{ld}$.

The set of all continuous functions on $\mathbb{T}$ contains both C_{rd} and C_{ld}.

Definition 1.14 *A function* $F : \mathbb{T} \to \mathbb{R}$ *is called a delta antiderivative of* $f : \mathbb{T} \to \mathbb{R}$ *if* $F^{\Delta}(t) = f(t)$ *for all* $t \in \mathbb{T}^{\kappa}$. *Then we define the delta integral by*

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

A function $G : \mathbb{T} \to \mathbb{R}$ *is called a nabla antiderivative of* $f : \mathbb{T} \to \mathbb{R}$ *if* $G^{\nabla}(t) = f(t)$ *for all* $t \in \mathbb{T}_{\kappa}$. *Then we define the nabla integral by*

$$\int_a^t f(s) \nabla s = G(t) - G(a).$$

The importance of rd-continuous and ld-continuous functions is revealed by the following result.

Theorem 1.6 (SEE [45, THEOREM 1.74, THEOREM 8.45]) *Every rd-continuous function has a delta antiderivative and every ld-continuous function has a nabla antiderivative.*

Definition 1.15 *Let $f : \mathbb{T} \to \mathbb{R}$ and $a, b \in \mathbb{T}$. Then the diamond-α integral of f from a to b is defined by*

$$\int_a^b f(t)\diamondsuit_\alpha t = \alpha \int_a^b f(t)\Delta t + (1-\alpha) \int_a^b f(t)\nabla t, \quad 0 \le \alpha \le 1.$$

Remark 1.12 Since the diamond-α integral is a combined Δ and ∇ integral, in general

$$\left(\int_a^s f(t)\diamondsuit_\alpha t \right)^{\diamondsuit_\alpha} \ne f(t), \quad t \in \mathbb{R}.$$

Example 1.2 Let $\alpha = \dfrac{1}{2}$ and $\mathbb{T} = \{0, 1, 2, 3\}$. Then the diamond-$\alpha$ derivative for a function on $\mathbb{T}$ is defined on the set $\mathbb{T}_\kappa^\kappa$ which is $\{1, 2\}$. Define the function $f(t) \equiv 0$. Next define functions F and G as follows:

$$\begin{aligned} F(0) &= 0, & G(0) &= 1; \\ F(1) &= 5, & G(1) &= -3; \\ F(2) &= 0, & G(2) &= 1; \\ F(3) &= 5, & G(3) &= -3. \end{aligned}$$

Then

$$\begin{aligned} F^{\diamondsuit_\alpha}(1) &= \frac{1}{2}\frac{F(2)-F(1)}{2-1} + \frac{1}{2}\frac{F(1)-F(0)}{1-0} \\ &= \frac{1}{2}(0-5) + \frac{1}{2}(5-0) = 0 = f(1) \end{aligned}$$

and

$$\begin{aligned} F^{\diamondsuit_\alpha}(2) &= \frac{1}{2}\frac{F(3)-F(2)}{3-2} + \frac{1}{2}\frac{F(2)-F(1)}{2-1} \\ &= \frac{1}{2}(5-0) + \frac{1}{2}(0-5) = 0 = f(2). \end{aligned}$$

Also

$$\begin{aligned} G^{\diamondsuit_\alpha}(1) &= \frac{1}{2}\frac{G(2)-G(1)}{2-1} + \frac{1}{2}\frac{G(1)-G(0)}{1-0} \\ &= \frac{1}{2}(1-(-3)) + \frac{1}{2}(-3-1) = 0 = f(1) \end{aligned}$$

and

$$\begin{aligned} G^{\diamondsuit_\alpha}(2) &= \frac{1}{2}\frac{G(3)-G(2)}{3-2} + \frac{1}{2}\frac{G(2)-G(1)}{2-1} \\ &= \frac{1}{2}(-3-1) + \frac{1}{2}(1-(-3)) = 0 = f(2). \end{aligned}$$

Thus

$$F^{\diamond\alpha(t)} = G^{\diamond\alpha}(t) = f(t)$$

on $\mathbb{T}_\kappa^\kappa$. We see that both F and G are diamond-α antiderivatives of f on $\mathbb{T}_\kappa^\kappa$. However,

$$F(2) - F(1) = -5 \neq 4 = G(2) - G(1).$$

Example 1.2 can be generalized for any fixed α strictly between 0 and 1, and for any purely discrete time scale, such as $\mathbb{T} = \mathbb{Z}$.

Next, we present an example where no diamond-α antiderivative exists.

