MATH 777 MARKOV STOCHASTIC PROCESSES SPRING 2009

Lecture 33. Itô's formula.

I was busy presenting Itô's and Gikhman's work. An important part in both works is what is called now *Itô's formula*, but should be called Itô – Gikhman's formula, or, in the alphabetical order, Gikhman – Itô's.

If a function x_t is differentiable in t:

$$dx_t = f(t)\,dt \tag{33.1}$$

($f(t)$ being its derivative), and $F(t,\, x)$ is a smooth function of two variables, we have:

$$dF(t,\, x_t) = \frac{\partial F}{\partial t}(t,\, x_t)\,dt + \frac{\partial F}{\partial x}(t,\, x_t)\,dx_t = \frac{\partial F}{\partial t}(t,\, x_t)\,dt + \frac{\partial F}{\partial x}(t,\, x_t)\cdot f(t)\,dt. \tag{33.2}$$

But this is not true for stochastic differentials: if

$$d\xi_t = f(t,\, \omega)\,dt + g(t,\, \omega)\,dW_t, \tag{33.3}$$

then, in general,

$$\begin{aligned} dF(t,\, \xi_t) &\ne \frac{\partial F}{\partial t}(t,\, \xi_t)\,dt + \frac{\partial F}{\partial x}(t,\, \xi_t)\,d\xi_t \\ &= \frac{\partial F}{\partial t}(t,\, \xi_t)\,dt + \frac{\partial F}{\partial x}(t,\, \xi_t)\cdot f(t,\, \omega)\,dt + \frac{\partial F}{\partial x}(t,\, \xi_t)\cdot g(t,\, \omega)\,dW_t. \end{aligned} \tag{33.4}$$

Why?

Differentials, as in (33.1), (33.2), or (33.3), can be considered as approximate expressions for an increment of the corresponding function of t over a time interval of a small length dt; and the accuracy of the approximation should be taken such as is sufficient for us to be able to reconstruct the increment of our function by integrating the differentials: adding up the approximate expressions for the increments over small time intervals of length, and taking the limit. E. g., the differential relation (33.1) can be interpreted as the statement that

$$x_b - x_a = \lim_{\max_{1\le i\le n}(t_i - t_{i-1})\to 0} f(t_i^*)\cdot(t_i - t_{i-1}). \tag{33.5}$$

How do we know that the linear approximate expression (33.1) for increments in small time intervals is enough to reconstruct the increment $x_b - x_a$ from it?

Take a little more precise approximation for the increment $F(t + dt,\, x_{t+dt}) - F(t,\, x_t)$, using the Taylor expansion for the function $F(t,\, x)$ up to second-order terms rather than just the linear approximation $F(t + dt,\, x + dx) - F(t,\, x) \approx \dfrac{\partial F}{\partial t}(t,\, x)\,dt + \dfrac{\partial F}{\partial x}(t,\, x)\,dx$:

$$\begin{aligned} F(t + dt,\, x + dx) - F(t,\, x) &\approx \frac{\partial F}{\partial t}(t,\, x)\,dt + \frac{\partial F}{\partial x}(t,\, x)\,dx \\ &\quad + \frac{1}{2}\,\frac{\partial^2 F}{\partial t^2}(t,\, x)\,dt^2 + \frac{\partial^2 F}{\partial t\partial x}(t,\, x)\,dt dx + \frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(t,\, x)\,dx^2. \end{aligned} \tag{33.6}$$

The term $\dfrac{1}{2}\,\dfrac{\partial^2 F}{\partial t^2}(t,\,x)\,dt^2$ is not needed, because when integrating, we would take sums with summands multiplied by $dt^2 = (t_i - t_{i-1})^2$, and

$$\sum_{i=1}^{n}(t_i - t_{i-1})^2 \le \max_{1\le i\le n}(t_i - t_{i-1}) \cdot \sum_{i=1}^{n}(t_i - t_{i-1}) = (b-a)\cdot \max_{1\le i\le n}(t_i - t_{i-1}) \to 0 \quad (33.7)$$

as the partition $\mathfrak{T}$ becomes infinitely small.

