Expansion of Quadratic Forms of Brownian Motion
with Econometric Application

Chaohua Dong
School of Economics  The University of Adelaide

Abstract
Since Merton’s well-known seminar work in 1970s and Black and Scholes’ ground-breaking paper in the same period were published, Brownian motion has been engaged to depict movements of natural phenomenon and human stochastic activities. In most cases, random variables involved are supposed to be adapted with respect to Brownian motion, that is, they are functions of Brownian motion. In both theoretical research and econometric application it is very convenient and much helpful if one has expansions of Brownian motion and its functions. Indeed in literature there are several versions of expansion of Brownian motion, for example, Paley-Wiener representation and Levy-Ciesielski representation. However, for functions of Brownian motion, there is no alike study. This partially brings about difficulty of application for Brownian motion. This study aims at finding an expansion of quadratic forms of Brownian motion, and for general forms their local approximations are given as well. Besides, An application of the results is discussed at the same time.

Key words: Brownian motion, expansion, orthonormal basis, function space, random variable space

1. Introduction

It is well known that at the moment the world economy is experiencing a global financial crisis. Most of countries have to spend a large sums of money to tackle the problem. For example, in the United States of America government issued a stimulus package of 787 billion dollars ([1]). In addition, president Obama said the world largest economy has lost 4.4 million jobs since recession began 14 months ago, so that unemployment rate rose sharply to 8.1% in February, the highest in a quarter century ([1]). More than one million homeowners have already lost their homes since mortgage crisis began last summer ([2]). Factories in Asia have
lost their orders, hence previous booming economies have slid into recession. For instance, in China government statistics showed that 67,000 factories of various sizes were shuttered in the first half of the year 2008, and by the end of the year, over 100,000 factories were closed ([3]). Financial crisis strikes every corner of the world. This international turmoil suggests a further examination is necessary.

It is clear that the leading factor in global financial crisis is in finance. Approximately 80% U.S. mortgages issued in the recent years to subprime borrowers were adjustable-rate mortgages. When America house prices began to decline in 2006-07, refinancing became more difficult, and as adjustable-rate mortgages began to reset at a higher rate, mortgage delinquencies soared. Securities backed with subprime mortgages, widely held by financial firms, lost most of their value. The result has been a large decline in capital of many banks and American government sponsored enterprises, tightening credit around the world. Economies in other part of the world suffer from a chain reaction of this crisis in America ([4]). Therefore, an investigation in financial market is required to explain the financial crisis.

Financial industry has been called the blood system in economy. With development of economy, financial industry definitely becomes prosperous. As can be seen that in New York, London, Frankfurt, Singapore, Tokyo, Hongkong, Shanghai there are considerable stock exchange institutes abounded with people in different ages, religions and colors. More dangerous is that there are sizeable financial derivatives available on the market and they are utilized by some laymen that may produce a disastrous consequences since financial derivatives are actually two-edged sword. To rein modern finance, so-called 'wild horse', the most efficient and powerful tool is mathematical finance, because this branch of mathematics unveils mathematical principle behind finance.

In the last half century, quantitative researches in finance have been getting more and more recognitions, especially in 1990s of 20 century Asian financial storm made people realize that to rein the wild horse, modern finance, is impossible if there was no quantitative methods. Apparently quantitative methods depend naturally on mathematics. Nowadays mathematical finance has entered into its maturity that makes it possible to apply mathematics extensively and profoundly in finance. In 1990 Harry M. Markowitz, Merton H. Miller and William F. Sharpe won Nobel prize in Economic Science because of their excellent contribution in the theory of portfolio selection; in 1997 Robert C. Merton and Nyron S.
Scholes won Nobel prize in Economic Science due to their famous formula of option pricing in financial market. These two works established corner stones for mathematical finance. Thus, as a science, mathematical finance emerges as times require and plays a central role in the field of finance.

The main difference between modern finance and the traditional one is that in modern finance there are a great number of continuous-time issues to be dealt with. Continuous-time modeling in modern finance, though introduced by Louise Bachelier’s thesis in 1900 on the theory of speculation, really started with Merton’s seminar work ([5]) in 1970s and Black and Scholes’ ground-breaking paper ([6]) in the same period. Since then, the continuous-time paradigm has prove to be an immensely useful tool in finance and more generally economies ([7],[8],[9]). Continuous-time models are widely employed to study issues that include the decision of optimal consumption, saving and investment, portfolio choice under a variety constraints, contingent claim pricing, capital accumulation, resource extraction, game theory and more recently contract theory ([10], [11], [12], [13]). A time series model used extensively in finance is continuous-time diffusion, or Ito process ([7], [17], [18]). In modeling the dynamics of the short-term riskless interest rate process \( X_t \), for example, the applicable diffusion process is

\[
dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dB_t
\]  

(1.1)

where \( B_t \) is standard Brownian motion and \( \mu(\cdot, \theta) \) and \( \sigma(\cdot, \theta) \) with unknown parameter \( \theta \) are the drift and volatility functions of the process respectively. Usually the process \( X_t \) is required to be adapted with respect to Brownian motion, that is, it is a function of Brownian motion up to the epoch in question. Hence drift function and volatility function are all of Brownian motion. To make use or to solve equation (1.1), one has to know the explicit or at least the approximate forms of the functions.

In literature, there are several versions of decomposition of Brownian motion ([16],[17]) due to different choice of bases in \( L^2 \) space. However, there is no expansion of functions of Brownian motion so far. Since expansions of functions of Brownian motion have potentially a variety application in solving equation (1.1) and in modeling two random variables with unknown relationship, this paper aims at finding such expansion. Nevertheless, the difficulty of decomposing function \( f(B_t) \) into orthogonal series is overwhelming because of the arbitrariness of \( f(\cdot) \). This paper starts with a simple function \( f(x) = x^2 \), then after getting the decomposition of \( B_t^2 \), a local approximation to general function of Brownian motion, \( f(B_t) \),
is given.

In the rest of the paper, it is organized as follows. Section 2 states all the preliminary definitions, critical existing results that will be used in later sections; in section 3, the expansion of $B^{2}\tau$ is derived, which is expanded in terms of orthonormal base in random variable space $L^{2}(\Omega)$. Also two examples are provided; approximation to general functions of Brownian motion by polynomial of Brownian motion with degree less than 3 is developed in section 4. An application of the results in econometrics is discussed in section 5 and section 6 is conclusion.

2. Preliminaries

In this section some existing results and pertinent contents are stated for completeness and concreteness, which will be used in later sections.

2.1 Measure Space and Probability Space

Some basic definitions are essential for following discussion, hence it is necessary to introduce here what definitions we need later.

**Definition 2.1.** Let $\mathcal{F}$ be a collection of subsets of a set $\Omega$. Then $\mathcal{F}$ is called a field iff $\Omega \in \mathcal{F}$ and $\mathcal{F}$ is closed under complementation and finite union, that is

1. $\Omega \in \mathcal{F}$;
2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$;
3. If $A_1, \ldots, A_n \in \mathcal{F}$, then $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$.

If condition (3) is replaced by closure under countable union, that is, if $A_1, A_2, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, we call $\mathcal{F}$ a $\sigma$-field.

It is clear that the largest $\sigma$-field of subsets of a fixed set $\Omega$ is the collection of all subsets of $\Omega$ (also called power set). The simplest $\sigma$-field is $\{\Omega, \emptyset\}$.

If $\mathcal{G}$ is a class of subsets of $\Omega$, the smallest $\sigma$-field containing all sets in $\mathcal{G}$ is called minimal $\sigma$-field over $\mathcal{G}$, denoted by $\sigma(\mathcal{G})$. It is also called the $\sigma$-field generated by $\mathcal{G}$. Obviously, $\{\Omega, \emptyset\}$ is the $\sigma$-field generated by $\Omega$ (or by $\emptyset$). If $A$ is a subset of $\Omega$ which is neither $\Omega$ nor $\emptyset$, the $\sigma$-field generated by $A$ is $\{\Omega, \emptyset, A, A^c\}$.
Example 2.1 Denote by $\mathcal{B}(\mathbb{R})$ the $\sigma$-field generated by collection of all intervals $(a, b], a, b \in \mathbb{R}$. Note that $\mathcal{B}(\mathbb{R})$ is guaranteed to exist; it may be described (admittedly in a rather ethereal way) an intersection of all $\sigma$-fields containing intervals $(a, b]$ (at least there is one such $\sigma$-field, power set). Also, if a $\sigma$-field contains all open intervals, it contains all intervals $(a, b]$ and conversely, since

$$(a, b] = \bigcap_{n=1}^{\infty} \left( a, b + \frac{1}{n} \right) \quad \text{and} \quad (a, b) = \bigcup_{n=1}^{\infty} \left( a, b - \frac{1}{n} \right)$$

Thus $\mathcal{B}(\mathbb{R})$ is the smallest $\sigma$-field containing all open intervals. \hfill \Box

If $\mathcal{F}$ is a $\sigma$-field of subsets of $\Omega$, $(\Omega, \mathcal{F})$ is called a measurable space; the sets in $\mathcal{F}$ are called measurable sets.

**Definition 2.2.** A measure on a $\sigma$-field $\mathcal{F}$ is a nonnegative, extended real-valued set function $\mu$ on $\mathcal{F}$ such that whenever $A_1, A_2, \ldots$ form a finite or countable infinite collection of disjoint sets in $\mathcal{F}$, we have

$$\mu \left( \bigcup_n A_n \right) = \sum_n \mu(A_n) \quad (2.1)$$

If $\mu(\Omega) = 1$, $\mu$ is called a probability measure.

A measure space is a triple $(\Omega, \mathcal{F}, \mu)$ where $\Omega$ is a set, $\mathcal{F}$ is a $\sigma$-field of subsets of $\Omega$, $\mu$ is a measure on $\mathcal{F}$. If $\mu$ is a probability measure, $(\Omega, \mathcal{F}, \mu)$ is called a probability space.

Note that countable additivity actually implies finite additivity. For if $\mu(A) = +\infty$ for all $A \in \mathcal{F}$, or $\mu(A) = -\infty$ for all $A \in \mathcal{F}$, the result is immediate; therefore assume $\mu(A)$ finite for some $A \in \mathcal{F}$. By considering sequence $A, \emptyset, \emptyset, \ldots$, we find that $\mu(\emptyset) = 0$, and finite additivity is now established by considering sequence $A_1, \ldots, A_n, \emptyset, \emptyset, \ldots$ where $A_1, \ldots, A_n$ are disjoint sets in $\mathcal{F}$.

