Introduction to Stochastic Calculus

Rajeeva L. Karandikar
Director
Chennai Mathematical Institute
rlk@cmi.ac.in
rkarandikar@gmail.com

The notion of Conditional Expectation of a random variable given a σ-field confuses many people who wish to learn stochastic calculus.

Let X be a random variable. Suppose we are required to make a *guess* for the value of X, we would like to be as close to X as possible. Suppose the penalty function is square of the error. Thus we wish to minimize

$$\mathbb{E}[(X-a)^2] \tag{1}$$

where a is the guess. The value of a that minimizes (**??**) is the mean $\mu = \mathbb{E}[X]$.

Let Y be another random variable which we can observe and we are allowed to use the observation Y while guessing X, *i.e.* our guess could be a function of Y. We should choose the function g such that

$$\mathbb{E}[(X - g(Y))^2] \tag{2}$$

takes the minimum possible value. When Y takes finitely many values, the function g can be thought of as a look-up table.

Assuming $E[X^2] < \infty$, it can be shown that there exists a function g (Borel measurable function from $\mathbb{R}$ to $\mathbb{R}$) such that

$$\mathbb{E}[(X - g(Y))^2] \leq \mathbb{E}[(X - f(Y))^2]. \tag{3}$$

$\mathscr{H} = L^2(\Omega, \mathscr{F}, \mathbb{P})$ - the space of all square integrable random variables on $(\Omega, \mathscr{F}, \mathbb{P})$ with norm $\|Z\| = \sqrt{\mathbb{E}[Z^2]}$ is an Hilbert space and

$$\mathscr{K} = \{f(Y) : f \text{ measurable}, \mathbb{E}[(f(Y))^2] < \infty\}$$

is a closed subspace of $\mathscr{H}$ and hence given $X \in \mathscr{H}$, there is a unique element in $g \in \mathscr{K}$ that satisfies

$$\mathbb{E}[(X - g(Y))^2] \leq \mathbb{E}[(X - f(Y))^2].$$

Further,

$$\mathbb{E}[(X - g(Y))f(Y)] = 0 \ \forall f \in \mathscr{K}.$$

For $X \in \mathscr{H}$, we define $g(Y)$ to be the conditional expectation of X given Y, written as $E[X \mid Y] = g(Y)$. One can show that for $X, Z \in \mathscr{H}$ and $a, b \in \mathbb{R}$, one has

$$\mathbb{E}[aX + bZ \mid Y] = a\mathbb{E}[X \mid Y] + b\mathbb{E}[Z \mid Y] \tag{4}$$

and

$$X \leq Z \text{ implies } \mathbb{E}[X \mid Y] \leq \mathbb{E}[Z \mid Y]. \tag{5}$$

Here and in this document, statements on random variables about equality or inequality are in the sense of almost sure, *i.e.* they are asserted to be true outside a null set.

If (X, Y) has bivariate Normal distrobution with means zero, variances 1 and correlation coefficient r, then

$$\mathbb{E}[X \mid Y] = rY.$$

Thus $g(y) = ry$ is the choice for conditional expectation in this case.

Now if instead of one random variable Y, we were to observe $Y_1, \ldots, Y_m$, we can similarly define

$$\mathbb{E}[X \mid Y_1, \ldots Y_m] = g(Y_1, \ldots Y_m)$$

where g satisfies

$$\mathbb{E}[(X - g(Y_1, \ldots Y_m))^2] \leq \mathbb{E}[(X - f(Y_1, \ldots Y_m))^2] = 0$$

$\forall$ measurable functions f on $\mathbb{R}^m$. Once again we would need to show that the conditional expectation is linear and monotone.

Also if we were to observe an infinite sequence, we have to proceed similarly, with f, g being Borel functions on $\mathbb{R}^\infty$, while if were to observe an $\mathbb{R}^d$ valued continuous stochastic process $\{Y_t : 0 \le t < \infty\}$, we can proceed as before with f, g being Borel functions on $C([0,\infty), \mathbb{R}^d)$.

In each case we will have to write down properties and proofs thereof and keep doing the same as the class of observable random variables changes.

For a random variable Y, (defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$), the smallest σ-field $\sigma(Y)$ with respect to which Y is measurable (also called the σ-field generated by Y) is given by

$$\sigma(Y) = \left\{A \in \mathscr{F} : A = \{Y \in B\},\ B \in \mathscr{B}(\mathbb{R})\right\}.$$

An important fact: A random variable Z can be written as $Z = g(Y)$ for a measurable function g if and only if Z is measurable with respect to $\sigma(Y)$. In view of these observations, one can define conditional expectation given a sigma field as follows.

Let X be a random variable on $(\Omega, \mathscr{F}, \mathbb{P})$ with $\mathbb{E}[|X|^2] < \infty$ and let $\mathscr{G}$ be a sub-σ field of $\mathscr{F}$. Then the *conditional expectation* of X given $\mathscr{G}$ is defined to be the $\mathscr{G}$ measurable random variable Z such that

$$\mathbb{E}[(X-Z)^2] \leq \mathbb{E}[(X-U)^2], \ \forall U \ -\mathscr{G}\text{-measurable r.v.} \tag{6}$$

The random variable Z is characterized via

$$\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Z\mathbf{1}_A] \ \forall A \in \mathscr{G} \tag{7}$$

It should be remembered that one mostly uses it when the sigma field $\mathscr{G}$ is generated by a bunch of observable random variables and then Z is a function of these observables.

We can see that information content in observing Y or observing Y^3 is the same and so intutively

$$\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid Y^3]$$

This can be seen to be true on observing that $\sigma(Y) = \sigma(Y^3)$ as Y is measurable w.r.t. $\sigma(Y^3)$ and Y^3 is measurable w.r.t $\sigma(Y)$.

We extend the definition of $\mathbb{E}[X \mid \mathscr{G}]$ to integrable X as follows. For $X \geq 0$, let $X^n = X \wedge n$. Then X^n is square integrable and hence $Z^n = \mathbb{E}[X^n \mid \mathscr{G}]$ is defined and $0 \leq Z^n \leq Z^{n+1}$ as $X^n \leq X^{n+1}$. So we define $Z = \lim Z^n$ and since

$$\mathbb{E}[X^n \mathbf{1}_A] = \mathbb{E}[Z^n \mathbf{1}_A] \;\; \forall A \in \mathscr{G}$$

monotone convergence theorem implies

$$\mathbb{E}[X \mathbf{1}_A] = \mathbb{E}[Z \mathbf{1}_A] \;\; \forall A \in \mathscr{G}$$

and it follows that $\mathbb{E}[Z] = \mathbb{E}[X] < \infty$ and we define

$$\mathbb{E}[X \mid \mathscr{G}] = Z.$$

Now given X such that $\mathbb{E}[|X|] < \infty$ we define

$$\mathbb{E}[X \mid \mathscr{G}] = \mathbb{E}[X^+ \mid \mathscr{G}] - \mathbb{E}[X^+ \mid \mathscr{G}]$$

where $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$. It follows that

$$\mathbb{E}\left[\mathbb{E}[X\mathscr{G}]\mathbf{1}_A\right] = \mathbb{E}\left[X\mathbf{1}_A\right]. \ \forall A \in \mathscr{G} \tag{8}$$

The property (**??**) characterizes the conditional expectation.

Let X, X_n, Z be integrable random variables on $(\Omega, \mathscr{F}, \mathbb{P})$ for $n \geq 1$ and $\mathscr{G}$ be a sub-σ field of $\mathscr{F}$ and $a, b \in \mathbb{R}$. Then we have

(i) $\mathbb{E}[aX + bZ \mid \mathscr{G}] = a\mathbb{E}[X \mid \mathscr{G}] + b\mathbb{E}[Z \mid \mathscr{G}]$.

(ii) If Y is $\mathscr{G}$ measurable and bounded then $\mathbb{E}(XY \mid \mathscr{G}) = Y\mathbb{E}(X \mid \mathscr{G})$.

(iii) $X \leq Z \Rightarrow \mathbb{E}(X \mid \mathscr{G}) \leq \mathbb{E}(Z \mid \mathscr{G})$.

(iv) $\mathbb{E}\big[\mathbb{E}[X \mid \mathscr{G}]\big] = \mathbb{E}[X]$.

(v) For a sub-*sigma* field $\mathscr{H}$ with $\mathscr{G} \subseteq \mathscr{H} \subseteq \mathscr{F}$ one has

$$\mathbb{E}\big[\mathbb{E}[X \mid \mathscr{H}] \mid \mathscr{G}\big] = \mathbb{E}[X \mid \mathscr{G}]$$

(vi) $|\mathbb{E}[X \mid \mathscr{G}]| \leq \mathbb{E}[|X| \mid \mathscr{G}]$

(vii) If $\mathbb{E}[|X_n - X|] \to 0$ then $\mathbb{E}\Big[|\mathbb{E}[X_n \mid \mathscr{G}] - \mathbb{E}[X \mid \mathscr{G}]|\Big] \to 0$.

Consider a situation from finance. Let S_t be a continuous process denoting the market price of shares of a company UVW. Let A_t denote the value of the assets of the company, B_t denote the value of contracts that the company has bid, C_t denote the value of contracts that the company is about to sign. The process S is observed by the public but the processes A, B, C are not observed by the public at large. Hence, while making a decision on investing in shares of the company UVW, an investor can only use information $\{S_u : 0 \leq u \leq t\}$.

