Financial Derivatives & Risk Management
Professor J. P. Singh
Department of Management Studies
Indian Institute of Technology, Roorkee
Lecture 41 - Stochastic Calculus: Ito's Equation

Calculus of stochastic variables

We, now, take up elementary calculus of functions of Brownian motion. The immediate question, then, is why we need a special or different calculus for functions of BM? Why is it that the conventional Newtonian calculus does not fit the bill?

Let us say, we have a simple function of Brownian motion $f(W_t)=W_t^2$ so that $df(W_t)=dW_t^2=2W_tdW_t$. Integrating conventionally, we get $W_t^2=\int_0^t d(W_s)^2=2\int_0^t W_s dW_s$ whence we must have $E(W_t^2)=2E\left(\int_0^t W_s dW_s\right)$. Let us see whether this holds in the context of BM function $f(W_t)=W_t^2$.

To evaluate the integral on the RHS, we partition the interval (0,t) into n parts (0, t/n, 2t/n,…,nt/n=t). WE, then, express the integrand as a sum of series with each term being the value at the beginning of the partition into increment over that partition as shown below:

Consider a function of Brownian motion say $f(W_t)=(W_t)^2$

By Newtonian calculus, we have $d(W_t)^2=2W_tdW_t$ *so that*

$$W_t^2=\int_0^t d(W_s)^2=2\int_0^t W_s dW_s$$

Let us calculate the expectations of both sides

For this purpose, we partition the time interval

$(0,t)$ *into a partition* $\left(0,\frac{t}{n},\frac{2t}{n},...,\frac{(n-1)t}{n},\frac{nt}{n}\right)$

$$Then, 2\int_0^t W_s dW_s=2\sum_{i=0}^{n-1} W\left(\frac{it}{n}\right)\left(W\left(\frac{(i+1)t}{n}\right)-W\left(\frac{it}{n}\right)\right)$$

Now, the difference term inside the bracket is just the increment of BM from one partition to another.

By definition, (i) it is independent of the history of BM upto that point i.e. the first term and (ii) it has zero mean. Thus, each term in the summation has zero mean and so does the summation i.e. $E\left(2\int_0^t W_s dW_s\right)=0$

$$Explicitly,\ \sum_{i=0}^{n\text{-}1} W\left(\frac{it}{n}\right)\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)$$

$$=W\left(0.\frac{t}{n}\right)\left[W\left(1.\frac{t}{n}\right)\text{-}W\left(0.\frac{t}{n}\right)\right]+W\left(1.\frac{t}{n}\right)\left[W\left(2.\frac{t}{n}\right)\text{-}W\left(1.\frac{t}{n}\right)\right]$$

$$+W\left(2.\frac{t}{n}\right)\left[W\left(3.\frac{t}{n}\right)\text{-}W\left(2.\frac{t}{n}\right)\right]+...+W\left(n-1.\frac{t}{n}\right)\left[W\left(n.\frac{t}{n}\right)\text{-}W\left(n-1.\frac{t}{n}\right)\right]$$

$$Now,\ for\ example: E\left\{W\left(2.\frac{t}{n}\right)\left[W\left(3.\frac{t}{n}\right)\text{-}W\left(2.\frac{t}{n}\right)\right]\right\}$$

$$=E\left[W\left(2.\frac{t}{n}\right)\right]E\left\{\left[W\left(3.\frac{t}{n}\right)\text{-}W\left(2.\frac{t}{n}\right)\right]\right\}=0$$

and this holds for each term in the summation.

