Lecture note 10. Itô's formula, proof.

Let us repeat the formulation of Theorem 9.1:

Theorem 9.1. *Let X_t, $t \geq t_0$, be a stochastic process with stochastic differential*

$$dX_t = f(t,\omega)\,dt + g(t,\omega)\,dW_t, \tag{9.5}$$

with random functions $f(t,\omega)$, $g(t,\omega)$ determined by the past of the Wiener process, with $|f(t,\omega)|$, $|g(t,\omega)| \leq C < \infty$, 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 and uniformly continuous. Then almost surely

$$\begin{aligned} &F(t,X_t) - F(t_0,X_{t_0}) \\ &= \int_{t_0}^{t} \Big[\frac{\partial F}{\partial t}(s,X_s) + \frac{\partial F}{\partial x}(s,X_s)\cdot f(s,\omega) + \frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(s,X_s)\cdot g(s,\omega)^2\Big]\,ds \\ &\quad + \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,X_s)\cdot g(s,\omega)\,dW_s. \end{aligned} \tag{9.17}$$

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 square integrable, and the function $F(t,x)$ having unbounded derivatives (only supposing that all integrals in (9.17) make sense. Indeed, we can introduce the random functions

$$f_C(t,\omega) = \begin{cases} f(t,\omega) & \text{if } -C \leq f(t,\omega) \leq C, \\ -C & \text{if } f(t,\omega) < -C, \\ C & \text{if } f(t,\omega) > C, \end{cases} \tag{10.1}$$

$$g_C(t,\omega) = \begin{cases} g(t,\omega) & \text{if } -C \leq g(t,\omega) \leq C, \\ -C & \text{if } g(t,\omega) < -C, \\ C & \text{if } g(t,\omega) > C \end{cases} \tag{10.2}$$

(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 $\omega \in \Omega$:

$$f_C(t,\omega) \to f(t,\omega), \qquad g_C(t,\omega) \to g(t,\omega). \tag{10.3}$$

It turns out that

$$\begin{aligned} \underset{C\to\infty}{\text{l.i.m.}} \int_{t_0}^{t} g_C(s,\omega)\;dW_s &= \int_{t_0}^{t} g(s,\omega)\;dW_s, \\ \underset{C\to\infty}{\text{l.i.m.}} \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,X_s)\cdot g_C(s,\omega)\;dW_s &= \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,X_s)\cdot g(s,\omega)\;dW_s, \end{aligned} \tag{10.4}$$

and similar statements hold for Riemann integrals in (9.17) and in the equality $X_t - X_{t_0} = \int_{t_0}^t f(s,\omega)\,ds + \int_{t_0}^t g(s,\omega)\,dW_s$, which is the meaning of the equality (9.5).

To prove (10.4), we have to establish that

$$\underset{C\to\infty}{\text{l.i.m.}}\, g_C(t,\omega) = g(t,\omega). \tag{10.5}$$

This is not the same as (10.3), 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 (Theorem 2.2). We have to prove that

$$\lim_{C\to\infty} E\Big(\big(g_C(t,\omega) - g(t,\omega)\big)^2\Big) = 0. \tag{10.6}$$

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:

$$\big(g_C(t,\omega) - g(t,\omega)\big)^2 \le 4g(t,\omega)^2, \tag{10.7}$$

the dominating random variable $4g(t,\omega)^2$ having finite expectation. So we have (10.6).

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.,

$$h_C(x) = \begin{cases} 1, & |x| \le C, \\ 1 - 10(|x|-C)^3 + 15(|x|-C)^4 - 6(|x|-C)^5, & C \le |x| \le C+1, \\ 0, & |x| \ge C+1. \end{cases} \tag{10.8}$$

Let us take

$$F_C(t,x) = F(t,x)\cdot h_C(x). \tag{10.9}$$

These functions are smooth, and the derivatives

$$\frac{\partial F_C}{\partial t}(t,x) = \frac{\partial F}{\partial t}(t,x)\cdot h_C(x), \quad \frac{\partial F_C}{\partial x}(t,x) = \frac{\partial F}{\partial x}(t,x)\cdot h_C(x) + F(t,x)\cdot h_C'(x),$$
$$\frac{\partial F_C}{\partial x}(t,x) = \frac{\partial^2 F}{\partial x^2}(t,x)\cdot h_C(x) + 2\frac{\partial F}{\partial x}(t,x)\cdot h_C'(x) + F(t,x)\cdot h_C''(x) \tag{10.10}$$