The increment $x_{t+dt} - x_t \approx f(t)\,dt$ has the same order as dt, so we don't need the terms with $dt\,dx$ and dx^2 either.

For stochastic differentials the situation is different: not because the functions are random, but because of the size of increments $dW_t = W_{t+dt} - W_t$ of the Wiener process. Let us write the approximation (33.6) taking $d\xi_t = f(t,\,\omega)\,dt + g(t,\,\omega)\,dW_t$ as dx:

$$\begin{aligned} F(t+dt,\,\xi_{t+dt}) - F(t,\,\xi_t) &\approx \frac{\partial F}{\partial t}(t,\,\xi_t)\,dt + \frac{\partial F}{\partial x}(t,\,\xi_t)\,d\xi_t \\ &\quad + \frac{1}{2}\,\frac{\partial^2 F}{\partial t^2}(t,\,\xi_t)\,dt^2 + \frac{\partial^2 F}{\partial t\partial x}(t,\,\xi_t)\,dt\,d\xi_t + \frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(t,\,\xi_t)\,(d\xi_t)^2. \end{aligned} \quad (33.8)$$

The term with dt^2, as we know, is not needed. As for the other two second-order terms, we need to know how large $d\xi_t$ is as compared with dt and its powers.

This "stochastic differential" (being an approximate expression for $\xi_{t+dt} - \xi_t$) is the sum of two random variables: the first, $f(t,\,\omega)$, multiplied by dt, and the second, $g(t,\,\omega)$, by $dW_t = W_{t+dt} - W_t$. This last increment is a normal random variable with expectation 0 and variance dt. The variance shows the "typical size" of a random variable *quadratically*; to understand of what order "typical values" of this random variable are, we take the standard deviation $\sqrt{\mathrm{Var}(dW_t)} = (dt)^{1/2}$. So: the stochastic differential $d\xi_t$ is the sum of two terms, the first one of order dt, and the second of order $\sqrt{dt}$. Of course, the *sum* is of order of $(dt)^{1/2}$.

So the term with $\dfrac{\partial^2 F}{\partial t\partial x}$ can be disregarded, because it is of order $(dt)^{3/2}$, which vanishes after we take the sum and let the partition become infinitely small: the number of summands is of order $(dt)^{-1}$, and $(dt)^{3/2}\cdot(dt)^{-1} = (dt)^{1/2}$ is infinitely small.

A more precise expression for the same reasoning (we don't need to be *too* precise, because all this is just a preliminary investigation): almost surely

$$\begin{aligned} \sum_{i=1}^{n}(t_i - t_{i-1})\cdot|\xi_{t_i} - \xi_{t_{i-1}}| &\le \max_{1\le i\le n}|\xi_{t_i} - \xi_{t_{i-1}}|\cdot\sum_{i=1}^{\infty}(t_i - t_{i-1}) \\ &= (b-a)\cdot\max_{1\le i\le n}|\xi_{t_i} - \xi_{t_{i-1}}| \;\to\; 0 \end{aligned} \quad (33.9)$$

as $\max_{1\le i\le n}(t_i - t_{i-1}) \to 0$, because almost surely the sample function $\xi_t(\omega)$ is continuous.

Now about the term with $(d\xi_t)^2$. Its order is that of $\big((dt)^{1/2}\big)^2 = dt$, i. e., exactly the same as that of the term with $\dfrac{\partial F}{\partial t}$: this term *cannot be disregarded*!

Should we take more terms in the Taylor expansion:

$$\begin{aligned} &+\frac{1}{6}\,\frac{\partial^3 F}{\partial t^3}(t,\,\xi_t)\,dt^3+\frac{1}{2}\,\frac{\partial^3 F}{\partial t^2\partial x}(t,\,\xi_t)\,dt^2\,d\xi_t \\ &\qquad+\frac{1}{2}\,\frac{\partial^3 F}{\partial t\partial x^2}(t,\,\xi_t)\,dt(d\xi_t)^2+\frac{1}{6}\,\frac{\partial^3 F}{\partial t^3}(t,\,\xi_t)\,(d\xi_t)^3\,? \end{aligned} \tag{33.10}$$

No, because these terms are of orders $(dt)^3$, $(dt)^{5/2}$, $(dt)^2$, and $(dt)^{3/2}$, all of which vanish at integration.

