Lecture 16:

CHAPTER 6: ITO'S FORMULA AND FIRST APPLICATIONS

In this chapter we will prove Ito's formula, which is a fundamental "change of variable formula" for stochastic integrals, and the source of explicit calculations. We will also discuss some of these applications. Throughout this chapter $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, P)$ will denote a filtered probability space, which satisfies the "usual conditions", cf. (4.5), (4.6).

Definition:

A continuous semimartingale $(Y_t)_{t\geq 0}$ on $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, P)$ is a continuous adapted process, which admits the decomposition

(6.1) $Y_t = Y_0 + M_t + A_t,\ t\geq 0,$

where $M_t, t\geq 0$, is a continuous local martingale such that $M_0=0$, and $A_t, t\geq 0$, is a continuous adapted process with bounded variation on finite intervals, such that $A_0=0$.

Remark:

The same argument used in the proof of (5.43), (5.44) shows that when $(Y_t)_{t\geq 0}$ is a continuous semimartingale,

(6.2) the decomposition (6.1) is essentially unique. □

Notation:

For $(Y_t)_{t\geq 0}$ a continuous semimartingale we will write

(6.3) $\Lambda(Y) = \{H$: $\mathcal{P}$-measurable on $\Omega\times\mathbb{R}_+$ such that P-a.s., for $t\geq 0$, $\int_0^t H_s^2\, d\langle M\rangle_s < \infty$ and $\int_0^t |H_s|\, d|A|_s < \infty\}$,

where M and A are as in (6.1) and $|A|_\cdot$ denotes the total variation process of A

Then for $H \in \Lambda(Y)$, we will use the notation

(6.4) $\int_0^t H_s\, dY_s = \int_0^t H_s\, dM_s + \int_0^t H_s\, dA_s, \quad t \geq 0,$

so that (6.4) defines $\int_0^t H_s\, dY_s$, $t \geq 0$, in an essentially unique fashion (with respect to the various versions and decompositions in (6.1)).

Example:

Any continuous adapted process $H_s(\omega)$ is automatically in $\Lambda(Y)$. In particular an expression such as $\int_0^t \exp\{\exp(Y_s^2 + s^2)\}\, dY_s$ (for instance) is well defined.

□

An important first step towards Ito's formula will be the next

Proposition: (Integration by parts formula)

If $Y_t, t \geq 0$, and $Z_t, t \geq 0$, are continuous semimartingales on $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, P)$, then P-a.s., for all $t \geq 0$,

(6.5) $Y_t Z_t = Y_0 Z_0 + \int_0^t Y_s\, dZ_s + \int_0^t Z_s\, dY_s + \langle Y, Z\rangle_t,$

where $\langle Y, Z\rangle_t, t \geq 0$, denotes the bracket of the local martingale parts of Y and Z.

Proof: We begin with several reductions.

It is enough to prove for Y as above that P-a.s.,

(6.6) $Y_t^2 = Y_0^2 + 2\int_0^t Y_s\, dY_s + \langle Y\rangle_t, \quad$ for $t \geq 0$,

(i.e. prove (6.5) when $Y = Z$). Indeed one then applies (6.6) to $(Y+Z)^2$ and $(Y-Z)^2$ and recovers (6.5).

Moreover we can also replace Y with $Y^n = 1\{T_n > 0\}\, Y^{T_n}$, where $Y^{T_n}_\cdot = Y_{\cdot \wedge T_n}$, $n \geq 1$, and $T_n \uparrow \infty$ is the sequence of stopping times

(6.7) $T_n = \inf\{s \geq 0, |Y_s| \geq n, |M_s| \geq n, |A|_s \geq n \text{ or } \langle M\rangle_s \geq n\}, \quad n \geq 1,$

and in this fashion we can assume that

$Y_\cdot$, $M_\cdot$, $|A|_\cdot$, $\langle Y\rangle_\cdot$ are continuous bounded processes, (so $M_\cdot$ is in fact a martingale)

Since all processes appearing in (6.6) are continuous, it is also sufficient to prove (6.6) for fixed t.

We thus pick a fixed $t \geq 0$, and define for $m \geq 1$,

(6.8) $t_i = \frac{i t}{m}$, $0 \leq i \leq m$.

We then write:

(6.9) $Y_t^2 = \big(Y_0 + \sum_{i=0}^{m-1} Y_{t_{i+1}} - Y_{t_i}\big)^2 = Y_0^2 + \sum_{i<m} (Y_{t_{i+1}} - Y_{t_i})^2 +$
$+ 2 \sum_{i<m} Y_{t_i}(Y_{t_{i+1}} - Y_{t_i})$.