Indeed, in trying to find an optimal investment policy $\pi = (\pi_t)$ (optimum under some criterion), the class of all investment strategy must be taken as all processes π such that for each t, π_t is a (measurable) function of $\{S_u : 0 \leq u \leq t\}$ *i.e.* for each t, π_t is measurable w.r.t. the σ-field $\mathscr{G}_t = \sigma(S_u : 0 \leq u \leq t)$.

In particular, the strategy cannot be a function of the unobserved processes A, B, C.

Thus it is useful to define for $t \geq 0$, $\mathscr{G}_t$ to be the σ- field generated by all the random variables observable upto time t and then require any action to be taken at time t (an estimate of some quantity or investment decision) should be measurable with respect to $\mathscr{G}_t$.

A filtration is any family $\{\mathscr{F}_t\}$ of sub-σ- fields indexted by $t \in [0,\infty)$ such that $\mathscr{F}_s \subseteq \mathscr{F}_t$ for $s \leq t$.

A process $X = (X_t)$ is said to be adapted to a filtration $\{\mathscr{F}_t\}$ if for each t, X_t is $\mathscr{F}_t$ measurable.

While it is not required in the definition, in most situations, the filtration $(\mathscr{F}_t)$ under consideration would be chosen to be $\{\mathscr{F}_t^Z : t \geq 0\}$ where

$$\mathscr{F}_t^Z = \sigma(Z_u : 0 \leq u \leq t)$$

for some process Z, which itself could be vector valued.

Sometimes, a filtration is treated as a mere technicality. We would like to stress that it is not so. It is a technical concept, but can be a very important ingredient of the analysis.

For example in the finance example, we must require any investment policy to be adapted to $\mathscr{F}_t = \sigma(S_u : 0 \le u \le t)$.

While for technical reasons we may also consider

$$\mathscr{H}_t = \sigma(S_u, A_u, B_u, C_u : 0 \le u \le t)$$

an $(\mathscr{H}_t)$ adapted process cannot be taken as an investment strategy.

Thus $(\mathscr{F}_t)$ is lot more than mere technicality.

A sequence of random variables $\{M_n : n \geq 1\}$ is said to be a *martingale* if

$$\mathbb{E}[M_{n+1} \mid M_0, M_1, \ldots, M_n] = M_n, \; n \geq 1.$$

The process $D_n = M_n - M_{n-1}$ is said to be a martingale difference sequence. An important property of martingale difference sequence:

$$\mathbb{E}[(D_1 + D_2 + \ldots + D_n)^2] = \mathbb{E}[(D_1)^2 + (D_2)^2 + \ldots + (D_n)^2]$$

as cross product terms vanish: for $i < j$,

$$\mathbb{E}[D_i D_j] = \mathbb{E}\big[D_i \mathbb{E}[D_j \mid M_1, M_2, \ldots, M_{j-1}]\big] = 0$$

as $\{D_j\}$ is a martingale difference sequence.

Thus martingale difference sequence shares a very important property with mean zero independent random variables:

Expected value of square of the sum
equals Expected value of sum of squares.

For iid random variables, this property is what makes the law of large numbers work. Thus martingales share many properties with sum of iid mean zero random variable, Law of large numbers, central limit theorem etc.

Indeed, Martingale is a single most powerful tool in modern probability theory.

Let (M_n) be a martingale and $D_n = M_n - M_{n-1}$ be the martingale difference sequence. For $k \geq 1$, let g_k be a bounded function on $\mathbb{R}^k$ and let

$$G_k = g_k(M_0, M_1, \ldots, M_{k-1})D_k.$$

Then it can be seen that G_k is also a martingale difference sequence and hence

$$N_n = \sum_{k=1}^n g_k(M_0, M_1, \ldots, M_{k-1})D_k$$

is a martingale. Note that the multiplier is a function of $M_0, M_1, \ldots, M_{k-1}$ and NOT of $M_0, M_1, \ldots, M_k$. In the later case, the result will not be true in general. Hence

$$\mathbb{E}[(G_1 + G_2 + \ldots + G_n)^2] = \mathbb{E}[(G_1)^2 + (G_2)^2 + \ldots + (G_n)^2]$$

The processes $\{N_n\}$ obtained via

$$N_n = \sum_{k=1}^{n} g_k(M_0, M_1, \ldots, M_{k-1})(M_k - M_{k-1})$$

is known as the martingale transform of M. Note that

$$\mathbb{E}(N_n)^2 = \sum_{k=1}^{n} \mathbb{E}[(g_k(M_0, M_1, \ldots, M_{k-1})(M_k - M_{k-1}))^2]$$

$$= \sum_{k=1}^{n} \mathbb{E}[(g_k(M_0, M_1, \ldots, M_{k-1})^2 \mathbb{E}[(M_k - M_{k-1})^2 \mid \sigma(M_1, \ldots, M_{k-1})]$$

and hence if g_k is bounded by 1, then

$$\mathbb{E}(N_n)^2 \leq \sum_{k=1}^{n} \mathbb{E}[(M_k - M_{k-1}))^2]$$

$$= \sum_{k=1}^{n} \mathbb{E}[\mathbb{E}[(M_k - M_{k-1})^2 \mid \sigma(M_1, \ldots, M_{k-1})]$$

A process (M_t) adapted to a filtration $(\mathscr{F}_t)$, (for $s < t$, $\mathscr{F}_s$ is sub-σ field of $\mathscr{F}_t$ and M_t is $\mathscr{F}_t$ measurable) is said to be a martingale if for $s < t$,

$$\mathbb{E}\big[M_t \mid \mathscr{F}_s\big] = M_s.$$

For $0 = t_0 < t_1 < \ldots < t_m$, in analogy with the discrete case one has

$$\mathbb{E}\big[(M_{t_m} - M_{t_0})^2\big] = \mathbb{E}\big[\textstyle\sum_{j=0}^{m-1}(M_{t_{j+1}} - M_{t_j})^2\big]$$

Under minimal conditions one can ensure that the *paths* $t \mapsto M_t(\omega)$ are right continuous functions with left limts (r.c.l.l.) for all $\omega \in \Omega$. So we will only consider r.c.l.l. martingales.

For any integrable random variable Z on $(\Omega, \mathscr{F}, \mathbb{P})$, the process

$$Z_t = \mathbb{E}[Z \mid \mathscr{F}_t]$$

is a martingale. This follows from the smoothing property of conditional expectation: for $\mathscr{G} \subseteq \mathscr{H}$

$$\mathbb{E}\big[\mathbb{E}[Z \mid \mathscr{H}] \mid \mathscr{G}\big] = \mathbb{E}[X \mid \mathscr{G}].$$

Here $\mathscr{H} = \mathscr{F}_t$ and $\mathscr{G} = \mathscr{F}_s$.

One of the most important result on Martingales is:

Doob's Maximal inequality

For an r.c.l.l. Martingale (M_t), for all $\lambda > 0$,

$$\mathbb{P}\big(\sup_{0\leq t\leq T}|M_t| > \lambda\big) \leq \frac{1}{\lambda^2}\mathbb{E}\big(M_T^2\big)$$

and

$$\mathbb{E}\big(\sup_{0\leq t\leq T}|M_t|^2\big) \leq 4\mathbb{E}\big(M_T^2\big).$$

It can be shown that for any t

$$\sum_{k=0}^{n-1}(M_{\frac{t(k+1)}{n}} - M_{\frac{t(k)}{n}})^2 \to A_t$$

where A_t is an increasing r.c.l.l. adapted process, called the quadratic variation of M. Further, when $\mathbb{E}[M_T^2] < \infty$ for all T, then

$$\sum_{k=0}^{n-1}\mathbb{E}\Big((M_{\frac{t(k+1)}{n}} - M_{\frac{t(k)}{n}})^2 \mid \mathscr{F}_{\frac{t(k)}{n}}\Big) \to B_t$$

where B_t is also an increasing r.c.l.l. adapted processand $M_t^2 - A_t$ and $M_t^2 - B_t$ are martingales.

The quadratic variation process (A_t) of M is written as $[M,M]_t$ and B_t is written as $\langle M,M\rangle_t$. $\langle M,M\rangle_t$ is called the the predictable quadratic variation of M.

When M has continuous paths ($t \mapsto M_t(\omega)$ is continuous for all ω) then $\langle M,M\rangle_t$=$[M,M]_t$.

For Brownian motion (β_t), $[\beta,\beta]_t = \sigma^2 t$ and $\langle \beta,\beta\rangle_t = \sigma^2 t$, while for a Poisson process (M_t) with rate λ, $N_t = M_t - \lambda t$ is a martingale with then $[N,N]_t = \lambda t$ and $\langle N,N\rangle_t = M_t$.

A stopping time τ is a $[0,\infty]$ valued random variable such that

$$\{\omega : \tau(\omega) \leq t\} \in \mathscr{F}_t, \ \forall t\}$$

The idea behind the concept: Whether the event $\{\tau \leq t\}$ has happened or not should be known at time t, when one has observed $\{X_u : u \leq t.\}$.

If (X_t) is an continuous adapted process then

$$\tau = \inf\{t : |X_t| \geq \theta\}$$

is a stopping time : for any $\theta \in \mathbb{R}$. For example, $\{\tau \leq 7\}$ has happened iff $\{\sup_{s \leq 7} |X_s| \geq \theta\}$.

In real life we are used to talking in terms of time in reference to time of occurence of other events, such as meet me within five minutes of your arrival or let us meet as soon as the class is over, or sell shares of company X as soon as its price falls below 100. Thse are stopping times

Of course a statement like buy shares of company X today if the price is going to increase tomorrow by 3 is meaningless. One can talk but cannot implement. The time specified in the statement is not a stopping time.