So it is likely that we should open all the parentheses in (33.8), and delete all terms with $(dt)^2$ and $dt\cdot d\xi_t$:

$$\begin{aligned} dF(t,\,\xi_t)&=\frac{\partial F}{\partial t}(t,\,\xi_t)\,dt+\frac{\partial F}{\partial x}(t,\,\xi_t)\,\big(f(t,\,\omega)\,dt+g(t,\,\omega)\,dW_t\big) \\ &\qquad+\frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(t,\,\xi_t)\cdot g(t,\,\omega)^2\,(dW_t)^2. \end{aligned} \tag{33.11}$$

Taking together all terms with dt, we get:

$$\begin{aligned} dF(t,\,\xi_t)&=\big[\frac{\partial F}{\partial t}(t,\,\xi_t)+\frac{\partial F}{\partial x}(t,\,\xi_t)\cdot f(t,\,\omega)\big]\,dt+\frac{\partial F}{\partial x}(t,\,\xi_t)\cdot g(t,\,\omega)\;dW_t \\ &\qquad+\frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(t,\,\xi_t)\cdot g(t,\omega)^2\;(dW_t)^2. \end{aligned} \tag{33.12}$$

How should we understand this relation in differentials? Yes, as an expression for an (almost-sure) equality between integrals. Something like this:

$$\begin{aligned} F(t,\,\xi_t)=F(t_0,\,\xi_{t_0})&+\int_{t_0}^{t}\big[\frac{\partial F}{\partial t}(s,\,\xi_s)+\frac{\partial F}{\partial x}(s,\,\xi_s)\cdot f(s,\,\omega)\big]\,ds \\ &+\int_{t_0}^{t}\frac{\partial F}{\partial x}(s,\,\xi_s)\cdot g(s,\,\omega)\;dW_s \\ &+\int_{t_0}^{t}\frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(s,\,\xi_s)\cdot g(s,\,\omega)^2\,(dW_s)^2; \end{aligned} \tag{33.13}$$

except that we haven't introduced any integral with respect to $(dW_t)^2$.

Well, we are not averse to introducing some new types of integrals (and this would be a new type of *stochastic* integrals). But it turns out that the types of integrals we are already using: Lebesgue (or Riemann) integrals, and stochastic integrals with respect to dW_t, are enough to handle this. Keeping in mind the fact that $\underset{\max_{1\le i\le n}(t_i-t_{i-1})\to 0}{\text{l.i.m.}}\sum_{i=1}^{n}(W_{t_i}-W_{t_{i-1}})^2=b-a$, it is easy to believe that $(dW_t)^2$ should be replaced with dt.

So, Itô's formula for the differential of a smooth function of t and the value of a stochastic process ξ_t having stochastic differential $d\xi_t=f(t,\,\omega)\,dt+g(t,\,\omega)\,dW_t$ is:

$$\begin{aligned} dF(t,\,\xi_t)&=\big[\frac{\partial F}{\partial t}(t,\,\xi_t)+\frac{\partial F}{\partial x}(t,\,\xi_t)\cdot f(t,\,\omega)+\frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(t,\,\xi_t)\cdot g(t,\,\omega)^2\big]\,dt \\ &\qquad+\frac{\partial F}{\partial x}(t,\,\xi_t)\cdot g(t,\,\omega)\,dW_t. \end{aligned} \tag{33.14}$$

This equality in differentials means the following equality involving both types of integrals (holding almost surely):

$$\begin{aligned} &F(t,\, \xi_t) - F(t_0,\, \xi_{t_0}) \\ &\quad = \int_{t_0}^{t} \big[\frac{\partial F}{\partial t}(s,\, \xi_s) + \frac{\partial F}{\partial x}(s,\, \xi_s)\cdot f(s,\, \omega) + \frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(s,\, \xi_s)\cdot g(s,\, \omega)^2\big]\; ds \\ &\qquad\qquad + \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,\, \xi_s)\cdot g(s,\, \omega)\; dW_s. \end{aligned} \tag{33.15}$$

Note that we haven't *proved* this formula.

