BROWNIAN MOTION AND ITO CALCULUS

K. Ito's **stochastic calculus** is a collection of tools which permit us to perform operations such as composition, integration and differentiation, on functions of Brownian paths and more general random functions known as *Ito processes.* As we shall see, Ito calculus and Ito processes are extremely useful in the formulation of financial risk-management techniques. These notes are intended to introduce the reader to stochastic calculus in a straightforward, intuitive way. For rigorous treatments of this rich subject the reader can consult, for instance, Ikeda and Watanabe (North Holland-Kodansha, 1989), Varadhan (1980) Karatzas and Shreve (Springer 1988).[1]

1. Brownian Motion

Intuitively, Brownian motion corresponds to the concept of a *homogeneous, continuous-time, continuous random walk.* One way to visualize Brownian paths is to consider a simple random walk on the real line, in which the walker starts at position $X_0 = 0$ and moves up or down by an amount $\sqrt{dt}$ after each time-interval of duration dt. If X_n denotes the position of the walker after the nth jump, we have

$$X_n = X_{n-1} \pm \sqrt{dt} \, , \quad n = 1, 2, \ldots \tag{1}$$

where the $+$ and $-$ signs occur with probability $1/2$. This process is called a simple random walk. Note that the magnitude of the jump and the lag between successive jumps are chosen so that the variance of the displacement of the walker after time T (where T is an integer multiple of dt) is exactly T.

[1] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd Ed., North Holland-Kodansha, Amsterdam and Tokyo, 1989; S.R.S. Varadhan *Lectures on Brownian Motion and Stochastic Differential Equations*, Tata Institute of Fundamental Research, Bombay; and I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.

A continuous path can be built from the variables X_n by interpolating linearly between the different points:

$$\overline{X}(t) = X_n + (t - n\,dt) \cdot (X_{n+1} - X_n) , \quad \text{for } n\,dt \le t \le (n+1)\,dt . \qquad (2)$$

These paths have the following properties:

1. If $t = n\,dt$ and $a > 0$, the increment $\overline{X}(t+a) - \overline{X}(t)$ is independent of the "past" $\{\overline{X}(s) , s \le t \}$;
2. $\mathbf{E}\left(\overline{X}(t)\right) = 0$;
3. $\mathbf{E}\left(\overline{X}(t)^2\right) = t$.

For small time-increments ($dt \ll 1$), these paths have nearly *independent increments* (neglecting the small persistence effect due to the linear interpolation). Moreover, the mean and variance of the walker's displacement are independent of dt.

For $dt \ll 1$, each increment of $\overline{X}(t)$ is a sum of many independent binomial random variables with mean zero and finite variance. Therefore, by the Central Limit Theorem, the limiting probability distribution of the increments of $\overline{X}(t)$ as $dt \to 0$ is Gaussian (or normal). More precisely, we have

$$\lim_{dt \to 0} \mathbf{P}\left\{\overline{X}(t+a) - \overline{X}(t) \ge x \right\} = \frac{1}{\sqrt{2\pi\, a}} \int_x^{+\infty} e^{-\frac{y^2}{2a}}\, dy$$

for all x. This property, together with the independence of the increments, characterizes the statistics of the paths $\overline{X}(t)$ in the limit.

This discussion motivates

Definition 1: *Brownian Motion is a probability distribution on the set of real-valued functions $Z(t)$,$0 \le t < \infty$ with the following properties*

1. $Z(0) = 0$ *with probability 1* ;
2. *For all $t > 0$ and $a > 0$, the increments $Z(t+a) - Z(t)$ are Gaussian with mean zero and variance a; and*
3. $Z(t+a) - Z(t)$ *is independent of* $\{ Z(s) , 0 \le s \le t \}$.

Items 1 – 3 completely specify the probability distribution of any n-tuple

$$(Z(t_1),\ Z(t_2),\ Z(t_3),\ \dots\ Z(t_n))$$

where $t_1 < t_2 < \ldots < t_n$ are arbitrary times. It is easy to see that this distribution is a multivariate Gaussian[2] with covariance

$$\mathbf{E}\,\{\, Z(t_i)\, Z(t_j)\,\} \;=\; \mathrm{Min}\,(t_i\,,\; t_j\,)\;.$$

2. Elementary properties of Brownian paths

The first important property is the **continuity of Brownian paths**, namely

$$\mathbf{P}\;\{Z(\cdot)\text{ is a continuous function }\} \;=\; 1\;.$$

The proof of this fact is mathematically non-trivial (see Ikeda and Watanabe or Varadhan).

Brownian paths are continuous but also very *irregular*. Any student of random walks (or, for the matter, of financial time series) has noticed that the sample paths of random walks are non-smooth and appear to have an infinite slope (making trends difficult to predict.) One way to see that Brownian paths are not differentiable is to consider their **quadratic variation**.