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 9.1. Since both Riemann 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:

$$f(t,\omega) = \sum_i Y_i(\omega) \cdot I_{(t_{i-1},\, t_i]}(t), \qquad g(t,\omega) = \sum_i Z_i(\omega) \cdot I_{(t_{i-1},\, t_i]}(t), \tag{10.11}$$

where the random variables Y_i, Z_i are determined by the past of the Wiener process up to time t_{i-1}: $Y_i = Y_i(W_s, s \le t_{i-1})$, and the same for Z_i. (You can see that I dropped the custom of taking the first interval *with* its left end: $[t_0, t_1]$ rather than $(t_0, t_1]$. This is because we have established that the value of the integrand at one point does not affect the stochastic integral – and we have known this for ages about Riemann integrals.)

The left-hand side of (9.17) can be represented as

$$\begin{aligned} F(t, X_t) - F(t_0, X_{t_0}) = [F(t_1, X_{t_1}) - F(t_0, X_{t_0})] + [F(t_2, X_{t_2}) - F(t_1, X_{t_1})] + ... \\ + [F(t_{j-1}, X_{t_{j-1}}) - F(t_{j-2}, X_{t_{j-2}})] + [F(t, X_t) - F(t_{j-1}, X_{t_{j-1}})], \end{aligned} \tag{10.12}$$

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

$$\begin{aligned} F(t_i, X_{t_i}) - F(t_{i-1}, X_{t_{i-1}}) = \int_{t_{i-1}}^{t_i} [\frac{\partial F}{\partial t}(s, X_s) + \frac{\partial F}{\partial x}(s, X_s) \cdot Y_i + \frac{\partial^2 F}{\partial x^2}(s, X_s) \cdot Z_i^2]\ ds \\ + \int_{t_{i-1}}^{t_i} \frac{\partial F}{\partial t}(s, X_s) \cdot Z_i\ dW_s, \end{aligned} \tag{10.13}$$

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

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

$$X_t = X_c + Y \cdot (t - c) + Z \cdot (W_t - W_c), \qquad c \le t \le d, \tag{10.14}$$

then

$$\begin{aligned} F(d, X_d) - F(c, X_c) = \int_c^d \frac{\partial F}{\partial t}(s, X_s)\ ds + Y \cdot \int_c^d \frac{\partial F}{\partial x}(s, X_s)\ ds \\ + Z \cdot \int_c^d \frac{\partial F}{\partial x}(s, X_s)\ dW_s + \frac{1}{2} Z^2 \cdot \int_0^d \frac{\partial^2 F}{\partial x^2}(s, X_s)\ ds. \end{aligned} \tag{10.15}$$

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:

$$F(d, X_d) - F(c, X_c) = \sum_{i=1}^{n} [F(s_i, X_{s_i}) - F(s_{i-1}, X_{s_{i-1}})]. \tag{10.16}$$

For the i-th difference we'll use a Taylor expansion based at the point $(s_{i-1}, X_{s_{i-1}})$; only, in contrast with Lecture note 9, where we wrote approximate equalities exploring the problem, we are going to write precise equalities with the derivatives taken at some intermediate points:

$$\begin{aligned} F(s_i, X_{s_i}) - F(s_{i-1}, X_{s_{i-1}}) & \\ &= [F(s_i, X_{s_i}) - F(s_{i-1}, X_{s_i})] + [F(s_{i-1}, X_{s_i}) - F(s_{i-1}, X_{s_{i-1}})] \\ &= \frac{\partial F}{\partial t}(s_i^*, X_{s_i}) \cdot (s_i - s_{i-1}) + \frac{\partial F}{\partial x}(s_{i-1}, X_{s_{i-1}}) \cdot (X_{s_i} - X_{s_{i-1}}) \\ &\qquad + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*) \cdot (X_{s_i} - X_{s_{i-1}})^2, \end{aligned} \tag{10.17}$$