Let me formulate the result about Itô's formula (one-dimensional). I'll formulate and prove it under some restrictions that are not necessary, but make the proof easier:

Theorem 33.1. *Let* $\boldsymbol{\xi}_t$, $t \geq t_0$, *be a stochastic process with stochastic differential* (33.3) *with progressively measurable random functions* $f(t,\, \omega)$, $g(t,\, \omega)$, *with* $|f(t,\, \omega)|$, $|g(t,\, \omega)|$ *being bounded by some constant* C, *mean-square continuous except at finitely many points, and with mean-square one-sided limits at these points*; *let the function* $F(t,\, x)$ *be once continuously differentiable in* t *and twice in* x, *and suppose that the partial derivatives* $\dfrac{\partial F}{\partial t}$, $\dfrac{\partial F}{\partial x}$, $\dfrac{\partial^2 F}{\partial x^2}$ *are bounded. Then almost surely* (33.15) *holds.*

We formulated Theorem 33.1 under very restrictive conditions, requiring some functions to be bounded. If we prove the theorem under these restrictions, we can also prove it in a more general situation, with the random functions $f(t,\, \omega)$, $g(t,\, \omega)$ not being uniformly bounded, but only integrable or square integrable, and the function $F(t,\, x)$ having unbounded derivatives (only supposing that all integrals in (33.15) make sense). Indeed, we can introduce the random functions

$$f_C(t,\, \omega) = \begin{cases} f(t,\, \omega) & \text{if } \; -C \leq f(t,\, \omega) \leq C, \\ -C & \text{if } \; f(t,\, \omega) < -C, \\ C & \text{if } \; f(t,\, \omega) > C, \end{cases} \tag{33.16}$$

$$g_C(t,\, \omega) = \begin{cases} g(t,\, \omega) & \text{if } \; -C \leq g(t,\, \omega) \leq C, \\ -C & \text{if } \; g(t,\, \omega) < -C, \\ C & \text{if } \; g(t,\, \omega) > C \end{cases} \tag{33.17}$$

(make a picture). It is clear that $|f_C(t,\, \omega)| \leq |f(t,\, \omega)|$, $|g_C(t,\, \omega)| \leq |g(t,\, \omega)|$. As $C \to \infty$, we have for every t and every $\omega \in \Omega$:

$$f_C(t,\, \omega) \to f(t,\, \omega), \qquad g_C(t,\, \omega) \to g(t,\, \omega). \tag{33.18}$$

It turns out that

$$\begin{gathered} \underset{C\to\infty}{\text{l.i.m.}} \int_{t_0}^{t} g_C(s,\, \omega)\; dW_s = \int_{t_0}^{t} g(s,\, \omega)\; dW_s, \\ \underset{C\to\infty}{\text{l.i.m.}} \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,\, \xi_s)\cdot g_C(s,\, \omega)\; dW_s = \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,\, \xi_s)\cdot g(s,\, \omega)\; dW_s, \end{gathered} \tag{33.19}$$

and similar statements hold for Lebesgue integrals in (34.15) and in the equality $\xi_t - \xi_{t_0} = \int_{t_0}^{t} f(s,\, \omega)\; ds + \int_{t_0}^{t} g(s,\, \omega)\; dW_s$, which is the meaning of the equality (34.3).

To prove (33.19), we have to establish that

$$\operatorname*{l.i.m.}_{C \to \infty} g_C(t,\, \omega) = g(t,\, \omega). \tag{33.20}$$

This is not the same as (33.18), and we know that, generally, almost-sure convergence, and even convergence for all ω, does not imply convergence in the square mean. But in this concrete case we can use the dominated-convergence theorem. We have to prove that

$$\lim_{C \to \infty} E\big(g_C(t,\, \omega) - g(t,\, \omega)\big)^2 = 0. \tag{33.21}$$

The random variable under the expectation sign (the square) converges to 0 as $C \to \infty$; and these random variables are dominated, for all C:

$$\big(g_C(t,\, \omega) - g(t,\, \omega)\big)^2 \le 4g(t,\, \omega)^2, \tag{33.22}$$

the dominating random variable $4g(t,\, \omega)^2$ having finite expectation. So we have (33.21).