Proposition 1. *Let* $0 = t_0 < t_1 < \ldots < t_n = T$ *represent a partition of the time interval* $[0,T]$ *and let* $dt = \max_j\,(t_j - t_{j-1}\,)$*. Set*

$$\Delta\, Z_j \;=\; Z(t_j) \;-\; Z(t_{j-1})\;.$$

With probability one, Brownian paths satisfy

$$\lim_{dt\to 0}\;\left\{\;\sum_{j=1}^{N}\,(\Delta\, Z_j\,)^2\;\right\} \;=\; T\;.$$

(The right-hand side of this last equation is known as the quadratic variation of the path $Z(\cdot)$ on the interval $[0,T]$.) This result is is a direct consequence of the Law of Large Numbers of probability theory. In fact, the random variables $(\Delta\, Z_j\,)^2$ are independent and

[2] Notice that the statistical distribution of the paths for finite dt is more complicated, because increments have multinomial distributions with different parameters, according to the number of elementary jumps between the times t and $t+a$. In this respect, Brownian motion is a simpler object than a random walk with a finite jump size.

have means $\mathbf{E}\{(\Delta Z_j)^2\} = t_j - t_{j-1}$. Therefore, as $dt \to 0$, the sum converges to its expected value, $\sum_j (t_j - t_{j-1}) = T$.

This result implies that Brownian paths are not differentiable. To see this, we note that if the function $f(t)$ is differentiable, then we have necessarily[3]

$$\lim_{dt \to 0} \left\{ \sum_{j=1}^{N} |\Delta f_j| \right\} \approx \int_0^T |f'(s)| \, ds \, .$$

But then, in view of the inequality

$$\sum_{j=1}^{N} (\Delta f_j)^2 \le \max_j |\Delta f_j| \cdot \sum_{j=1}^{N} |\Delta f_j| \, ,$$

and the fact that

$$\max_j |\Delta f_j| = O(dt) \, ,$$

we conclude that differentiable functions f satisfy

$$\sum_{j=1}^{N} (\Delta f_j)^2 \approx \left(\text{constant} \cdot \int_0^T |f'(s)| \, ds \right) \cdot dt \ll 1 \, .$$

In other words, differentiability implies that the quadratic variation must vanish. The finiteness of the quadratic variation of Brownian Motion implies therefore that its paths are not differentiable.

Another remarkable property of Brownian motion is statistical **self-similarity**. For any parameter $\lambda > 0$, the transformation

$$Z(t) \mapsto \lambda^{-\frac{1}{2}} Z(\lambda t)$$

maps Brownian paths into Brownian paths. This means that if one considers, for instance, the ensemble of Brownian paths on the interval $[0, 1]$ and "stretches" them according to the above transformation with $\lambda = 1/2$, the result is Brownian motion on the interval $[0, 2]$. This self-similarity is consistent with the fact that Brownian paths are "fractal" objects.

[3] Using the Intermediate Value Theorem.

3. Stochastic Integrals

Definition 2. *A function $f(t)$ is said to be non-anticipative with respect to the Brownian motion $Z(t)$ if, for all $t > 0$,*

$$f(t) = \tilde{f}(\{Z(s);\ s \le t\}, t),$$

i.e. if the value of the function at time t is determined by the values taken by the history of the path $Z(\cdot)$ up to time t.

We also include in this definition deterministic functions – these have a "trivial" dependence on the random paths. The point of introducing the concept of a non-anticipative function is to distinguish between general functions of Brownian paths and those which are determined by the natural "flow of information" associated with the path $Z(\cdot)$ as time progresses.

Examples. The function

$$f_1(t) = \begin{cases} 0 & \text{if} \quad \max\limits_{0 \le s \le t} Z(s) < 5 \\ 1 & \text{if} \quad \max\limits_{0 \le s \le t} Z(s) \ge 5 \end{cases}$$

is non-anticipative. (This function is equal to zero at time t if the walk has not reached the value 5 by time t and is equal to 1 otherwise.) On the other hand,

$$f_2(t) = \begin{cases} 0 & \text{if} \quad \max\limits_{0 \le s \le 1} Z(s) < 5 \\ 0 & \text{if} \quad \max\limits_{0 \le s \le 1} Z(s) \ge 5 \end{cases}$$

is not. The reason is that the value of the latter function at any time $t < 1$ is determined by the realization of the path $Z(\cdot)$ over the entire interval $[0, 1]$ – the information gained by knowing $Z(s)$ for $s \le t$ is insufficient to determine $f_2(t)$.

Non-anticipative functions are the "natural" objects to perform integration with repect Brownian increments.

Proposition 2. *Let $f(\cdot)$ be a continuous, non-anticipative function such that*

$$\mathbf{E}\left\{ \int_0^T |f(t)|^2\, dt \right\} < \infty .$$

Then, given any sequence of partitions of the interval $[0, T]$ *with mesh size* $\Delta t \to 0$*, we have*

$$\lim_{\Delta t \to 0} \sum_{j=1}^{N} f(t_{j-1}) \cdot (Z(t_j) - Z(t_{j-1})) =$$

$$\lim_{\Delta t \to 0} \sum_{j=1}^{N} f(t_{j-1}) \cdot \Delta Z(t_j)$$

exists and is independent of the sequence used to take the limit. This limit is, by definition, the **Ito integral of f**. *It is denoted by*