**Example 2.2** Let $\Omega = \{x_1, x_2, \ldots\}$ be a finite or countable infinite set, and let $p_1, p_2, \ldots$ be nonnegative numbers. Take $\mathcal{F}$ as all subsets of $\Omega$ and define for all $A \in \mathcal{F}$

$$\mu(A) = \sum_{x_i \in A} p_i.$$ 

Thus if $A = \{x_1, x_2, \ldots\}$, then $\mu(A) = p_{i_1} + p_{i_2} + \cdots$. The set function $\mu$ therefore is a measure on $\mathcal{F}$ and $\mu(\{x_i\}) = p_i$ for $i = 1, 2, \ldots$. A probability measure will be obtained iff $\sum_i p_i = 1$; if all $p_i = 1$, the measure is counting measure. \hfill \Box

2.2 Hilbert Space $L^2[a, b]$ and $L^2(\Omega)$

**Definition 2.3.** Given a set $S$, a pseudometric for $S$ is a function $d$ from $S \times S$ into $[0, \infty)$ such that
• for all $x \in S$, $d(x, x) = 0$;

• for all $x, y \in S$, $d(x, y) = d(y, x)$ (symmetry);

• for all $x, y$ and $z$ in $S$, $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality); if also

• $d(x, y) = 0$ implies $x = y$, then $d$ is called a metric.

Let $f$ be a function from a set $S$ into a set $Y$. Then for any subset $A$ of $Y$, let

\[ f^{-1}(A) := \{ x \in S : f(x) \in A \}. \]

This $f^{-1}(A)$ sometimes is called the inverse image of $A$ under $f$. Note that $f$ is not necessary to be $1-1$, hence $f^{-1}$ need not be a function. The inverse image preserves all unions and intersections: for any nonempty collection \( \{ B_i \}_{i \in I} \) of subsets of $B$,

\[ f^{-1} \left( \bigcap_{i \in I} B_i \right) = \bigcap_{i \in I} f^{-1}(B_i) \quad \text{and} \quad f^{-1} \left( \bigcup_{i \in I} B_i \right) = \bigcup_{i \in I} f^{-1}(B_i). \]

**Definition 2.4.** Suppose that $(S, \mathcal{S})$ and $(Y, \mathcal{Y})$ are two measurable spaces and $f$ is a function from $S$ into $Y$. $f$ is called measurable if $f^{-1}(A) \in \mathcal{S}$ for all $A \in \mathcal{Y}$.

Now we are in the position to introduce some function spaces.

**Definition 2.5.** For any measure space $(S, \mathcal{S}, \mu)$ and $0 < p < \infty$, $L^p(S, \mathcal{S}, \mu) = L^p(S, \mathcal{S}, \mu, \mathbb{R})$ denotes the set of all measurable functions $f$ on $X$ such that

\[ \int |f|^p \, d\mu < \infty \quad (2.2) \]

and the values of $f$ are real numbers except possibly on a set of measure 0, where $f$ may be undefined or infinite. For $1 \leq p < \infty$, let

\[ \|f\|_p := \left( \int |f|^p \, d\mu \right)^{1/p}, \quad (2.3) \]

called the "$L^p$ norm" or $p$-norm of $f$.

$L^p$ spaces will also be defined for $p = +\infty$. A measurable function $f$ is called essentially bounded if for some $M > 0$, $|f| \leq M$ almost everywhere. The set of all essentially bounded real functions is called $L^\infty(S, \mathcal{S}, \mu)$. Let $\|f\|_\infty := \inf \{ M : |f| \leq M, \text{a.e.} \}$. A pseudometric on $L^p$ for $p \geq 1$ will be defined by $d(f, g) := \|f - g\|_p$.

**Definition 2.6.** Let $S$ be a real vector space. A seminorm on $S$ is a function $\| \cdot \|$ from $S$ into $[0, \infty)$ such that

• $\|cx\| = |c|\|x\|$ for $\forall c \in \mathbb{R}$ and $\forall x \in S$;
\[ \|x + y\| \leq \|x\| + \|y\| \text{ for } \forall x, y \in S \]

A seminorm \( \| \cdot \| \) is called norm iff \( \|x\| = 0 \) only for \( x = 0 \).

For any set \( S \) with a pseudometric \( d \), let \( x \sim y \) iff \( d(x, y) = 0 \). Let \( S^\sim \) be the set of all equivalent classes \( x^\sim \) for all \( x \in S \). Let \( d(x^\sim, y^\sim) := d(x, y) \) for all \( x, y \in X \). It is easy to see that \( d \) is well-defined and a metric on \( S^\sim \). If \( S \) is a vector space with a seminorm \( \| \cdot \| \), then \( \{ x \in S : \|x\| = 0 \} \) is a vector subspace \( Z \) of \( S \). For each \( x \in S \) let \( x^\sim := \{ y : y \sim x \} \). Let \( S^\sim = \{ x^\sim : x \in S \} \), often called the quotient space or factor space \( S/Z \), is then in a natural way also a vector space, on which \( \| \cdot \| \) defines a norm. For each space \( L^p \), the factor space of equivalent classes defined in this way is called \( L^p \), that is, \( L^p(S, S, \mu) := \{ f^\sim : f \in L^p(S, S, \mu) \} \).

If \( S \) is a vector space and \( \| \cdot \| \) is a norm on it, then \( (S, \| \cdot \|) \) is called a normed linear space. A Banach space is a normed linear space which is complete, for the metric defined by the norm.

**Theorem 2.1 (R.M.Dudley, theorem 5.2.1).** For any measure space \((S, S, \mu)\) and \(1 \leq p \leq \infty\), \((L^p(S, S, \mu), \| \cdot \|)\) is a Banach space.

Let \( H \) be a vector space over a field \( K \), where \( K = \mathbb{R} \) or \( \mathbb{C} \).

**Definition 2.7.** A semi-inner product on \( H \) is a function \((\cdot, \cdot)\) from \( H \times H \) into \( K \) such that

- \((cf + g, h) = c(f, h) + (g, h)\) for all \( c \in K \) and \( f, g, h \in H \);
- \((f, g) = (g, f)\) for complex numbers;
- \((\cdot, \cdot)\) is nonnegative definite, i.e. \((f, f) \geq 0\) for all \( f \in K \).

A semi-inner product is called an inner product if \((f, f) = 0\) implies \( f = 0 \). A (semi-)inner product space is a pair \((V, (\cdot, \cdot))\) where \( V \) is a vector space over \( K \) and \((\cdot, \cdot)\) is a (semi-)inner product on \( V \).

Suppose \((H, (\cdot, \cdot))\) is an inner product space. Then \( H \) is called Hilbert space if it is complete for the metric defined by the norm \( d(x, y) = \|x - y\| = (x - y, x - y)^{1/2} \). Let \((S, S, \mu)\) be a measure space. Then \( L^2(S, S, \mu) \) is a Hilbert space with the inner product \((f^\sim, g^\sim) = (f, g) = \int fg d\mu\). Particularly, \( S = [a, b] \), \( S \) the class of Borel sets of \([a, b]\) and \( \mu \) being Lebesgue measure give function space \( L^2[a, b] \), which is a Hilbert space consisting
of all square integrable functions on \([a, b]\) with inner product \((f, g) = \int_a^b fg\) and norm \(\|f\|_2 = \left(\int_a^b f^2 dx\right)^{1/2} \).

Another concrete example of Hilbert space is \(L^2(\Omega) = L^2(\Omega, \mathcal{F}, P)\), where \((\Omega, \mathcal{F}, P)\) is a probability space. The elements of \(L^2(\Omega)\) are all equivalent classes and each equivalent class consists of random variables which agree almost everywhere. The inner product and norm in \(L^2(\Omega)\) are denoted by particular notations
\[
\langle X, Y \rangle = E[X Y], \quad \|X\|_2(\Omega) = \left(E[\|X\|^2]\right)^{1/2}
\]
for \(X, Y \in L^2(\Omega)\).

The \(\Omega\) following \(\| \cdot \|\) only appears in the place to distinguish from other norm in different space, hence it will be dropped if not necessary.

It would follow quite easily from Cauchy-Bunyakovsky-Schwarz inequality that the inner product \((x, y)\) is continuous in \(x\) for fixed \(y\), and continuous in \(y\) for fixed \(x\). Following theorem is for joint continuity.

**Theorem 2.2** (R.M.Dudley, Theorem 5.3.5). For any inner product space \((H, (\cdot, \cdot))\), the inner product is joint continuous, that is, continuous for the product topology from \(H \times H\) into \(K\).

**Proof.** Given \(u\) and \(v\) in \(K\), we have as \(x \to u\) and \(y \to v\)
\[
|\langle x, y \rangle - \langle u, v \rangle| = |\langle x, y \rangle - \langle u, y \rangle + \langle u, y \rangle - \langle u, v \rangle| = |\langle x - u, y \rangle + \langle u, y - v \rangle|
\leq \|x - u\|\|y\| + \|u\|\|y - v\|
\leq \|x - u\|(|y - v| + \|v\|) + \|u\|\|y - v\|
\to 0
\]
since \(\|x - u\| \to 0\) and \(\|y - v\| \to 0\).

We say two elements \(f\) and \(g\) are orthogonal or perpendicular, written \(f \perp g\), iff \((f, g) = 0\).

A set \(\{e_i\}_{i \in \mathbb{N}}\) in semi-inner product space \(H\) is called orthonormal iff \((e_i, e_j) = 1\) whenever \(i = j\), and 0 otherwise. Given an inner product space \((H, (\cdot, \cdot))\), an orthonormal basis for \(H\) is an orthonormal set \(\{e_\alpha\}\) such that for each \(x \in H\), \(x = \sum_\alpha \langle x, e_\alpha \rangle e_\alpha\).

A significant result is that every Hilbert space has an orthonormal basis. Two examples of orthonormal bases in \(L^2[0, \pi]\) and \(L^2[0, T]\) respectively will be given in Example 3.1 and 3.2.
2.3 Stochastic Process

A stochastic process is by definition a collection of random variables \((X_t, t \in D)\) defined on a probability space \((\Omega, \mathcal{F}, P)\) where the index set \(D\) is a subset of real line \(\mathbb{R}\). Thus a stochastic process \(X\) is a real-valued function \(X(t, \omega)\) on \(D \times \Omega\) which is a \(\mathcal{F}\)-measurable function on \(\Omega\) for each \(t \in D\). For our purpose, \(D\) is often an interval, \(D = [a, b]\) or \([a, \infty)\) for \(a < b\). In this case we then call \(X\) a continuous-time process. If \(D\) is finite set or \(\mathbb{N}\), the process is simply a sequence \(X_n\), which is classified as discrete-time process. For obvious reasons, the index \(t\) of random variable \(X_t\) is frequently referred to as time.

