ON BELATED DIFFERENTIATION AND A CHARACTERIZATION OF HENSTOCK-KURZWEIL-ITO INTEGRABLE PROCESSES

TIN-LAM TOH, TUAN-SENG CHEW, Singapore

(Received March 29, 2004)

Abstract. The Henstock-Kurzweil approach, also known as the generalized Riemann approach, has been successful in giving an alternative definition to the classical Itô integral. The Riemann approach is well-known for its directness in defining integrals. In this note we will prove the Fundamental Theorem for the Henstock-Kurzweil-Itô integral, thereby providing a characterization of Henstock-Kurzweil-Itô integrable stochastic processes in terms of their primitive processes.

Keywords: belated differentation, Henstock-Kurzweil-Itô integral, integrable processes
MSC 2000: 26A39, 60H05

1. Introduction

The generalized Riemann approach, more commonly known as the Henstock-Kurzweil approach, has been successful in giving an alternative definition to the classical Itô integral, see [1], [5], [7], [8], [9], [10]. The advantage of using the Henstock-Kurzweil approach has been its explicitness and intuitiveness in giving a direct definition of the integral rather than the classical non-explicit $L^2$-procedure.

It is also well-known from the classical non-stochastic integration theory that all integrable functions can be characterized in terms of their primitives, that is, a function $f$ is Lebesgue (Henstock-Kurzweil) integrable on a compact interval $[a, b]$ if and only if there exists a function $F$ which is absolutely continuous (respectively, generalized absolutely continuous) there such that $F' = f$ a.e. on $[a, b]$, where $F'$ is the usual derivative of $F$, see for example [4].

In this paper, we will define the “belated derivative” of a stochastic process and thereby characterize the class of all Henstock-Kurzweil-Itô integrable processes on $[a, b]$ by its primitive process.
2. Setting

Let \( \Omega \) denote the set of all real-valued continuous functions on \([a, b]\) and \( \mathbb{R} \) the set of all real numbers.

The class of all Borel cylindrical sets \( B \) in \( \Omega \), denoted by \( \mathcal{C} \), is a collection of all sets \( B \) in \( \Omega \) of the form

\[
 B = \{ w: (w(t_1), w(t_2), \ldots, w(t_n)) \in E \}
\]

where \( 0 \leq t_1 < t_2 < \ldots < t_n \leq 1 \) and \( E \) is any Borel set in \( \mathbb{R}^n \) (\( n \) is not fixed). The Borel \( \sigma \)-field of \( \mathcal{C} \) is denoted by \( \mathcal{F} \), i.e., it is the smallest \( \sigma \)-field which contains \( \mathcal{C} \). Let \( P \) be the Wiener measure defined on \((\Omega, \mathcal{F})\). Then \((\Omega, \mathcal{F}, P)\) is a probability space, that is, a measure space with \( P(\Omega) = 1 \).

A stochastic process \( \{ \varphi(t, \omega): t \in [a, b] \} \) on \((\Omega, \mathcal{F}, P)\) is a family of \( \mathcal{F} \)-measurable functions (which are called random variables) on \((\Omega, \mathcal{F}, P)\). Very often, \( \varphi(t, \omega) \) is denoted by \( \varphi_t(\omega) \). Now we shall consider a very special and important process, namely, the Brownian motion denoted by \( W \).

Let \( W = \{W_t(\omega)\}_{a \leq t \leq b} \) be a canonical Brownian motion, that is, it possesses the following properties:

1. \( W_a(\omega) = 0 \) for all \( \omega \in \Omega \);
2. it has Normal Increments, that is, \( W_t - W_s \) has a normal distribution with mean \( 0 \) and variance \( t - s \) for all \( t > s \) (which naturally implies that \( W_t \) has a normal distribution with mean \( 0 \) and variance \( t \));
3. it has Independent Increments, that is, \( W_t - W_s \) is independent of its past, that is, of \( W_u, 0 \leq u < s < t \); and
4. its sample paths are continuous, i.e., for each \( \omega \in \Omega \), \( W_t(\omega) \) as a function of \( t \) is continuous on \([a, b]\).