As for why the conditions we imposed on the derivatives $\dfrac{\partial F}{\partial t}$, $\dfrac{\partial F}{\partial x}$, $\dfrac{\partial^2 F}{\partial x^2}$ can also be eliminated:

If we "cut" the function F at the levels $-C$ and C, the result will be a non-smooth function (draw a picture, or look at your old one). So we use another way. Let $h_C(x)$ be a twice continuously differentiable function that is equal to 1 for $|x| \le C$, and to 0 for $|x| \ge C + 1$; e. g.,

$$h_C(x) = \begin{cases} 1, & |x| \le C, \\ 1 - 10(|x| - C)^3 + 15(|x| - C)^4 - 6(|x| - C)^5, & C \le |x| \le C + 1, \\ 0, & |x| \ge C + 1. \end{cases} \tag{33.23}$$

Let us take

$$F_C(t,\, x) = F(t,\, x) \cdot h_C(x). \tag{33.24}$$

These functions are smooth, and the derivatives

$$\frac{\partial F_C}{\partial t}(t,\, x) = \frac{\partial F}{\partial t}(t,\, x) \cdot h_C(x), \quad \frac{\partial F_C}{\partial x}(t,\, x) = \frac{\partial F}{\partial x}(t,\, x) \cdot h_C(x) + F(t,\, x) \cdot h'_C(x),$$
$$\frac{\partial F_C}{\partial x}(t,\, x) = \frac{\partial^2 F}{\partial x^2}(t,\, x) \cdot h_C(x) + 2\frac{\partial F}{\partial x}(t,\, x) \cdot h'_C(x) + F(t,\, x) \cdot h''_C(x) \tag{33.25}$$

are bounded and uniformly continuous; and as $C \to \infty$, the function F_C and its derivatives not only converge to F and its derivatives, but even are equal to them for sufficiently large C (for $C \ge |x|$).

Now let's go to the

Proof of Theorem 33.1. Since both Lebesgue and stochastic integrals are limits of integrals of step functions, it's enough to prove the statement for $f(t,\, \omega)$ and $g(t,\, \omega)$ in the stochastic differential being step random functions:

$$f(t,\, \omega) = \sum_i \eta_i(\omega) \cdot I_{[t_{i-1},\, t_i)}(t), \qquad g(t,\, \omega) = \sum_i \zeta_i(\omega) \cdot I_{[t_{i-1},\, t_i)}(t), \tag{33.26}$$

where the random variables η_i, ζ_i are measurable with respect to $\mathcal{F}_{t_{i-1}}$.

The left-hand side of (34.15) can be represented as

$$\begin{aligned} F(t,\, \xi_t) - F(t_0,\, \xi_{t_0}) = [F(t_1,\, \xi_{t_1}) - F(t_0,\, \xi_{t_0})] + [F(t_2,\, \xi_{t_2}) - F(t_1,\, \xi_{t_1})] + ... \\ + [F(t_{j-1},\, \xi_{t_{j-1}}) - F(t_{j-2},\, \xi_{t_{j-2}})] + [F(t,\, \xi_t) - F(t_{j-1},\, \xi_{t_{j-1}})], \end{aligned} \tag{33.27}$$

where t_{j-1} is the last point of the partition that is smaller than t. So what we need to prove is that almost surely

$$\begin{aligned} F(t_i,\, \xi_{t_i}) - F(t_{i-1},\, \xi_{t_{i-1}}) = \int_{t_{i-1}}^{t_i} \Big[\frac{\partial F}{\partial t}(s,\, \xi_s) + \frac{\partial F}{\partial x}(s,\, \xi_s) \cdot \eta_i + \frac{\partial^2 F}{\partial x^2}(s,\, \xi_s) \cdot \zeta_i^2\Big]\, ds \\ + \int_{t_{i-1}}^{t_i} \frac{\partial F}{\partial t}(s,\, \xi_s) \cdot \zeta_i\, dW_s, \end{aligned} \tag{33.28}$$

or the same with i replaced with j and t_i with t.