We summarize the definition of stochastic process as follows. A stochastic process \(X\) is a function of two variables. For a fixed instant time \(t\), it is a random variable \(X_t = X_t(\omega), \omega \in \Omega\); for a fixed random outcome \(\omega \in \Omega\), it is a function of time \(t, X(\omega) = X_t(\omega), t \in D\). This function is called a realization, a trajectory, or a sample path of process of \(X\).

**Definition 2.8.** Two stochastic processes \(X\) and \(Y\) on probability space \((\Omega, \mathcal{F}, P)\) and index set \(D\) are called equivalent if for every \(t \in D\), \(X(t, \omega) = Y(t, \omega)\) for a.e. \(\omega\), that is, there exists a set \(\Lambda_t \in \mathcal{F}\) with \(P(\Lambda_t) = 0\) such that \(X(t, \omega) = Y(t, \omega)\) for \(\omega \in \Lambda_t^c\). \(X\) and \(Y\) are said to be almost surely equal if the sample functions of \(X\) and \(Y\) are identical for a.e. \(\omega\), that is, there exists a set \(\Lambda \in \mathcal{F}\) with \(P(\Lambda) = 0\) such that \(X(\cdot, \omega) = Y(\cdot, \omega)\) for \(\omega \in \Lambda^c\).

**Definition 2.9.** A stochastic process \(X_t, t \in [a, b]\), on a probability space \((\Omega, \mathcal{F}, P)\) is said to be stochastically continuous at \(t_0 \in [a, b]\) if \(X_t(\cdot)\) converges to \(X_{t_0}(\cdot)\) in probability as \(t \to t_0\) in the sense that

\[
\lim_{t \to t_0} P\{\omega \in \Omega; |X_t(\omega) - X_{t_0}(\omega)| \geq \varepsilon\} = 0, \quad \forall \varepsilon > 0.
\]

We shall write \(P - \lim_{t \to t_0} X_t(\cdot) = X_{t_0}(\cdot)\) when this condition is satisfied. We say that \(X\) is stochastically continuous when it is stochastically continuous at each point in the interval.

**Remark 2.1.** \(P - \lim_{t \to t_0} X_t(\cdot) = X_{t_0}(\cdot)\) iff for every sequence \(\{t_n\}\) which converges to \(t_0\), we have \(P - \lim_{n \to \infty} X_{t_n}(\cdot) = X_{t_0}(\cdot)\)

**Definition 2.10.** The finite-dimensional distributions of the process \(X\) are distributions of the finite-dimensional vectors

\[
(X_{t_1}, \ldots, X_{t_n}), \quad t_1, \ldots, t_n \in D
\]

for all possible choices of times \(t_1, \ldots, t_n \in D\) and every \(n \geq 1\).
A stochastic process is called Gaussian if all its finite-dimensional distributions are multivariate Gaussian. Another way of classifying stochastic processes consists of imposing a special dependence structure. The process $X$ is said strictly stationary if its finite-dimensional distributions are invariant under the shifts of time $t$, i.e.

$$(X_{t_1}, \ldots, X_{t_n}) \overset{d}{=} (X_{t_1+h}, \ldots, X_{t_n+h})$$

for all possible choices of indices $t_1, \ldots, t_n \in D$, $n \geq 1$ and $h$ such that $t_1+h, \ldots, t_n+h \in D$. If for a process $X$ its expectation function $\mu_X(t)$ and covariance function $c_X(s, t)$ are invariant under shifts of times, i.e.

$$\mu_X(t) = \mu_X(t+h), \quad \text{and} \quad c_X(s, t) = c_X(s+h, t+h)$$

for all $s, t \in D$ such that $s+h, t+h \in D$, we say the process is stationary.

Stationarity also can be imposed on the increments of a process. The process itself is then not necessarily to be stationary. Let $X = (X_t, t \in D)$ be a process and $D$ be an interval. $X$ is said to have stationary increments if

$$X_t - X_s \overset{d}{=} X_{t+h} - X_{s+h}$$

for all $t, s \in D$ and $h$ such that $t+h, s+h \in D$. $X$ is said to have independent increments if for every choice of $t_i \in D$ such that $t_1 < t_2 < \cdots < t_n$ and $n > 1$,

$$X_{t_2-t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$$

are independent random variables.

**Definition 2.11.** A stochastic process $B = (B_t, t \in [0, \infty))$ is called a standard Brownian motion or a Wiener process if it satisfies

- It starts at zero, $B_0 = 0$.
- It has stationary, independent increments.
- For every $t > 0$, $B_t$ has a normal $N(0, t)$ distribution.
- Its sample paths are continuous.

A standard Brownian motion starts at zero. In some cases we need a generalized definition for it (See Yeh(1973)). For example, we may need Brownian motion defined on $[a, b]$ starting at $a$. 

10
Definition 2.12. By Brownian motion we mean a stochastic process \( B = (B_t) \) with independent increments defined on a probability space \((\Omega, \mathcal{F}, P)\) and an interval \(D\) in which the probability distribution of random variable \(B_t - B_s, t, s \in D, t > s\), is normal distribution \(N(0, t - s)\).

2.4 Stochastic integral

One possible mathematical model for the price of shares in the stock market is the Brownian motion: the price of one share of a particular stock is assumed to equal \( a + \lambda B_t(\omega) \) at time \( t \). At times \( 0 = t_0 < t_1 < \cdots < t_n = t \), an investor decides the amount \( F(t_i, \omega) \) of shares of the stock to have during time \((t_i, t_{i+1})\). The gain in the time interval \((0, t)\) is

\[
\lambda \sum_{i=0}^{n-1} F(t_i, \omega)(B_{i+1}(\omega) - B_i(\omega)).
\]

At time \( t_i \), the investor only knows what has happened up to time \( t_i \). Therefore each \( F(t_i, \omega) \) should be \( \mathcal{F}_{t_i} \)-measurable. What happens when the time intervals between consecutive decisions become smaller? Does the above sum has a limit? That is, can we define integrals of the form \( \int F(t, \omega) dB_t(\omega) \).

2.4.1. Ito Stochastic Integral for Simple Processes

We start the investigation of Ito integral for a class of processes whose paths assume only finite number of values, As usual, \( B = (B_t, t \geq 0) \) stands for standard Brownian motion, and \( \mathcal{F}_t = \sigma(B_s, s \leq t) \) is the corresponding natural filtration. Recall that a process is adapted to Brownian motion if it is adapted to \( \mathcal{F}_t \) ([16]). That means it is a function of Brownian motion. In what follows, we concentrate on processes that is defined in fixed interval \([0, T]\).

Definition 2.13. The stochastic process \( C = (C_t, t \in [0, T]) \) is said to be simple if it satisfies following properties:

There exist a partition

\[
\tau_n: \quad 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T,
\]

and a sequence \((Z_i, i = 1, \ldots n)\) of random variables such that

\[
C_t = \begin{cases} 
  Z_n, & \text{if } t = T \\
  Z_i, & \text{if } t_{i-1} \leq t < t_i, \quad i = 1, \ldots n
\end{cases}
\]

The sequence \((Z_i)\) is adapted to \( \mathcal{F}_{t_i-1} \), i.e. \( Z_i \) is the function of Brownian motion up to time \( t_{i-1} \) and satisfies \( EZ_i^2 < \infty \) for all \( i \).
**Definition 2.14.** The Ito stochastic integral of simple process $C$ on $[0, T]$ is given by

$$
\int_0^T C_s dB_s = \sum_{i=1}^n C_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) = \sum_{i=1}^n C_{t_{i-1}} \Delta_i B
$$

The Ito stochastic integral of a simple process $C$ on $[0, t], t \leq T$, is given by

$$
\int_0^t C_s dB_s := \int_0^T C_s X_{[0,t]}(s) d\mathcal{B}_s
$$

where $X_{[0,t]}$ is the indicator function on $[0, t]$.

### 2.4.2. The General Ito Stochastic Integral

As the example shown, it is necessary to impose some conditions for integrand stochastic process.

**Assumptions on the integrand process $C$**

1. $C$ is adapted to Brownian motion on $[0, T]$;
2. the integral $\int_0^T E C_s^2 ds$ is finite.

Note that the Assumptions are trivially satisfied for a simple process. Thus the Ito integral for simple process is a particular case of the general Ito integrals.

**Theorem 2.3** (Thomas Mikosch, p109). Let $C$ be a process satisfying the Assumptions. Then there exists a sequence $(C_n)$ of simple processes such that

$$
\int_0^T E[C_s - C_n(s)]^2 ds \to 0, \quad \text{as} \quad n \to \infty
$$

Now we can define Ito stochastic integral for general process.

**Definition 2.15.** Suppose that $C$ is a process on $[0, T]$ satisfying the Assumption. Let $C_n$ be the sequence of simple processes that converges to $C$ in the mean square sense of (2.6). The Ito stochastic integral of $C$ with respect to Brownian motion is defined as the mean square limit of the sequence of stochastic integrals of $C_n$, that is,

$$
\int_0^t C_s dB_s = \lim_{n \to \infty} \int_0^t C_n(s) dB_s \quad t \in [0, T]
$$

**Proposition 1** (Thomas Mikosch, p111). Ito stochastic integral has following properties:

1. The stochastic process $I_t(C) = \int_0^t C_s dB_s$, $t \in [0, T]$, is a martingale with respect to natural Brownian filtration $(\mathcal{F}_t, t \in [0, T])$. 