By the properties of Brownian motion, Brownian motion is a Markov process. The evolution of a Brownian motion does not depend on its past history. Therefore, the initial term $W\left(\frac{it}{n}\right)$ and its increment $\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)$ are independent of each other. $\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)$ is the further evolution of Brownian motion from this point $W\left(\frac{it}{n}\right)$ onwards. Therefore, this evolution $\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)$is independent of what has happened earlier i.e. up to $W\left(\frac{it}{n}\right)$. Therefore, because these two are independent the expected value $E\left[W\left(\frac{it}{n}\right)\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]=E\left[W\left(\frac{it}{n}\right)\right]E\left[\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]$. But the expected value of a Brownian motion increment is 0. Therefore, $E\left[\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]=0$ and hence $E\left[W\left(\frac{it}{n}\right)\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]$ $=E\left[W\left(\frac{it}{n}\right)\right]E\left[\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]=0$. Now, $E\left[W\left(\frac{it}{n}\right)\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]$ is any general term of the sum. Hence, very term corresponding to each value of I between 0 and n-1 vanishes. It follows that the entire sum $\sum_{i=0}^{n-1} E\left[W\left(\frac{it}{n}\right)\left(W\left(\frac{(i+1)t}{n}\right)\text{-}W\left(\frac{it}{n}\right)\right)\right]=0$ and hence, the expectation value $2E\left(\int_0^t W_s dW_s\right)=0$.

Now, let us look at the LHS. We write this again as the sum of a series $W_t^2 = \sum_{i=0}^{n-1}\left[W^2\left(i+1\frac{t}{n}\right) - W^2\left(i\frac{t}{n}\right)\right]$. Consider the general term $W^2\left(i+1\frac{t}{n}\right)$. We know that W(t) is normally distributed with a mean of zero and a variance of t and can therefore be represented in terms of the standard normal variate as: W(t)= z√t where z is N(0,1) so that $W^2(t) = z^2 t$ where $E(z^2)=1$. We use this expression to write $W^2\left(i+1\frac{t}{n}\right)$ as $\left(i+1\frac{t}{n}\right)z_{i+1}^2$ and $W^2\left(i\frac{t}{n}\right)$ as $\left(i\frac{t}{n}\right)z_i^2$ where z_{i+1} and z_i are both N(0,1) with $E(z_{i+1}^2)= E(z_i^2)=1$. Further, $E\left[\left(i+1\frac{t}{n}\right)z_{i+1}^2\right] = \left(i+1\frac{t}{n}\right)E\left[z_{i+1}^2\right] = \left(i+1\frac{t}{n}\right)$ and similarly $E\left[\left(i\frac{t}{n}\right)z_{i+1}^2\right] = \left(i\frac{t}{n}\right)$ so that

$E\left(W_t^2\right) = E\left\{\sum_{i=0}^{n-1}\left[W^2\left(i+1\frac{t}{n}\right) - W^2\left(i\frac{t}{n}\right)\right]\right\} = \sum_{i=0}^{n-1}E\left[W^2\left(i+1\frac{t}{n}\right) - W^2\left(i\frac{t}{n}\right)\right]$

$= \sum_{i=0}^{n-1}E\left[\left(i+1\frac{t}{n}\right)z_{i+1}^2 - \left(i\frac{t}{n}\right)z_{i+1}^2\right] = \sum_{i=0}^{n-1}\left[\left(i+1\frac{t}{n}\right) - \left(i\frac{t}{n}\right)\right] = \sum_{i=0}^{n-1}\left(\frac{t}{n}\right) = t$. This is shown below:

$$W_t^2 = \sum_{i=0}^{n-1}\left[W^2\left(i+1\frac{t}{n}\right) - W^2\left(i\frac{t}{n}\right)\right] = \sum_{i=0}^{n-1}\left[\left(i+1\frac{t}{n}\right)z_{i+1}^2 - \left(i\frac{t}{n}\right)z_i^2\right]$$

$$\textit{so that } E\left(W_t^2\right) = E\left\{\sum_{i=0}^{n-1}\left[W^2\left(i+1\frac{t}{n}\right) - W^2\left(i\frac{t}{n}\right)\right]\right\} = \sum_{i=0}^{n-1}E\left\{\left[\left(i+1\frac{t}{n}\right)z_{i+1}^2 - \left(i\frac{t}{n}\right)z_i^2\right]\right\}$$

$$= \sum_{i=0}^{n-1}E\left[\left(i+1\frac{t}{n}\right)z_{i+1}^2\right] - \sum_{i=0}^{n-1}E\left[\left(i\frac{t}{n}\right)z_i^2\right] = \sum_{i=0}^{n-1}\left(i+1\frac{t}{n}\right) - \sum_{i=0}^{n-1}\left(i\frac{t}{n}\right) = t \ \sin ce \ E\left(z^2\right)=1$$