The time at which an americal option can be exercised has to be a stopping time as the decision to exercise the option at a given time or not has to depend upon the observations upto that time.

Martingales and stopping times:

Let (M_t) be a martingale, τ be a stopping time and $(X_t),(Y_t)$ be r.c.l.l. adapted processes and Y be bounded.

1. (X_t) is a martingale if and only if $\mathbb{E}[X_\tau]=\mathbb{E}[X_0]$ for all bounded stopping times τ
2. $N_t = M_{t\wedge\tau}$ is a martingale.
3. $U_t = Y_{t\wedge\tau}\big(M_t - M_{t\wedge\tau}\big)$ is also a martingale.

A martingale (M_t) is said to be a square integrable martingale if $\mathbb{E}[M_t^2]<\infty$ for all $t<\infty$.

A process (X_t) is said to be *locally bounded* if there exists an increasing sequence of stopping times $\{\tau^n\}$ increasing to ∞ such that $X_t^n = X_{t\wedge\tau^n}$ is bounded for every n.

It can be seen that every (X_t) such that X_0 is bounded and (X_t) is continuous is locally bounded.

Likewise, an adapted process (M_t) is said to be a *local martingale* (or locally square integrable martingale) if there exists an increasing sequence of stopping times $\{\tau^n\}$ increasing to ∞ such that $M_t^n = M_{t\wedge\tau^n}$ is a martingale (is a square integrable martingale) for every n.

The sequence $\{\tau^n\}$ is called a localizing sequence.

Let M be a local martingale with localizing sequence $\{\tau^n\}$. Thus

$$\mathbb{E}[M_{t\wedge\tau^n} \mid \mathscr{F}_s] = M_{s\wedge\tau^n}.$$

So if M is bounded then M is a martingale. Other sufficient conditions for a local martingale to be a martingale are

- $\mathbb{E}\left[\sup_{0\leq s\leq T}|M_s|\right] < \infty$
- $\mathbb{E}\left[|M_T|^2\right] < \infty, \ \forall T < \infty$
- $\mathbb{E}\left[|M_T|^{1+\delta}\right] < \infty, \ \forall T < \infty,$ for some $\delta > 0$.

Let M be a $[0,\infty)$ valued local martingale with localizing sequence $\{\tau^n\}$. Then since

$$\mathbb{E}[M_{t\wedge\tau^n}] = \mathbb{E}[M_0]$$

using Fatou's lemma it follows that for all $t > 0$,

$$\mathbb{E}[M_t] \leq \mathbb{E}[M_0].$$

If $\mathbb{E}[M_t] = \mathbb{E}[M_0]$ it is a martingale.

When (M_t) is a locally square integrable martingale, it can be shown that for any t

$$\sum_{k=0}^{n-1}(M_{\frac{t(k+1)}{n}} - M_{\frac{t(k)}{n}})^2 \to A_t$$

where again A_t is an increasing adapted process such that $M_t^2 - A_t$ is a local martingale. A is again called quadratic variation process, written as $[M, M]_t$. One can also show that there is an increasing adapted r.c.l.l. process B such that

$$M_t^n = M_{t\wedge\tau^n}, \;\; \langle M^n, M^n\rangle_t = B_{t\wedge\tau^n}.$$

and then B_t is written as $\langle M, M\rangle_t$.

Let (M_t) be a locally square integrable martingale with $M_0 = 0$ and let $\{\tau^n\}$ be a localizing sequence such that $M_t^n = M_{t\wedge\tau^n}$ is a martingale for every n.

Then, using $\mathbb{E}[M^2_{t\wedge\tau^n}] = \mathbb{E}[\langle M, M\rangle_{t\wedge\tau^n}] = \mathbb{E}[[M, M]_{t\wedge\tau^n}]$ and Doob's inequality, one has

$$\mathbb{E}\Big(\sup_{0\le s\le t\wedge\tau^n} |M_s|^2\Big) \le 4\mathbb{E}[\langle M, M\rangle_{t\wedge\tau^n}] \tag{9}$$

as well as

$$\mathbb{E}\Big(\sup_{0\le s\le t\wedge\tau^n} |M_s|^2\Big) \le 4\mathbb{E}[[M, M]_{t\wedge\tau^n}] \tag{10}$$

A Brownian motion (β_t) is a continuous process with $\beta_0 = 0$ and such that for $t_1 < t_2 < \ldots < t_m$; $m \geq 1$;

$$(\beta_{t_1}, \beta_{t_2}, \ldots, \beta_{t_m})$$

has a multivariate Normal distribution with mean vector 0 and variance co-variance matrix $\Sigma = ((\sigma_{ij}))$ given by

$$\sigma_{ij} = \min(t_i, t_j)$$

It can be checked that (with $t_0 = 0$)

$$\beta_{t_1} - \beta_{t_0}, \beta_{t_2} - \beta_{t_1}, \ldots, \beta_{t_m} - \beta_{t_{m-1}}$$

are independent Normal random variables, with mean 0 and $Var(\beta_{t_j} - \beta_{t_{j-1}}) = t_j - t_{j-1}$.

One of the most important properties of a Brownian motion is that (β_t) is a martingale with

$$[\beta, \beta]_t = \langle \beta_t, \beta_t \rangle_t = t.$$

Thus, $\beta_t^2 - t$ is a martingale and

$$Q_t^n = \sum_{k=1}^{\infty} \left(\beta_{\frac{k}{2^n} \wedge t} - \beta_{\frac{k-1}{2^n} \wedge t} \right)^2 \rightarrow t$$

where the convergence is in probability.

The Q_t^n for a fixed n, t is a finite sum, indeed it has $j = [2^n t] + 1$ terms, where $[2^n t]$ is the integer part of $2^n t$ since for k such that $[2^n t] + 1 < k$, $\frac{k-1}{2^n} \wedge t = t$ and $\frac{k}{2^n} \wedge t = t$.

Let $U_k^n = \left(\beta_{\frac{k}{2^n}\wedge t} - \beta_{\frac{k-1}{2^n}\wedge t}\right)^2 - \left(\frac{k}{2^n}\wedge t - \frac{k-1}{2^n}\wedge t\right)$

Then $W_t^n = \sum_{k=1}^{[2^n t]+1} U_k^n = Q_t^n - t$ and using properties of Normal distribution, it follows that for each n, $\{U_k^n : 1 \leq k \leq [2^n t]\}$ are i.i.d. random variables, with mean 0 and variance $\frac{2}{2^{2n}}$. Hence

$$\text{Variance}(W_t^n) \leq ([2^n t] + 1)\frac{2}{2^{2n}} \leq \frac{2t+1}{2^n}$$

and hence W_t^n converges to zero in probability and so $Q_t^n \to t$ in probability.

With little work, it can be shown that Q_t^n converges to t, uniformly on $t \in [0, T]$ for all $T < \infty$ almost surely.

In fact, for a continuous process (X_t), if $X_0 = 0$ and if

$$(X_t) \text{ and } (Y_t = X_t^2 - t) \text{ are martingales}$$

then (X_t) is Brownian motion (Levy's characterization).

For this, continuity of (X_t) is very important. Result not true for r.c.l.l. martinagles. $X_t = N_t - t$ is a counter example where (N_t) is a poisson process with parameter 1.

Let us fix an observable process (Z_t) (it could be vector valued) with Z_t denoting the observation at time t and let

$$\mathscr{F}_t = \sigma(Z_u : u \leq t)$$

denote the filtration generated by observations.

A Brownian motion w.r.t. this filtration $(\mathscr{F}_t)$ means that β_t and $\beta_t^2 - t$ are martingales w.r.t. this filtration $(\mathscr{F}_t)$.

Let (β_t) be a Brownian motion w.r.t. a filtration $(\mathscr{F}_t)$.

There are occasions when one gets expressions that are similar to Riemann-Steltjes sums for the integral $\int_0^t Y_s d\beta_s$ where (Y_s) is say a continuous $(\mathscr{F}_t)$ adapted process:

$$R_t^n = \sum_{k=1}^{\infty} Y_{s_{k,n}\wedge t} \left(\beta_{\frac{k}{2^n}\wedge t} - \beta_{\frac{k-1}{2^n}\wedge t}\right)$$

where $\frac{k-1}{2^n} \leq s_{k,n} \leq \frac{k}{2^n}$.

For Riemann-Steltjes integral $\int_0^t Y_s d\beta_s$ to exist, the above sums (R_t^n) should converge (as $n \to \infty$) for all choices of $s_{k,n}$ satisfying $\frac{k-1}{2^n} \leq s_{k,n} \leq \frac{k}{2^n}$.

Let $Y = \beta$. Let us choose $s_{k,n}$ to be the lower end point in every interval : $s_{k,n} = \frac{k-1}{2^n}$ and for this choice let us denote the Riemann sums R_t^n by A_t^n.

Thus

$$A_t^n = \sum_{k=1}^{\infty} \beta_{\frac{k-1}{2^n} \wedge t} \left(\beta_{\frac{k}{2^n} \wedge t} - \beta_{\frac{k-1}{2^n} \wedge t} \right)$$

Let us now choose $s_{k,n}$ to be the upper end point in every interval : $s_{k,n} = \frac{k}{2^n}$ and for this choice let us denote the Riemann sums R_t^n by B_t^n.