In general, we want to prove that if for some $t_0 \le c < d$

$$\xi_t = \xi_c + \eta \cdot (t - c) + \zeta \cdot (W_t - W_c), \qquad c \le t \le d, \tag{33.29}$$

then

$$\begin{aligned} F(d,\, \xi_d) - F(c,\, \xi_c) = \int_c^d \frac{\partial F}{\partial t}(s,\, \xi_s)\, ds + \eta \cdot \int_c^d \frac{\partial F}{\partial x}(s,\, \xi_s)\, ds \\ + \zeta \cdot \int_c^d \frac{\partial F}{\partial x}(s,\, \xi_s)\, dW_s + \frac{1}{2}\zeta^2 \cdot \int_0^d \frac{\partial^2 F}{\partial x^2}(s,\, \xi_s)\, ds. \end{aligned} \tag{33.30}$$

Let us take a partition $\mathfrak{S}$ (capital Gothic "S") of the interval from c to d with partition points $s_0 = c < s_1 < ... < s_n = d$ (note that it gets a little elaborate: the points c and d used to be some partition points of our original partition $\mathfrak{T}$; so we have a partition within a partition: $\mathfrak{S}$ within $\mathfrak{T}$). We have:

$$F(d,\, \xi_d) - F(c,\, \xi_c) = \sum_{i=1}^n [F(s_i,\, \xi_{s_i}) - F(s_{i-1},\, \xi_{s_{i-1}})]. \tag{33.31}$$

For the i-th difference we'll use a Taylor expansion based at the point $(s_{i-1},\, \xi_{s_{i-1}})$; only (33.8) was written with approximate equalities while exploring the problem, and now we are going to write precise equalities with the derivatives taken at some intermediate points:

$$\begin{aligned} F(s_i,\, \xi_{s_i}) - {}& F(s_{i-1},\, \xi_{s_{i-1}}) \\ &= [F(s_i,\, \xi_{s_i}) - F(s_{i-1},\, \xi_{s_i})] + [F(s_{i-1},\, \xi_{s_i}) - F(s_{i-1},\, \xi_{s_{i-1}})] \\ &= \frac{\partial F}{\partial t}(s_i^*,\, \xi_{s_i}) \cdot (s_i - s_{i-1}) + \frac{\partial F}{\partial x}(s_{i-1},\, \xi_{s_{i-1}}) \cdot (\xi_{s_i} - \xi_{s_{i-1}}) \\ &\qquad + \frac{1}{2}\, \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, \xi_i^*) \cdot (\xi_{s_i} - \xi_{s_{i-1}})^2, \end{aligned} \tag{33.32}$$

where s_i^* is some point between s_{i-1} and s_i, and ξ_i^* between $\xi_{s_{i-1}}$ and ξ_{s_i}. Using formula (33.29), we can replace $\xi_{s_i} - \xi_{s_{i-1}}$ with $\eta \cdot (s_i - s_{i-1}) + \zeta \cdot (W_{s_i} - W_{s_{i-1}})$. Then the right-hand side of (33.32) is expressed as the sum of 6 summands (after we open the outer parentheses in $\big(\eta \cdot (s_i - s_{i-1}) + \zeta \cdot (W_{s_i} - W_{s_{i-1}})\big)^2$). So the increment (33.31) is represented as the sum of six sums over i. We are going to prove that these six sums converge (as $\max_{1 \le i \le n}(s_i - s_{i-1}) \to 0$) to the following limits:

$$\sum_{i=1}^{n} \frac{\partial F}{\partial t}(s_i^*,\, \xi_{s_i}) \cdot (s_i - s_{i-1}) \to \int_c^d \frac{\partial F}{\partial t}(s,\, \xi_s)\; ds, \tag{33.33}$$