A stochastic process \( \{ \varphi_t(\omega): t \in [a, b] \} \) is said to be adapted to the standard filtering space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)\) if \( \varphi_t \) is \( \mathcal{F}_t \)-measurable for each \( t \in [a, b] \). We always assume that \( W = \{W_t(\omega)\} \) is adapted to \( \{\mathcal{F}_t\} \). For example, if \( \{\mathcal{F}_t\} \) is the smallest \( \sigma \)-field generated by \( \{W_s(\omega): s \leq t\} \), then \( W = \{W_t(\omega)\} \) is adapted to \( \{\mathcal{F}_t\} \).

A stochastic process \( X = \{X_t(\omega): t \in [a, b]\} \) on the standard filtering space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)\) is called a martingale if

1. \( X \) is adapted to \( \{\mathcal{F}_t\} \), that is, \( X_t \) is \( \mathcal{F}_t \)-measurable for each \( t \in [a, b] \);
2. \( \int_{\Omega} |X_t| dP \) is finite for almost all \( t \in [a, b] \); and
3. \( E(X_t | \mathcal{F}_s) = X_s \) for all \( t \geq s \), where \( E(X_t | \mathcal{F}_s) \) is the conditional expectation of \( X_t \) given \( \mathcal{F}_s \).
If in addition
\[ \sup_{t \in [a,b]} \int_{\Omega} |X_t|^2 \, dP \]
is finite, we say that \( X \) is an \( L_2 \)-martingale.

In the following we define \( E(f) \) to be \( \int_{\Omega} f \, dP \) for any random variable \( f \).

It is well-known, see for example [6, P239], that the following assertions hold. The details are given for the convenience of readers who are not familiar with stochastic analysis.

(i) \[ E[X_s] = E[E[X_t|\mathcal{F}_s]] = E[X_t] \]
for any \( t \geq s \), that is, \( E[X_s] \) is a constant for all \( s \in [a,b] \).

(ii) For any \( a \leq u < v \leq s < t \leq b \), we have
\[ E[(X_t - X_s)(X_v - X_u)] = 0, \]
that is, a martingale has orthogonal increments.

(iii) From (ii) we get
\[ E\left(D \sum (X_v - X_u)^2 \right)^2 = (D) \sum E(X_v - X_u)^2 \]
for any partial partition \( D = \{[u,v] \} \) of \( [a,b] \).

(iv) For any \( u < v \) we have
\[ E[X_vX_u] = E[E[X_vX_u|\mathcal{F}_u]] = E[X_uE[X_v|\mathcal{F}_u]] = E[X_u^2] \]
and hence
\[ E(X_v - X_u)^2 = E(X_v^2 - X_u^2). \]

It is also well-known, see for example [6, P28], that a canonical Brownian motion is a martingale. In fact, it is an \( L_2 \)-martingale with \( E(W_t^2) = t \), see property 2 of a Brownian motion.
3. Differentiation

In this section we define our belated derivative and state its basic properties.

**Definition 1.** Let $F = \{F_t : t \in [a, b]\}$ be an $L^2$-martingale. A stochastic process $F$ is said to be *belated differentiable* at $t \in [a, b)$ if there exists a random variable $f_t$ such that for any $\varepsilon > 0$, there exists a positive number $\delta(t)$ on $[a, b]$ such that whenever $[t, v] \subset [t, t + \delta(t)]$, we have

$$E(|f_t(W_v - W_t) - (F_v - F_t)|^2) \leq \varepsilon E(W_v - W_t)^2 = \varepsilon |v - t|.$$ 

The random variable $f_t$ is called the *belated derivative* of $F$ at the point $t$. We will denote $f_t$ by $D\beta F_t$ in our subsequent presentation. It is also easily checked that the belated derivative of $F$ is defined uniquely up to a set of probability measure zero. The proof is omitted.

The $L^2$-martingale $F$ is said to be belated differentiable at $t \in [a, b)$ if $f_t$ in the above definition exists.