2. The Ito stochastic integral has expectation zero.

3. Ito stochastic integral satisfies the isometry property:

\[ E \left( \int_0^t C_s dB_s \right)^2 = \int_0^t EC^2_s ds, \quad t \in [0, T]. \]

4. Ito stochastic integral shares some properties with Riemann and Riemann-Stieltjes integrals:

- Ito stochastic integral is linear: For constants \( c_1, c_2 \) and processes \( C^{(1)} \) and \( C^{(2)} \) on \([0, T] \), satisfying the Assumptions,

\[ \int_0^T \left[ c_1 C^{(1)}_s + c_2 C^{(2)}_s \right] dB_s = c_1 \int_0^T C^{(1)}_s dB_s + c_2 \int_0^T C^{(2)}_s dB_s \]

- The Ito stochastic integral is linear on adjacent intervals, i.e. for \( 0 \leq t \leq T \)

\[ \int_0^T C_s dB_s = \int_0^t C_s dB_s + \int_t^T C_s dB_s. \]

5. The process \( I(C) \) has continuous sample path.

**Definition 2.16.** Suppose that \((X_t)\) and \((Y_t), t \in [a, b]\), are stochastic processes on \((\Omega, \mathcal{F}, P)\). We say \((X_t)\) is a version of (or a modification of) \((Y_t)\) if

\[ P\{ \omega \in \Omega; X_t(\omega) = Y_t(\omega) \} = 1, \quad \forall t \in [a, b] \]

**Remark 2.2.** Note that if \(X_t\) is a version of \(Y_t\), then \(X_t\) and \(Y_t\) have the same finite dimensional distributions. Thus from the point of view that a stochastic process is a probability law on \((\mathbb{R}^n)^{[0, \infty)}\) two such processes are the same, but nevertheless their path properties may be different.

**Theorem 2.4** (Bernt, Theorem 3.2.5). Suppose that \( F(t, \omega) \) suppose the Assumptions of integrability on \([0, T]\). Then there exists a \( t \)-continuous version of

\[ \int_0^t f(s, \omega) dB_s(\omega), \quad t \in [0, T]; \]

namely, there exists a \( t \)-continuous stochastic process \( X_t \) such that

\[ P\{ \omega \in \Omega; X_t = \int_0^t f dB \} = 1, \quad \forall t \in [0, T] \]

**Remark 2.3.** Similarly, the integral

\[ \int_\varepsilon^T f(s, \omega) dB_s(\omega), \quad \varepsilon \in (0, T] \]

is also continuous with respect to \( \varepsilon \).
2.5 Expansion of Brownian motion

Orthogonal decomposition of Brownian motion gives a convenient way to simulate Brownian motion. And theoretically it is helpful to understand what the Brownian motion is. In [16] and [17], there are two versions of Brownian motion expansion. One is Paley-Wiener representation

\[ B_t(\omega) = Z_0(\omega) \frac{t}{\sqrt{2\pi}} + \frac{2}{\sqrt{\pi}} \sum_{n=1}^{\infty} Z_n(\omega) \frac{\sin(nt/2)}{n}, \quad t \in [0, 2\pi] \]  

where \( Z_n \) are i.i.d. \( N(0, 1) \) random variables.

Another one, more generally, is Levy-Ciesielski representation

\[ B_t(\omega) = \sum_{n=1}^{\infty} Z_n(\omega) \int_0^t \phi_n(x)dx, \quad t \in [0, 1] \]

where \( Z_n \) are i.i.d \( N(0, 1) \) random variables and \( (\phi_n) \) is a complete orthogonal system on \([0, 1]\).

3. Expansion of Quadratic Form of Brownian Motion

Suppose that Brownian motion is defined on probability space \((\Omega, \mathcal{F}, P)\) and interval \([0, \infty)\). To begin with, we introduce following definition of a kind of stochastic integral.

**Definition 3.1.** Suppose \( B_t \) is the standard Brownian motion on \((\Omega, \mathcal{F}, P)\). Let \( F(x, \cdot) \) be a function defined on \([a, b] \times \Omega\) satisfying integrability conditions

- \( F(x, B_x) \) is \( \mathcal{B} \times \mathcal{F} \)-measurable, where \( \mathcal{B} \) denotes the Borel \( \sigma \)-field on \([a, b]\);
- \( F(x, \cdot) \) is adapted with natural filtration \( \mathcal{F}_t = \sigma(B_t) \) generated by Brownian motion;
- \( \int_a^b E[F^2(x, B_x)]dx = \int_a^b \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi x}} e^{-\frac{y^2}{2x}} F^2(x, y)dydx < \infty \)

Suppose that an arbitrary partition is described by a sequence of points on \([a, b]\), \( \tau_n[a, b] \): \( a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b \). Denote by \( \Delta x_i = x_i - x_{i-1} \) and \( \Delta x = \max_i \{\Delta x_i\} \). Define a stochastic integral

\[ I(F) := \int_a^b F(x, B_x)dB_x \]  

as a mean square limit of following summation

\[ \sum_{i=1}^{n} F(x_{i-1}, B_{x_{i-1}})(B_{x_i} - B_{x_{i-1}}) \]
Remark 3.1. The integrability conditions are normal requirements in usual text book about Ito integral ([18]). In addition, this definition is almost the same as that in general stochastic integral text book ([17]); the only difference is that here the integrand stochastic process is of the particular form $F(x, B_x)$. This is only in sake of later use in this study. Furthermore, it can be seen that this integral of $F$ with respect to Brownian motion is actually a mean square limit of stochastic integrals of a sequence of simple processes.

This section aims at finding a transformation between $L^2[a, b]$ and $L^2(\Omega)$ and then studying expansion of $B_x^2$, a particular function in Brownian motion. To this end, we will focus on the case of $F(x, B_x) = f(x)\sqrt{x}B_x$ and introduce subsequent definition.

Definition 3.2. Suppose $f(x) \in L^2[a, b]$ and $a > 0$. Define a transformation $\mathcal{T}$ between $L^2[a, b]$ and $L^2(\Omega)$ as

$$\mathcal{T}(f; [a, b]) = \int_a^b \frac{f(x)}{\sqrt{x}} B_x dB_x \quad (3.2)$$

Remark 3.2. When it is necessary to stress the interval $[a, b]$ we put it into the notation, but most of time we use the notation without interval. It is crucial to ensure that the transformation we defined exists for every $f \in L^2[a, b]$. We actually can verify this fact by checking whether it satisfies the three integrability conditions in definition 3.1. Obviously we only need to check the last one:

$$\int_a^b E\left(\frac{f(x)}{\sqrt{x}} B_x \right)^2 dx = \int_a^b \frac{f^2(x)}{x} E[B_x^2]dx = \int_a^b f^2(x)dx < \infty$$

Therefore the transformation $\mathcal{T}(f; [a, b])$ is a mapping from $L^2[a, b]$ into random variables. It can be seen that $\mathcal{T}(f) \in L^2(\Omega)$ from following theorem.

Theorem 3.1. Suppose $f(x), g(x) \in L^2[a, b]$ and $a > 0$. The transformation in definition (3.2) satisfies

1. $E[\mathcal{T}(f)] = 0$;
2. $\langle \mathcal{T}(f), \mathcal{T}(g) \rangle = (f, g)$;
3. $\|\mathcal{T}(f)\|_2^2(\Omega) = \|f\|_2^2$. hence $\mathcal{T}(f) \in L^2(\Omega)$;
4. $\mathcal{T}(c_1 f + c_2 g)(\omega) = c_1 \mathcal{T}(f)(\omega) + c_2 \mathcal{T}(g)(\omega)$ for every $\omega \in \Omega$, where $c_1, c_2$ are real constants.

Proof. Denote by $\tau_n[a, b]$ an arbitrary partition on interval $[a, b]$: $a = x_0 < x_1 < \cdots < x_n = b$, where $n$ is any positive integer number. Let $\Delta x = \max_i \{x_i - x_{i-1}\}$. For $f \in L^2[a, b]$, denote
by \( S^\tau_n(f) \) the sum corresponding to the partition \( \tau_n[a, b] \)

\[
S^\tau_n(f) := \sum_{i=1}^{n} \frac{f(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}})
\]

By the independence of increments of Brownian motion and its distribution, it follows easily that

\[
E[S^\tau_n(f)] = \sum_{i=1}^{n} \frac{f(x_{i-1})}{\sqrt{x_{i-1}}} E[B_{x_{i-1}}] E[(B_{x_{i}} - B_{x_{i-1}})] = 0.
\]

1. Since \( E[S^\tau_n(f)] = 0 \), by Jensen inequality we have

\[
\{E[\mathcal{F}(f)]\}^2 = \{E[\mathcal{F}(f)] - E[S^\tau_n(f)]\}^2 \leq E[\mathcal{F}(f) - S^\tau_n(f)]^2 \to 0, \quad \text{as} \quad n \to \infty,
\]

which implies \( E[\mathcal{F}(f)] = 0 \).

2. Since \( \mathcal{F}(f) \) and \( \mathcal{F}(g) \) are mean square limit of sequence \( S^\tau_n(f) \) and \( S^\tau_n(g) \) respectively, it follows from continuity of inner product in Hilbert space \( L^2(\Omega) \) that

\[
\langle \mathcal{F}(f), \mathcal{F}(g) \rangle = \left\langle \lim_{\Delta x \to 0} S^\tau_n(f), \lim_{\Delta x \to 0} S^\tau_n(g) \right\rangle
\]

\[
= \lim_{\Delta x \to 0} \langle S^\tau_n(f), S^\tau_n(g) \rangle = \lim_{\Delta x \to 0} E[S^\tau_n(f)S^\tau_n(g)]
\]

\[
= \lim_{\Delta x \to 0} E \left( \sum_{i=1}^{n} \frac{f(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}}) \cdot \sum_{i=1}^{n} \frac{g(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}}) \right)
\]

\[
= \lim_{\Delta x \to 0} E \left( \sum_{i=1}^{n} \frac{f(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}}) \cdot \frac{g(x_{j-1})}{\sqrt{x_{j-1}}} B_{x_{j-1}}(B_{x_{j}} - B_{x_{j-1}}) \right)
\]

\[
= \lim_{\Delta x \to 0} \left( \sum_{i=1}^{n} \frac{f(x_{i-1})g(x_{i-1})}{x_{i-1}} E[B_{x_{i-1}}^2] E[(B_{x_{i}} - B_{x_{i-1}})^2] \right)
\]

\[
= \lim_{\Delta x \to 0} \left( \sum_{i=1}^{n} f(x_{i-1})g(x_{i-1})(x_{i} - x_{i-1}) \right)
\]

\[
= \int_{a}^{b} f(x)g(x)dx = (f, g)
\]

3. This is a particular case of property 2 when \( f = g \).
Lemma 3.1. As has been shown, i.e. for $h$ from $T$ directly. It is therefore reasonable to apply improper integration to it. The transformation

Now concentration moves to the interval $[0, T]$.