$$\int_0^T f(t)\, dZ(t) \ .$$

Stochastic integration is a natural operation associated with Brownian paths: a path is "sliced" into consecutive Gaussian increments, each increment is multiplied by a random variable and these numbers are then added together again to reconstruct the stochastic integral. Thus, the stochastic integral can be viewed as a random walk with increments which have different amplitudes, or *conditional variances* – a sort of "inhomogeneous random walk". In this respect, it is important to emphasize the role of the non-anticipative assumption. Consider the j^{th} increment after multiplication by the random variable $f(t_{j-1})$:

$$f(t_{j-1})\, dZ_j = f(t_{j-1}) \cdot (Z(t_j) - Z(t_{j-1})) \ . \tag{3}$$

Once the history of the path up to time t_{j-1} is revealed, the value of $f(t_{j-1})$ is also known. Therefore, the increment of the stochastic integral over the next period *conditional on the past up to time* t is Gaussian with mean zero and variance $f(t_{j-1})^2 \cdot (t_j - t_{j-1})$. If the function $f(t)$ was anticipative (for lack of a better word), the two factors in (3) need not be conditionally independent. In the latter case, the mean of the increment may not be zero necessarily and the variance cannot be calculated in explicit form. Thus, non-anticipative functions are the correct functions to define a continuous, inhomogeneous random walk through the stochastic integral.

Some basic properties of the stochastic integral are given in

Proposition 3: *Under the above assumptions, we have*

$$\mathbf{E} \left\{ \int_0^T f(t)\, dZ(t) \right\} = 0 \ , \tag{4}$$

$$\mathbf{E}\left\{\left(\int_0^T f(t)\,dZ(t)\right)^2\right\} = \mathbf{E}\left\{\int_0^T |f(t)|^2\,dt\right\} . \tag{5}$$

Moreover, $t \mapsto \int_0^t f(s)\,dZ(s)$ *is a non-naticipative process which is continuous with probability 1.*

Proofs of Propositions 2 and 3 are sketched in the Appendix. The main idea behind the definition of the stochastic integral is the fact that Brownian motion increments dZ point "to the future" of f at each discretization time.

Example 1. Consider a function $f(t) = \sigma(t)$ which is deterministic (i.e. independent of the Brownian motion).[4] In this case, the stochastic integral

$$X(t) = \int_0^t \sigma(s)\,dZ(s) \quad , \quad t \geq 0 \ , \tag{6}$$

is a Gaussian random process, in the sense that $(X(t_1),\ X(t_2),\ \ldots X(t_N)\ )$ is a multivariate normal vector for any set of times $t_1 < t_2 < \ldots < t_N$. The reason is that the stochastic integral is a limit (in the sense of the mean-square norm) of sums of independent Gaussian random variables. Thus, if the integrand $\sigma(t)$ is deterministic, the stochastic integral is a Gaussian random walk with *time-dependent local variance.* The variance of any increment is given by the formula

$$\mathbf{E}\left\{\ (X(t+a) - X(t)\,)^2\ \right\} = \int_t^{t+a} \sigma^2(s)\,ds \ .$$

In Options Theory, stochastic processes of type (6) are often used to model the term structure of volatility.

Mathematicians have a different way of thinking about $X(t)$: in fact, suppose that we define a *new time scale* $\theta(t)$ by the equation

$$\theta(t) = \int_0^t \sigma^2(s)\,ds \ .$$

The reader may think of θ as the time that a special clock would give you whenever the "real' time was t. Defining a new process $\tilde{X}(\theta) \equiv X(t)$ (the displacement with repect to the "rubber clock") we have,

[4] We use the notation $\sigma(t)$ to suggest the notion of a "local variance".

$$\mathbf{E}\left\{ \left(\tilde{X}(\theta)\right)^2 \right\} = \mathbf{E}\left\{ (X(t))^2 \right\}$$

$$= \int_0^t \sigma^2(s)\, ds$$

$$= \theta .$$

Thus the process induced by the stochastic integral can be regarded as a Brownian motion with respect to the new clock. This is a nice result, because it characterizes the probability distribution of Ito integrals in the case of deterministic integrands.[5]

Example 2. Suppose that $f(t) = f(Z(t))$, i.e. that f depends on the current value of the Brownian path. This is, of course, a non-anticipative function so the Ito integral can be defined. Intuitively, the stochastic integral corresponds to a walk in which the local variance of increments is a function of the auxiliary Brownian path. The stochastic integral

$$\int_0^t f(Z(s))\, dZ(s)$$

will be non-Gaussian in general. We should make at this point an important remark, which is related to the results of the previous section and to what lies ahead. If Brownian paths were smooth, or rather, if one ignored the fact that they are non-smooth, then one might be tempted to consider the anti-derivative (primitive) of f and to write

$$dF(Z(t)) = f(Z(t)) \cdot dZ(t)$$

whereby

$$\int_0^t f(Z(s))\, dZ(s) = F(Z(t)) - F(0) . \tag{7}$$

[5] This trick of changing time has also made its way into financial modeling. Some exchanges close while others remain open and interest accrual on deposits may take into account time at which thereis no trading. Because of this, the local volatility of certain assets over holidays and quiet periods can be lowered to reflect the lack of strong trading activity. Similarly, the short-term volatility parameter might be increased on the date of an important economic or political announcement which will have a large impact on prices or rates. To take into account these effects in valuation models it is useful to introduce the notion of “nonlinear” time.