**Remark.** The word belated is used in the above definition because the point of differentiation $t$ is always the left end point of the interval $[t, v]$. This is motivated by the use of belated division in the definition of Henstock-Kurzweil-Itô integrals, see [1].

Next we shall state the standard properties of belated differentiation.

**Theorem 2.** Let $X$ and $Y$ be two $L^2$-martingales which are belated differentiable at $t \in [a, b)$ and let $\alpha \in \mathbb{R}$. Then

(a) $X + Y$ is belated differentiable at $t$ and

$$D\beta(X + Y)_t = (D\beta X)_t + (D\beta Y)_t,$$

(b) $\alpha X$ is belated differentiable at $t$ and

$$(D\beta(\alpha X))_t = \alpha(D\beta X)_t.$$

**Proof.** The proof of Theorem 2 is straightforward and hence omitted. \qed

**Example 3.** Let $X = \{X_t : t \in [0, 1]\}$ be the stochastic process $X_t = \frac{1}{2}W_t^2 - \frac{1}{2}t$, where $W$ is the Brownian motion, over the standard filtering space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$. Then it is easy to verify that $X$ is in fact an $L^2$-martingale with respect to the standard filtering space. Furthermore, it can be proved that

$$D\beta X_t = W_t$$

for all $t \in [a, b]$. 

66
That $X$ is a martingale follows from the fact that

$$E(W_b^2 - W_a^2 | \mathcal{F}_a) = b - a$$

where $0 \leq a \leq b$. Furthermore,

$$E(|X_t|^2) = \frac{1}{2} t^2 \leq \frac{1}{2} b^2$$

for all $t \in [a, b]$, thereby showing that $X$ is in fact an $L^2$-martingale. We next show that $D_\beta X_t = W_t$ for all $t \in [a, b]$.

Given $\varepsilon > 0$, let $\delta(t) \leq 2\varepsilon$ for all $t \in [a, b)$. Consider a $\delta$-fine interval-point pair $([t, v], t)$ such that $[t, v] \subset [t, t + \delta(t)]$ so that $|v - t| \leq 2\varepsilon$. Then

$$E(W_t(W_v - W_t) - (X_v - X_t))^2 = E\left(W_t(W_v - W_t) - \frac{1}{2}(W_v^2 - W_t^2) - \frac{1}{2}(t - v)\right)^2$$

$$= E\left(\frac{1}{2}(W_v - W_t)^2 + \frac{1}{2}(t - v)\right)^2$$

$$= \frac{1}{4} E((W_v - W_t)^2 - (v - t))^2$$

$$= \frac{1}{4} E((W_v - W_t)^4 - 2(W_v - W_t)^2(v - t) + (v - t)^2)$$

$$= \frac{1}{2}(v - t)^2 \leq \frac{1}{2} \cdot 2\varepsilon(v - t) = \varepsilon(v - t),$$

which completes our proof. \qed

By Definition 1, belated differentiation is defined for $L^2$-martingales in our context. If we were to allow the belated differentiation to be defined for more general stochastic processes, we could even have $D_\beta(\frac{d}{dt} W_t^2) = W_t$. However, in this sense, the anti-derivative of $W_t$ would not be uniquely defined. Hence we restrict ourselves to the belated differentiation of $L^2$-martingales.

**Definition 4.** A stochastic process $X = \{X_t : t \in [a, b]\}$ on $(\Omega, \mathcal{F}, P)$ is said to be $AC^2$ on $[a, b]$ if given any $\varepsilon > 0$, there exists $\eta > 0$ such that

$$E\left(\sum_{i=1}^n (X_{v_i} - X_{u_i})^2\right) \leq \varepsilon$$

for any finite collection $D = \{[u_i, v_i]\}_{i=1}^n$ of non-overlapping intervals for which $\sum_{i=1}^n |v_i - u_i| \leq \eta$. 67
Example 5. The stochastic process \( X = \{X_t: t \in [a, b]\} \), where

\[
X_t = \frac{1}{2} W_t^2 - \frac{1}{2} t
\]

in Example 3, is \( AC^2 \) on \([a, b]\). The proof is easy and hence omitted.

