STEFFENSEN’S INEQUALITY ON TIME SCALES FOR
CONVEX FUNCTIONS

MOHAMMED MUNIRU IDDRISU

Abstract. The Steffensen’s Inequality was discovered in 1918 by Johan Frederic Steffensen (1873-1961). This inequality is very popular in the research environment and attracted the attention of many people working in similar area. Various extensions and generalisations have been provided concerning the inequality. This paper presents some further refinements of the Steffensen’s Inequality on Time scales using methods of convexity, differentiability and monotonicity.

1. Introduction

The well-known Steffensen’s inequality is presented as

\[ \int_{b-\lambda}^{b} g(x)dx \leq \int_{a}^{b} g(x)f(x)dx \leq \int_{a}^{a+\lambda} g(x)dx, \]

where \( \lambda = \int_{a}^{b} f(x)dx \), \( f \) and \( g \) are integrable functions defined on \((a, b)\), \( g \) is decreasing and \( 0 \leq f(x) \leq 1 \) for each \( x \in (a, b) \). This inequality was initially not popular in the research environment until 1947 when it appeared in [18]. Subsequently, it became known in the literature as one of the most useful inequalities in mathematical analysis. A number of research papers have been written on it in terms of providing extensions, generalisations and numerous applications (see [11, pp. 311-312], [12, 14, 15] and the references cited therein). The generalisation of (1) was first given in [2] and the result was detected in [5] to be incorrect. However, the correction was done over a decade later by Pečarić in [14] as

\[ \left( \int_{0}^{1} f(x)g(x)dx \right)^{p} \leq \int_{0}^{\lambda} g(x)^{p}dx \]

Received February 22, 2018. Accepted January 14, 2019.
2010 Mathematics Subject Classification. 26D15, 26E70, 34N05.
Key words and phrases. Steffensen’s Inequality, Convex functions, Time Scales.
where $\lambda = \left(\int_0^1 f(x)dx\right)^p$, $g : [0, 1] \to \mathbb{R}$ is a non-negative and non-increasing function and $f : [0, 1] \to \mathbb{R}$ is an integrable function such that $0 \leq f(x) \leq 1 \,(\forall x \in [0, 1])$ for $p \geq 1$. An analogous inequality to (2) was further given in [14] as

$$\frac{\int_0^1 f(x)g(x)dx}{\int_0^1 f(x)dx} \leq \frac{1}{\lambda} \int_0^\lambda g(x)dx.$$  \hspace{1cm} (3)

Following (1) and using the substitution $f(x) = \frac{\lambda F(x)}{\int_a^x F(x)dx}$, the inequality

$$\frac{1}{\lambda} \int_{b-\lambda}^b g(x)dx \leq \frac{\int_a^b g(x)F(x)dx}{\int_a^b F(x)dx} \leq \frac{1}{\lambda} \int_a^{a+\lambda} g(x)dx$$  \hspace{1cm} (4)

was further established in [15] with the assumption that $f(x)$ and $g(x)$ being integrable functions defined on $[a, b]$ and that $g(x)$ never increases with

$$0 \leq \lambda F(x) \leq \int_a^b F(x)dx, \quad (\forall x \in [a, b]),$$

where $\lambda$ is a positive number.

Furthermore, the inequality (1) was extended by Mercer in [10] but his generalisation was later detected to be incorrect in [20] (see also [9] and [16]) and the modified version presented as

$$\int_a^b g(x)f(x)dx \leq \int_a^{a+\lambda} g(x)h(x)dx,$$  \hspace{1cm} (5)

where $\lambda$ is given by

$$\int_a^{a+\lambda} h(x)dx = \int_a^b f(x)dx,$$

with $f$, $g$ and $h$ being integrable functions on $(a, b)$, $g$ decreasing and $0 \leq f \leq h$.

The second inequality of (1) was also modified as

$$\int_{b-\lambda}^b g(x)h(x)dx \leq \int_a^b g(x)f(x)dx,$$  \hspace{1cm} (6)

where $\lambda$ is given by

$$\int_{b-\lambda}^b h(x)dx = \int_a^b f(x)dx,$$
with \(f\), \(g\) and \(h\) being integrable functions on \((a, b)\), \(g\) decreasing and \(0 \leq f \leq h\).