$$\sum_{i=1}^{n} \frac{\partial F}{\partial x}(s_{i-1},\, \xi_{s_{i-1}}) \cdot \eta \cdot (s_i - s_{i-1}) \to \int_c^d \frac{\partial F}{\partial x}(s,\, \xi_s) \cdot \eta\; ds, \tag{33.34}$$

$$\sum_{i=1}^{n} \frac{\partial F}{\partial x}(s_{i-1},\, \xi_{s_{i-1}}) \cdot \zeta \cdot (W_{s_i} - W_{s_{i-1}}) \to \int_c^d \frac{\partial F}{\partial x}(s,\, \xi_s) \cdot \zeta\; dW_s, \tag{33.35}$$

$$\sum_{i=1}^{n} \frac{1}{2}\, \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, \xi_i^*) \cdot \eta^2 \cdot (s_i - s_{i-1})^2 \to 0, \tag{33.36}$$

$$\sum_{i=1}^{n} \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, \xi_i^*) \cdot \eta\zeta \cdot (s_i - s_{i-1})(W_{s_i} - W_{s_{i-1}}) \to 0, \tag{33.37}$$

$$\sum_{i=1}^{n} \frac{1}{2}\, \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, \xi_i^*) \cdot \zeta^2 \cdot (W_{s_i} - W_{s_{i-1}})^2 \to \int_c^d \frac{1}{2}\, \frac{\partial^2 F}{\partial x^2}(s,\, \xi_s) \cdot \zeta^2\; ds. \tag{33.38}$$

Then limit passage proves (33.30).

In what probabilistic sense are we going to prove that the limits (33.33) – (33.38) take place? For some of them almost surely, for some, in the mean squares; but this is OK, because convergence in probability serves as "the common denominator" for both these types of convergence.

The sum in (33.33) is not a Riemann sum for the integral in the right-hand side; but we can do without it. The function $\dfrac{\partial F}{\partial t}$ is uniformly continuous, so for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|s - s'| < \delta$, $|x - y| < \delta \;\Rightarrow\; \bigl|\frac{\partial F}{\partial t}(s,\, x) - \frac{\partial F}{\partial t}(s',\, y)\bigr| < \varepsilon$. Let us define a step function

$$h^*_{\mathfrak{S}}(s,\, \omega) = \sum_{i=1}^{n} \frac{\partial F}{\partial t}(s^*_i,\, \xi_{s_i}) \cdot I_{[s_{i-1},\, s_i)}(s); \tag{33.39}$$

the difference of both sides of (33.33) is equal to

$$\int_c^d \bigl[h^*_{\mathfrak{S}}(s,\, \omega) - \frac{\partial F}{\partial t}(s,\, \xi_s)\bigr]\, ds. \tag{33.40}$$

For every $\omega \in \Omega$ such that the trajectory $\xi_t(\omega)$ is continuous (for every non-exceptional ω, the exceptional ones forming a set of zero probability) for the positive δ introduced above there exists a positive $\delta' \le \delta$ such that $|s - s'| < \delta' \;\Rightarrow\; |\xi_s(\omega) - \xi_{s'}(\omega)| < \delta$. For such ω, for every partition with $\max_{1 \le i \le n}(s_i - s_{i-1}) < \delta'$ we have:

$$|h^*_{\mathfrak{S}}(s,\, \omega) - \frac{\partial F}{\partial t}(s,\, \xi_s)| < \varepsilon, \qquad \Bigl|\int_c^d \bigl[h^*_{\mathfrak{S}}(s,\, \omega) - \frac{\partial F}{\partial t}(s,\, \xi_s)\bigr]\, ds\Bigr| < \varepsilon \cdot (d - c). \tag{33.41}$$

This proves (33.33), almost surely.

The same way, using the uniform continuity of $\frac{\partial F}{\partial x}$, (33.34) is proved, only the difference of the integrals has to be multiplied by $\eta(\omega)$ (which is $\le C$ in absolute value).