where s_i^* is some point between s_{i-1} and s_i, and x_i^* between $X_{s_{i-1}}$ and X_{s_i}. Using formula (10.14), we can replace $X_{s_i} - X_{s_{i-1}}$ with $Y \cdot (s_i - s_{i-1}) + Z \cdot (W_{s_i} - W_{s_{i-1}})$. Then the right-hand side of (10.17) is expressed as the sum of 6 summands (after we open the outer parentheses in $\big(Y \cdot (s_i - s_{i-1}) + Z \cdot (W_{s_i} - W_{s_{i-1}})\big)^2$). So the increment (10.16) 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:

$$\sum_{i=1}^n \frac{\partial F}{\partial t}(s_i^*, X_{s_i}) \cdot (s_i - s_{i-1}) \to \int_c^d \frac{\partial F}{\partial t}(s, X_s)\, ds, \tag{10.18}$$

$$\sum_{i=1}^n \frac{\partial F}{\partial x}(s_{i-1}, X_{s_{i-1}}) \cdot Y \cdot (s_i - s_{i-1}) \to \int_c^d \frac{\partial F}{\partial x}(s, X_s) \cdot Y\, ds, \tag{10.19}$$

$$\sum_{i=1}^n \frac{\partial F}{\partial x}(s_{i-1}, X_{s_{i-1}}) \cdot Z \cdot (W_{s_i} - W_{s_{i-1}}) \to \int_c^d \frac{\partial F}{\partial x}(s, X_s) \cdot Z\, dW_s, \tag{10.20}$$

$$\sum_{i=1}^n \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*) \cdot Y^2 \cdot (s_i - s_{i-1})^2 \to 0, \tag{10.21}$$

$$\sum_{i=1}^n \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*) \cdot YZ \cdot (s_i - s_{i-1})(W_{s_i} - W_{s_{i-1}}) \to 0, \tag{10.22}$$

$$\sum_{i=1}^n \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*) \cdot Z^2 \cdot (W_{s_i} - W_{s_{i-1}})^2 \to \int_c^d \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(s, X_s) \cdot Z^2\, ds. \tag{10.23}$$

Then the limit passage proves (10.15).

In what probabilistic sense are we going to prove that the limits (10.18) – (10.23) 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 (10.18) 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 $|s - s'| < \delta$, $|x - y| < \delta \;\Rightarrow\; \left|\dfrac{\partial F}{\partial t}(s, x) - \dfrac{\partial F}{\partial t}(s', y)\right| < \varepsilon$. Let us define a step function

$$h^*_{\mathfrak{S}}(s, \omega) = \sum_{i=1}^{n} \frac{\partial F}{\partial t}(s^*_i, X_{s_i}) \cdot I_{(s_{i-1},\, s_i]}(s); \tag{10.24}$$

the difference of both sides of (10.18) is equal to

$$\int_c^d \left[h^*_{\mathfrak{S}}(s, \omega) - \frac{\partial F}{\partial t}(s, X_s)\right]\, ds. \tag{10.25}$$

For every $\omega \in \Omega$ such that the trajectory $X_t(\omega)$ 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 $|s - s'| < \delta' \;\Rightarrow\; |X_s(\omega) - X_{s'}(\omega)| < \delta$. For such ω, for every partition with $\max_{1 \le i \le n}(s_i - s_{i-1}) < \delta'$ we have:

$$|h^*_{\mathfrak{S}}(s, \omega) - \frac{\partial F}{\partial t}(s, X_s)| < \varepsilon, \qquad \left|\int_c^d \left[h^*_{\mathfrak{S}}(s, \omega) - \frac{\partial F}{\partial t}(s, X_s)\right]\, ds\right| < \varepsilon \cdot (d - c). \tag{10.26}$$

This proves (10.18), almost surely.

The same way, using the uniform continuity of $\dfrac{\partial F}{\partial x}$, (10.19) is proved, only the difference of the integrals has to be multiplied by $Y(\omega)$ (which is $\le C$ in absolute value).