For arbitrary constants $c_1$, $c_2$,

$$\mathcal{T}(c_1 f + c_2 g) = \lim_{\Delta x \to 0} S^{\tau_n}(c_1 f + c_2 g)$$

$$= \lim_{\Delta x \to 0} \sum_{i=1}^{n} \frac{c_1 f(x_{i-1}) + c_2 g(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}})$$

$$= \lim_{\Delta x \to 0} \sum_{i=1}^{n} \frac{c_1 f(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}}) + \lim_{\Delta x \to 0} \sum_{i=1}^{n} \frac{c_2 g(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}})$$

$$= c_1 \lim_{\Delta x \to 0} \sum_{i=1}^{n} \frac{f(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}}) + c_2 \lim_{\Delta x \to 0} \sum_{i=1}^{n} \frac{g(x_{i-1})}{\sqrt{x_{i-1}}} B_{x_{i-1}}(B_{x_{i}} - B_{x_{i-1}})$$

$$= c_1 \int_{a}^{b} \frac{f(x)}{\sqrt{x}} B_x dB_x + c_2 \int_{a}^{b} \frac{g(x)}{\sqrt{x}} B_x dB_x$$

$$= c_1 \mathcal{T}(f) + c_2 \mathcal{T}(g)$$

Apart from the properties listed in previous theorem, $\mathcal{T}(f)$ possesses other characteristics of usual integral, for example, linearity on adjacent intervals and continuity about lower limit $a$ and upper limit $b$. Linearity on adjacent intervals says that if $a < c < b$, $\mathcal{T}(f; [a, b]) = \mathcal{T}(f; [a, c]) + \mathcal{T}(f; [c, b])$; while continuity about $a$ means that if $a' \to a$ the integral $\mathcal{T}(f; [a', b])$ will converge to the integral $\mathcal{T}(f; [a, b])$ in mean square sense. Now concentration moves to the interval $[0, T], T > 0$ and fixed. The main reason why we shift to this case is simply that customarily the standard Brownian motion starts at point zero. Moreover this relaxes the restriction that $a > 0$. However, the movement is not trivial. Since $1/\sqrt{x}$ is undefined at the point zero, the mapping $\mathcal{T}$ can not be used on interval $[0, T]$ directly. It is therefore reasonable to apply improper integration to it. The transformation $\mathcal{T}$ from $L^2[0, T]$ to $L^2(\Omega)$ is defined as

$$\mathcal{T}(f; [0, T]) = \lim_{\epsilon \to 0^+} \mathcal{T}(f; [\epsilon, T]) = \lim_{\epsilon \to 0^+} \int_{\epsilon}^{T} \frac{f(x)}{\sqrt{x}} B_x dB_x \quad \text{(in norm)}$$

(3.3)

It raises that whether the improper integrations exist for the functions we are studying, i.e. for $h(x) \in L^2[0, T]$, does $\mathcal{T}(h; [0, T])$ exist? To answer this question, let us define

$$\mathcal{T}_\epsilon(h) = \mathcal{T}(h; [\epsilon, T]) = \int_{\epsilon}^{T} \frac{h(x)}{\sqrt{x}} B_x dB_x$$

As has been shown, $\mathcal{T}_\epsilon(h)$ is always well defined for $0 < \epsilon < T$, and $\mathcal{T}_\epsilon(h) \in L^2(\Omega)$.

**Lemma 3.1.** For $\forall h(x) \in L^2[0, T]$, the improper integral $\mathcal{T}(h; [0, T])$ exists.
Proof. To begin with, suppose \( \{\delta_n\} \) is a positive sequence which converges to zero. Consequently there is a sequence \( \{\mathcal{T}_{\delta_n}(h)\} \) in \( L^2(\Omega) \) for \( h \in L^2[0,T] \). The aim is to show that \( \{\mathcal{T}_{\delta_n}(h)\} \) is a Cauchy sequence. Actually,

\[
\|\mathcal{T}_{\delta_n}(h) - \mathcal{T}_{\delta_m}(h)\|_2^2(\Omega) = \left\| \int_{\delta_n}^T \frac{h(x)}{\sqrt{x}} B_x dB_x - \int_{\delta_m}^T \frac{h(x)}{\sqrt{x}} B_x dB_x \right\|_2^2 = \left\| \int_{\delta_n \vee \delta_m}^{\delta_n \wedge \delta_m} \frac{h(x)}{\sqrt{x}} B_x dB_x \right\|_2^2
\]

which implies the sequence \( \mathcal{T}_{\delta_n}(h) \) is a Cauchy sequence in \( L^2(\Omega) \).

Since \( L^2(\Omega) \) is a Hilbert space, every Cauchy sequence has a limit in the space. Suppose \( \lim_{n \to \infty} \mathcal{T}_{\delta_n}(h) = X \) in \( L^2 \)-norm for the function \( h \) and the sequence \( \{\delta_n\} \), where \( X \) is a random variable in \( L^2(\Omega) \). The problem is, for any other positive sequence, \( \{\epsilon_n\} \) say, which converges to zero, whether the corresponding sequence \( \mathcal{T}_{\epsilon_n}(h) \) converges in norm to another random variable \( Y \)? In fact,

\[
\|X - Y\|_2(\Omega) \leq \|X - I_{\delta_n}(h)\|_2 + \|I_{\delta_n}(h) - I_{\epsilon_n}(h)\|_2 + \|I_{\epsilon_n}(h) - Y\|_2
\]

\[
= \|X - I_{\delta_n}(h)\|_2 + \int_{\delta_n \vee \epsilon_n}^{\delta_n \wedge \epsilon_n} h^2(x) dx + \|I_{\epsilon_n}(h) - Y\|_2
\]

\[
\to 0 \text{ (as } n \to \infty),
\]

which implies \( X = Y \) in \( L^2(\Omega) \). Now that for every positive sequence \( \delta_n \) which converges to zero, the integral sequence \( I_{\delta_n}(h) \) converges the same random variable, \( X \) say, the limit of \( \mathcal{T}(h) \) as \( \epsilon \to 0 \) also should be \( X \). Actually it follows from following inequality,

\[
\|I_{\epsilon}(h) - X\|_2(\Omega) \leq \|I_{\epsilon}(h) - I_{\delta_n}(h)\|_2 + \|I_{\delta_n}(h) - X\|_2 = \int_{\delta_n \wedge \epsilon}^{\delta_n \vee \epsilon} h^2(x) dx + \|I_{\delta_n}(h) - X\|_2 \to 0.
\]

as \( \epsilon \to 0, n \to 0 \). This finishes the proof. \( \square \)

**Theorem 3.2.** For any \( f, g \in L^2[0,T] \) and constants \( c_1, c_2 \), the following hold:

1. \( E[\mathcal{T}(f;[0,T])] = 0 \)
2. \( \langle \mathcal{T}(f;[0,T]), \mathcal{T}(g;[0,T]) \rangle = (f, g) \)
3. \( \|\mathcal{T}(f)\|_2^2(\Omega) = \|f\|_2^2 \)
4. \( \mathcal{T}(c_1 f + c_2 g;[0,T])(\omega) = c_1 \mathcal{T}(f;[0,T])(\omega) + c_2 \mathcal{T}(g;[0,T])(\omega) \) for every \( \omega \in \Omega \).
Proof. For $\forall \varepsilon : 0 < \varepsilon < T$, $\mathcal{T}_\varepsilon(\cdot)$ is defined as before, and $\mathcal{T}(\cdot)$ is the mean square limit of $\mathcal{R}_\varepsilon(\cdot)$ as $\varepsilon \to 0$.

1. It follows from $E[\mathcal{R}_\varepsilon(f)] = 0$ and Jensen inequality that

$$\{E[\mathcal{T}(f)]\}^2 = \{E[\mathcal{T}(f)] - E[\mathcal{R}_\varepsilon(f)]\}^2 = \{E[\mathcal{T}(f) - \mathcal{R}_\varepsilon(f)]\}^2 \leq E[\mathcal{T}(f) - \mathcal{R}_\varepsilon(f)]^2 \to 0 \quad \text{as} \quad \varepsilon \to +0,$$

hence $E[I(f)] = 0$.

2. Using the definition of transformation $\mathcal{T}(\cdot)$ and the continuity of inner product in Hilbert space, we have

$$\langle \mathcal{T}(f), \mathcal{T}(g) \rangle = \left\langle \lim_{\varepsilon \to +0} \mathcal{R}_\varepsilon(f), \lim_{\varepsilon \to +0} \mathcal{R}_\varepsilon(g) \right\rangle = \lim_{\varepsilon \to +0} \langle \mathcal{R}_\varepsilon(f), \mathcal{R}_\varepsilon(g) \rangle = \lim_{\varepsilon \to +0} \int_0^T f(x)g(x)dx = \langle f, g \rangle$$

where the property 2 in theorem 3.1 is used to derive $\langle \mathcal{R}_\varepsilon(f), \mathcal{R}_\varepsilon(g) \rangle$.

3. This equality is the particular case of (2).

4. This property can be proved similarly as the counterpart in theorem 3.1 except that here it is needed to take limit for $\varepsilon \to +0$. Therefore the proof is neglected. \qed

In order to continue the investigation on the transformation $\mathcal{T}$, denote by $\Theta$ the image of $L^2[0, T]$ under the mapping. Then the properties of $\mathcal{T}$ mapping $L^2[0, T]$ onto $\Theta$ are studied in the sequel.

**Theorem 3.3.** The stochastic integral $\mathcal{T}$ defines a one-to-one mapping from $L^2[0, T]$ to $\Theta$. The image $\Theta$ of $L^2[0, T]$ under the mapping $\mathcal{T}$ is a closed linear subspace of $L^2(\Omega)$. Furthermore, the transformation $\mathcal{T}$ is an isomorphism as well between $L^2[0, T]$ and Hilbert space $\Theta$.

Proof. To begin with, if $f, g \in L^2[0, T]$ and $f \neq g$, then $\|f - g\|_2 \neq 0$. This implies $\mathcal{T}(f - g) = \mathcal{T}(f) - \mathcal{T}(g) \neq 0$ since $\mathcal{T}(f - g)$ is a random variable with zero mean and variance $\|f - g\|_2$, so that $\mathcal{T}(f) \neq \mathcal{T}(g)$. On the other hand, if $X \in \Theta$ and $\mathcal{T}(f) = X$ and $\mathcal{T}(g) = X$ for $f, g \in L^2[0, T]$. Because $\mathcal{T}(f - g) = \mathcal{T}(f) - \mathcal{T}(g) = X - X = 0$, the variance of $\mathcal{T}(f - g)$ is definitely zero, i.e. $\|f - g\|_2 = 0$ and this means $f = g$. Therefore $\mathcal{T}$ is one-to-one.