We now analyze the convergence of the last two terms in the right-hand side of (6.9), as $m \to \infty$. We have:

(6.10) $\sum_{i<m} Y_{t_i}(Y_{t_{i+1}} - Y_{t_i}) = \sum_{i<m} Y_{t_i}(M_{t_{i+1}} - M_{t_i}) + \sum_{i<m} Y_{t_i}(A_{t_{i+1}} - A_{t_i})$

$$= \int_0^t Y_s^m \, dM_s + \int_0^t Y_s^m \, dA_s,$$

with

(6.11) $Y_s^m(\omega) = \sum_{0 \leq i < m} Y_{t_i}(\omega) 1_{(t_i, t_{i+1}]}(s)$.

Clearly using dominated convergence we find that:

$E[\int_0^t (Y_s - Y_s^m)^2 \, d\langle Y \rangle_s] \underset{m\to\infty}{\longrightarrow} 0$, and

$E[\int_0^t |Y_s - Y_s^m| \, d|A|_s] \underset{m\to\infty}{\longrightarrow} 0$.

It thus follows that

(6.12) $\int_0^t Y_s^m \, dM_s \xrightarrow[m\to\infty]{L^2(P)} \int_0^t Y_s \, dM_s$ and $\int_0^t Y_s^m \, dA_s \xrightarrow[m\to\infty]{L^1(P)} \int_0^t Y_s \, dA_s$.

This shows that

(6.13) $\sum_{i<m} Y_{t_i}(Y_{t_{i+1}} - Y_{t_i}) \xrightarrow[m\to\infty]{L^1(P)} \int_0^t Y_s \, dY_s$.

We now come back to the second term of the right-hand side of (6.9).

We write for $0 \leq i < m$,

(6.14) $\Delta_i^m \overset{\text{def}}{=} (M_{t_{i+1}} - M_{t_i})^2 - (\langle Y \rangle_{t_{i+1}} - \langle Y \rangle_{t_i})$.

The calculation below resembles what we did in (3.4). For $m \geq 1$,

(6.15) $A_m \overset{\text{def}}{=} E\big[\big(\sum_{i<m} (M_{t_{i+1}} - M_{t_i})^2 - \langle Y \rangle_t\big)^2\big] = E\big[\big(\sum_{i<m} \Delta_i^m\big)^2\big] =$
$\sum_{i<m} E[(\Delta_i^m)^2] + 2 \sum_{i<j<m} E[\Delta_i^m \Delta_j^m]$.

As we now explain the Δ_i^m, $1 \leq i < m$, are pairwise orthogonal.

Indeed by (5.9), we have for $j < m$

(6.16) $E[\Delta_j^m | \mathcal{G}_{t_j}] = E[M_{t_{j+1}}^2 - 2M_{t_{j+1}} M_{t_j} + M_{t_j}^2 - (\langle Y\rangle_{t_{j+1}} - \langle Y\rangle_{t_j}) | \mathcal{G}_{t_j}] =$

$E[M_{t_{j+1}}^2 - \langle M\rangle_{t_{j+1}} | \mathcal{G}_{t_j}] - 2M_{t_j} E[M_{t_{j+1}} | \mathcal{G}_{t_j}] + M_{t_j}^2 + \langle Y\rangle_{t_j} \overset{(5.9)}{=}$

$M_{t_j}^2 - \langle M\rangle_{t_j} - 2M_{t_j}^2 + M_{t_j}^2 + \langle Y\rangle_{t_j} = 0.$

Thus the last term of (6.15) vanishes. Moreover we have

$$A_m = \sum_{i<m} E[(\Delta_i^m)^2] \leq \sum_{i<m} 2E[(M_{t_{i+1}} - M_{t_i})^4] + 2E[(\langle Y\rangle_{t_{i+1}} - \langle Y\rangle_{t_i})^2]$$

$$\leq 2\, E[\sup_{i<m} |M_{t_{i+1}} - M_{t_i}|^2 |\sum_{i<m} \Delta_i^m|]$$

$$+ 2\, E[(\sup_{i<m} |M_{t_{i+1}} - M_{t_i}|^2 + \sup_{i<m} |\langle Y\rangle_{t_{i+1}} - \langle Y\rangle_{t_i}|)\, \langle Y\rangle_t]$$

$m \to \infty \downarrow$ 0 by dominated convergence.