Moreover, the double inequality was thus presented as

\[\int_{b-\lambda}^{b} g(x)h(x)dx \leq \int_{a}^{b} g(x)f(x)dx \leq \int_{a}^{a+\lambda} g(x)h(x)dx,\]

provided that there exists \(\lambda \in [0, b - a]\) such that

\[\int_{b-\lambda}^{b} h(x)dx = \int_{a}^{b} f(x)dx = \int_{a}^{a+\lambda} h(x)dx,\]

with \(f\), \(g\) and \(h\) being integrable functions on \((a, b)\), \(g\) decreasing and \(0 \leq f \leq h\).

In [8], a refinement of (1) was established also for convex functions as

\[\phi\left(\int_{0}^{1} f(t)dt\right) \leq \int_{0}^{1} f(t)\phi'(t)dt\]

where \(f : [0, 1] \rightarrow \mathbb{R}\) is a continuous function with \(0 \leq f \leq 1\) and \(\phi : [0, 1] \rightarrow \mathbb{R}\) is a convex and differentiable function with \(\phi(0) = 0\) for all \(t \in [0, 1]\).

With the discovery of the calculus on time scales, various papers presented new versions of the Steffensen’s inequality on time scales. The reader is referred to the papers [1, 13, 21].

The purpose of this paper is thus to present further refinements of the Steffensen’s inequality for convex functions on Time scales but with emphasis on \(T = \mathbb{R}\).

2. Preliminaries on Time scales

The following presentations in relation to the notion of time scales can be seen in [3, 6, 7] and the references cited therein. A time scale is an arbitrary nonempty closed subset of the real numbers. For example, the set of real numbers \((\mathbb{R})\), the set of integers \((\mathbb{Z})\), the natural numbers \((\mathbb{N})\) and the set of nonnegative integers \((\mathbb{N}_0)\) are all time scales while the sets of rational numbers \((\mathbb{Q})\), irrational numbers \((\mathbb{R} \setminus \mathbb{Q})\) and complex numbers \((\mathbb{C})\) are not time scales. The theory of Time Scales was
introduced in [6] by Stefan Hilger in his PhD thesis. This concept was introduced with the object of unifying continuous and discrete analysis and has since gained much attention in the area of Mathematical Analysis in recent times.

Furthermore, the definitions for the forward and backward jump operators are presented as follows:

**Definition 2.1.** Let $\mathbb{T}$ be a time scale and let $t \in \mathbb{T}$. The mapping 
$$\sigma : \mathbb{T} \rightarrow [0, \infty)$$

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

is called the forward jump operator and the mapping 
$$\rho : \mathbb{T} \rightarrow [0, \infty)$$

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

is called the backward jump operator. 
The convention is:
$$\inf \emptyset := \sup \mathbb{T} \quad \text{and} \quad \sup \emptyset := \inf \mathbb{T}.$$ 

If $\sigma(t) > t$, $t$ is right-scattered 
If $\rho(t) < t$, $t$ is left-scattered. 
If $\sigma(t) = t$, $t$ is right-dense 
If $\rho(t) = t$, $t$ is left-dense

Points that are right-scattered and left-scattered at the same time are called isolated. Also, points that are right-dense and left-dense at the same time are called dense. If $\mathbb{T}$ has a right-scattered minimum $m$, define $\mathbb{T}_k := \mathbb{T} \setminus \{m\}$; otherwise, set $\mathbb{T}_k = \mathbb{T}$. If $\mathbb{T}$ has a left-scattered maximum $M$, define $\mathbb{T}^k := \mathbb{T} \setminus \{M\}$; otherwise, set $\mathbb{T}^k = \mathbb{T}$. Finally, put $\mathbb{T}^k_k = \mathbb{T}_k \cap \mathbb{T}^k$. (see also [1, 4]).

**Definition 2.2.** The mapping $\mu : \mathbb{T} \rightarrow [0, \infty)$ defined by 
$$\mu(t) = \sigma(t) - t$$

is called the forward graininess function whiles the mapping $\nu(t) : \mathbb{T} \rightarrow [0, \infty)$ defined by
$$\nu(t) = t - \rho(t)$$

is called the backward graininess function.
Definition 2.3. Let $f : \mathbb{T} \rightarrow \mathbb{R}$ be a function. Then $f^\sigma : \mathbb{T} \rightarrow \mathbb{R}$ is defined by

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

i.e., $f^\sigma = f \circ \sigma$.