This is wrong !! (Unless, f is constant.) For instance, if the primitive $F(Z)$ is positive and vanishes at $Z = 0$ we would conclude from (7) that the stochastic integral is positive. This cannot be, since we know from Proposition 3 that stochastic integrals have mean zero. Consider, for instance, the elementary case $f(Z) = Z$ (and hence $F(Z) = Z^2/2$). Using the definition the stochastic integral and standard notation, we have

$$\begin{aligned}
\int_0^T Z(s)\,dZ(s) &\approx \sum_j Z_{j-1}\,dZ_j \\
&= \sum_j Z_{j-1}\,(Z_j - Z_{j-1}) \\
&= -\sum_j (Z_j - Z_{j-1})^2 + \sum_j Z_j\,(Z_j - Z_{j-1}) \\
&= -\sum_j (Z_j - Z_{j-1})^2 + \sum_j (Z_j^2 - Z_{j-1}^2) \\
&\qquad\qquad - \sum_j Z_{j-1}\,(Z_j - Z_{j-1}) \\
&\approx -T + Z^2(T) - \int_0^T Z(s)dZ(s)\ ,
\end{aligned} \tag{8}$$

where we used the result on the quadratic variation of Brownian motion (Proposition 1). Thus, we conclude that

$$\int_0^T Z(s)\,dZ(S) = \frac{Z(T)^2}{2} - \frac{T}{2}\ . \tag{9}$$

The right-hand side now has expectation zero (at it should). The additional term $\frac{T}{2}$, which is missing in the "naive" formula (7) comes from the quadratic variation $\sum (dZ_j)^2$.

Example 3. Suppose that one wants to implement the idea of a random walk with conditionally Gaussian increments (i.e. a stochastic integral) in which *the local variance depends on the position of the walk*. This means that the "elementary increment" of the stochastic integral should have the form

$$\Delta X(t) \; = \; \sigma\left(X(t)\right) \cdot \Delta Z(t) \; .$$

More formally, we would like to define a process that satisfies

$$X(t) \; = \; \int_0^t \sigma\left(X(s) \right) \, dZ(s) \; . \tag{10}$$

As opposed to Example 3, the stochastic integral on the right-hand side depends on the left-hand side of the equation. This is therefore an **integral equation** for $X(t)$, which is often written in differential form

$$dX(t) \; = \; \sigma\left(X(t) \right) \, dZ(t) \; , \tag{11}$$

in which case it is termed a **stochastic differential equation**. If $Z(t)$ were smooth, then (11) has a clear meaning and can be solved by standard methods of Ordinary Differential Equations. In the case of Brownian differentials, (11) should be interpreted as the integral equation (10). The existence and uniqueness of solutions of stochastic differential or integral equations such as (11) and (10) is treated in most books on stochastic calculus (Ikeda and Watanabe, Varadhan and Karatzas and Shreve, among many others.)

The Ito integral, or stochastic integral, is a powerful analytic tool for constructing stochastic processes which are similar to Brownian motion but have local characteristics which depend on time, the value of the process itself or more general non-anticipative factors. The implications for financial modeling are very interesting: stochastic integrals and stochastic differential equations can be used to model heterogeneity of the local price volatility, a key theme in lectures to follow.

4. Ito's Lemma

We now discuss a systematic approach for evaluating Ito integrals and functions of Brownian motion. This result can be viewed as the analogue of of the Fundamental Theorem of Calculus for functions of Brownian motion.

Proposition 4. *Let $F(Z,t)$ be a smooth function of two real variables Z and t with bounded derivatives of all orders. Then*

$$F(Z(T),\,T) \;=\; F(0,\,0) \;+\; \int_0^T \frac{\partial F}{\partial Z}(Z(s),\,s)\,dZ(s) \;+$$

$$\int_0^T \left\{ \;\frac{\partial F}{\partial t}(Z(s),\,s) \;+\; \frac{1}{2}\frac{\partial^2 F}{\partial Z^2}(Z(s),\,s)\right\} ds\;. \tag{12}$$

This proposition is known as **Ito's Lemma**. It provides the correction to formula (7) that would result from a naive application of standard Calculus. The additional term term is

$$\int_0^T \frac{1}{2}\frac{\partial^2 F}{\partial Z^2}(Z(s),\,s)\,ds\;. \tag{13}$$

(Compare with (9).) The Fundamental Theorem of Calculus (cf. (7)) does not involve second derivatives, the reason for this being that the contribution to the integral due to quadratic and higher-order terms is negligible. (The quadratic variation of a smooth function is zero). In contrast, the $(dZ)^2$-terms contribute to the differential in the case of Brownian motion because the quadratic variation of the paths is non-trivial. The beauty of Ito's Lemma is that it provides a "closed-form" expression for $(F(Z(t),\,t)$ for any reasonable smooth function F, elucidating the effect of the quadratic variation. This avoids having to go through manipulations of sums like the ones done in the Example 2 of the previous section.