Thus

$$B_t^n = \sum_{k=1}^{\infty} \beta_{\frac{k}{2^n} \wedge t} \left(\beta_{\frac{k}{2^n} \wedge t} - \beta_{\frac{k-1}{2^n} \wedge t} \right)$$

Recall

$$A_t^n = \sum_{k=1}^{\infty} \beta_{\frac{k-1}{2^n} \wedge t} \left(\beta_{\frac{k}{2^n} \wedge t} - \beta_{\frac{k-1}{2^n} \wedge t} \right)$$

while

$$B_t^n = \sum_{k=1}^{\infty} \beta_{\frac{k}{2^n} \wedge t} \left(\beta_{\frac{k}{2^n} \wedge t} - \beta_{\frac{k-1}{2^n} \wedge t} \right)$$

and hence

$$B_t^n - A_t^n = \sum_{k=1}^{\infty} \left(\beta_{\frac{k}{2^n} \wedge t} - \beta_{\frac{k-1}{2^n} \wedge t} \right)^2 = Q_t^n.$$

Since $Q_t^n \to t$, A_t^n and B_t^n cannot converge to the same limit and hence the integral $\int_0^t Y_s d\beta_s$ cannot be defined as a Riemann-Steltjes integral.

Indeed, Brownian motion is not an exception but the rule. It can be shown that if a continuous process is used to model stock price movements, then it can not be a process with bounded variation as that would imply existence of arbitrage opportunities. Hence, we cannot use Riemann-Steltjes integrals in this context. The Riemann like sums appear when one is trying to compute the value function for a trading strategy, where the process used to model the stock price appears as an integrator.

It turns out that in the context of construction of diffusion processes, Riemann sums, where one evaluates the integrand at the lower end point, are appropriate. So taking that approach, Ito defined stochastic integral (in '40s).

It so happens that in the context of applications to finance, the same is natural and so Ito Calculus could be used as it is. This was observed in the early '80s.

Let X_t denote the stock price at time t. Consider a simple trading strategy, where buy-sell takes place at fixed times $0 = s_0 < s_1 < \ldots < s_{m+1}$. Suppose the number of shares held by the investor during $(s_j, s_{j+1}]$ is a_j. The transaction takes place at time s_j and so the investor can use information available to her/him before time t, so a_j should be $\mathscr{F}_{s_j}$ measurable. So the trading strategy is given by: at time t hold f_t shares where

$$f_s = a_0 \mathbf{1}_{\{0\}}(s) + \sum_{j=0}^{m} a_j \mathbf{1}_{(s_j, s_{j+1}]}(s)$$

and then the net gain/loss for the investor due to trading in shares upto time t is given by

$$V_t = \sum_{j=0}^{m} a_j (X_{s_{j+1} \wedge t} - X_{s_j \wedge t}) = \int_0^t f_u dX_u$$

Now suppose (X_t) is a square integrable martingale with r.c.l.l. paths (right continuous with left limits). Recalling

$f_s = a_0\mathbf{1}_{\{0\}}(s) + \sum_{j=0}^{m} a_j \mathbf{1}_{(s_j,s_{j+1}]}(s)$, $V_t = \sum_{j=0}^{m} a_j(X_{s_{j+1}\wedge t} - X_{s_j\wedge t})$.

it follows that (V_t) is also a martingale and

$$\begin{aligned}[V,V]_t &= \sum_{j=0}^{m-1} a_j^2([X,X]_{s_{j+1}\wedge t} - [X,X]_{s_j\wedge t}) \\ &= \int_0^t f_u^2 d[X,X]_u\end{aligned}$$

Hence, using (**??**)

$$\mathbb{E}\left[\sup_{t\le T}(\int_0^t f_u dX_u)^2\right] \le 4\mathbb{E}\left[\int_0^T f_u^2 d[X,X]_u\right].$$

The estimate

$$\mathbb{E}\left[\sup_{t\le T}(\int_0^t f_u dX_u)^2\right] \le 4\mathbb{E}\left[\int_0^T f_u^2 d[X,X]_u\right]$$

for *simple* integrands f can be used to extend the integral to integrands g that can be approximated by simple functions in the norm $\|\cdot\|_X$ given by RHS above:

$$\|f\|_X^2 = \mathbb{E}\left[\int_0^T f_u^2 d[X,X]_u\right]$$

Let (Y_t) be a bounded adapted continuous process and (X_t) be a square integrable martingale. Let $t_m^n = \frac{m}{n}$. Then it can be shown that (Y_t) can be approximated in the *norm* $\|\cdot\|_X$ by simple processes Y^n defined by

$$Y_t^n = \sum_{m=0}^{\infty} Y_{t_m^n \wedge t} \mathbf{1}_{(t_m^n,\, t_{m+1}^n]}(t)$$

and hence the stochastic integral $\int_0^t Y_u dX_u$ is defined and is given by

$$\int_0^t Y_u dX_u = \lim_{n\to\infty} \sum_{m=0}^{\infty} Y_{t_m^n \wedge t}\Big(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\Big).$$

The two infinite sums above are actually finite sums for every n, t (if $m > tn$ the terms are zero). Also, the limit above is in the sense of convergence in probability, uniformly in $t \in [0, T]$ (follows from Doob's inequality).

In other words, for every $\varepsilon > 0$, as n tends to infinity

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|\int_0^t Y_u dX_u - \sum_{m=0}^{\infty} Y_{t_m^n\wedge t}\big(X_{t_{m+1}^n\wedge t} - X_{t_m^n\wedge t}\big)\right| > \varepsilon\right] \to 0.$$

This shows that the stochastic integral $\int_0^t Y_u dX_u$ is a process with r.c.l.l.paths. Further, if X is a continuous process then so is $\int_0^t Y_u dX_u$ continuous process. Of course, when X is a process with bounded variation the integral $\int_0^t Y_u dX_u$ is defined as the Riemann-Steltjes integral.

One can directly verify that when X is a square integrable martingale,

$$\int_0^t Y_u^n dX_u = \sum_{m=0}^{\infty} Y_{t_m^n \wedge t}\left(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\right)$$

is a square integrable martingale. Then convergence of $\int_0^t Y_u^n dX_u$ in L^2 to $\int_0^t Y_u dX_u$ for every t implies that

$$M_t = \int_0^t Y_u dX_u$$

is also a square integrable martingale.

One of the important properties of the integral $Z_t = \int_0^t Y_u dX_u$ so defined is that for a stopping time τ, defining $X_t^{[\tau]} = X_{t\wedge\tau}$,

$$\int_0^t Y_u dX_u^{[\tau]} = Z_{t\wedge\tau}.$$

This can be seen to be true for simple integrands and hence for bounded continuous processes Y. Thus if X is a locally square integrable martingale so that for stopping times $\{\tau^n\}$ increasing to ∞, $X^n = X^{[\tau^n]}$ are square integrable martingales, then $Z^n = \int Y dX^n$ satisfy

$$Z_t^n = Z_{t\wedge\tau^n}^m, \text{ for } m \geq n.$$

Thus we can define $Z_t = Z_t^n$ for $\tau^{n-1} < t \leq \tau^n$ and we would have

$$Z_t^n = Z_{t \wedge \tau^n}.$$

Thus we can define $Z = \int Y dX$. Thus for a locally square integrable martingale X, for all continuous adapted processes Y we continue to have for every $\varepsilon > 0$, as n tends to infinity

$$\mathbb{P}\left[\sup_{0 \leq t \leq T} \left| \int_0^t Y_u dX_u - \sum_{m=0}^{\infty} Y_{t_m^n \wedge t}\left(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\right) \right| > \varepsilon \right] \to 0.$$

By construction, each Z^n is a martingale and hence Z is a local martingale.

If X were a process with bounded variation paths on $[0, T]$ for every T, of course $\int_0^t YdX$ is defined as Riemann Steltjes integral and we have

$$\mathbb{P}\left[\sup_{0\le t\le T}\left|\int_0^t Y_u dX_u - \sum_{m=0}^{\infty} Y_{t_m^n\wedge t}\left(X_{t_{m+1}^n\wedge t} - X_{t_m^n\wedge t}\right)\right| > \varepsilon\right] \to 0$$

and then $\int_0^t Y_u dX_u$ is a process with bounded variation paths.

An r.c.l.l. process X is said to be a *semimartingale* if it admits a decomposition

$$X_t = M_t + A_t,$$

where (M_t) is a locally square integrable martingale and (A_t) is an adapted process whose paths $t \mapsto A_t(\omega)$ have bounded variation on $[0, T]$ for every $T < \infty$ for every ω.

Such a process A will be called process with bounded variation paths.

Let X be a semimartingale, with $X = M + A$, M a locally square integrable martingale and A a process with bounded variation paths.

For a bounded continuous adapted process Y, the integral $\int_0^t Y_s dX_s$ can be defined as the sum of $\int_0^t Y_s dM_s$ and $\int_0^t Y_s dA_s$ and again one has every $\varepsilon > 0$, as n tends to infinity

$$\mathbb{P}\left[\sup_{0\le t\le T}\left|\int_0^t Y_u dX_u - \sum_{m=0}^{\infty} Y_{t_m^n\wedge t}\left(X_{t_{m+1}^n\wedge t} - X_{t_m^n\wedge t}\right)\right| > \varepsilon\right] \to 0.$$

Thus the integral $\int Y dX$ does not depend upon the decomposition $X = M + A$.

We have seen that for a semimartingale X and a bounded continuous adapted process Y, the integral $\int_0^t Y_s dX_s$ can be defined on the lines of Riemann-Steltjes integral, provided we always evaluate the integrand at the lower end point while taking the Riemann sums:

$$\int_0^t Y_u dX_u = \lim_{n\to\infty} \sum_{m=0}^{\infty} Y_{t_m^n \wedge t}\big(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\big).$$

Further, if X is a local martingale, then so is $Z_t = \int_0^t Y_u dX_u$ and if X is a process with bounded variation paths, then so is Z and as a result, Z is always a semimartingale.