It is evident that $\Theta$ is a linear subspace of $L^2(\Omega)$ as the transformation $\mathcal{T}$ is linear.

To show $\Theta$ is closed, suppose that $\{X_n\}$ is a sequence in $\Theta$ and $X$ is an element in $L^2(\Omega)$ such that $\|X_n - X\|_2 \to 0$. As $\mathcal{T}$ is one-to-one there is a sequence $\{f_n\} \in L^2[0, T]$ such that $\mathcal{T}(f_n) = X_n$. Moreover, $\mathcal{T}$ preserves metric, which indicates $\|f_n\|_2$ is a Cauchy sequence since
\[\|f_n - f_m\|_2 = \|\mathcal{T}(f_n - f_m)\|_2(\Omega) = \|\mathcal{T}(f_n) - \mathcal{T}(f_m)\|_2(\Omega) = \|X_n - X_m\|_2(\Omega)\] and \(\{X_n\}\) is of Cauchy. Thus there exists a function \(f \in L^2[0, T]\) such that \(\|f_n - f\|_2 \to 0\). Once again by the property of the mapping we have \(\|X_n - \mathcal{T}(f)\|_2(\Omega) = \|\mathcal{T}(f_n) - \mathcal{T}(f)\|_2(\Omega) = \|f_n - f\|_2 \to 0\). Then the inequality
\[
\|X - \mathcal{T}(f)\|_2(\Omega) \leq \|X_n - X\|_2(\Omega) + \|X_n - \mathcal{T}(f)\|_2(\Omega) \to 0
\]
as \(n \to \infty\) implies \(X = \mathcal{T}(f) \in \Theta\). This proves that \(\Theta\) is closed.

Finally, as a closed linear subspace of Hilbert space \(L^2(\Omega)\), \(\Theta\) is Hilbert space as well. It follows from (2) and (4) in theorem 3.2 that \(\mathcal{T}\) is isomorphism. \(\square\)

**Corollary 3.1.** If \(\{f_n\}\) is a full orthonormal system in \(L^2[0, T]\), then \(\{\mathcal{T}(f_n)\}\) is a full orthonormal system in \(\Theta\). The inverse is also true.

**Proof.** What needs to prove is merely that when \(\{f_n\}\) is full in \(L^2[0, T]\), \(\{\mathcal{T}(f_n)\}\) is complete in \(L^2(\Omega)\). Denote \(X_n = \mathcal{T}(f_n), n = 1, \ldots, \infty\). For any \(X \in \Theta\), there is one and only one \(f \in L^2[0, T]\) such that \(X = \mathcal{T}(f)\). This \(f\) can be uniquely represented as \(f = \sum_{n=1}^{\infty} c_n f_n\), so that \(X = \mathcal{T}(f) = \sum_{n=1}^{\infty} c_n \mathcal{T}(f_n) = \sum_{n=1}^{\infty} c_n X_n\) and the representation is unique. Thus, \(\{X_n\}\) is a complete orthogonal system.

The inverse is true as \(\mathcal{T}\) is an one-to-one mapping. \(\square\)

**Theorem 3.4.** Let \(\mathcal{T}\) be a transformation from \(L^2[0, T]\) into \(L^2(\Omega)\) defined by equation (3.3) and \(\{f_n\}\) be an orthonormal system in \(L^2[0, T]\). Let \(X_n = \mathcal{T}(f_n), n = 1, 2, \ldots\). For any \(X \in \Theta\) with \(X = \mathcal{T}(f)\), we have
\[
X = \sum_{n=1}^{\infty} \langle X, X_n \rangle X_n = \sum_{n=1}^{\infty} (f, f_n) X_n,
\]
where the convergence of the infinite series is in the sense of norm in \(L^2(\Omega)\). Furthermore,
\[
X(\omega) = \sum_{n=1}^{\infty} \langle X, X_n \rangle X_n(\omega) = \sum_{n=1}^{\infty} (f, f_n) X_n(\omega), \quad \text{in probability}
\]

**Proof.** According to the theory of Fourier expansion, the representation is evident and the coefficients \(\langle X, X_n \rangle\) are called Fourier coefficients. Since \(\mathcal{T}\) is isomorphism, \(\langle X, X_n \rangle = (f, f_n)\).

Because convergence in norm implies convergence in probability, the second expression is valid. \(\square\)

**Theorem 3.5.** Suppose \(B_t\) is standard Brownian motion on \([0, \infty)\) and \(\{f_n\}\) is an arbitrary full orthonormal system in \(L^2[0, T]\) where \(T > 0\) is an finite real number. The stochastic
process $B_t^2$ in $[0,T]$ can be expanded as

$$B_t^2 = t + 2 \sum_{n=1}^{\infty} (\sqrt{x} \chi_{[0,t]}, f_n) X_n,$$

(3.4)

where $\chi_{[0,t]}$ is the indicator function on $[0,t]$, and $X_n = \mathcal{J}(f_n)$.

**Proof.** This is a particular case of Theorem 3.4. For any $0 < t \leq T$, let $f(x) = \sqrt{x} \chi_{[0,t]}(x)$. Then

$$\mathcal{J}(f) = \int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t).$$

On the other hand, $\mathcal{J}(f) = \sum_{n=1}^{\infty} (f, f_n) X_n = \sum_{n=1}^{\infty} (\sqrt{x} \chi_{[0,t]}, f_n) X_n$. Thus the expansion follows.

**Example 3.1** Let $T = \pi$. On $[0, \pi]$, a full orthonormal system is $\{f_n\}$ where

$$f_0(x) = \frac{1}{\sqrt{\pi}}, \quad f_n(x) = \sqrt{\frac{2}{\pi}} \cos(nx), \quad n = 1, 2, \cdots.$$

To get the decomposition of $B_t^2(0 < t \leq \pi)$, let $f(x) = \sqrt{x} \chi_{[0,t]}(x)$ for $0 < t \leq \pi$. Compute

$$c_0 = c_0(t) = (f, f_0) = \int_0^\pi f(x) f_0(x) dx = \int_0^t \sqrt{x} \frac{1}{\sqrt{\pi}} dx = \frac{2}{3\sqrt{\pi}} t^{3/2},$$

$$c_n = c_n(t) = (f, f_n) = \int_0^\pi f(x) f_n(x) dx = \sqrt{\frac{2}{\pi}} \int_0^t \sqrt{x} \cos(nx) dx$$

The corresponding orthonormal system in $L^2(\Omega)$ is

$$X_0 = \mathcal{J}(f_0) = \int_0^\pi \frac{1}{\sqrt{\pi x}} B_x dB_x,$$

$$X_n = \mathcal{J}(f_n) = \sqrt{\frac{2}{\pi}} \int_0^\pi \frac{\cos(nx)}{\sqrt{x}} B_x dB_x.$$

Thus

$$B_t^2 = t + 2 \sum_{n=0}^{\infty} c_n(t) X_n$$

**Example 3.2** In $L^2[0,T]$ there is an orthonormal system consisting of

$$f_n(x) = \sqrt{\frac{2}{T}} \sin \left( \left( n + \frac{1}{2} \right) \frac{\pi}{T} x \right), \quad x \in [0,T], \quad n = 0, 1, 2, \cdots$$

Again, to expand $B_t^2(0 < t \leq T)$, let $f(x) = \sqrt{x} \chi_{[0,t]}(x)$, where $0 < t \leq T$.

$$c_0 = c_0(t) = (f, f_0) = \int_0^T f(x) f_0(x) dx = \sqrt{\frac{2}{T}} \int_0^t \sqrt{x} \sin \left( \frac{\pi}{2T} x \right) dx$$

$$c_n = c_n(t) = (f, f_n) = \int_0^T f(x) f_n(x) dx = \sqrt{\frac{2}{T}} \int_0^t \sqrt{x} \sin \left( \left( n + \frac{1}{2} \right) \frac{\pi}{T} x \right) dx$$
The corresponding orthonormal system in $L^2(\Omega)$ is

$$X_0 = \mathcal{T}(f_0) = \sqrt{\frac{2}{T}} \int_0^T \frac{1}{\sqrt{x}} \sin \left( \frac{\pi}{2T x} \right) B_x dB_x$$

$$X_n = \mathcal{T}(f_n) = \sqrt{\frac{2}{T}} \int_0^T \frac{1}{\sqrt{x}} \sin \left( \left( n + \frac{1}{2} \right) \frac{\pi}{T x} \right) B_x dB_x$$

Thus

$$B_t^2 = t + 2 \sum_{n=0}^\infty c_n(t) X_n$$ (3.5)

**Note 1.** All studies on interval $[0, T]$ in this section can revert to $[a, b]$. This means all results about interval $[0, T]$ are valid for $[a, b], a > 0$, under mapping $\mathcal{T}(f; [a, b])$ from $L^2[a, b]$ into $L^2(\Omega)$. The only change is that Brownian motion starts at $a$ and almost surely is zero at point $a$.

### 4. Local Approximation

Although decomposition of $B_t^k$ ($k \geq 3$) can be obtained similarly, the bases in $L^2(\Omega)$ for expansion of $B_t^k$ are different from each other. This sets up an impediment for getting decomposition of $f(B_x)$ for general function $f$ by Taylor series. Thus one has to steer to finding approximation of $f(B_x)$ by lower order of $B_x$. Following is a result of approximation to $f(B_x)$ by polynomial in Brownian motion with degree less than three.