Sketch of the proof of Ito's Lemma.[6]

We consider the Taylor expansion of of F about some point (Z,t). Formally, we have

$$\Delta F \;=\; F_Z\,\Delta Z \;+\; F_t\,\Delta t \;+\; \frac{1}{2}\,F_{ZZ}\;(\Delta Z\,)^2 \;+\; F_{Zt}\,\Delta Z\,\Delta t \;+\; \frac{1}{2}\,F_{tt}\;\;(\Delta t\,)^2 \;+\; ... \tag{14}$$

[6] Later on, in the study of hedging in imperfect markets with transaction costs, we will need to revisit the mathematical derivation of Ito's Lemma.

Now, assume that we have a partition of the interval $[0,\, T]$, $\{t_j\}$, and that

$$Z \;= Z(t_{j-1})\;,\; t \;=\; t_{j-1}\;,$$

$$\Delta t \;=\; t_j \;-\; t_{j-1}\;,$$

and

$$\Delta Z_j \;=\; Z_{t_j} \;-\; Z_{t_{j-1}}\;.$$

Considering the Taylor expansion (14) and adding up the successive increments, we find that the first term in the right-hand side of (14) gives the contribution $\int F_Z(Z(s),\, s)\, ds$ to (12). Similarly, the second term in the Taylor expansion (14) contributes to the integral $\int F_t\, dt$ in (12). These are the two terms that we expect from standard differential Calculus.

Let us turn to the higher-order terms in the Taylor expansion. Since we have

$$\mathbf{E}\,\{\,|\Delta Z\,|^p\;(\,\Delta t\,)^q\,\}\;\propto\;(\Delta t\,)^{p/2\,+\,q}\;,$$

the contribution to $\sum dF_j$ which arises from adding the $(N \;=\; T/\Delta t)$ terms proportional to $(\Delta Z_j)^p\,(\Delta t)^q$ has order $(\,\Delta t\,)^{p/2\,+\,q\,-\,1}$. The conclusion is that the only terms with $p\,+\,q \geq\; 2$ which contribute are those with $p \;=\; 2$ and $q \;=\; 0$, i.e.the $(\Delta Z)^2$ terms. All other terms vanish asymptotically as $\Delta t\; \to 0$.

This means that we should study the asymptotic behavior of the sums

$$\frac{1}{2}\sum_{j=1}^{N} F_{ZZ}\,(Z(t_{j-1},\, t_{j-1}\,)\;(\Delta Z_j\,)^2\;, \qquad (15)$$

recalling that each increment DZ_j points to the future of t_{j-1}. This sum is equal to the Riemann sum

$$\frac{1}{2}\sum_{j=1}^{N} F_{ZZ}\,(Z(t_{j-1},\, t_{j-1}\,)\;\Delta t\;,$$

approximating the integral in (13). To establish Ito's Lemma, it suffices to prove that the sums

$$\frac{1}{2}\sum_{j=1}^{N} F_{ZZ}\,(Z(t_{j-1},\, t_{j-1}\,)\;[\;\Delta Z_j\,)^2\;-\;\Delta t\;] \qquad (16)$$

converge to zero as $\Delta t \to 0$ in a suitable sense. To see this, we analyze the mean and the variance of the sum. This analysis is similar to the one of the proof of Propositions 2 and 3 given in the Appendix. The crucial point is that the increments in (16) have mean zero and point towards the future. Because of the latter property, we have

$$\mathbf{E}\ \{\ F_{ZZ}\,(Z(t_{j-1},\,t_{j-1}\,)\ \left[\ \Delta Z_j\,)^2\ -\ \Delta t\ \right]\ \}\ =\ 0\ .$$

The expected value of the sums (16) is therefore zero. Consider now the variance of (16),

$$\frac{1}{4}\,\mathbf{E}\left\{\left(\sum_{j=1}^{N} F_{ZZ}\,(Z(t_{j-1},\,t_{j-1}\,)\ \left[\ \Delta Z_j\,)^2\ -\ \Delta t\ \right]\right)^2\right\}\ .$$

We observe two things: first, the variables $F_{ZZ}\,(Z(t_{j-1},\,t_{j-1}\,)\ \left[\ \Delta Z_j\,)^2\ -\ \Delta t\ \right]$ and $F_{ZZ}\,(Z(t_{k-1},\,t_{k-1}\,)\ \left[\ \Delta Z_k\,)^2\ -\ \Delta t\ \right]$ are uncorrelated for $j\ \neq\ k$. The reason is that if, say, $t_j\ <\ t_k$,

$$F_{ZZ}\,(Z(t_{j-1},\,t_{j-1}\,)\ \left[\ \Delta Z_j\,)^2\ -\ \Delta t\ \right]\cdot F_{ZZ}\,(Z(t_{k-1},\,t_{k-1}\,)$$

and

$$(\Delta Z_k\,)^2\ -\ \Delta t$$

are independent conditionally on the past until time t_{k_1}. Second, the latter random variable has expectation zero. This guarantees that the expectation of their product is zero. Consequently, the variance is given, to leading order, by