Example 1.3 Let $\alpha = \dfrac{1}{2}$. Let $\mathbb{T} = (-\infty, 1] \cup [2, \infty)$. Set

$$f(t) = \begin{cases} -1, & x \leq 1, \\ \\ 5, & x \geq 2. \end{cases}$$

Assume a diamond-α antiderivative F of f exists on $\mathbb{T}_\kappa^\kappa$. On $(-\infty, 1]$, F must be of the form $-t + C_1$, where C_1 is a constant. On $[2, \infty)$, F must be of the form $5t + C_2$. It follows therefore

$$F^{\diamond\alpha}(1) = \frac{1}{2}F^\Delta(1) + \frac{1}{2}F^\nabla(1) = f(1).$$

Thus

$$\frac{1}{2}[(5(2) + C_2) - (-1(1) + C_1)] + \frac{1}{2}(-1) = -1. \tag{1.12}$$

Also,

$$F^{\diamond\alpha}(2) = \frac{1}{2}F^\Delta(2) + \frac{1}{2}F^\nabla(2) = f(2).$$

Thus

$$\frac{1}{2}(5) + \frac{1}{2}[(5(2) + C_2) - (-1(1) + C_1)] = 5. \tag{1.13}$$

From (1.12) and (1.13), we obtain the system of equations

$$C_1 - C_2 = 12,$$
$$C_1 - C_2 = 6,$$

which has no solution. Thus for the function f, which is continuous on $\mathbb{T}$, no diamond-α antiderivative exists on $\mathbb{T}_\kappa^\kappa$.

Now we give some properties of the delta integral.

Theorem 1.7 (SEE [45, THEOREM 1.77]) *If $a, b, c \in \mathbb{T}$, $\alpha \in \mathbb{R}$, and $f, g \in C_{rd}$, 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) $\int_b^a f(t)\Delta t = -\int_a^b f(t)\Delta t,$

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

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

(vi) *if* $f(t) \geq 0$ *for all t, then* $\int_a^b f(t)\Delta t \geq 0.$

A similar theorem holds for the nabla integral, for $f, g \in C_{ld}$, and for the diamond-α integral, for $f, g \in C$.

Regarding integral calculus on time scales, the literature includes, among others, the Cauchy delta integral [45, 70], the Cauchy nabla integral [23, 45], the Riemann delta integral [46, 63, 64], the Riemann nabla integral [63]; the Cauchy diamond-α integral [16, 123], the Riemann diamond-α integral [97], the Lebesgue delta and nabla integrals [46, 62], the multiple Riemann and multiple Lebesgue delta, nabla and diamond-α integrals [42, 43].

Let $n \in \mathbb{N}$ be fixed. For each $i \in \{1, \ldots, n\}$, let $\mathbb{T}_i$ denote a time scale and

$$\Lambda^n = \mathbb{T}_1 \times \ldots \times \mathbb{T}_n = \{t = (t_1, \ldots, t_n) : t_i \in \mathbb{T}_i,\ 1 \leq i \leq n\} \tag{1.14}$$

an n-dimensional time scale. Let μ_Δ be the σ-additive Lebesgue Δ-measure on Λ^n and $\mathscr{F}$ be the family of Δ-measurable subsets of Λ^n. Let $\mathscr{E} \subset \Lambda^n$ and $(\mathscr{E}, \mathscr{F}, \mu_\Delta)$ be a time scale measure space. Then for a Δ-measurable function $f : \mathscr{E} \to \mathbb{R}$, the corresponding Δ-integral of f over $\mathscr{E}$ is denoted according to [43, (3.18)] by

$$\int_{\mathscr{E}} f(t_1, \ldots, t_n) \Delta_1 t_1 \ldots \Delta_n t_n, \quad \int_{\mathscr{E}} f(t) \Delta t, \quad \int_{\mathscr{E}} f \mathrm{d}\mu_\Delta, \quad \text{or} \quad \int_{\mathscr{E}} f(t) \mathrm{d}\mu_\Delta(t).$$

By [43, Section 3], all theorems of the general Lebesgue integration theory, including the Lebesgue dominated convergence theorem, hold also for Lebesgue Δ-integrals on Λ^n. Here we state Fubini's theorem for multiple Lebesgue Δ-integrals on time scales. It is used in Chapter 9.

Theorem 1.8 *Let* $(X, \mathscr{K}, \mu_\Delta)$ *and* $(Y, \mathscr{L}, \nu_\Delta)$ *be two finite-dimensional time scale measure spaces. If* $f : X \times Y \to \mathbb{R}$ *is a Δ-integrable function and if we define the functions*

$$\varphi(y) = \int_X f(x, y) \mathrm{d}\mu_\Delta(x) \quad \text{for a.e.} \quad y \in Y$$

and

$$\psi(x) = \int_Y f(x, y) \mathrm{d}\nu_\Delta(y) \quad \text{for a.e.} \quad x \in X,$$

then φ is Δ-integrable on Y and ψ is Δ-integrable on X and

$$\int_X \mathrm{d}\mu_\Delta(x) \int_Y f(x, y) \mathrm{d}\nu_\Delta(y) = \int_Y \mathrm{d}\nu_\Delta(y) \int_X f(x, y) \mathrm{d}\mu_\Delta(x).$$