Question: Just like one extends Riemann integral $\int_0^1 f(x)dx$ defined for say continuous functions g on $[0,1]$ to Lebesgue integral $\int_0^1 g d\lambda$ for all bounded measurable functions, can we extend $\int_0^t Y_s dX_s$ defined for all bounded continuous adapted processes to a larger class of integrands?

The natural class of processes is then the class that is closed under bounded pointwise convergence and contains all bounded continuous adapted processes.

If we think of every process Y_t as a function on $\bar{\Omega} = [0, \infty) \times \Omega$, then the required class is the collection of all $\mathscr{P}$-measurable bounded functions, where $\mathscr{P}$ is the σ-field generated by bounded continuous adapted processes.

$\mathscr{P}$ is called the predictable σ-field and $\mathscr{P}$ measurable processes are called predictable processes.

A left continuous adapted process (Z_t) is predictable, but an r.c.l.l. (right continuous with left limits) adapted process (V_t) need not be predictable. (V_t) is predictable if for all t, V_t is $\mathscr{F}_{t-\varepsilon}$ measurable.

If (Z_t) is predictable with respect to a filtration $(\mathscr{F}_t)$ then for each t, Z_t is measurable w.r.t.

$$\mathscr{F}_{t-} = \sigma\left(\cup_{s<t}\mathscr{F}_s\right).$$

This justifies the use of the name predictable.

Thus we require a mapping J_X from the class of bounded predictable process f into r.c.l.l. processes such that for adapted continuous processes f, $J_X(f)_t = \int_0^t f dX$ (as defined earlier) and if f^n converges to f bounded pointwise, then

$$\mathbb{P}\left[\sup_{0 \le t \le T} |J_X(f^n)_t - J_X(f)_t| > \varepsilon\right] \to 0.$$

When X is a semimartingale, then such an extension exists and we denote it as $J_X(f)_t = \int_0^t f dX$.

The Bichteller-Dellacherie-Meyer-Mokodobodski Theorem states that for an r.c.l.l. process X, existence of mapping J_X from the class of bounded predictable process f into r.c.l.l. processes such that for adapted continuous processes f, $J_X(f)_t = \int_0^t f dX$ (as defined earlier) and if f^n converges to f bounded pointwise, then

$$\mathbb{P}\left[\sup_{0\leq t\leq T} |J_X(f^n)_t - J_X(f)_t| > \varepsilon\right] \rightarrow 0$$

implies that X s a semimartingale. We will not discuss proof of this.

Let X be a semimartingale. It can be shown that

$$(f, X) \mapsto \int_0^{\cdot} f dX \text{ is linear.}$$

Important fact: If τ is a stopping time, then

$$g_t = \mathbf{1}_{[0,\tau]}(t)$$

is predictable (as it is adapted, left continuous). Now for a bounded predictable process (f_t), $h_t = f_t \mathbf{1}_{[0,\tau]}(t)$ is predictable and

$$\int_0^t h_s dX_s = \int_0^{t \wedge \tau} f_s dX_s = \int f_s dX_s^{[\tau]}$$

where $X_s^{[\tau]} = X_{s \wedge \tau}$.

Localization: An adapted process (f_t) is said to be locally bounded if there exists an increasing sequence of stopping times $\{\tau^n\}$ increasing to ∞ such that $f_t^n = f_t \mathbf{1}_{[0,\tau^n]}(t)$ is bounded for every n.

Here, for a stopping times τ, the stochastic intervals $[0, \tau]$, play the role of compact sets $[-K, K]$ in $\mathbb{R}$. Recall definition of local martingale.

Given f, $\{\tau^n\}$ as above, for a semimartingale X, let

$$Z_t^n = \int_0^t f \mathbf{1}_{[0,\tau^n]}(s) dX_s$$

Then for $m \geq n$

$$Z_t^n = Z_{t \wedge \tau^n}^m$$

Hence we piece together the pieces to define

$$\int_0^t f_s dX_s = Z_t^n \text{ for } \tau^{n-1} < t \leq \tau^n.$$

Thus for a semimartingale X and for locally bounded process f, we have defined stochastic integral $Z_t = \int_0^t f_s dX_s$. It is an r.c.l.l. process.

Moreover,

$$(\Delta Z)_t = f_t (\Delta X)_t.$$

One can verify this for simple f and then for general f via approximation.

If X is a square integrable local martingale, then so is Z, if X has bounded variation paths then so does Z and in general, Z is a semimartingale with decomposition

$$Z_t = \int_0^t f_s dX_s = \int_0^t f_s dM_s + \int_0^t f_s dA_s = N_t + B_t$$

where $X = M + A$, M a locally square integrable martingale, A a process with bounded variation paths and $N_t = \int_0^t f_s dM_s$ is a locally square integrable martingale and $B_t = \int_0^t f_s dA_s$ is a process with bounded variation paths.

Further if g is a locally bounded predictable process, then

$$\int_0^t g dZ = \int_0^t g f dX.$$

One can verify this first when f, g are simple functions and then for a fixed g simple one can first verify it for all f and then the general case.

An r.c.l.l. adapted process Y need not be predictable. However, the process Y^- defined by $Y_t^- = Y_{t-}$ (left limit at t for $t > 0$), $Y_0^- = 0$ is predictable. It is also locally bounded and with $t_m^n = \frac{m}{n}$ one has

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} Y_{t_m^n \wedge t}\left(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\right) = \int_0^t Y_u^- dX_u$$

Note that (using $b^2 - a^2 = 2a(b-a) + (b-a)^2$)

$$X_t^2 - X_0^2 = \sum_{m=0}^{\infty} \left(X^2_{t^n_{m+1} \wedge t} - X^2_{t^n_m \wedge t} \right)$$

$$= \sum_{m=0}^{\infty} 2X_{t^n_m \wedge t} \left(X_{t^n_{m+1} \wedge t} - X_{t^n_m \wedge t} \right) + \sum_{m=0}^{\infty} \left(X_{t^n_{m+1} \wedge t} - X_{t^n_m \wedge t} \right)^2$$

Hence for a semimartingale X,

$$\textstyle\sum_{m=0}^{\infty} \left(X_{t^n_{m+1} \wedge t} - X_{t^n_m \wedge t} \right)^2 \to X_t^2 - X_0^2 - 2 \int_0^t X_u^- \, dX_u.$$

We denote the limit of $\sum_{m=0}^{\infty} \left(X_{t^n_{m+1} \wedge t} - X_{t^n_m \wedge t} \right)^2$ by $[X,X]_t$ and note that it is an increasing adapted process. Further we have

$$X_t^2 = X_0^2 + 2 \int_0^t X_u^- \, dX_u + [X,X]_t.$$

For semimartingales X, Y let us define $[X, Y]$ by the parallelogram identity

$$[X,Y]_t = \frac{1}{4}([X+Y,X+Y]_t - [X-Y,X-Y]_t)$$

It follows that $[X, Y]_t$ is a process with bounded variation paths and

$$[X,Y]_t = \lim_{n\to\infty}\sum_{m=0}^{\infty}\big(X_{t^n_{m+1}\wedge t} - X_{t^n_m\wedge t}\big)\big(Y_{t^n_{m+1}\wedge t} - Y_{t^n_m\wedge t}\big)$$

Also, it follows that

$$X_tY_t = X_0Y_0 + \int_0^t Y_s^- dXs + \int_0^t X_s^- dY_s + [X,Y]_t.$$

This is called *the integration by parts formula*.

Like Lebesgue integrals, we have a Dominated Convergence Theorem for stochastic integrals:

Let (X_t) be a semimartingale and let f^n, f be predictable processes such that

$$|f^n| \leq g$$

where g is some locally bounded predictable process and such that f^n converges pointwise to f. Then

$$\mathbb{P}\left[\sup_{0\leq t\leq T}\left|\int_0^t f_u^n dX_u - \int_0^t f_u dX_u\right| > \varepsilon\right] \to 0.$$

Some Facts: Let X be a semimartingale, with $X_t = M_t + A_t$, where (M_t) is a locally square integrable martingale and (A_t) is a process with bounded variation paths. Let $(\Delta X)_t = X_t - X_{t-}$, (jump at t)

$$\Delta[X,X]_t = (\Delta X)_t^2,$$

$$\sum_{0<s\leq t} (\Delta X)_s^2 \leq [X,X]t.$$

If A has bounded variation paths then

$$[A,A]t = \sum_{0<s\leq t} (\Delta A)_s^2.$$

Moreover, if X is a semimartingale and A is a continuous process with bounded variation paths, then

$$[X,A]_t = 0.$$

To see this

$$\int_0^t X_s dA_s = \sum_{m=0}^{\infty} X_{t_m^n \wedge t}\left(A_{t_{m+1}^n \wedge t} - A_{t_m^n \wedge t}\right)$$

and

$$\int_0^t X_s dA_s = \sum_{m=0}^{\infty} X_{t_{m+1}^n \wedge t}\left(A_{t_{m+1}^n \wedge t} - A_{t_m^n \wedge t}\right)$$

as the integral $\int_0^t X_s dA_s$ is the Riemann-Steltjes and thus

$$[X,A]_t = \sum_{m=0}^{\infty} \left(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\right)\left(A_{t_{m+1}^n \wedge t} - A_{t_m^n \wedge t}\right) = 0.$$

As a result

$$[A,A]_t = 0.$$

For a continuous semimartingale (X_t) the decomposition

$$X_t = M_t + A_t$$

into a continuous local martingale (M_t) with $M_0 = 0$ and a continuous process A with bounded variation paths is unique. The uniqueness is proven as follows. If

$$X_t = M_t + A_t = N_t + B_t$$

with M, N continuous local martingales, $M_0 = N_0 = 0$ and A, B processes with bounded variation, then $U_t = M_t - N_t = B_t - A_t$ is a continuous local martingale with $U_0 = 0$, and has bounded variation paths. Thus, $[U, U]_t = 0$.