In (33.35), we want to establish mean-square convergence. Let us introduce

$$h_{\mathfrak{S}}(s,\, \omega) = \sum_{i=1}^{n} \frac{\partial F}{\partial x}(s_{i-1},\, \xi_{s_{i-1}}) \cdot I_{[s_{i-1},\, s_i)}(s); \tag{33.42}$$

the difference of both sides of (33.35) is equal to

$$\int_c^d \zeta \cdot \bigl[h_{\mathfrak{S}}(s,\, \omega) - \frac{\partial F}{\partial x}(s,\, \xi_s)\bigr]\, dW_s, \tag{33.43}$$

and the expectation of the square of this difference is equal to

$$\int_c^d E\bigl(\zeta^2 \cdot \bigl(h_{\mathfrak{S}}(s,\, \omega) - \frac{\partial F}{\partial x}(s,\, \xi_s)\bigr)^2\bigr)\, ds. \tag{33.44}$$

Let a positive δ be chosen as above with $\frac{\partial F}{\partial t}$ replaced by $\frac{\partial F}{\partial x}$. We have almost surely

$$\lim_{\delta' \downarrow 0}\, \max_{|s - s'| \le \delta'} |\xi_s(\omega) - \xi_{s'}(\omega)| = 0. \tag{33.45}$$

From almost-sure convergence convergence in probability follows, so there exists a positive $\delta' \le \delta$ such that the probability of the event

$$A_{\delta,\,\delta'} = \{ \max_{|s-s'|\le\delta'} |\xi_s(\omega) - \xi_{s'}(\omega)| \ge \delta \} \tag{33.46}$$

is less than ε. We have for $\max_{1\le i\le n}(s_i - s_{i-1}) < \delta'$:

$$\begin{aligned} E\big(\zeta^2 \cdot \big(h_{\mathfrak{S}}(s,\,\omega) - \frac{\partial F}{\partial x}(s,\,\xi_s)\big)^2\big) =& E\big(I_{A_{\delta,\,\delta'}} \cdot \zeta^2 \cdot \big(h_{\mathfrak{S}}(s,\,\omega) - \frac{\partial F}{\partial x}(s,\,\xi_s)\big)^2\big) \\ &+ E\big(I_{A^c_{\delta,\,\delta'}} \cdot \zeta^2 \cdot \big(h_{\mathfrak{S}}(s,\,\omega) - \frac{\partial F}{\partial x}(s,\,\xi_s)\big)^2\big) \\ \le& C^2 \big[\varepsilon \cdot 4\big(\sup\big|\frac{\partial F}{\partial x}\big|\big)^2 + \varepsilon^2\big]. \end{aligned} \tag{33.47}$$

Since this expression can be made arbitrarily small (we could have managed to take our previous estimates so that it would be less than ε exactly), we have the convergence (33.35).

Limits (33.36), (33.37) are very easy:

$$\begin{aligned} &\big|\frac{\eta^2}{2}\sum_{i=1}^{n}\frac{\partial^2 F}{\partial x^2}(s_{i-1},\,\xi_i^*) \cdot (s_i - s_{i-1})^2\big| \\ &\qquad \le \frac{C^2 \cdot \big(\sup\big|\frac{\partial^2 F}{\partial x^2}(s_{i-1},\,\xi_i^*)\big|\big)^2}{2} \cdot \max_{1\le i\le n}(s_i - s_{i-1}) \cdot \sum_{i=1}^{n}(s_i - s_{i-1}), \end{aligned} \tag{33.48}$$

which is some constant times $\max_{1\le i\le n}(s_i - s_{i-1})$, and goes to 0;

$$\begin{aligned} &\big|\eta\zeta\sum_{i=1}^{n}\frac{\partial^2 F}{\partial x^2}(s_{i-1},\,\xi_i^*) \cdot (s_i - s_{i-1})(W_{s_i} - W_{s_{i-1}})\big| \\ &\qquad \le \text{const} \cdot \max_{1\le i\le n}|W_{s_i} - W_{s_{i-1}}| \cdot \sum_{i=1}^{n}(s_i - s_{i-1}), \end{aligned} \tag{33.49}$$

and this goes to 0 almost surely (for those ω for which the trajectory $W_t(\omega)$ of the Wiener process is continuous).

The most complicated is (33.38).