Definition 2.4. For $f : \mathbb{T} \rightarrow \mathbb{R}$ and $t \in \mathbb{T}_k$, $f$ is delta differentiable at $t$ (or simply differentiable) denoted by $f^\Delta(t)$ (provided it exists) with the property that given any $\epsilon > 0$, there is a neighbourhood $U$ of $t$ such that

$$|f(\sigma(t)) - f(s) - f^\Delta(t)\sigma(t) - s| < \epsilon|\sigma(t) - s|$$

for all $s \in U$.

For $f : \mathbb{T} \rightarrow \mathbb{R}$ and $t \in \mathbb{T}_k$ the nabla derivative of $f$ at $t$ is defined to be the number denoted by $f^\nabla(t)$ (provided it exists), with the property that, given any $\epsilon > 0$, there is a neighbourhood $V$ of $t$ such that

$$|[f(\rho(t)) - f(s)] - f^\nabla(t)[\rho(t) - s]| < \epsilon|\rho(t) - s|$$

for all $s \in V$.

Note that for $\mathbb{T} = \mathbb{R}$,

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

and for $\mathbb{T} = \mathbb{Z}$

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

is the forward difference operator, while

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

is the backward difference operator.

Now, in [4, 13, 17] and the references cited therein we have the following properties for $f : \mathbb{T} \rightarrow \mathbb{R}$ and $t \in \mathbb{T}_k$.

- If $f$ is (delta) differentiable at $t$, then $f$ is continuous at $t$.
- If $f$ is left continuous at $t$ and $t$ is right-scattered, then $f$ is delta differentiable at $t$ with

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

- If $t$ is right-dense, then $f$ is delta differentiable at $t$, if and only if, \[ \lim_{s \rightarrow t} \frac{f(t) - f(s)}{t - s} \] exists as a finite number. In this case,

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

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

The following properties also exist for $t \in T_k$.

- If $f$ is nabla differentiable at $t$, then $f$ is continuous at $t$.
- If $f$ is right continuous at $t$ and $t$ is left-scattered, then $f$ is nabla differentiable at $t$ with
  \[
  f^\nabla(t) = \frac{f(t) - f(\rho(t))}{t - \rho(t)}.
  \]
- If $t$ is left-dense, then $f$ is nabla differentiable at $t$, if and only if, \(\lim_{s \to t} \frac{f(t) - f(s)}{t - s}\) exists as a finite number. In this case,
  \[
  f^\nabla(t) = \lim_{s \to t} \frac{f(t) - f(s)}{t - s}.
  \]
- If $f$ is nabla differentiable at $t$, then $f(\rho(t)) = f(t) - \nu(t)f^\nabla(t)$.

Note the following special cases:

- If $T = \mathbb{R}$, then $\sigma(t) = t$, $\mu(t) = 0$, $f^\Delta(t) = f'(t)$, $\int_a^b f(t) \Delta t = \int_a^b f(t) dt$.
- If $T = \mathbb{Z}$, then $\sigma(t) = t + 1$, $\mu(t) = 1$, $f^\Delta(t) = \Delta f(t)$, $\int_a^b f(t) \Delta t = \sum_{t=a}^{t=a+1} f(t)$.
- If $T = h\mathbb{Z}$, $h > 0$, then $\sigma(t) = t + h$, $\mu(t) = h$ and
  \[
  y^\Delta(t) = \Delta_h y(t) = \frac{y(t+h) - y(t)}{h}, \quad \int_a^b f(t) \Delta t = \sum_{k=0}^{b-a} f(a + kh)h.
  \]

**Definition 2.5.** [4] A function $f : T \to \mathbb{R}$ is called rd-continuous, if it is continuous at all right-dense points in $T$ and its left-sided limits are finite at all left-dense points in $T$. The set of all rd-continuous functions are denoted by $C_{rd}$.

A function $F : T \to \mathbb{R}$ is called ld-continuous, if it is continuous at all left-dense points in $T$ and its right-sided limits are finite at all right-dense points in $T$. The set of all ld-continuous functions are denoted by $C_{ld}$.

(See also [19]).