**Theorem 4.1.** Suppose that $f(\cdot)$ is continuous on $\mathbb{R}$. Let $B_x$ be the standard Brownian motion in probability space $(\Omega, \mathcal{F}, P)$ and interval $[0, \infty)$. $x_0 \in (0, \infty)$. Then under one of following conditions

1. there are a neighborhood of $x_0$, $N_\delta(x_0) = \{ x \in \mathbb{R} : |x - x_0| < \delta \}$, $\delta > 0$, and some integer $n \geq 3$ such that $f(\cdot) \in C^{(n)}(\mathbb{R})$ and $f^{(n)}(B_x)$ is bounded almost everywhere in $N_\delta(x_0)$ in the sense that there exists a constant $K(x_0, \delta) > 0$ such that

   $$P \left( |f^{(n)}(B_x)| > K(x_0, \delta) \right) = 0. \quad \forall x \in N_\delta(x_0)$$ (4.1)

2. there are a neighborhood of $x_0$, $N_\delta(x_0) = \{ x \in \mathbb{R} : |x - x_0| < \delta \}$, $\delta > 0$, and some integer $n \geq 3$ such that $f(\cdot) \in C^{(n)}(\mathbb{R})$ and $E \left[ f^{(n)}(B_x) \right]^3$ exists and is bounded in $N_\delta(x_0)$, that is, there exists a constant $K(x_0, \delta) > 0$ such that

   $$\left| E \left[ f^{(n)}(B_x) \right]^3 \right| \leq K(x_0, \delta), \quad \forall x \in N_\delta(x_0)$$ (4.2)
we have approximation of $f(B_x)$:

$$f(B_x) = f(B_{x_0}) + f'(B_{x_0})(B_x - B_{x_0}) + \frac{1}{2}f''(B_{x_0})(B_x - B_{x_0})^2 + O(|x - x_0|^{3/2}) \ a.s. \ (4.3)$$

**Proof.** Note that in any case $f$ has the third order derivative. Let $Q(B_x) = f(B_{x_0}) + f'(B_{x_0})(B_x - B_{x_0}) + \frac{1}{2}f''(B_{x_0})(B_x - B_{x_0})^2$.

Firstly suppose $x \in N_{\delta}(x_0)$ and $x > x_0$.

In the first case, if $n = 3$, it follows from Taylor expansion that

$$f(B_x) = f(B_{x_0}) + f'(B_{x_0})(B_x - B_{x_0}) + \frac{1}{2}f''(B_{x_0})(B_x - B_{x_0})^2$$

$$+ \frac{1}{3!}(3f'''(B_{\xi})(B_x - B_{x_0})^3$$

$$= Q(B_x) + \frac{1}{3!}f(3\xi)(B_x - B_{x_0})^3 \ (4.4)$$

where $\xi$ is in between $x_0$ and $x$. Denote by $A(x) = \{ \omega : |f^{(3)}(B_x)| \leq K(x_0, \delta) \}, x \in N_{\delta}(x_0)$ and $\mathcal{X}_A$ the indicator function on $A$. Thus

$$\|f(B_x) - Q(B_x)\| = \left\| \frac{1}{3!}f^{(3)}(B_{\xi})(B_x - B_{x_0})^3 \right\| = \left\| \frac{1}{3!}f^{(3)}(B_{\xi})(\mathcal{X}_A(\xi) + \mathcal{X}_A(B_x)) (B_x - B_{x_0})^3 \right\|$$

$$= \left\| \frac{1}{3!}f^{(3)}(B_{\xi})\mathcal{X}_A(\xi)(B_x - B_{x_0})^3 + \frac{1}{3!}f^{(3)}(B_{\xi})\mathcal{X}_A(B_x)(B_x - B_{x_0})^3 \right\|$$

$$\leq \left\| \frac{1}{3!}f^{(3)}(B_{\xi})\mathcal{X}_A(\xi)(B_x - B_{x_0})^3 \right\| + \left\| \frac{1}{3!}f^{(3)}(B_{\xi})\mathcal{X}_A(B_x)(B_x - B_{x_0})^3 \right\|$$

$$= \left\| \frac{1}{3!}f^{(3)}(B_{\xi})\mathcal{X}_A(\xi)(B_x - B_{x_0})^3 \right\| \leq \frac{1}{6}K(x_0, \delta) \| (B_x - B_{x_0})^3 \|$$

$$= \frac{1}{6}K(x_0, \delta) (E(B_x - B_{x_0})^6)^{1/2} = \frac{1}{6}K(x_0, \delta) (E(B_{x-x_0})^6)^{1/2}$$

$$= \frac{1}{6}K(x_0, \delta) (E(\sqrt{|x - x_0|}B_1)^6)^{1/2} = \frac{1}{6}K(x_0, \delta) \sqrt{|x - x_0|}^3 [E(B_1^6)]^{1/2}$$

$$= \frac{\sqrt{15}}{6}K(x_0, \delta) \sqrt{|x - x_0|}^3 = O \left( |x - x_0|^{3/2} \right)$$

If $n > 3$, again applying the Taylor expansion until the $n$-th derivative yields

$$f(B_x) - Q(B_x) = \sum_{j=3}^{n-1} \frac{1}{j!}f^{(j)}(B_{x_0})(B_x - B_{x_0})^j + \frac{1}{n!}f^{(n)}(B_{\xi})(B_x - B_{x_0})^n \ (4.5)$$

where $\xi$ is in between $x_0$ and $x$. It follows from the triangle inequality of norm and indepen-
dence of increments of Brownian motion that
\[
\|f(B_x) - Q(B_x)\| = \left\| \sum_{j=3}^{n-1} \frac{1}{j!} f^{(j)}(B_{x_0})(B_x - B_{x_0})^j + \frac{1}{n!} f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\|
\leq \sum_{j=3}^{n-1} \frac{1}{j!} \left\| f^{(j)}(B_{x_0})(B_x - B_{x_0})^j \right\| + \frac{1}{n!} \left\| f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\|
= \sum_{j=3}^{n-1} \frac{1}{j!} \left\| f^{(j)}(B_{x_0}) \right\| \| (B_x - B_{x_0})^j \| + \frac{1}{n!} \left\| f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\| \quad (4.6)
\]

It is evident that for \(3 \leq j \leq n - 1\), \(\| (B_x - B_{x_0})^j \| = O(|x - x_0|^{j/2})\), and similar argument as in the case of \(n = 3\) gives
\[
\frac{1}{n!} \left\| f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\| = O(|x - x_0|^{n/2})
\]

To sum up, the norm of difference between \(f(B_x)\) and \(Q(B_x)\) is of order \(O(|x - x_0|^{3/2})\). This finishes the proof of case one.

In the second case, when \(n = 3\) the same formula of Taylor expansion (4.4) is used, but the reminder term by Holder inequality with \(p = 3/2\) and \(q = 3\) satisfies
\[
\|f(B_x) - Q(B_x)\| = \left\| \frac{1}{3!} f^{(3)}(B_\xi)(B_x - B_{x_0})^3 \right\| = \frac{1}{6} \left\{ E \left[ \left( f^{(3)}(B_\xi) \right)^2 (B_x - B_{x_0})^6 \right] \right\}^{1/2}
\leq \frac{1}{6} \left\{ \left( E \left( f^{(3)}(B_\xi) \right)^2 \right)^{2/3} \left[ E (B_x - B_{x_0})^{18} \right]^{1/3} \right\}^{1/2}
= \frac{1}{6} \left[ E \left( f^{(3)}(B_\xi) \right)^3 \right]^{1/3} \left[ E (B_x - B_{x_0})^{18} \right]^{1/6}
\leq \frac{1}{6} K^{1/3}(x_0, \delta)|x - x_0|^{3/2} \left[ E(B_1^{18}) \right]^{1/6} = \frac{(17!!)^{1/6}}{6} K^{1/3}(x_0, \delta)|x - x_0|^{3/2}
= O(|x - x_0|^{3/2}) \quad (4.7)
\]

If \(n > 3\), similar to (4.6), Holder inequality gives
\[
\|f(B_x) - Q(B_x)\| = \left\| \sum_{j=3}^{n-1} \frac{1}{j!} f^{(j)}(B_{x_0})(B_x - B_{x_0})^j + \frac{1}{n!} f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\|
\leq \sum_{j=3}^{n-1} \frac{1}{j!} \left\| f^{(j)}(B_{x_0}) \right\| \| (B_x - B_{x_0})^j \| + \frac{1}{n!} \left\| f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\|
\leq \sum_{j=3}^{n-1} \frac{1}{j!} \left\| f^{(j)}(B_{x_0}) \right\| \| (B_x - B_{x_0})^j \| + \frac{1}{n!} \left\| f^{(n)}(B_\xi)(B_x - B_{x_0})^n \right\| \quad (4.8)
\]
whereas
\[
\frac{1}{n!} \left\| f^{(n)}(B_{\xi})(B_x - B_{x_0})^n \right\| = \frac{1}{n!} \left\{ E \left[ \left( f^{(n)}(B_{\xi}) \right)^2 (B_x - B_{x_0})^{2n} \right] \right\}^{1/2}
\leq \frac{1}{n!} \left\{ \left[ E \left( f^{(n)}(B_{\xi}) \right)^3 \right]^{2/3} \left[ E (B_x - B_{x_0})^{6n} \right]^{1/3} \right\}^{1/2}
= \frac{1}{n!} \left[ E \left( f^{(n)}(B_{\xi}) \right)^3 \right]^{1/3} \left[ E (B_x - B_{x_0})^{6n} \right]^{1/6}
\leq \frac{1}{n!} K^{1/3}(x_0, \delta) |x - x_0|^{n/2} \left[ E (B_1^{6n}) \right]^{1/6}
= \frac{[(6n - 1)!!]^{1/6}}{n!} K^{1/3}(x_0, \delta) |x - x_0|^{n/2}
= O(|x - x_0|^{n/2}) \tag{4.9}
\]

We conclude \(\| f(B_x) - Q(B_x) \| = O(|x - x_0|^{3/2})\).