As a result (recall integration by parts formula)

$$U_t^2 = \int_0^t U_s dU_s$$

and hence U_t^2 is a continuous local martingale with $U_0^2 = 0$. If $\{\tau^n\}$ is a localizing sequence, then $U^2_{t\wedge\tau^n}$ is a martingale and so

$$\mathbb{E}[U^2_{t\wedge\tau^n}] = 0, \ \forall n$$

and as a result $U^2_{t\wedge\tau^n} = 0$ for all t, n. This proves (X_t) the decomposition

$$X_t = M_t + A_t$$

into a continuous local martingale (M_t) with $M_0 = 0$ and a continuous process A with bounded variation paths is unique.

If X is a semimartingale and f is a locally bounded predictable process, then $V_t = \int_0^t f_s dX_s$ is also a semimartingale with decomposition $V_t = N_t + B_t$ where $X_t = M_t + A_t$ is a decomposition of X, with M a locally square integrable martingale, A a process with bounded variation paths,

$$N_t = \int_0^t f_s dN_s \text{ is a locally square integrable martingale}$$

and

$$B_t = \int_0^t f_s dA_s \text{ is a process with bounded variation paths.}$$

Further,

$$[V, V]_t = \int_0^t f_s^2 d[X, X]_s.$$

Ito formula or Change of variable formula

Let f be a smooth function on $\mathbb{R}$ and let (G_t) be a continuous function with bounded variation paths. If G were continuously differentiable with derivate g, then

$$h(t) = f(G(t))$$

is differentiable with continuous derivative $f'(G(t))g(t)$ and hence

$$\begin{aligned} f(G(t)) =& f(G(0)) + \int_0^t f'(G(s))g(s)ds \\ =& h(0) + \int_0^t f'(G(s))dG(s) \end{aligned}$$

Even if G is not differentiable, one still has

$$f(G(t)) = f(G(0)) + \int_0^t f'(G(s))dG(s).$$

The proof uses taylor's expansion:

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2}f''(c)(b-a)^2$$

where c belongs to the interval $\{a,b\}$, i.e. $c \in (a,b)$ or $c \in (b,a)$ according as $a < b$ or $b < a$

Another fact that is needed is: If G is a continuous function with bounded variation: then there exists a function $H(t)$ such that for any partition for $0 = t_0 < t_1 < \ldots < t_m = T$ of $[0, T]$, $m \geq 1$

$$\left(\sum_{i=0}^{m} |G(t_{i+1}) - G(t_i)|\right) \leq H(T) < \infty,$$

then with $t_m^n = \frac{m}{n}$

$$\left(\sum_m G(t_m^n \wedge t) - G(t_{m-1}^n \wedge t)\right)^2 \leq H(t) \sup_m \left(|G(t_m^n \wedge t) - G(t_{m-1}^n \wedge t)|\right)$$

$$\to 0 \text{ as } n \to \infty$$

as G being continuous on $[0, t]$ is uniformly continuous on $[0, t]$ and hence $\sup_m \left(|G(t_m^n \wedge t) - G(t_{m-1}^n \wedge t)|\right)$ converges to 0.

Observe that for every n with $t_m^n = \frac{m}{n}$ one has

$$f(G(t)) - f(G(0)) = \sum_{m=1}^{\infty} \big(f(G(t_m^n \wedge t)) - f(G(t_{m-1}^n \wedge t))\big)$$

$$= \sum_{m=1}^{\infty} f'(G(t_{m-1}^n \wedge t)) \big(G(t_m^n \wedge t) - G(t_{m-1}^n \wedge t)\big)$$

$$+ \sum_{m=1}^{\infty} f''(\theta_{n,m}) \big(G(t_m^n \wedge t) - G(t_{m-1}^n \wedge t)\big)^2$$

Here $\theta_{n,m}$ is some point in the interval $\{G(t_{m-1}^n \wedge t), G(t_m^n \wedge t)\}$. Since G is a continuous function with bounded variation, the second term goes to zero and first term goes to

$$\int_0^t f'(G(s))dG(s).$$

Ito's formula

Let X be a continuous semimartingle and let f be a thrice continuously differentiable function on $\mathbb{R}$ with f''' bounded. Then

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d[X,X]_s$$

This time we need Taylor's expansion with two terms:

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2}f''(a)(b-a)^2 + \frac{1}{6}f'''(c)(b-a)^3$$

where c belongs to the interval $\{a,b\}$, i.e. $c \in (a,b)$ or $c \in (b,a)$ according as $a < b$ or $b < a$.

We will also need that since X has finite quadratic variation, with $t^n_m = \frac{m}{n}$ one has

$$\sum_m \left(X_{t^n_m \wedge t} - X_{t^n_{m-1} \wedge t}\right)^3 \leq \sum_m \left(X_{t^n_m \wedge t} - X_{t^n_{m-1} \wedge t}\right)^2 \sup_m \left(|X_{t^n_m \wedge t} - X_{t^n_{m-1} \wedge t}|\right)$$
$$\leq [X, X]_t \sup_m \left(|X_{t^n_m \wedge t} - X_{t^n_{m-1} \wedge t}|\right)$$
$$\to 0 \text{ as } n \to \infty$$

as uniformly continuity of paths $s \mapsto X_s(\omega)$ of X on $[0, t]$ implies $\sup_m \left(|X_{t^n_m \wedge t}(\omega) - X_{t^n_{m-1} \wedge t}(\omega)|\right)$ converges to 0 for every ω.

Recall: we have seen that for every continuous adapted process Y

$$\lim_{n\to\infty}\sum_{m=0}^{\infty} Y_{t^n_m\wedge t}\big(X_{t^n_{m+1}\wedge t}-X_{t^n_m\wedge t}\big)=\int_0^t Y_u^- dX_u \tag{11}$$

Also that,

$$\sum_{m=0}^{\infty}\big(X_{t^n_{m+1}\wedge t}-X_{t^n_m\wedge t}\big)^2\to[X,X]_t. \tag{12}$$

On similar lines it can be shown that

$$\lim_{n\to\infty}\sum_{m=0}^{\infty} Y_{t^n_m\wedge t}\big(X_{t^n_{m+1}\wedge t}-X_{t^n_m\wedge t}\big)^2=\int_0^t Y_u^- d[X,X]_u. \tag{13}$$

The relation (**??**) is a weighted version of (**??**).

As in deterministic case, let us write

$$f(X_t) - f(X_0) = \sum_m f'(X_{t_m^n \wedge t})\big(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\big) + \frac{1}{2}\sum_m f''(X_{t_m^n \wedge t})\big(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\big)^2 + \frac{1}{6} f'''(\theta_{n,m})\big(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\big)^3$$

As noted earlier, the first term converges to $\int_0^t f'(X_s)dX_s$ and the second term to $\frac{1}{2}\int_0^t f''(X_s)d[X,X]_s$. Using boundedness of f''' and the fact (proven earlier) that $\sum\big(X_{t_{m+1}^n \wedge t} - X_{t_m^n \wedge t}\big)^3$ goes to zero, it follows that third term goes to zero. This completes the proof.

The Ito formula is true even when f is just twice continuously differentiable function. For the proof, one just uses a version of Taylor's theorem with a different remainder form:

$$f(b)=f(a)+f'(a)(b-a)+\frac{1}{2}f''(a)(b-a)^2+\frac{1}{2}(f''(c)-f''(a))(b-a)^2$$

where c belongs to the interval $\{a,b\}$, i.e. $c\in(a,b)$ or $c\in(b,a)$ according as $a<b$ or $b<a$.

Ito's formula for Brownian motion

Let (β_t) be a standard Brownian motion. Then we have noted that $[\beta, \beta]_t = t$ and hence for a twice continuously differentiable function f, one has

$$f(\beta_t) = f(\beta_0) + \int_0^t f'(\beta_s) d\beta_s + \frac{1}{2} \int_0^t f''(\beta_s) ds$$

Taking $f(x) = \exp(x)$ we get, writing $Y_t = \exp(\beta_t)$

$$Y_t = 1 + \int_0^t Y_s dY_s + \frac{1}{2} \int_0^t Y_s ds.$$

The last term is the Ito correction term.