$$\frac{1}{4}\sum_{j=1}^{N}\mathbf{E}\left\{\ (F_{ZZ}\,(Z(t_{j-1},\,t_{j-1}\,)\,)^2\right\}\ \mathbf{E}\left\{\ (\,(\Delta Z_j\,)^2\ -\ \Delta t\,)^2\ \right\}\ \approx$$

$$\frac{1}{2}\ \Delta t\ \cdot\ \left(\int_0^t \mathbf{E}\ \left\{\ (F_{ZZ}\,(Z(t),\,t\,)\,)^2\right\}\ dt\right)\ ,$$

a negligible quantity as $\Delta t\ \to\ 0$. Here, we used the fact that

$$\mathbf{E}\ \left\{\,(\Delta Z^2\ -\ \Delta t\,)^2\ \right\}\ =\ (\Delta t)^2\ \mathbf{E}\ \left\{\,(N^2\ -\ 1\,)^2\,\right\}\ =\ 2\,(\Delta t)^2\ ,$$

where N is a standard normal.[7] We have thus shown that the sums in (15) converge to the integral (13) (the "Ito correction term") in the sense that the difference has mean zero and variance converging to zero as $\Delta t \to 0$.

Example. Consider the following function of Brownian motion:

$$S(t) = S_0 \, e^{\sigma Z(t) + \mu t} \quad , \quad t \geq 0 \, ,$$

where σ and μ are constants. This random process is sometimes called **geometric Brownian motion**. Let us apply Ito's Lemma to $S(t)$, with the object of finding the integral equation satisfied by it. Accordingly, applying (12) to the function $F(Z, t) = S_0 \exp \{ \sigma Z + \mu t \}$, we find that

$$S(t) - S(0) = \int_0^t S(\tau) \, \sigma \, dZ(\tau) + \int_0^t S(\tau) \left(\mu + \frac{\sigma^2}{2} \right) d\tau \, . \tag{18}$$

If we use differential notation, we obtain

$$dS(t) = S(t) \, \sigma \, dZ(t) + S(t) \left(\mu + \frac{\sigma^2}{2} \right) dt \, . \tag{19}$$

This stochastic differential equation states that geometric Brownian motion has the property that the infinitesimal *relative increments*

$$\frac{dS(t)}{S(t)} = \sigma \, dZ(t) + \left(\mu + \frac{\sigma^2}{2} \right) dt$$

are normal with mean $\mu + \frac{\sigma^2}{2}$ and variance σ^2. This is the Black & Scholes "world" for option pricing theory, which we will discuss shortly using stochastic calculus ideas.

[7] From the explicit form of the moment-generating function for the standard normal distribution, $\mathbf{E} \left\{ e^{\lambda N} \right\} = e^{\frac{\lambda^2}{2}}$, it follows that $\mathbf{E} \left\{ N^{2k} \right\} = (2k)! / \, 2^k k!$ for all integers $k > 0$.

4. Ito Processes and Ito Calculus

What is the most general class of random processes that can be described in terms of sums of stochastic integrals and standard integrals? The answer is the class of **Ito processes**.

Definition: *We say that a random process* $X(t)$, $t \geq 0$ *is an* Ito Process *if there exists a Brownian motion measure and two non-anticipative functions* $\sigma(t)$ *and* $b(t)$, $t \geq 0$ *such that*

$$X(t) = X(0) + \int_0^t \sigma(s)\, dZ(s) + \int_0^t b(s)\, ds , \quad t > 0 , \tag{20}$$

or, in terms of differentials,

$$dX(t) = \sigma(t)\, dZ(t) + b(t)\, dt . \tag{21}$$

Intuitively speaking, Ito processes are continuous random functions with infinitesimal increments which, conditionally on the past until time t, are Gaussian with mean $b(t)$ and variance $\sigma^2(t)$.[8]

An important class of Ito processes which we already encountered in the lecture are functions of Brownian motion, i.e.

$$X(t) = F(Z(t), t) . \tag{22}$$

In fact, Ito's Lemma shows that the **local parameters** associated to this process are

$$\sigma(t) = \frac{\partial F(Z(t), t)}{\partial Z}$$

and

$$b(t) = \frac{\partial F(Z(t), t)}{\partial t} + \frac{1}{2} \frac{\partial^2 F(Z(t), t)}{\partial Z^2} .$$

The next result is very important for applications. It states that any smooth function of an Ito process is also an Ito process and gives a formula for the local parameters.

[8] is important to observe that, in general, the increments will not be Gaussian over a *finite* time interval.

Proposition 5 (Generalized Ito Lemma) *Suppose that $X(t)$ is an Ito process with local parameters $\sigma(t)$ and $b(t)$, i.e that (21) or (22) hold. Let $F(X,\, t)$ be smooth function of (X, t) with bounded derivatives of all orders. Then*

$$Y(t) \;=\; F(X(t),\, t)$$

is an Ito process with local parameters $\Sigma(t)$ and $B(t)$ given by

$$\Sigma(t) \;=\; \frac{\partial F(X(t),\, t)}{\partial X} \cdot \sigma(t) \tag{23}$$

and

$$B(t) \;=\; \frac{\partial F(X(t),\, t)}{\partial t} \;+\; \frac{1}{2}\,\sigma^2(t)\,\frac{\partial^2 F(X(t),\, t)}{\partial X^2} \;+\; b(t)\,\frac{\partial F(X(t),\, t)}{\partial X}\,. \tag{24}$$