4. Antiderivative and Henstock-Ito integral

In this section we will characterize the class of all Henstock-Itô adapted processes in terms of their derivatives.

Let \( \delta \) be a positive function on \([a, b]\). A finite collection \( D \) of interval-point pairs \( \{([\xi_i, v_i], \xi_i), i = 1, 2, \ldots, n\} \) is called a \( \delta \)-fine belated partial division of \([a, b]\) if

1. \( \{[\xi_i, v_i], i = 1, 2, \ldots, n\} \) is a collection of non-overlapping subintervals of \([a, b]\); and
2. \( [\xi_i, v_i] \subseteq [\xi_i, \xi_i + \delta(\xi_i)] \) for each \( i = 1, 2, 3, \ldots, n \).

In the sequel we will denote \( \{([\xi_i, v_i], \xi_i), i = 1, 2, 3, \ldots, n\} \) by \( \{(\xi, v, \xi)\} \).

Definition 6 (See [1, Definition 2]). Let \( f = \{f_t: t \in [a, b]\} \) be an adapted process on the standard filtering space \( (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P) \). Then \( f \) is said to be Henstock-Kurzweil-Itô integrable on \([a, b]\) if there exists a process \( F = \{F_t: t \in [a, b]\} \) which is an \( L^2 \)-martingale and \( AC^2 \) on \([a, b]\) such that for any \( \varepsilon > 0 \) there exists a positive function \( \delta \) on \([a, b]\) such that

\[
E\left( \sum_{D} \{f_t(W_v - W_u) - (F_v - F_u)\}^2 \right) \leq \varepsilon
\]

whenever \( D = \{([\xi, v, \xi])\} \) is a \( \delta \)-fine belated partial division of \([a, b]\).

It follows from Vitali’s Covering Lemma that given any positive function \( \delta \) there exists a belated partial division of \([a, b]\) covering this interval up to a set of arbitrarily small positive measure, hence the uniqueness of the integral process \( F \) follows.

It was also proved in [1] that the standard properties of integrals (such as uniqueness of the integral, additivity of the integral, integrability over subintervals) hold true for the Henstock-Kurzweil-Itô integral. The proofs are similar to the classical integration theory, see [2], [3], [4]. In fact, it has been proved in Theorem 9 of [1] that the integral defined by this new approach is equivalent to the classical Itô integral.

We have a class of stochastic processes which are Henstock-Kurzweil-Itô integrable on \([a, b]\).
Example 7. Let $L_2$ denote the class of all adapted stochastic processes $\varphi = \{ \varphi_t: t \in [a, b] \}$ on the standard filtering space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ such that

$$\int_a^b E|\varphi(t, \omega)|^2 \, dt$$

is finite. Then any adapted process from $L_2$ is Henstock-Kurzweil-Itô integrable on $[a, b]$.

In fact, $L_2$ is the class of all classical Itô integrable functions. We have proved in [1] that if $f$ is classical Itô integrable, then $f$ is also Henstock-Kurzweil-Itô integrable and the two integrals coincide.

Theorem 8. Let an adapted process $f$ be Henstock-Kurzweil-Itô integrable on $[a, b]$ and let $F_t = \int_a^t f_s \, dW_s$. Then $D_\beta F_t = f_t$ a.e. on $[a, b]$.

Proof. The idea of this proof is motivated by that of the Henstock integration theory. We need to show that the set of points $B$ of $[a, b]$ for which $D_\beta F_t$ does not exist or is unequal to $f$ is of Lebesgue measure zero. Let $t \in B$. By definition, there exists $\gamma(t) > 0$ such that for any positive number $\delta(t)$, there exists $[t, v] \subset [t, t + \delta(t)]$ such that

$$E(\|f_s(W_v - W_t) - (F_v - F_t)^2\| > \gamma(t)(v - t)).$$

From the definition of the Henstock-Kurzweil-Itô integral (see Definition 6), given $\varepsilon > 0$, there exists a positive function $\beta$ on $[a, b]$ such that whenever $D = \{(\xi_i, v_i, \xi_i)\}_{i=1}^n$ is a $\beta$-fine belated partial division of $[a, b]$, we have

$$E\left(\sum_{i=1}^n |f_{\xi_i} W_{v_i} - W_{\xi_i} - (F_{v_i} - F_{\xi_i})^2| \right) \leq \varepsilon.$$