The fact that $E(z^2)=1$ is easily established. We have:

$$\textit{Now, if } z \xrightarrow{distribution} N(0,1), \textit{ then}$$
$$E\left(z^2\right) = \left[E(z)\right]^2 + Var(z) = 0+1 = 1$$

Thus, we conclude that $E\left(W_t^2\right) = 2E\left(\int_0^t W_s dW_s\right)$ does not hold. It is not correct for functions of Brownian motion. In other words, the Newtonian calculus is not the correct description of Brownian motion calculus. We need to revise Newtonian calculus in order to do calculus with functions of Brownian motion. The Newtonian calculus that we have does not fit in with the Brownian motion, why? For this purpose, let us work out the value of $\int_0^t (dW_s)^2$. We have,

$\int_0^t (dW_s)^2 = \sum_{i=0}^{n-1} \left[W\left(i+1.\frac{t}{n} \right) \text{-} W\left(i.\frac{t}{n} \right) \right]^2$

Now, each $\left[W\left(i+1.\frac{t}{n} \right) \text{-} W\left(i.\frac{t}{n} \right) \right]$ *is a BM increment*

and hence, is $N(0, t/n)$*.Thus,* $z_{n,i} = \dfrac{W\left(i+1.\frac{t}{n} \right) \text{-} W\left(i.\frac{t}{n} \right)}{\sqrt{t/n}}$

is $N(0,1)$ *distributed.*

Hence, $z_{n,i}^2$ *is distributed with mean* 1.

Sin*ce, the sequence* $z_{n,i}^2$ *is a sequence of*

n independent var*iables each with mean* 1,

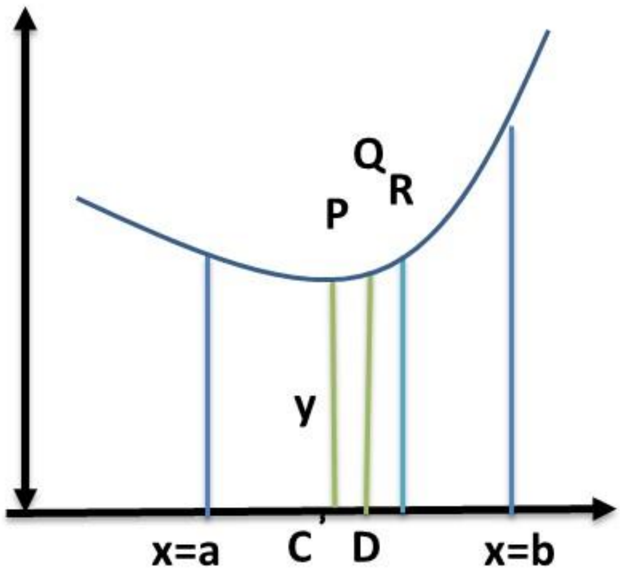
$\sum_{i=1}^{n} z_{n,i}^2$ *has mean n so that* $\int_0^t (dW_s)^2 = \frac{t}{n} \sum_{i=1}^{n} z_{n,i}^2$ *has the mean t.*

*To unders***tan***d, what happens in case of a BM increment* dW_t *instead of dx, we do a Taylor* **exp***ansion in both cases:*

For $y = f(x)$*, we have,* $dy = df(x) = f'(x)dx + \frac{1}{2} f''(x)(dx)^2 + \ldots$

Thus, when we talk about Newtonian differentiation, it is clear that we truncate the Taylor **exp***ansion at the first order term i.e. dx and we write* $dy = f'(x)dx$

In other words, we asume dx to be so small that $(dx)^2$ *and higher powers can be ignored. To justify this we add the* **lim***iting prescription*
$Lim_{dx \to 0} f(x) = f'(x)$



$$L = \int_{x=a}^{x=b} ds = \int_{x=a}^{x=b} \sqrt{(dx^2 + dy^2)} = \int_{x=a}^{x=b} dx \sqrt{\left[1 + \left(\frac{dy}{dx} \right)^2 \right]} = \mathbf{lim}_{n\to\infty} \sum_{i=1}^{n} \sqrt{\left[1 + \left(\frac{dy_i}{dx_i} \right)^2 \right]} dx_i$$