Further note that $2ab \leq 2a^2 + \frac{b^2}{2}$, for a, b in $\mathbb{R}$.

(6.17) $2E[\sup_{i<m} |M_{t_{i+1}} - M_{t_i}|^2 |\sum_{i<m} \Delta_i^m|] \overset{(6.15)}{\leq} \frac{A_m}{2} + 2\, E[\sup_{i<m} |M_{t_{i+1}} - M_{t_i}|^2]$ $\searrow_0^{m\to\infty}$ dominated convergence

Thus coming back to the line, we have shown

(6.18) $A_m = \sum_{i<m} E[(\Delta_i^m)^2] \xrightarrow[m\to\infty]{} 0.$

By (6.15) this means that

(6.19) $\sum_{i<m} (M_{t_{i+1}} - M_{t_i})^2 \xrightarrow[m\to\infty]{L^2(P)} \langle Y\rangle_t.$

To prove an analogous statement for $\sum_{i<m} (Y_{t_{i+1}} - Y_{t_i})^2$, which is our main object of interest in view of (6.9), we write:

(6.20) $|\sum_{i<m} (Y_{t_{i+1}} - Y_{t_i})^2 - \sum_{i<m} (M_{t_{i+1}} - M_{t_i})^2| \leq$

$$2 \sum_{i<m} |A_{t_{i+1}} - A_{t_i}| |M_{t_{i+1}} - M_{t_i}| + |A_{t_{i+1}} - A_{t_i}|^2 \leq$$

$$2 \left(\sum_{i<m} |A_{t_{i+1}} - A_{t_i}|^2\right)^{1/2} \left(\sum_{i<m} (M_{t_{i+1}} - M_{t_i})^2\right)^{1/2} + \sum_{i<m} (A_{t_{i+1}} - A_{t_i})^2$$

and observe that by dominated convergence and continuity,

$$\sum_{i<m} (A_{t_{i+1}} - A_{t_i})^2 \leq \sup_{i<m} |A_{t_{i+1}} - A_{t_i}|\, |A|_t \xrightarrow[m\to\infty]{L^2(P)} 0.$$

Thus coming back to the last line of (6.20) we see using Cauchy-Schwarz's inequality for the first term and (6.19) that all terms converge to 0 in $L^2(P)$. We have thus shown that

(6.21) $\sum_{i<m} (Y_{t_{i+1}} - Y_{t_i})^2 \xrightarrow[m\to\infty]{L^2(P)} \langle Y\rangle_t.$

Together with (6.13) this concludes the proof of (6.6) for fixed t, and thus our general claim (6.5) has been established. $\square$

We now turn to the main result.

<u>Theorem:</u> (Itô's formula)

Let F be a C^2-function on $\mathbb{R}^d$, and $Y^1_\cdot, \dots, Y^d_\cdot$ be continuous semimartingales on $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, P)$. Then writing $Y_\cdot = (Y^1_\cdot, \dots, Y^d_\cdot)$, the real-valued process $F(Y_t)$, $t \geq 0$, is a continuous semimartingale and P-a.s., for all $t \geq 0$,

$$(6.22)\quad F(Y_t) = F(Y_0) + \sum_{i=1}^d \int_0^t \partial_i F(Y_s)\, dY^i_s + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial^2_{ij} F(Y_s)\, d\langle Y^i, Y^j\rangle_s .$$

<u>Proof:</u>

The formula (6.22) shows that $F(X_t)$, $t \geq 0$, is a continuous semimartingale. We use several reductions to prove (6.22).

- first reduction:

using "localization", just as explained above (6.7), we can assume that Y^i_t, $|\langle Y^i, Y^j\rangle|_t$ and the total variation of the bounded variation processes entering the decomposition (6.1) of the Y^i_t, $t \geq 0$, are uniformly bounded processes.

- second reduction:

we can assume that $F(.)$ is C^2 with compact support.

- third reduction:

we can assume that $F(.)$ is a polynomial.