The same proof as outlined earlier, yields for a continuous semimartingale X the second version of Ito formula. Here one uses, for an r.c.l.l. process Z, and semimartingales X, Y

$$\int_0^t Z_u^- d[X,Y]_u = \lim_{n\to\infty} \sum_{m=0}^{\infty} Z_{t_m^n\wedge t}\big(X_{t_{m+1}^n\wedge t} - X_{t_m^n\wedge t}\big)\big(Y_{t_{m+1}^n\wedge t} - Y_{t_m^n\wedge t}\big)$$

Ito's formula-2

Let X be a continuous semimartingle and let f function on $[0,\infty)\times\mathbb{R}$ such that $f_t=\frac{\partial}{\partial t}f$, $f_x=\frac{\partial}{\partial x}f$ and $f_{xx}=\frac{\partial^2}{\partial^2 x}f$ exist and are continuous. Then

$$\begin{aligned} f(t,X_t)=&f(0,X_0)+\int_0^t f_t(s,X_s)ds+\int_0^t f_x(s,X_s)dX_s \\ &+\frac{1}{2}\int_0^t f_{xx}(s,X_s)d[X,X]_s \end{aligned}$$

Let (β_t) be a Brownian motion. Let $f(t,x) = \exp(\sigma x - \frac{1}{2}\sigma^2 t)$. Let $Y_t = f(t,\beta_t)$. Then noting that $f_t = \frac{\partial}{\partial t} f = -\frac{1}{2}\sigma^2 f$ and $f_{xx} = \frac{\partial^2}{\partial^2 x} f = \sigma^2 f$, one has

$$Y_t = Y_0 + \sigma \int_0^t Y_s d\beta_s.$$

In particular, (Y_t) is a martingale.

If (X_t) takes values in an open convex set U and f is a twice continuously differentiable function, then the same version of Ito formula is valid.
In particular, if (X_t) is a $(0,\infty)$ valued continuous semimartingale, then taking $f(x) = \log(x)$ we get

$$\log(X_t) = \log(X_0) + \int_0^t X_s^{-1} dX_s - \frac{1}{2}\int_0^t X_s^{-2} d[X,X]_s.$$

Once again the last term is the Ito correction term.

Let us write $Y_t = \int_0^t X_s^{-1} dX_s$. Then

$$[Y,Y]_t = \int_0^t X_s^{-2} d[X,X]_s.$$

Hence, recalling,

$$\log(X_t) = \log(X_0) + \int_0^t X_s^{-1} dX_s - \frac{1}{2}\int_0^t X_s^{-2} d[X,X]_s$$

we get

$$\log(X_t) = \log(X_0) + Y_t - \frac{1}{2}[Y,Y]_t$$

and hence

$$X_t = X_0 \exp\big(Y_t - \frac{1}{2}[Y,Y]_t\big)$$

Further, if X is a $(0, \infty)$ valued continuous local martingale, then

$$M_t = \int_0^t X_s^{-1} dX_s$$

is also a local martingale and then X admits a representation

$$X_t = X_0 \exp\big(M_t - \frac{1}{2}[M, M]_t\big)$$

So every positive continuous local martingale is of the form given above for a local martingale M.

Ito's formula-3

Let X be a continuous semimartingle, A be a continuous process with bounded variation and let $f = f(a,x)$ function on $\mathbb{R} \times \mathbb{R}$ such that $f_a = \frac{\partial}{\partial a} f$, $f_x = \frac{\partial}{\partial x} f$ and $f_{xx} = \frac{\partial^2}{\partial^2 x} f$ exist and are continuous. Then

$$\begin{aligned} f(A_t, X_t) = & f(A_0, X_0) + \int_0^t f_a(A_s, X_s) dAs + \int_0^t f_x(A_s, X_s) dX_s \\ & + \frac{1}{2} \int_0^t f_{xx}(A_s, X_s) d[X, X]_s \end{aligned}$$

In typical applications of Ito's formula, one needs to show a certain process is a martingale. First one shows that it is a local martingale by expressing it as an integral w.r.t. a locally square integrable martingale and then uses other methods to show it is a martingale.

Let (M_t) be a continuous locally square integrable martingale and let (h_s) be a locally bounded predictable process.

Question: for what values of λ is Z a martingale where

$$Z_t = \exp\left\{\sigma \int_0^t h_s dM_s + \lambda \int_0^t h_s^2 d\langle M, M\rangle_s\right\}.$$

Let

$$Y_t = \int_0^t h_s dM_s \text{ and } V_t = \int_0^t h_s^2 d\langle M, M\rangle_s.$$

Note that (Y_t) is a square integrable local martingale and $V_t = [Y, Y]_t$.

Let $f(a, x) = \exp(\sigma x + \lambda a)$. Then $Z_t = f(V_t, Y_t)$. Now noting that $f_a = \frac{\partial}{\partial a} f = \lambda f$ and $f_{xx} = \frac{\partial^2}{\partial^2 x} f = \sigma^2 f$, one has

$$Z_t = Z_0 + \sigma \int_0^t Z_s dY_s + (\lambda + \frac{1}{2}\sigma^2) \int_0^t Z_s dV_s.$$

Hence if $\lambda = -\frac{1}{2}\sigma^2$, it follows that

$$Z_t = Z_0 + \sigma \int_0^t Z_s dY_s.$$

Hence for $\lambda = -\frac{1}{2}\sigma^2$, Z is a local martingale. It can be shown that this is a necessary and sufficient condition.

Of course, we will need some other way to prove that the local martingale is a martingale.

Levy's characterization of Brownian motion.

Suppose M is a continuous local martingale with $M_0 = 0$ and such that $Z_t = (M_t)^2 - t$ is also a local martingale, *or* $[M, M]_t = t$, then M is a brownian motion.

Apply Ito formula to (for a fixed real number λ)

$$f(t, x) = \exp\left\{ i\lambda x + \frac{\lambda^2 t}{2} \right\}.$$

We note that $\frac{1}{2} f_{xx} = -f_t$ and hence $Z_t = \exp\left\{ i\lambda M_t + \frac{\lambda^2 t}{2} \right\}$ is a local martingale. Since Z is bounded, it is a martingale.

As a result, $\mathbb{E}[Z_t \mid \mathscr{F}_s] = Z_s$ *i.e.*

$$\mathbb{E}\Big[\exp\big\{i\lambda M_t + \frac{\lambda^2 t}{2}\big\} \mid \mathscr{F}_s\Big] = \exp\big\{i\lambda M_s + \frac{\lambda^2 s^2}{2}\big\}.$$

Hence

$$\mathbb{E}\Big[\exp\big\{i\lambda(M_t - M_s) + \frac{\lambda^2(t-s)}{2}\big\} \mid \mathscr{F}_s\Big] = 1.$$

or

$$\mathbb{E}\Big[\exp\big\{i\lambda(M_t - M_s)\big\} \mid \mathscr{F}_s\Big] = \exp\big\{-\frac{\lambda^2(t-s)}{2}\big\}.$$

Thus

$$\mathbb{E}\Big[\exp\big\{i\lambda(M_t - M_s)\big\}\Big] = \exp\big\{-\frac{\lambda^2(t-s)}{2}\big\}$$

and so $(M_t - M_s)$ has normal distribution, mean 0 Variance $t-s$.

Further,

$$\mathbb{E}\Big[\exp\big\{i\lambda(M_t - M_s)\big\} \mid \mathscr{F}_s\Big] = \mathbb{E}\Big[\exp\big\{i\lambda(M_t - M_s)\big\}\Big]$$

implies that $(M_t - M_s)$ is independent of any set of $\mathscr{F}_s$ measurable random variables. In particular, $(M_t - M_s)$ is independent of $M_{u_1}, M_{u_2}, \ldots, M_{u_k}$. Thus M_t is a Brownian motion.

Sometimes, given a process (Z_t) one needs to find a process (A_t) with bounded variation paths such that

$$M_t = Z_t - A_t$$

is a local martingale or a martingale (U_t) such that

$$N_t = Z_t U_t$$

is a local martingale.
Example: Let (β_t) be the standard Brownian motion and for $\sigma > 0$ and $\mu \in \mathbb{R}$ let

$$S_t = S_0 \exp\{\sigma\beta_t + \mu t\}.$$

We need to find a local martingale (U_t) such that

$$Z_t = S_t U_t$$

is a martingale. Let us try $U_t = \exp\{\lambda\beta_t - \frac{1}{2}\lambda^2 t\}$. We know this is a martingale. Now

$$Z_t = S_0 \exp\left\{(\sigma+\lambda)\beta_t - \frac{1}{2}(\lambda^2 - 2\mu)\right\}$$

Thus (Z_t) would be a martingale if $(\sigma+\lambda)^2 = (\lambda^2 - 2\mu)$. Thus $\sigma^2 + 2\sigma\lambda = -2\mu$ and hence

$$\lambda = -\frac{\mu + \frac{1}{2}\sigma^2}{\sigma}$$

.

In this case, (U_t) and (Z_t) turn out to be martingales.

Ito's formula-4

Let $X^1, X^2, ..., X^d$ be a continuous semimartingles and let f function on $[0,\infty) \times \mathbb{R}^d$ such that $f_t = \frac{\partial}{\partial t} f$, $f_i = \frac{\partial}{\partial x_i} f$ and $f_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} f$ exist and are continuous. Let $X_t = (X_t^1, \ldots, X_t^d)$. Then

$$f(t,X_t) = f(0,X_0) + \int_0^t f_t(s,X_s)ds + \sum_{i=1}^d \int_0^t f_i(s,X_s)dX_s^i + \frac{1}{2}\sum_{i,j=1}^d \int_0^t f_{ij}(s,X_s)d[X^i,X^j]_s$$

Often one needs to compute the part with bounded variation paths separately. In light of this, we can recast the formula showing the (local) martingale part and the part with bounded variation paths separately.

Let $X^1, \ldots, X^d$ be continuous semimartingales with $X^i = M^i + A^i$ where M^i are continuous local martingales and A^i be processes with bounded variation paths.