Now we consider a special $D$ such that each $[\xi_i, v_i]$ satisfies (1) and (2). Let us denote $B_m = \{ t \in [a, b]: \gamma(t) \geq \frac{1}{m} \}$, $m = 1, 2, 3, \ldots$, and fix $B_m$. Suppose each $\xi_i \in B_m$. Then by (1) and (2), we have

$$\sum_{i=1}^n (v_i - \xi_i) \leq m \varepsilon.$$

Let $\mathcal{G}$ be the family of collections of intervals $[\xi, v]$ induced from all $\beta$-fine belated partial divisions with $\xi \in B_m$ satisfying (1). Then $\mathcal{G}$ covers $B_m$ in Vitali’s sense. Applying the Vitali Covering Theorem, there exists a finite collection of intervals $\{[\xi_i, v_i], i = 1, 2, 3, \ldots, q\}$ such that

$$\mu(B_m) \leq \sum_{i=1}^q |v_i - \xi_i| + \varepsilon \leq (m + 1) \varepsilon.$$

Hence $\mu(B_m) = 0$ and so $\mu(B) = 0$. Thus our proof is completed. \qed
Theorem 9. Let $f$ be an adapted process on $[a, b]$ such that

(i) $F$ is an $L^2$-martingale with $F_a = 0$ a.e.;

(ii) $F$ has the $AC^2$ property;

(iii) $D\beta F_t = f_t$ a.e. on $[a, b]$; then $f$ is Henstock-Kurzweil-Itô integrable on $[a, b]$ with $F_t = \int_a^t f_s \, dW_s$.

The reader is reminded that (iii) means that $D\beta F_t(\omega) = f_t(\omega)$ for almost all $\omega \in \Omega$ for a.e. $t \in [a, b]$.

Proof. Let $D\beta F_t = f_t$ for all $t \in [a, b]$ except possibly for a set $B$ which has Lebesgue measure zero. Let $\xi \in [a, b] \setminus B$. Given $\varepsilon > 0$, there exists a positive function $\delta$ on $[a, b]$ such that whenever $(\xi, v)$ is $\delta$-fine, we have

$$E(|f_\xi(W_v - W_\xi) - (F_v - F_\xi)|^2) \leq \varepsilon |v - \xi|.$$

Let $D = \{(\xi_i, v_i, \xi_i), i = 1, 2, 3, \ldots, n\}$ be a $\delta$-fine belated partial division of $[a, b]$ with all $\xi_i \in [a, b] \setminus B$. Then

$$E\left(\left|\sum_{i=1}^{n} f_\xi(W_{v_i} - W_{\xi_i}) - (F_{v_i} - F_{\xi_i})\right|^2\right)$$

$$= E\left(\sum_{i=1}^{n} |f_\xi(W_{v_i} - W_{\xi_i}) - (F_{v_i} - F_{\xi_i})|^2\right)$$

$$\leq \varepsilon \sum_{i=1}^{n} |v_i - \xi_i| \leq \varepsilon (b - a).$$

Thus if $B = \varnothing$, it is clear from the above that $f$ is Itô integrable with $F_t = \int_a^t f_s \, dW_s$. In general, $B$ is nonempty with $\mu(B) = 0$.

Now let

$$B_m = \{t \in [a, b]: m - 1 < E[f_t^2] \leq m\},$$

where $\mu(B_m) = 0$ and $B = \bigcup_{m=1}^{\infty} B_m$.

Since $F$ has the $AC^2$ property, given any positive integer $m$, there exists $\eta_m > 0$ with $\eta_m \leq (\varepsilon/2m)^2 \cdot m^{-2}$ such that whenever $\{(u_i, v_i)\}$ is a finite collection of disjoint left-open subintervals of $[a, b]$ with $\sum |v_i - u_i| \leq \eta_m$, we have

$$E\left(\left|\sum [F_{v_i} - F_{u_i}]\right|^2\right) \leq \left(\frac{\varepsilon}{2m}\right)^2.$$ 

Take an open set $G_m \supset B_m$ such that $\mu(G_m) \leq \eta_m$.