Secondly, consider \(x \in N_\delta(x_0)\) and \(x < x_0\). The discussions for this situation are all the same as its counterpart except that in equation (4.6) \(f^{(j)}(B_{x_0})\) and \((B_x - B_{x_0})^j\) are no longer independent. But \(f^{(j)}(B_{x_0}), 3 \leq j < n\), can be expanded by Taylor formula as
\[
f^{(j)}(B_{x_0}) = f^{(j)}(B_x) + f^{(j+1)}(B_x)(B_{x_0} - B_x),
\]
where the higher-order term is neglected since it does not affect the result. Therefore the summands in the sum sign of equation (4.6) can be estimated
\[
\frac{1}{j!} \left\| f^{(j)}(B_{x_0})(B_x - B_{x_0})^j \right\| = \frac{1}{j!} \left\| \left( f^{(j)}(B_x) + f^{(j+1)}(B_x)(B_{x_0} - B_x) \right) (B_x - B_{x_0})^j \right\|
\leq \frac{1}{j!} \left\| f^{(j)}(B_x)(B_x - B_{x_0})^j \right\| + \frac{1}{j!} \left\| f^{(j+1)}(B_x)(B_x - B_{x_0})^{j+1} \right\|
= \frac{1}{j!} \left\| f^{(j)}(B_x) \right\| \left\| (B_x - B_{x_0})^j \right\| + \frac{1}{j!} \left\| f^{(j+1)}(B_x) \right\| \left\| (B_x - B_{x_0})^{j+1} \right\|
= O(|x - x_0|^{j/2}),
\]
the same result as the case \(x > x_0\). Hence the proof has been finished completely. \(\square\)

Remark 4.1. Any polynomial \(P_n(x) = a_0x^n + \cdots + a_n\) belongs to the case one of Theorem 4.1 since its \(n\)-th derivative is constant \(a_0n!\), hence it can be approached by polynomial in Brownian motion with degree less than three. Alternatively, according to Taylor formula, \(P_n(x)\) can be exactly written as a polynomial of \((x - x_0)\), i.e.
\[
P_n(x) = P_n(x_0) + P'_n(x_0)(x - x_0) + \cdots + \frac{1}{n!} P^{(n)}(x_0)(x - x_0)^n.
\]
Then
\begin{align*}
P_n(B_x) &= P_n(B_{x_0}) + P_n'(B_{x_0})(B_x - B_{x_0}) + \frac{1}{2} P_n''(B_{x_0})(B_x - B_{x_0})^2 \\
&\quad + \cdots + \frac{1}{n!} P_n^{(n)}(B_{x_0})(B_x - B_{x_0})^n \\
&= Q(B_x) + \frac{1}{3!} P_n^{(3)}(B_{x_0})(B_x - B_{x_0})^3 + \cdots + \frac{1}{n!} P_n^{(n)}(B_{x_0})(B_x - B_{x_0})^n,
\end{align*}
where \(Q(B_x)\) is as defined in the theorem. Therefore, the conclusion \(\|P_n(B_x) - Q(B_x)\| = O(|x - x_0|^{3/2})\) is obvious.

**Corollary 4.1.** There are two classes of functions for which the result (4.3) is valid:

1. If the third-order derivative \(f^{(3)}(\cdot)\) of function \(f\) has bound globally, then (4.3) is valid.

2. If the third-order derivative \(f^{(3)}(\cdot)\) is continuous in \(\mathbb{R}\) and \(E\left[f^{(3)}(B_x)\right]^3\) exists in neighborhood of \(x_0\), then (4.3) holds.

**Proof.** In case 1, since global bounded implies that the function is bounded locally, the condition (4.1) is automatically satisfied. Some examples are \(\sin x, \cos x\) and \(\exp(-x^2)\), and their any combination by +, −, ×.

In case 2, because, for any \(\omega \in \Omega\), Brownian motion \(B_x(\omega)\) is a continuous function of \(x\), the composite function \(f^{(3)}(B_x)\) is a continuous function of \(x\) as well. In addition, for \(x > 0\),
\begin{align*}
\lim_{x' \to x} E\left[f^{(3)}(B_{x'})\right]^3 &= \lim_{x' \to x} \frac{1}{\sqrt{2\pi x'}} \int_{-\infty}^{+\infty} \left[f^{(3)}(y)\right]^3 \exp\left(-\frac{y^2}{2x'}\right) dy \\
&= \frac{1}{\sqrt{2\pi x}} \int_{-\infty}^{+\infty} \left[f^{(3)}(y)\right]^3 \exp\left(-\frac{y^2}{2x}\right) dy = E\left[f^{(3)}(B_x)\right]^3 \quad (4.10)
\end{align*}
which implies \(E\left[f^{(3)}(B_x)\right]^3\) is a continuous function of \(x\), hence it has minimum and maximum in any closed interval. Accordingly, a closed interval which contains \(N_\delta(x_0)\) guarantees the condition (4.2) holds.

\[\square\]

### 5. Application in Econometric Estimation

In both economic and scientific world researchers often face a problem that how to describe two random variables related each other with unknown relationship? For example, if there are observations of \(Y_t\) and \(U_t\), can we depict them by equation:
\[Y_t = f(U_t) + e_t \quad (5.1)\]
where $e_t$, $t = 1, \ldots, T$, are error terms, which form a stationary time series with mean zero and variance $\sigma^2$.

In economic world (for instance, financial market), $U_t$ are generated by random walk, i.e.

$$U_t = U_{t-1} + u_t, \quad t = 1, 2, \ldots,$$

where $U_0 = 0$ and $u_t$ are i.i.d. random variables following normal distribution $N(0, \sigma^2_u)$. Thus $U_t = \sum_{s=1}^t u_s$ is a Brownian motion and $f(U_t)$ in equation (5.1) becomes a function of Brownian motion.

In order to estimate $f(U_t)$, suppose $f(\cdot)$ has continuous third-order derivative, then by theorem 4.1, $f(U_t) \approx f(U_{t_0}) + f'(U_{t_0})(U_t - U_{t_0}) + f''(U_{t_0})(U_t - U_{t_0})^2$. (5.3)

Using expansions (2.8) and (3.5) and taking sufficient large $q$, we have

$$U_t - U_{t_0} = U_{t-t_0} = Z_0 t - t_0 \sqrt{2\pi} + \frac{2}{\sqrt{\pi}} \sum_{n=1}^q Z_n \frac{\sin(n(t-t_0)/2)}{n},$$

$$U_t^2 - U_{t_0}^2 = U_{t}^2 - t_0 = (t - t_0) + 2 \sum_{n=1}^q c_n(t - t_0)X_n,$$

where $Z_n, n = 0, 1, \ldots, q$, are the first $q+1$ terms of orthonormal in $L^2(\Omega)$ for expansion of Brownian motion; $X_n, n = 1, \ldots, q$, are the first $q$ terms of the orthonormal in $L^2(\Omega)$ for expansion of quadratic form of Brownian motion. Plugging above two expressions into (5.3) gives

$$f(U_t) \approx f(U_{t_0}) + f'(U_{t_0})(U_t - U_{t_0}) + f''(U_{t_0})(U_t - U_{t_0})^2$$

$$= f(U_{t_0}) + f'(U_{t_0}) \left( Z_0 \frac{t - t_0}{\sqrt{2\pi}} + \frac{2}{\sqrt{\pi}} \sum_{n=1}^q Z_n \frac{\sin(n(t-t_0)/2)}{n} \right)$$

$$+ f''(U_{t_0}) \left( (t - t_0) + 2 \sum_{n=1}^q c_n(t - t_0)X_n \right)$$

$$:= \alpha_0 + \left( \beta_0 \frac{t - t_0}{\sqrt{2\pi}} + \frac{2}{\sqrt{\pi}} \sum_{n=1}^q \beta_n \frac{\sin(n(t-t_0)/2)}{n} \right)$$

$$+ \left( \gamma_0(t - t_0) + 2 \sum_{n=1}^q \gamma_n c_n(t - t_0) \right)$$

27
where

\[
\begin{align*}
\alpha_0 &= f(U_0), \\
\beta_n &= f'(U_0)Z_n, \quad n = 0, 1, \ldots, q, \\
\gamma_0 &= f''(U_0), \quad \gamma_n = f''(U_0)X_n, \quad n = 1, \ldots, q.
\end{align*}
\]  

(5.6)

Given observations \((Y_t, U_t), t = 1, \ldots, T\), these three classes of parameters \(\alpha, \beta, \gamma\) can be regressed by OLS method, yielding \(\hat{\alpha}, \hat{\beta}, \hat{\gamma}\). Finally, \(f(U_t)\) is estimated by

\[
\hat{f}(U_t) = \hat{\alpha}_0 + \left(\frac{\hat{\beta}_0 t - t_0}{\sqrt{2\pi}} + \frac{2}{\sqrt{\pi}} \sum_{n=1}^q \hat{\beta}_n \frac{\sin(n(t-t_0)/2)}{n}\right) + \left(\hat{\gamma}_0 (t-t_0) + 2 \sum_{n=1}^q \hat{\gamma}_n c_n (t-t_0)\right).
\]

Then estimation of equation (5.1) is given by \(Y_t = \hat{f}(U_t) + \epsilon_t\).

### 6. Conclusion and discussion

The expansions of Brownian motion and its functions can be easily used to estimate unknown relationship between variables if observations are available. In order to estimate equation (1.1), discretize the differential equation

\[
X_{t+\Delta t} - X_t = \mu(X_t)\Delta t + \sigma(X_t)(B_{t+\Delta t} - B_t).
\]  

(6.1)

Rewrite (6.1)

\[
\frac{X_{t+\Delta t} - X_t}{\Delta t} = \mu(X_t) + \sigma(X_t)\frac{B_{t+\Delta t} - B_t}{\Delta t}.
\]  

(6.2)

Denote by \(Y_t = (X_{t+\Delta t} - X_t)/\Delta t\) and \(m(X_t) = \sigma(X_t)(B_{t+\Delta t} - B_t)/\Delta t\). Equation (6.2) becomes

\[
Y_t = \mu(X_t) + E[m(X_t)] + m(X_t) - E[m(X_t)] := f(X_t) + \epsilon_t,
\]  

(6.3)

where \(f(X_t) = \mu(X_t) + E[m(X_t)]\), \(\epsilon_t = m(X_t) - E[m(X_t)]\). Thus equation (6.3) is in the same form as equation (5.1), hence it can be estimated. Another particular case is that when in practice \(\mu(\cdot) = 0\), then

\[
Y_t = \sigma(X_t)\frac{B_{t+\Delta t} - B_t}{\Delta t}
\]  

(6.4)

Squaring equation (6.4) and taking logarithm yields

\[
\log(Y_t^2) = \log(\sigma^2(X_t)) + \log\left(\frac{B_{t+\Delta t} - B_t}{\Delta t}\right)^2
\]  

(6.5)
This equation again can be written as

\[ \tilde{Y}_t = f(X_t) + e_t \]  

(6.6)

where

\[ \tilde{Y}_t = \log (Y_t^2) \]
\[ f(X_t) = \log (\sigma^2(X_t)) + E \left[ \log \left( \frac{B_{t+\Delta t} - B_t}{\Delta t} \right)^2 \right] \]
\[ e_t = \log \left( \frac{B_{t+\Delta t} - B_t}{\Delta t} \right)^2 - E \left[ \log \left( \frac{B_{t+\Delta t} - B_t}{\Delta t} \right)^2 \right] \]

Therefore it can be estimated given observations of \( X_t \). To conclude, the expansions of Brownian motion and its functions are immensely helpful for econometric estimations.

References