In other terms, the Generalized Ito Lemma states that, for all $t \;>\; 0$, the integral equation

$$F(X(t),\, t) \;=\; F(X(0),\, 0) \;+\; \int_0^t \frac{\partial F(X(s),\, s)}{\partial X} \cdot \sigma(s)\, dZ(s) \;+$$

$$\int_0^t \left\{ \frac{\partial F(X(s),\, s)}{\partial t} \;+\; \frac{1}{2}\,\sigma^2(s)\,\frac{\partial^2 F(X(s),\, s)}{\partial X^2} \;+\; b(s)\,\frac{\partial F(X(s),\, s)}{\partial X} \right\} ds \tag{25}$$

holds for all sufficiently smooth functions $F(X, t)$. It is convenient to express equation (25) in differential form, like we did in for the case of Brownian motion. Accordingly,

$$dF(X(t),\, t) \;=\; \frac{\partial F(X(t),\, t)}{\partial X}\,\sigma(t)\, dZ(t) \;+\; \frac{\partial F(X(t),\, t)}{\partial X}\; b(t)\, dt \;+$$

$$\frac{\partial F(X(t),\, t)}{\partial t}\, dt \;+ \frac{1}{2}\,\sigma^2(t)\,\frac{\partial^2 F(X(t),\, t)}{\partial X^2}\, dt$$

$$= \frac{\partial F(X(t),\, t)}{\partial X} \; (\, \sigma(t)\, dZ(t) \; + \; b(t)\, dt \;) \; +$$

$$\frac{\partial F(X(t),\, t)}{\partial t}\, dt \; + \; \frac{1}{2}\, \sigma^2(t)\, \frac{\partial^2 F(X(t),\, t)}{\partial X^2}\, dt$$

$$= \frac{\partial F(X(t),\, t)}{\partial X} \cdot dX(t) \; + \; \frac{\partial F(X(t),\, t)}{\partial t}\, dt \; + \; \frac{1}{2}\, \sigma^2(t)\, \frac{\partial^2 F(X(t),\, t)}{\partial X^2}\, dt$$

$$\equiv \frac{\partial F(X(t),\, t)}{\partial X} \cdot dX(t) \; + \; \frac{\partial F(X(t),\, t)}{\partial t}\, dt \; + \; \frac{1}{2}\, \frac{\partial^2 F(X(t),\, t)}{\partial X^2}\, (dX(t)\,)^2 \; , \tag{26}$$

where we used equation (21) and the formal relation

$$(dX(t)\,)^2 \; \equiv \; \sigma^2(t)\, dt \; . \tag{27}$$

The latter equation is a convention: it expresses in synthetic form the contributions from quadratic terms ($(dX(t))^2$ or $(dZ(t))^2$-terms). Thus, the Generalized Ito Lemma can be written concisely in the form

$$dF(X(t),\, t) \; == \; \frac{\partial F(X(t),\, t)}{\partial X} \cdot dX(t) \; + \; \frac{\partial F(X(t),\, t)}{\partial t}\, dt \; + \; \frac{1}{2}\, \frac{\partial^2 F(X(t),\, t)}{\partial X^2}\, (dX(t)\,)^2 \; .$$

The result is simple to remember: *the infinitesimal increment of a smooth function of an Ito process is obtained by making a Taylor expansion of order 1 in* dt*, of order 2 in* dX *and using the convention (27).* This convention often referred to as the **Ito multiplication rule**. Note that if $\sigma(t) \equiv 1$ and $b(t) \equiv 0$, we recover the "little" Ito Lemma (Proposition 4) of the previous section.

In summary, *stochastic differential calculus is an extension of standard calculus in which we use the Ito multiplication rule (27) to account for the effect of non-trivial quadratic variations.*

The proof of the Generalized Ito Lemma is very similar to the proof of Proposition 4. We omit it. The interested reader should consult for instance Ikeda and Watanabe, Varadhan, or Karatzas and Shreve.

APPENDIX: PROPERTIES OF THE ITO INTEGRAL

The aim of this Appendix is to provide additional technical elements to understand the basic properties of the stochastic integral stated in Propositions 1 and 2.[9]

Let us consider Proposition 2. Notice that any "approximating sum"

$$s_N \;=\; \sum_{j=1}^{N} f(t_{j-1})\,(\,Z(t_j) \,-\, Z(t_{j-1})\,) \;\equiv\; \sum_{j=1}^{N} f_{j-1}\, dZ_j \tag{28}$$

has mean zero and uniformly bounded variance. In fact, since $f(t)$ is non-anticipative, we have

$$\begin{aligned}
\mathbf{E}\,\{f_{j-1}\, dZ_j\,\} \;&=\; \mathbf{E}\,\{\,\mathbf{E}\,\{f_{j-1}\, dZ_j\,|Z(s),\ s\,\le\, t_{j-1}\,\}\,\} \\
&=\; \mathbf{E}\,\{\,f_{j-1}\,\mathbf{E}\,\{dZ_j\,|Z(s),\ s\,\le\, t_{j-1}\,\}\,\} \\
&=\; 0\,,
\end{aligned}$$