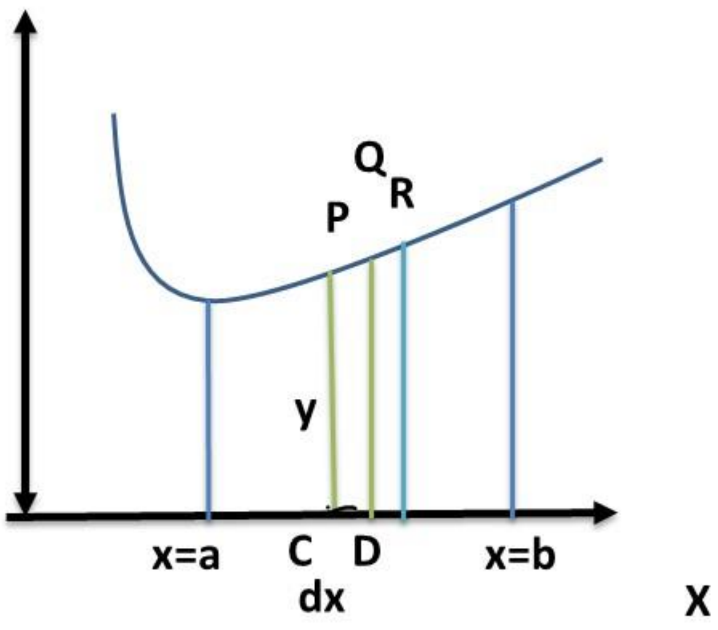
Let us further explore it geometrically. When we try to work out the length of an arc, we have a given function y=f(x) and we are required to work out the length of the arc that this function maps between the points x=a and x=b. We integrate $ds=\sqrt{(dx^2+dy^2)}$ between these limits x=a,b where ds is an infinitesimal length e.g. PQ and we integrate it to obtain the entire length of the arc.

The points P, Q on the curve y=f(x) are assumed so close to each other, that PQ approx. a straight line. This is integrated to obtain the arc length L. Thus, at the infinitesimal level, we assume that the curve is constructed by an assortment of infinitesimal straight lines.

We write this infinitesimal length PQ in terms of dx and dy as $\sqrt{(dx^2+dy^2)}$. Now, the very fact that we are writing $ds=\sqrt{(dx^2+dy^2)}$ conveys a very important geometrical meaning. The geometrical meaning is, that we are assuming that in the region CD which is of length dx, the curve is a straight line, the interval dx is so small that in this small infinitesimal interval, the curve $PQ=ds= \sqrt{(dx^2+dy^2)}$ can be approximated by a straight line and therefore we can apply the Pythagoras theorem and write ds as $\sqrt{(dx^2+dy^2)}$.

The points P and Q on the curve y=f(x) are so close to each other that PQ approximates a straight line and therefore we can write the distance $PQ = \sqrt{(dx^2+dy^2)}$ i.e. in accordance with the pythagoras theorem. The fallout is that at the infinitesimal level when we talk about Newtonian calculus, we assume that the entire curve is constituted of infinitesimal straight lines combined together in an appropriate pattern.

At the infinitesimal level, if I were to zoom in very, very high, I would see only a straight line and that is the important thing which is to which is conveyed by the Newtonian calculus. This philosophy is violated in the calculus of Brownian motion.



$$A = \int_{x=a}^{x=b} y dx = \mathbf{lim}_{n\to\infty} \sum_{i=1}^{n} y_i dx_i$$

Another example relates to the area computations. To compute areas, we integrate ydx. When we integrate ydx, it is obvious that y the length i.e. the initial value of the function is multiplied by the width, dx, of a thin strip of at that point. In other words, we assume that CPQD is a rectangle and hence, its area is given by CP x CD=ydx. Equivalently, we premise that PQ length is a straight line in this small domain dx.