In (10.20), we want to establish mean-square convergence. Let us introduce

$$h_{\mathfrak{S}}(s, \omega) = \sum_{i=1}^{n} \frac{\partial F}{\partial x}(s_{i-1}, X_{s_{i-1}}) \cdot I_{(s_{i-1},\, s_i]}(s); \tag{10.27}$$

the difference of both sides of (10.20) is equal to

$$\int_c^d Z \cdot \left[h_{\mathfrak{S}}(s, \omega) - \frac{\partial F}{\partial x}(s, X_s)\right]\, dW_s, \tag{10.28}$$

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

$$\int_c^d E\big(Z^2 \cdot \big(h_{\mathfrak{S}}(s, \omega) - \frac{\partial F}{\partial x}(s, X_s)\big)^2\big)\, ds. \tag{10.29}$$

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

$$\lim_{\delta' \downarrow 0} \max_{|s - s'| \le \delta'} |X_s(\omega) - X_{s'}(\omega)| = 0. \tag{10.30}$$

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

$$A_{\delta,\,\delta'} = \{\max_{|s-s'|\le\delta'} |X_s(\omega) - X_{s'}(\omega)| \ge \delta\} \tag{10.31}$$

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

$$\begin{aligned} E\big(Z^2\cdot\big(h_{\mathfrak{S}}(s,\omega) - \frac{\partial F}{\partial x}(s, X_s)\big)^2\big) &= E\big(I_{A_{\delta,\,\delta'}}\cdot Z^2\cdot\big(h_{\mathfrak{S}}(s,\omega) - \frac{\partial F}{\partial x}(s, X_s)\big)^2\big) \\ &\quad + E\big(I_{A^c_{\delta,\,\delta'}}\cdot Z^2\cdot\big(h_{\mathfrak{S}}(s,\omega) - \frac{\partial F}{\partial x}(s, X_s)\big)^2\big) \\ &\le C^2\big[\varepsilon\cdot 4\big(\sup\big|\frac{\partial F}{\partial x}\big|\big)^2 + \varepsilon^2\big]. \end{aligned} \tag{10.32}$$

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 (10.20).

Limits (10.21), (10.22) are very easy:

$$\begin{aligned} &\Big|\frac{Y^2}{2}\sum_{i=1}^{n}\frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*)\cdot(s_i - s_{i-1})^2\Big| \\ &\qquad\le \frac{C^2\cdot\big(\sup\big|\frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*)\big|\big)^2}{2}\cdot\max_{1\le i\le n}(s_i - s_{i-1})\cdot\sum_{i=1}^{n}(s_i - s_{i-1}), \end{aligned} \tag{10.33}$$

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

$$\begin{aligned} &\Big|YZ\sum_{i=1}^{n}\frac{\partial^2 F}{\partial x^2}(s_{i-1},\, x_i^*)\cdot(s_i - s_{i-1})(W_{s_i} - W_{s_{i-1}})\Big| \\ &\qquad\qquad\le \text{const}\cdot\max_{1\le i\le n}|W_{s_i} - W_{s_{i-1}}|\cdot\sum_{i=1}^{n}(s_i - s_{i-1}), \end{aligned} \tag{10.34}$$

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 (10.23).

We have already considered a smaller partition $\mathfrak{S}$ within a larger, $\mathfrak{T}$ (even if they were partitions of different time intervals, $\mathfrak{S}$ of only a small interval $[c,d]$). Now we are going to consider two partitions of the same interval $[c,d]$: a larger one, $\mathfrak{U}$ (capital Gothic "U"), with partition points $u_0 = c < u_1 < u_2 < ... < u_m = d$, and a smaller one, $\mathfrak{S}$, with all partition points u_j of the larger partition included in the smaller one: $u_j = s_{i_j}$. So we have:

$$\begin{aligned} c = u_0 = s_0 < s_1 < ... < s_{i_1} = u_1 < s_{i_1+1} < ... < s_{i_2-1} < s_{i_2} = u_2 < s_{i_2+1} < ... \\ ... < s_{i_m-1} < s_{i_m} = u_m = d. \end{aligned} \tag{10.35}$$