where $\mathbf{E}\,\{\,\bullet\,|Z(s),\ s\,\le\, t_{j-1}\,\}$ represents the conditional expectation operator given the history of the path up to time t_{j-1}. The conditional expectation of dZ_j vanishes because the increments of $Z(t)$ are independent of their past. We conclude that the sum in (28) has mean zero. To compute its variance, we must evaluate the sum of terms of the form

$$\mathbf{E}\,\{f_{j-1}\, dZ_j\,\cdot\,f_{k-1}\, dZ_k\,\}\;,$$

where $1\,\le\, j,k\,\le\, N$. The key observation is that this expectation vanishes if $j \,\ne\, k$. In fact, if $j\,<\,k$, we have

$$\begin{aligned}
\mathbf{E}\,\{f_{j-1}\, dZ_j\,\cdot\,f_{k-1}\, dZ_k\,\} \;&=\; \mathbf{E}\,\{\,\mathbf{E}\,\{f_{j-1}\, dZ_j\,\cdot\,f_{k-1}\, dZ_k\,|Z(s),\ s\,\le\, t_{k-1}\,\}\,\} \\
&=\; \mathbf{E}\,\{\,f_{j-1}\; dZ_j\; f_{k-1}\;\mathbf{E}\,\{dZ_k\,|Z(s),\ s\,\le\, t_{k-1}\,\}\,\} \\
&=\; 0
\end{aligned}$$

[9] These are in lieu of formal proofs, which would take us too far from the subject of these lectures into Measure Theory. I encourage the interested reader to consult the aforementioned references to obtain mathematically precise hypotheses on $f(t)$, on the types of convergence of random variables used in the proofs, etc. So much to say, so little time...

The variance on (28) is therefore equal to the sum of the "diagonal" terms

$$\sum_{j=1}^{N} \mathbf{E} \left\{ (f_{j-1} \ dZ_j)^2 \right\} \ = \ \sum_{j=1}^{N} \mathbf{E} \left\{ (f(t_{j-1}))^2 \right\} (t_j \ - \ t_{j-1})$$

$$\approx \int_0^T \mathbf{E} \left\{ (f(t))^2 \right\} dt \ .$$

(This assumes implicitly that the sum can be approximated by the integral, i.e. that $\mathbf{E} \left\{ (f(t))^2 \right\}$ is continuous.)

We have established that the random variables s_N have bounded mean and variance. Next, we show that the sums *converge to a limit* as the partitions become finer and finer.

Consider a sequence of partitions $\left\{ t_j^{(\nu)} \right\}_{j=1}^{N_\nu}$, where

$$\max_{1 \ \leq \ j \ \leq \ N_\nu} \left(t_j^{(\nu)} \ - \ t_{j-1}^{(\nu)} \right) \ \rightarrow \ 0 \ \text{ as } \ \nu \ \rightarrow \ \infty \ ,$$

and the the corresponding sequence of sums s_{N_ν}. We claim that

$$\lim_{\nu, \, \nu' \ \rightarrow \infty} \mathbf{E} \left\{ \ (s_{N_\nu} \ - \ s_{N_{\nu'}} \ )^2 \ \right\} \ = \ 0 \ .$$

Consider two such sums, s_{N_ν} and $s_{N_{\nu'}}$, corresponding to different partitions. The idea is to "merge" the two partitions to form a finer one which combines the nodes of each of them. A moment of reflection shows that, with respect this refined partition (which we call $\{t_j\}$ to fix ideas) both sums can be written in the form

$$\sum_{j=1}^{N} f(t^*_{j-1}) \ (\ Z(t_j) \ - \ Z(t_{j-1}) \ )$$

and

$$\sum_{j=1}^{N} f(t^{**}_{j-1}) \ (\ Z(t_j) \ - \ Z(t_{j-1}) \ )$$

where the t^*_j s and the t^{**}_j are times (of the original partitions) such that

$$t_j^* \leq t_j \quad , \quad t_j^{**} \leq t_j \tag{29}$$

and the difference $t_j^{**} - t_j^*$ converges to zero as $\nu \to \infty$. The difference of the two sums can therefore be expressed in the form

$$\sum_{j=1}^{N} \left(f(t_{j-1}^*) - f(t_{j-1}^{**}) \right) \cdot \left(Z(t_j) - Z(t_{j-1}) \right)$$

Taking into account (29), it is easy to show that this sum has mean zero and variance

$$\sum_{j=1}^{N} \mathbf{E} \left\{ \left(f(t_{j-1}^*) - f(t_{j-1}^{**}) \right)^2 \right\} \cdot \left(t_j - t_{j-1} \right) . \tag{30}$$

If $f(t)$ satisfies suitable boundedness and continuity assumptions, e.g, if $f(t)$ is bounded and

$$\lim_{h \to 0} \mathbf{E} \left\{ \left(f(t+h) - f(t) \right)^2 \right\} = 0 ,$$

the sum in (30) tends to zero with the mesh size.

This argument shows the existence of the Ito stochastic integral. The properties stated in Proposition 3 also follow from these calculations.