Because of assuming PQ as a straight line, CPQD is nothing but a rectangle and because the strip is a rectangle the area of the strip is given by the initial value (y) and the forward increment (dx). So that is another example where conventional calculus assumes that, at the infinitesimal level, deterministic curves are constituted of assortment of small straight lines.

But that does not happen in the case of Brownian motion, let us see why.

Let us consider the case of BM. We have,

$$df\left(W_t\right)=f'\left(W_t\right)dW_t+\frac{1}{2}f''\left(W_t\right)dW_t.dW_t+\dots$$

$$dW_t.dW_t=z^2dt \text{ so } E\left(z^2\right)=1, Var\left(z^2\right)=2$$

$$E\left(dW_t.dW_t\right)=dt, Var\left(dW_t.dW_t\right)=2\left(dt\right)^2$$

so that $dW_t.dW_t$ *has value dt* **&** *is not stochastic.*

Hence, it cannot be IGNORED.

$$\text{From above}: E\left(2\int_0^t W_s dW_s\right)=0;\ E\left[\int_0^t \left(dW_s\right)^2\right]=t$$

By Taylor expansion

$$d\left(W_t^2\right)=2W_t dW_t+\left(dW_s\right)^2+\dots$$

$$t=E\left[\int_0^t d\left(W_s\right)^2\right]=E\left(2\int_0^t W_s dW_s\right)+E\left[\int_0^t \left(dW_s\right)^2\right]=0+t=t$$

$$\text{Thus, } E\left[\int_0^t d\left(W_s\right)^2\right]\neq E\left(2\int_0^t W_s dW_s\right)$$

Why? because $E\left[\int_0^t \left(dW_s\right)^2\right]=t\neq 0$. *If we include this second order term in the Taylor expansion, we get consistent results, at least, at the level of expectations:* $t=E\left[\int_0^t d\left(W_s\right)^2\right]=E\left(2\int_0^t W_s dW_s\right)+E\left[\int_0^t \left(dW_s\right)^2\right]=0+t=t$

$$\text{Compare}: d\left(x^2\right)=2xdx+\left(dx\right)^2\cong 2xdx$$

We introduced a very simple example $f(W_t)=W_t^2$ and showed that Newtonian calculus leads us to inconsistent results. The reason is that in the case of Newtonian calculus, we can safely ignore second and higher order terms of the Taylor expansion on the premise that we are taking infinitesimal limits on dx i.e. $dx\to 0$ so that $(dx)^2$ and higher orders can be ignored. This does not happen to be the case when we work with infinitesimal increments of BM because squares of BM increments i.e. $(dW_t)^2$ have a mean of dt (which is of first order in dt) and a variance of $2(dt)^2$. Thus, (i) the mean is of first order in dt and (ii) the variance is of second order in dt which means that it is much too small to be of relevance. In other words, we can assume that the mean of $(dW_t)^2$ is of first order in dt and is non-stochastic (as its variance is too small). It follows that $(dW_t)^2$ cannot be ignored in the Taylor expansion and we need to retain terms up to at least second order to get a mathematically consistent calculus.

This is what we did in the in our example when we calculated the Newtonian derivative of $f(x)=x^2$, $f'(x)$ was 2x, so 2xdx was what we got of $d(x^2)$ and retained only the first order term in dx where dx small enough that higher orders of dx could be safely ignored. But when we talk about Brownian motion, the expression $dW_t\, dW_t$ contributes to the second order term. Now, $dW_t=z\sqrt{dt}$ where z is standard normal variable, so $dW_t\, dW_t=z^2dt$. Now, z^2 has a mean value of 1 and a variance of 2 whence $dW_t\, dW_t=z^2dt$ has a mean of dt and variance of $2dt^2$. So the first thing we get is that $dW_t\, dW_t$ is at least in the mean, at least in expectation is a first order term in dt and because it is the first order term in dt, it cannot prime facie be ignored. Thus, unlike $(dx)^2$, we cannot simply throw away $dW_t\, dW_t$.