Clearly, the set of all continuous functions on $T$ contains both $C_{rd}$ and $C_{ld}$. 
Definition 2.6. A function $F : \mathbb{T} \rightarrow \mathbb{R}$ is called a delta antiderivative of $f : \mathbb{T} \rightarrow \mathbb{R}$ if $F(\Delta(t)) = f(t)$, for all $t \in \mathbb{T}^k$, and the delta integral is defined as
\[ \int_a^t f(s) \Delta s = F(t) - F(a). \]

A function $G : \mathbb{T} \rightarrow \mathbb{R}$ is called a nabla antiderivative of $f : \mathbb{T} \rightarrow \mathbb{R}$ if $G(\nabla(t)) = f(t)$, for all $t \in \mathbb{T}^k$, and the nabla integral is defined as
\[ \int_a^t f(s) \nabla s = G(t) - G(a). \]

Theorem 2.7. [3, Theorem 1.74] Every rd-continuous function has an antiderivative. In particular if $t_0 \in \mathbb{T}$ then $F$ defined by
\[ F(t) = \int_{t_0}^t f(\eta) \Delta \eta \quad \text{for} \quad t \in \mathbb{T} \]
is an antiderivative of $f$.

Similarly, every ld-continuous function has an antiderivative.

Theorem 2.8. [3, Theorem 1.75]
(a) If $f \in C_{rd}$ and $t \in \mathbb{T}^k$, then
\[ \int_{\sigma(t)}^t f(s) \Delta s = \mu(t)f(t). \]
(b) If $f \in C_{ld}$ and $t \in \mathbb{T}^k$, then
\[ \int_{\rho(t)}^t f(s) \nabla s = \nu(t)f(t). \]

Theorem 2.9. [3, Theorem 1.76] If $f^\Delta \geq 0$ or $f^\nabla \geq 0$, then $f$ is nondecreasing.

Definition 2.10. [4] A function $\varphi : \mathbb{T} \rightarrow \mathbb{R}$ is said to be convex on $I_{\mathbb{T}}$ if
\[ \varphi(\lambda t_1 + (1 - \lambda)t_2) \leq \lambda \varphi(t_1) + (1 - \lambda)\varphi(t_2), \]
for all $t_1, t_2 \in I_{\mathbb{T}}$ with $\lambda t_1 + (1 - \lambda)t_2 \in I_{\mathbb{T}}$ and all $\lambda \in [0, 1]$.

A convex function is necessarily continuous for $t_1, t_2 \in I_{\mathbb{T}}$. A function $\varphi$ is said to be strictly convex if for all $t_1 \neq t_2$, $\varphi$ is said to be strictly convex.
3. Results and Discussions

This section presents the time scale version of the Steffensen’s inequality for convex functions. In [19], Jensen’s Inequality on time scale is presented.

Lemma 3.1. Let \( g : [0,1]_\mathbb{T} \rightarrow \mathbb{R} \) be a non-negative and non-increasing function and \( f : [0,1]_\mathbb{T} \rightarrow \mathbb{R} \) is an integrable function such that \( 0 \leq f(t) \leq 1 \) for all \( t \in [0,1] \), then
\[
\left( \int_0^1 f(t)g(t) \Delta t \right)^p \leq \int_0^1 g(t)^p \Delta t
\]
where \( \lambda = \left( \int_0^1 f(t) \Delta t \right)^p \) for \( p \geq 1 \).

Proof. Following the proof in [14, Theorem 1] in the context of time scale yields the result.

Theorem 3.2. Let the function \( f : [0,1]_\mathbb{T} \rightarrow \mathbb{R} \) be continuous with \( 0 \leq f(t) \leq 1 \). If \( \varphi : [0,1]_\mathbb{T} \rightarrow \mathbb{R} \) is a convex and delta differentiable function on \( \mathbb{T}^k \) with \( \varphi(0) = 0 \), then
\[
\varphi \left( \int_0^1 f(t) \Delta t \right) \leq \int_0^1 f(t) \varphi^\Delta(t) \Delta t
\]
for all \( t \in [0,1] \).