Let us take together the summands in the sum (10.23) corresponding to the same interval from u_{j-1} to u_j. Then we can write the difference of both sides in (10.23) as

$$\sum_{j=1}^{m}\Big[\sum_{i=i_{j-1}+1}^{i_j}\frac{Z^2}{2}\cdot\frac{\partial^2 F}{\partial x^2}(s_{i-1},\,x_i^*)\cdot(W_{s_i}-W_{s_{i-1}})^2-\int_{u_{j-1}}^{u_j}\frac{Z^2}{2}\cdot\frac{\partial^2 F}{\partial x^2}(s,X_s)\;ds\Big]. \quad (10.36)$$

If we change in this formula the arguments in $\dfrac{\partial^2 F}{\partial x^2}$ to $u_{j-1},\,X_{u_{j-1}}$, we obtain

$$\sum_{j=1}^{m}\big[\sum_{i=i_{j-1}+1}^{i_j}\frac{Z^2}{2}\cdot\frac{\partial^2 F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\cdot(W_{s_i}-W_{s_{i-1}})^2-\frac{Z^2}{2}\cdot\frac{\partial^2 F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\cdot(u_j-u_{j-1})\big], \quad (10.37)$$

and the $\|\ \ \|_2$-norm of this random variable is less or equal than

$$\sum_{j=1}^{m}\frac{C^2\cdot\big(\sup\big|\frac{\partial^2 F}{\partial x^2}\big|\big)^2}{2}\cdot\sqrt{E\Big(\Big(\sum_{i=i_{j-1}+1}^{i_j}(W_{s_i}-W_{s_{i-1}})^2-(u_j-u_{j-1})\Big)^2\Big)}. \quad (10.38)$$

We have calculated the expectation under the square root sign: it is equal to $2\sum_{i=i_{j-1}+1}^{i_j}(s_i-s_{i-1})^2$ (see formula (4.16)). This sum goes to 0 as the partition $\mathfrak{S}$ becomes infinitely small, and the sum (10.37) converges to 0 in the mean square.

To get that (10.36) also converges to 0, we have to prove that the differences between the corresponding terms in (10.36) and in (10.37):

$$\sum_{j=1}^{m}\sum_{i=i_{j-1}+1}^{i_j}\frac{Z^2}{2}\cdot\big[\frac{\partial^2 F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2 F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\big]\cdot(W_{s_i}-W_{s_{i-1}})^2, \quad (10.39)$$

$$\sum_{j=1}^{m}\int_{u_{j-1}}^{u_j}\frac{Z^2}{2}\cdot\big[\frac{\partial^2 F}{\partial x^2}(s,X_s)-\frac{\partial^2 F}{\partial x^2}(u_{j-1},X_{u_{j-1}})\big]\;ds \quad (10.40)$$

converge to 0 almost surely, or in the mean squares.

The sum (10.40) is easy. We choose a $\delta>0$ so that $|s-s'|<\delta$, $|x-y|<\delta\ \Rightarrow\ \big|\frac{\partial^2 F}{\partial x^2}(s,x)-\frac{\partial^2 F}{\partial x^2}(s',y)\big|<\varepsilon$. For almost all $\omega\in\Omega$, there exists a positive $\delta'\le\delta$ such that $|s-s'|<\delta'\ \Rightarrow\ |X_s(\omega)-X_{s'}(\omega)|<\delta$; and for such ω, if $\max_{1\le i\le n}(s_i-s_{i-1})<\delta'$, we have:

$$\Big|\sum_{j=1}^{m}\int_{u_{j-1}}^{u_j}\frac{Z^2}{2}\cdot\big[\frac{\partial^2 F}{\partial x^2}(s,X_s)-\frac{\partial^2 F}{\partial x^2}(u_{j-1},X_{u_{j-1}})\big]\;ds\Big|\le\frac{C^2}{2}\cdot\varepsilon. \quad (10.41)$$

This means that the sum (10.40) converges to 0 almost surely.