Let us look at the variance, the variance of $dW_t\, dW_t$ is $2(dt)^2$. Now, the variance of a stochastic process is of the order of dt i.e. is in first order in dt, Brownian motion is a typical example. Brownian motion has a variance which is of order dt, scaled Brownian motion also has a variance of order dt. However, in this case, we are finding that the variance is second order in dt. If dt is small that means the variance is very, very small and if the variance is very, very small we can approximate it by a non-stochastic variable because small or negligible variance means that fluctuations are negligible or that the underlying variable's evolution is deterministic. Thus, $dW_t\, dW_t$ is deterministic and of order dt.

We find that $dW_t\, dW_t$ not only has the mean dt but has also got a very, very small variance and therefore is almost deterministic i.e. does not have fluctuations. Thus, if it has expectation of order dt, then it is throughout of order dt and therefore obviously it cannot be ignored.

Thus, the second order term $dW_t\, dW_t$ which we ignore in conventional calculus, has a finite deterministic value at least at the first order in the case of stochastic calculus and because it is of the first order in dt we need to take it into account in working out the differentials.

Mean & variance of z^2

We have used two results in the above section viz. $E(z^2)=1$ and $Var(z^2)=2$. The former has already been established. We prove the latter here:

$$E\left(z^4\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^4e^{-\frac{1}{2}z^2}dz=2\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}z^4e^{-\frac{1}{2}z^2}dz\quad\left(Even\ \ function\right)$$
$$Let\ \frac{1}{2}z^2=t\ \ so\ that\ \ zdz=dt$$
$$E\left(z^4\right)=2\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\left(2t\right)^{3/2}e^{-t}dt=\frac{4}{\sqrt{\pi}}\int_{0}^{\infty}t^{3/2}e^{-t}dt=\frac{4}{\sqrt{\pi}}\Gamma\left(5/2\right)$$
$$Now,\ \Gamma\left(n+1\right)=n\Gamma\left(n\right);\qquad\Gamma\left(1/2\right)=\sqrt{\pi}\ \ so\ that$$
$$E\left(z^4\right)=\frac{4}{\sqrt{\pi}}\frac{3}{2}\Gamma\left(3/2\right)=\frac{4}{\sqrt{\pi}}\frac{3}{2}\frac{1}{2}\Gamma\left(1/2\right)=\frac{4}{\sqrt{\pi}}\frac{3}{2}\frac{1}{2}\sqrt{\pi}=3;$$
$$E\left(z^2\right)=1;\ Var\left(z^2\right)=E\left(z^4\right)-\left[E\left(z^2\right)\right]^2=3-1=2$$

In the above we have used the fact that $\Gamma(1/2)=\sqrt{\pi}$ which is proved below:

$$\Gamma(1/2)=\int_0^\infty t^{-1/2}e^{-t}dt\ \ Let\ t=u^2,\ dt=2udu$$
$$\Gamma(1/2)=\int_0^\infty t^{-1/2}e^{-t}dt=2\int_0^\infty e^{-u^2}du=\int_{-\infty}^\infty e^{-u^2}du$$
$$Let\ u=\frac{v}{\sqrt{2}},\ \ du=\frac{dv}{\sqrt{2}}\ \ so\ that$$
$$\Gamma(1/2)=\int_{-\infty}^\infty e^{-u^2}du=\frac{1}{\sqrt{2}}\int_{-\infty}^\infty e^{-\frac{1}{2}v^2}dv=\frac{1}{\sqrt{2}}\sqrt{2\pi}=\sqrt{\pi}$$
$$\textbf{sin}\,ce\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{1}{2}v^2}dv=1\ (CDF\ of\ normal\ distribution)$$

Ito calculus

So far the problems and the issues arising from the calculus of functions of Brownian motion have been discussed. In other words, we have discussed how the Newtonian framework fails when we talk about calculus of functions of stochastic variables. Now, let us see how do we remedy the situation.