Indeed note that it suffices now to prove (6.22) for $F \in C^\infty_c(\mathbb{R}^d, \mathbb{R})$, since any $F \in C^2_c(\mathbb{R}^d, \mathbb{R})$ is approximated, for instance by convolution, in C^2-topology by such functions, and (6.22) remains true in the limit. But $F \in C^\infty_c(\mathbb{R}^d, \mathbb{R})$ is approximated in C^2-topology on any compact by linear combinations of $e^{i\xi \cdot x}$, with $\xi \in \mathbb{R}^d$, (for instance using Fourier series, when L is large enough so that support$(f) \subset (-\frac{L}{2}, \frac{L}{2})^d$, one has:

$F(x) = \sum_{k \in \mathbb{Z}^d} a_k \, e^{i 2\pi \frac{k}{L} \cdot x}$, for $x \in (-\frac{L}{2}, \frac{L}{2})^d$, with $a_k = \frac{1}{L^d} \int_{(-\frac{L}{2}, \frac{L}{2})^d} f(z) e^{-i 2\pi \frac{k}{L} \cdot z} dz$).

Now we have the expansion $e^{i\xi \cdot x} = \sum_{n \geq 0} \frac{1}{n!} (i\xi \cdot x)^n$, which shows that $e^{i\xi \cdot x}$ is approximated in C^2-topology on compact sets by polynomials and the third reduction follows.

We are thus reduced to proving (6.22) for $F(\cdot)$ a polynomial in the coordinate variables. We will now prove that:

(6.23) the validity of (6.22) for the polynomial F implies the validity of (6.22) for the polynomials $G(x_1, \dots, x_d) = x_{i_0} F(x_1, \dots, x_d)$, for any $1 \leq i_0 \leq d$.

Indeed we apply (6.5) to Y^{i_0} and $F(Y)$, and find that

$$G(Y_t) = G(Y_0) + \int_0^t Y_s^{i_0} \, dF(Y)_s + \int_0^t F(Y_s) \, dY_s^{i_0} + \langle Y^{i_0}, F(Y) \rangle_t$$

and since (6.22) holds true for $F(Y)$

$$G(Y_t) = G(Y_0) + \sum_{i=1}^d \int_0^t Y_s^{i_0} \partial_i F(Y_s) \, dY_s^i + \int_0^t F(Y_s) \, dY_s^{i_0}$$

$$+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t Y_s^{i_0} \partial_{ij}^2 F(Y_s) \, d\langle Y^i, Y^j \rangle_s + \langle Y^{i_0}, F(Y) \rangle_t .$$

Now with (5.82) and (6.22) we also have

$$\langle Y^{i_0}, F(Y) \rangle_t = \sum_{i=1}^d \int_0^t \partial_i F(Y_s) \, d\langle Y^{i_0}, Y^i \rangle_s .$$

Inserting this identity in the last term of the previous formula now yields that:

$$G(Y_t) = G(Y_0) + \sum_{i=1}^d \int_0^t \partial_i G(Y_s) \, dY_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{ij}^2 G(Y_s) \, d\langle Y^i, Y^j \rangle_s ,$$

and (6.23) is proved.

Since (6.22) clearly holds when F = constant, it follows by (6.23), that (6.22) holds for all polynomials F. This as explained above yields the general claim. $\square$

<u>Example:</u> (canonical d-dimensional Brownian motion)

We can apply the above theorem in the special case where $Y_. = X_. (= (X^1_., \dots, X^d_.))$ is the canonical d-dimensional Brownian motion, and $(C, \overline{\mathcal{F}}, (\mathcal{F}_t)_{t \geq 0}, W_0)$, cf. (4.7), is the filtered probability space, (which as we have seen satisfies the usual conditions).

In this example each X^i_t, $t \geq 0$, are $(\mathcal{F}_t)$-martingales, and

(6.24) $X^i_t X^j_t - \delta_{ij}\, t$, $t \geq 0$, are $(\mathcal{F}_t)$-martingales, for $1 \leq i, j \leq d$,

$\uparrow$ Kronecker's symbol.

(when $i = j$, see (4.2), (4.18), the case $i \neq j$ simply uses a simple modification of (4.4)). In view of (5.43), (5.44), this means that

(6.25) $\langle X^i, X^j \rangle_t = \delta_{ij}\, t$, $t \geq 0$, for $1 \leq i, j \leq d$.

In particular, in the rightmost term of Ito's formula, only terms with $i = j$ are present, and hence for $F \in C^2(\mathbb{R}^d, \mathbb{R})$, W_0-a.s., for all $t \geq 0$,

(6.26) $F(X_t) = F(0) + \sum_{i=1}^d \int_0^t \partial_i F(X_s)\, dX^i_s + \frac{1}{2} \int_0^t \Delta F(X_s)\, ds$,

where $\Delta F(x) \overset{\text{def.}}{=} \sum_{i=1}^d \partial_i^2 F(x)$ is the Laplacian of F.

$\square$