70
Fix a positive integer $m$. Let $D = \{((\xi_i, v_i), \xi_i)\}$ be a $\beta$-fine belated partial division of $[a, b]$ such that $\xi_i \in B_m$ for all $i$. Then we have

$$E\left(\left| \sum_i f_{\xi_i} [W_{v_i} - W_{\xi_i}] - (F_{v_i} - F_{\xi_i}) \right|^2 \right)$$

$$\leq 2E\left(\left| \sum_i f_{\xi_i} (W_{v_i} - W_{u_i}) \right|^2 \right) + 2E\left(\left| \sum_i (F_{v_i} - F_{\xi_i}) \right|^2 \right)$$

$$\leq 2\sum_i E[f_{\xi_i}^2](v_i - u_i) + 2\left( \frac{\epsilon}{2m} \right)^2 \leq 4\left( \frac{\epsilon}{2m} \right)^2.$$ 

So, considering any $\beta$-fine belated partial division over $[a, b]$, denoted by $D_1 = \{((\xi_i, v_i), \xi_i)\}$, we have

$$E\left(\left| \sum_i f_{\xi_i} (W_{v_i} - W_{\xi_i}) - (F_{v_i} - F_{\xi_i}) \right|^2 \right)$$

$$\leq 2E\left(\left\{ \sum_{\xi \in [a, b]/B} f_{\xi_i} (W_{v_i} - W_{\xi_i}) - (F_{v_i} - F_{\xi_i}) \right\}^2 \right)$$

$$+ 2E\left(\left\{ \sum_{m=1}^{\infty} \sum_{\xi_i \in B_m} f_{\xi_i} (W_{v_i} - W_{\xi_i}) - (F_{v_i} - F_{\xi_i}) \right\}^2 \right)$$

$$\leq 2\varepsilon(b - a) + 2\varepsilon,$$

thereby showing that $f$ is Itô integrable with $F_t = \int_a^t f_t \, dW_t$. 

Combining Theorems 8 and 9, we have the following characterization of all Henstock-Kurzweil-Itô integrable stochastic processes:

**Theorem 10.** Let $f$ be an adapted process on $[a, b]$. Then $f$ is Henstock-Kurzweil-Itô integrable on $[a, b]$ if and only if there exists an $L^2$-martingale $F$ on $[a, b]$ with $F_a = 0$ a.s. and $AC^2$ on $[a, b]$ such that $D_\beta F_t = f_t$ almost everywhere on $[a, b]$.

**Example 11.** From Example 3, $X_t = \frac{1}{2}W_t^2 - \frac{1}{4}t$ is an $L^2$-martingale on $[a, b]$. Hence the process

$$F_t = \frac{1}{2}W_t^2 - \frac{1}{2}W_a^2 - \frac{1}{2}(t - a),$$

where $F_a = 0$, is an $L^2$-martingale on $[a, b]$. It can be also easily verified that $F$ is $AC^2$ on $[a, b]$. Furthermore, it was shown that

$$D_\beta X_t = W_t$$

71
on \([a, b]\), hence

\[ D_B F_t = W_t \]
on \([a, b]\). By Theorem 10, we have

\[ \int_a^b W_t \, dW_t = F_b = \frac{1}{2} \left( W_b^2 - W_a^2 \right) - \frac{1}{2} (b - a). \]

**Example 12.** Let \( f \in \mathcal{L}_2 \), the class of all classical Itô integrable adapted processes on the standard filtering space. Then there exists an \( L^2 \)-martingale \( F \) on \([a, b]\) which is also \( AC^2 \) on \([a, b]\), such that \( D_B F_t = f_t \) a.e. on \([a, b]\).

**References**


**Author’s address:** Tin-Lam Toh, National Institute of Education, 1, Nanyang Walk, Singapore 637616, Republic of Singapore, e-mail: tltoh@nie.edu.sg.