Let $f \in C^{1,2}([0, \infty) \times \mathbb{R}^d_b)$. Then

$$f(t, X_t) = N_t + B_t$$

where

$$N_t = \sum_{j=1}^{d} \int_0^t f_j(s, X_s) dM_s^j$$

and

$$B_t = f(0, X_0) + \int_0^t f_t(s, X_s) ds + \textstyle\sum_{j=1}^{d} \int_0^t f_j(s, X_s) dA_s^j$$
$$+ \frac{1}{2} \textstyle\sum_{j=1}^{d} \sum_{k=1}^{d} \int_0^t f_{jk}(s, X_s) d\langle M^j, M^k \rangle_s$$

Ito's formula for r.c.l.l. semimartingales.

Let $X^1,\ldots,X^d$ be semimartingales, $X_t := (X_t^1,\ldots,X_t^d)$. Let $f \in C^{1,2}([0,\infty)\times\mathbb{R}^d)$.

$$
\begin{aligned}
f(t,X_t) =& f(0,X_0) + \int_0^t f_t(s,X_s)ds \\
& + \textstyle\sum_{j=1}^d \int_0^t f_j(s,X_{s-})dX_s^j \\
& + \frac{1}{2}\textstyle\sum_{j=1}^d \sum_{k=1}^d \int_0^t f_{jk}(s,X_{s-})d[X^j,X^k] \\
& + \textstyle\sum_{0<s\le t}\Big\{ f(s,X_s) - f(s,X_{s-}) - \sum_{j=1}^d f_j(s,X_{s-})\Delta X_s^j \\
& \qquad - \textstyle\sum_{j=1}^d \sum_{k=1}^d \frac{1}{2} f_{jk}(s,X_{s-})(\Delta X_s^j)(\Delta X_s^k)\Big\}.
\end{aligned}
$$

Martingale representation theorem

Let (β_t) be a standard Brownian motion. Let $\mathscr{F}_t = \sigma\{\beta_u : u \leq t\}$ Let (M_t) be a martingale w.r.t. $(\mathscr{F}_t)$. Then M admits a representation

$$M_t = M_0 + \int_0^t h_s d\beta_s, \; 0 \leq t \leq T$$

for some predictable process (h_s) such that

$$\int_0^T h_s^2 ds < \infty \; a.s.$$

Martingale representation theorem- multidimensional

Let $(\beta_t^1,\ldots,\beta_t^d)$ be d- independent standard Brownian motions.
Let $\mathscr{F}_t = \sigma\{\beta_u^j :\ u \leq t, 1 \leq j \leq d\}$
Let (M_t) be a martingale w.r.t. $(\mathscr{F}_t)$. Then M admits a representation

$$M_t = M_0 + \sum_{j=1}^{d} \int_0^t h_s^j d\beta_s^d,\ 0 \leq t \leq T$$

for some predictable process (h_s) such that $\int_0^T (h_s^j)^2 ds < \infty$ *a.s.*, $1 \leq j \leq d$.

Suppose X is a continuous process on $(\Omega, \mathscr{F}, \mathbb{P})$ and let $\mathbb{Q}$ be an equivalent probability measure on $(\Omega, \mathscr{F})$ such that X is a martingale on $(\Omega, \mathscr{F}, \mathbb{Q})$. Equivalent means $\mathbb{P}(A) = 0$ if and only if $\mathbb{Q}(A) = 0$.

Such a $\mathbb{Q}$ is called an Equivalent Martingale Measure, written as EMM, for X and is of great importance in Mathematical finance.

Let R denote the Radon-Nikodym derivative of $\mathbb{Q}$ w.r.t. $\mathbb{P}$, em i.e.

$$\mathbb{Q}(A) = \int_A R d\mathbb{P}.$$

Equivalence of $\mathbb{P}$ and $\mathbb{Q}$ means $\mathbb{P}(R > 0) = 1$. Let $R_t = \mathbb{E}_{\mathbb{P}}[R \mathscr{F}_t]$. Then as noted earlier (R_t) is a martingale. Moreover, for $A \in \mathscr{F}_t$

$$\begin{aligned} \mathbb{Q}(A) &= \int_A R d\mathbb{P} \\ &= \int_A \mathbb{E}[R \mid \mathscr{F}_t] d\mathbb{P}. \\ &= \int_A R_t d\mathbb{P} \end{aligned}$$

Thus R_t is the R-N derivative of $\mathbb{Q}$ w.r.t. $\mathbb{P}$ on $\mathscr{F}_t$.

Now, (X_t) being a martingale on $(\Omega, \mathscr{F}, \mathbb{Q})$ means

$$\int_A X_t d\mathbb{Q} = \int_A X_s d\mathbb{Q}$$

for $s < t$ and $a \in \mathscr{F}_s$. But this is same as

$$\int_A X_t R_t d\mathbb{P} = \int_A X_s R_s d\mathbb{P}$$

since R_u is the R-N derivative of $\mathbb{Q}$ w.r.t. $\mathbb{P}$ on $\mathscr{F}_u$.

Hence, $(X_t R_t)$ is a martingale on $(\Omega, \mathscr{F}, \mathbb{P})$. Further, R_t being a positive martingale implies R_t^{-1} is a semimartingale and thus

$$X_t = X_t R_t . R_t^{-1}$$

is a semimartingale.

Suppose $X_t = N_t + B_t$ is the decomposition of X with local martingale N. We have seen that

$$R_t = R_0 \exp(M_t - \frac{1}{2}[M, M]_t)$$

where $M_t = \int_0^t R_s^{-1} dR_s$.

Hence $Z_t = (N_t + B_t)\exp(M_t - \frac{1}{2}[M,M]_t)$ is a martingale.
So we see that if an EMM exists for X, then X must be a semimartingale and further the R-N derivative for the EMM w.r.t. the model probability measure on $\mathscr{F}_t$ is

$$R_t = R_0 \exp(M_t - \frac{1}{2}[M,M]_t)$$

where M is a local martingale such that

$$Z_t = (N_t + B_t)\exp(M_t - \frac{1}{2}[M,M]_t)$$

is a martingale.

(Girsanov's Theorem) Let $\{W_t^1, W_t^2, \ldots, W_t^d : 0 \le t \le T\}$ be independent Brownian motions on $(\Omega, \mathscr{F}, \mathbb{P})$. Let $f^1, f^2, \ldots, f^d$ be predictable processes such that

$$\int_0^T \sum_{j=1}^d |f_s^j|^2 ds < \infty.$$

Let

$$L_t := \exp\Big\{ \sum_{j=1}^d \int_0^t f^j dW^j - \frac{1}{2} \sum_{j=1}^d \int_0^t |f_s^j|^2 ds \Big\}.$$

Suppose that $\mathbb{E}[L_T] = 1$. Let $\mathbb{Q}$ be the probability measure defined by $d\mathbb{Q} = L_T d\mathbb{P}$. Then the processes

$$\hat{W}_t^j := W_t^j - \int_0^t f_s^j ds, \quad j = 1, \ldots, d, \; 0 \le t \le T$$

are independent Brownian motions under $\mathbb{Q}$.

Let $W_t = (W_t^1 \cdots, W_t^d)$ be a d-dimensional Brownian motion. For a function f on $[0, T] \times \mathbb{R}^d$ for which $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ exist, by Ito's formula one has

$$f(t, W_t) - \int_0^t (\frac{\partial f}{\partial s} + \frac{1}{2} \triangle f)(s, W_s) ds$$

equals a stochastic integral w.r.t. Brownian motion and hence is a local martingale.
Here $\triangle f(x) = \sum_{j=1}^d \frac{\partial^2 f}{\partial x_j \partial x_j}$.

Thus if g satisfies

$$\frac{\partial g}{\partial t} + \frac{1}{2}\triangle g = 0,$$

taking $f(t,y) = g(s+t, x+y)$, one gets (s and x fixed and Δ is w.r.t. y)

$$M_t^{g,x,s} = g(s+t, x+W_t)$$

is a local martingale.

If we have an appropriate integrability condition on g that implies that this local martingale is indeed a martingale, it would follow that (equality of expectation at $t=0$ and $t=T-s$)

$$g(s,x)=\mathbb{E}(g(T,x+W_{T-s})),$$

giving us a representation of g in terms of its *boundary* value $g(T,\cdot)$.

Similarly, if h satisfies

$$\frac{\partial h}{\partial t} = \frac{1}{2}\triangle h,$$

taking $g(t,y) = h(T-t, x+y)$ and using the heuristics given above we could conclude (with $u = T-t$)

$$h(u,x) = \mathbb{E}(h(0, x+W_u))$$

provided we have conditions that ensure that the local martingale $h(u-t, x+W_t)$ is a martingale.

Suppose f is a continuous function on $\mathbb{R}^d$ and

$$|f(x)| \leq C_1 \exp(a_1|x|^2),\ \ x \in \mathbb{R}^d \tag{14}$$

for $C_1 < \infty$ and $a_1 < \infty$. Then the Cauchy problem (for $u \in C^{1,2}((0,T) \times \mathbb{R}^d)$) for the heat equation

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) &= (\frac{1}{2}\triangle u)(t,x) \quad : \quad (t,x) \in (0,T) \times \mathbb{R}^d \\ \lim_{t \downarrow 0} u(t,x) &= f(x) \end{aligned}$$

admits a unique solution in the class of functions g satisfying

$$|u(t,x)| \leq C_2 \exp(a_2|x|^2),\ \ (t,x) \in (0,T) \times \mathbb{R}^d \tag{15}$$

for some $C_2 < \infty, a_2 < \infty$.

The unique solution is given by

$$u(t,x) = E(f(x + W_t)) \tag{16}$$

where $W_t = (W_t', \cdots, W_t^d)$ is a d-dimensional Brownian motion.