The solution is provided by a very celebrated Lemma which carries the name of the person who has propounded it, the Ito's Lemma. It states that:

Ito's Lemma: Let G(x,t) be continuous & at least twice differentiable function of a stochastic variable x and time t and let x be defined as the stochastic (Ito) process: dx = a(x,t)dt + b(x,t)dW Then, we have:

$$dG=\left(a.\frac{\partial G}{\partial x}+\frac{\partial G}{\partial t}+\frac{1}{2}b^2.\frac{\partial^2 G}{\partial x^2}\right)dt+b.\frac{\partial G}{\partial x}dW_t$$

A skeleton proof of the lemma runs on the following lines:

$$Consider\ a\ continuous\ and\ differentiable\ function\ G(x,t)\ of\ a$$
$$stochastic\ \text{var}\,iable\ x\ and\ t\ where\ x\ satisfies$$
$$dx=a(x,t)dt+b(x,t)dW.\ Taylor\ \exp ansion\ of\ G(x,t)\ is$$
$$dG=\frac{\partial G}{\partial x}dx+\frac{\partial G}{\partial t}dt+\frac{1}{2}\frac{\partial^2 G}{\partial x^2}dx^2+\frac{\partial^2 G}{\partial x\partial t}dxdt+\frac{1}{2}\frac{\partial^2 G}{\partial t^2}dt^2+...$$
$$=\frac{\partial G}{\partial x}\left[a(x,t)dt+b(x,t)dW\right]+\frac{\partial G}{\partial t}dt+\frac{1}{2}\frac{\partial^2 G}{\partial x^2}\begin{bmatrix}a(x,t)dt+\\ b(x,t)dW\end{bmatrix}^2$$
$$+\frac{\partial^2 G}{\partial x\partial t}\left[a(x,t)dt+b(x,t)dW\right]dt+\frac{1}{2}\frac{\partial^2 G}{\partial t^2}dt^2+...$$

$$We\ have\ dx^2 = \left[a(x,t)dt + b(x,t)dW\right]^2$$
$$= \left[a(x,t)dt + b(x,t)z\sqrt{dt}\right]^2 = b^2z^2dt + h.o\ terms\ in\ dt$$
$$Now,\ E(z^2) = 1\ whence\ E(b^2z^2dt) = b^2dt.$$
$$Also\ Var\ (z^2) = 2\ whence\ Var(b^2z^2dt) = 2b^4dt^2 = 0$$
if higher order terms than dt are neglected.
Now the Variance of a stochastic variable in time dt is proportional to dt not dt^2.
Since $Var(b^2z^2dt) = 2b^4dt^2$ is proportional to dt^2 it is too small to have a stochastic component.
Hence, in the limit that higher order terms in dt are neglected, (b^2z^2dt) and hence, dx^2 may be considered non-stochastic with a value of b^2dt.
Thus, we get the Ito Lemma as:
$$dG = \left(a.\frac{\partial G}{\partial x} + \frac{\partial G}{\partial t} + \frac{1}{2}b^2.\frac{\partial^2 G}{\partial x^2}\right)dt + b.\frac{\partial G}{\partial x}dW$$

The important thing in the above derivation is again, as alluded to earlier, that dW_tdW_t becomes deterministic in t. In fact, it has an expected value of dt and a negligible variance. Hence, it may be approximated by a deterministic dt with no stochastic or random component. Thus, dW_tdW_t cannot be ignored in the Taylor series expansion.

Hence, we cannot altogether ignore the term $(dx)^2$. When we do the expansion of $(dx)^2$, we will have to retain the term involving dW_tdW_t which is equal to dt. Hence, we write:

$$(dx)^2 \cong b^2dW_tdW_t = b^2dt$$

Drift rate and variance rate

From Ito's Lemma, function $G(x,t)$ of a stochastic variable x, satisfies the Ito equation
$$dG = \left(a.\frac{\partial G}{\partial x} + \frac{\partial G}{\partial t} + \frac{1}{2}b^2.\frac{\partial^2 G}{\partial x^2}\right)dt + b.\frac{\partial G}{\partial x}dW$$
From this eq. we see that in an infinitesimal time interval dt, the process G is stochastic with a drift rate $\left(a.\frac{\partial G}{\partial x} + \frac{\partial G}{\partial t} + \frac{1}{2}b^2.\frac{\partial^2 G}{\partial x^2}\right)$, variance rate $\left(b.\frac{\partial G}{\partial x}\right)^2$