Proof. The convex function \( \varphi(t) \) is delta differentiable on the interval \( [0,1] \) and so \( \varphi^\Delta(t) \) is non-decreasing for all \( t \in [0,1] \). This means that \( -\varphi^\Delta(t) \) is non-increasing. Putting \( g(t) = -\varphi^\Delta(t) \), \( a = 0 \) and \( b = 1 \), inequality (1) yields
\[
\int_0^\lambda \varphi^\Delta(t) \Delta t \leq \int_0^1 f(t) \varphi^\Delta(t) \Delta t \leq \int_0^1 \varphi^\Delta(t) \Delta t.
\]
By Theorems (2.7) and (2.8), this simplifies to
\[
\varphi(\lambda) - \varphi(0) \leq \int_0^1 f(t) \varphi^\Delta(t) \Delta t \leq \varphi(1) - \varphi(1 - \lambda).
\]
Since \( \lambda = \int_0^1 f(t) \Delta t \) and \( \varphi(0) = 0 \), the left side of the double inequality (13) yields
\[
\varphi \left( \int_0^1 f(t) \Delta t \right) \leq \int_0^1 f(t) \varphi^\Delta(t) \Delta t.
\]
\[\square\]
Theorem 3.3. Let \( f : [0, 1]_T \rightarrow \mathbb{R} \) be continuous with \( 0 \leq f(t) \leq 1 \) for all \( t \in [0, 1] \). If \( \varphi : [0, 1]_T \rightarrow \mathbb{R} \) is convex and delta differentiable function on \( T_k \) with \( p \geq 1 \), then

\[
(14) \quad \left( \int_0^\lambda (\varphi^\Delta(t))^\frac{1}{p} \Delta t \right)^p \leq \left( \int_0^1 f(t)\varphi^\Delta(t) \Delta t \right) .
\]

Proof. In (10), replace \( g(t) \) with \(-\varphi^\Delta(t)\) (i.e. negative of the delta derivative of the convex function \( \varphi(t) \)) yields

\[
\left( -\int_0^1 f(t)\varphi^\Delta(t) \Delta t \right)^p \leq \int_0^\lambda (-\varphi^\Delta(t))^p \Delta t.
\]

That is

\[
\left( -\int_0^1 f(t)\varphi^\Delta(t) \Delta t \right)^p \leq (-1)^p \int_0^\lambda (\varphi^\Delta(t))^p \Delta t.
\]

Raising both sides of the inequality to the power of \((\frac{1}{p})\) and dividing through by \((-1)^p\) yields

\[
\left( \int_0^\lambda (\varphi^\Delta(t))^p \Delta t \right)^{\frac{1}{p}} \leq \int_0^1 f(t)\varphi^\Delta(t) \Delta t
\]
as required. \( \square \)

Remark 3.4. For a particular case of \( p = 1 \) with \( \varphi(0) = 0 \) yields

\[
\int_0^\lambda \varphi^\Delta(t) \Delta t \leq \int_0^1 f(t)\varphi^\Delta(t) \Delta t
\]

This simplifies to

\[
(15) \quad \varphi(\lambda) - \varphi(0) \leq \int_0^1 f(t)\varphi^\Delta(t) \Delta t.
\]

Since \( \lambda = \int_0^1 f(t) \Delta t \), thus

\[
(16) \quad \varphi \left( \int_0^1 f(t) \Delta t \right) \leq \int_0^1 f(t)\varphi^\Delta(t) \Delta t
\]
as required by Theorem 3.2.

Example 3.5. Consider a convex function \( \varphi(t) = t^2 \). Since \( \varphi^\Delta(t) = (t^2)^\Delta = t + \sigma(t) \), thus (16) yields

\[
\left( \int_0^1 f(t) \Delta t \right)^2 \leq \int_0^1 f(t)(t + \sigma(t)) \Delta t.
\]
But \( \sigma(t) = t \) for \( T = \mathbb{R} \). Hence

\[
(17) \quad \left( \int_0^1 f(t) \Delta t \right)^2 \leq 2 \int_0^1 tf(t) \Delta t.
\]

Example 3.6. Let

\[
f(t) = \begin{cases} 
  t^q & \text{for} \quad 0 < t \leq 1 \\
  0 & \text{elsewhere}
\end{cases}
\]

Thus from (17) we obtain a valid inequality

\[
\left( \int_0^1 t^q \Delta t \right)^2 \leq 2 \int_0^1 t^{q+1} \Delta t
\]

for all \( q \geq 0 \).

4. Conclusion

The paper presented the background together with extensions or generalisations of Steffensen’s Inequality. Time scales analysis was also presented. Moreover, further refinements of the Steffensen’s Inequality were established with illustrative examples in the context of time scales.

References


Mohammed Muniru Iddrisu  
Department of Mathematics,  
Faculty of Mathematical Sciences,  
University for Development Studies,  
P. O. Box 24,  
Navrongo Campus,  
Navrongo, Ghana.  
E-mail: mmuniru@uds.edu.gh