$$\textit{Let } F(x,t) \textit{ be continuous } \& \textit{ at least twice diff.}$$
$$\textit{Then, by Taylor } \mathbf{exp}\textit{ansion:}$$
$$dF(W_t,t) = F'(W_t,t)dW_t + \frac{1}{2}F''(W_t,t)dW_t.dW_t + \frac{\partial F(W_t,t)}{\partial t}dt + ...$$
$$= F'(W_t,t)dW_t + \left[\frac{1}{2}F''(W_t,t) + \frac{\partial F(W_t,t)}{\partial t}\right]dt$$
$$\textit{For non}-\mathbf{exp}\textit{licit time dependent } F(W_t):$$
$$\int_{t=t_1}^{t=t_2} F'(W_t)dW_t = F(W_{t_2}) - F(W_{t_i}) - \frac{1}{2}\int_{t=t_1}^{t=t_2} F''(W_t)dt$$

This is a simpler form of Ito’s lemma. In the earlier case, we had function of a stochastic variable x that had both a drift and a diffusion term i.e. dx=adt+bdW and time t. Now, we consider a function of Brownian motion dW and time alone. Thus, there is no drift term and no scaling of BM. Hence, we consider $F(W_t,t)$ instead of G(x,t).

So, $\frac{1}{2}F''(W_t,t)dt$ is the piece that is additional when we talk about stochastic calculus. And now, let us use this expression to evaluate $F(W_t,t)=W_t^2$. We have:

$$\textit{For non}-\mathbf{exp}\textit{licit time dependent } F(W_t):$$
$$\int_{t=t_1}^{t=t_2} F'(W_t)dW_t = F(W_{t_2}) - F(W_{t_i}) - \frac{1}{2}\int_{t=t_1}^{t=t_2} F''(W_t)dt$$
$$\textit{Hence, } u\mathbf{sin}g \;\; F'(W_t) = W_t, \textit{ we get:}$$
$$\int_{t=t_1}^{t=t_2} W_t dW_t = \frac{1}{2}(W_{t_2})^2 - \frac{1}{2}(W_{t_i})^2 - \frac{1}{2}(t_2 - t_1)$$
$$= \frac{1}{2}(t_2 - t_1) - \frac{1}{2}(t_2 - t_1) = 0$$

showing the Ito framework to be consistent.

The result that we you arrived at using summation by series is consistent with the result that we arrived through Ito's lemma.

Example

Let $G(S,t) = \ln S$ where $dS = \mu S dt + \sigma S dW$. Using Ito’s Lemma find the drift and diffusion terms and the distribution of G(S,t).

Solution

Ito’s Lemma states that if G(x,t) is a twice differentiable function of x where x is given by dx=adt+bdW, then

$$dG = \left(a.\frac{\partial G}{\partial x} + \frac{\partial G}{\partial t} + \frac{1}{2}b^2.\frac{\partial^2 G}{\partial x^2}\right)dt + b.\frac{\partial G}{\partial x}dW_t$$

In our problem $G(S,t) = \ln S$, $\frac{\partial G}{\partial S} = \frac{\partial \ln S}{\partial S} = \frac{1}{S}$, $\frac{\partial^2 G}{\partial S^2} = \frac{\partial}{\partial S}\left(\frac{1}{S}\right) = -\frac{1}{S^2}$, $\frac{\partial G}{\partial t} = \frac{\partial \ln S}{\partial t} = 0$, $a = \mu S$, $b = \sigma S$. Putting the values, we get:

$$dG = d\left(\ln S\right) = \left(\mu S\frac{1}{S} - \frac{1}{2}\sigma^2 S^2\frac{1}{S^2}\right)dt + \sigma S\frac{1}{S}dW_t = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t$$

This shows that the drift term of $G\left(S,t\right) = \ln S$ is $\left(\mu - \frac{1}{2}\sigma^2\right)$ and the diffusion term is σ.

Further, $d\left(\ln S\right)$ is normally distribution with a mean of $\left(\mu - \frac{1}{2}\sigma^2\right)dt$ and a variance of $\sigma^2 dt$

.

Equivalently $\ln S_T$ is $N\left[\ln S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)T, \sigma^2 T\right]$.