For (10.39), we prove mean-square convergence. We have:

$$\Big\|\sum_{j=1}^{m}\sum_{i=i_{j-1}+1}^{i_j}\frac{Z^2}{2}\cdot\Big[\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big]\cdot(W_{s_i}-W_{s_{i-1}})^2\Big\|_2$$
$$\le\frac{C^2}{2}\sum_{j=1}^{m}\sum_{i=i_{j-1}+1}^{i_j}\Big\|\Big(\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big)\cdot(W_{s_i}-W_{s_{i-1}})^2\Big\|_2$$
$$=\frac{C^2}{2}\sum_{j=1}^{m}\sum_{i=i_{j-1}+1}^{i_j}\sqrt{E\Big(\Big(\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big)^2\Big)\cdot E\big((W_{s_i}-W_{s_{i-1}})^4\big)}.\tag{10.42}$$

By Schwarz's inequality, the expectation under the square root sign is less or equal than

$$\sqrt{E\Big(\Big(\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big)^4\Big)}\cdot\sqrt{E\big((W_{s_i}-W_{s_{i-1}})^8\big)}.\tag{10.43}$$

The expectation of the eighth power of a normal random variable with parameters $(0, s_i - s_{i-1})$

$$E\big((W_{s_i}-W_{s_{i-1}})^8\big)=1\cdot3\cdot5\cdot7\cdot(s_i-s_{i-1})^4.\tag{10.44}$$

As for $E\Big(\Big(\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big)^4\Big)$, we estimate it just as in the proof of (10.20) (see formula (10.32)): for $\max_{1\le j\le m}(u_j-u_{j-1})$ less than some positive δ' we have:

$$E\Big(\Big(\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big)^4\Big)\le\varepsilon\cdot\Big(2\sup\Big|\frac{\partial^2F}{\partial x^2}\Big|\Big)^4+\varepsilon^4.\tag{10.45}$$

So for such partitions $\mathfrak{U}$, and partitions $\mathfrak{S}$ with some partition points s_i added

$$\Big\|\sum_{j=1}^{m}\sum_{i=i_{j-1}+1}^{i_j}\frac{Z^2}{2}\cdot\Big[\frac{\partial^2F}{\partial x^2}(s_{i-1},\,x_i^*)-\frac{\partial^2F}{\partial x^2}(u_{j-1},\,X_{u_{j-1}})\Big]\cdot(W_{s_i}-W_{s_{i-1}})^2\Big\|_2$$
$$\le\frac{C^2}{2}\Big(\varepsilon\cdot\Big(2\sup\Big|\frac{\partial^2F}{\partial x^2}\Big|\Big)^4+\varepsilon^4\Big)^{1/4}\cdot(105)^{1/4}\cdot\sum_{i=1}^{n}(s_i-s_{i-1})\tag{10.46}$$
$$=\text{const}\cdot\Big(\varepsilon\cdot\Big(2\sup\Big|\frac{\partial^2F}{\partial x^2}\Big|\Big)^4+\varepsilon^4\Big)^{1/4}.$$

This proves our statement, the last one in the theorem we wanted to prove.

We have proved our Itô's formula under some restrictions; let us formulate the result under minimum restrictions (but we are not going to give the proof; its ideas are outlined at the beginning of this lecture note):

Theorem 10.1. *Let X_t, $t\ge t_0$, be a stochastic process with stochastic differential*

$$dX_t=f(t,\omega)\,dt+g(t,\omega)\,dW_t,\tag{10.47}$$

with random functions $f(t,\omega)$, $g(t,\omega)$ *determined by the past of the Wiener process,* $f(t,\omega)$ *Riemann integrable,* $g(t,\omega)$ *with finite* $\int_{t_0}^{t} E\big(g(t,\omega)^2\big)\,ds$; *let the function* $F(t,x)$ *be once continuously differentiable in* t *and twice in* x, *and suppose that*

$$\int_{t_0}^{t} E\big(\big(\frac{\partial F}{\partial x^2}(s,X_s)\cdot g(s,\omega)\big)^2\big)\,ds < \infty. \tag{10.48}$$

Then almost surely

$$\begin{aligned} F(t,X_t) &- F(t_0,X_{t_0}) \\ &= \int_{t_0}^{t} \big[\frac{\partial F}{\partial t}(s,X_s) + \frac{\partial F}{\partial x}(s,X_s)\cdot f(s,\omega) + \frac{1}{2}\,\frac{\partial^2 F}{\partial x^2}(s,X_s)\cdot g(s,\omega)^2\big]\,ds \\ &\qquad + \int_{t_0}^{t} \frac{\partial F}{\partial x}(s,X_s)\cdot g(s,\omega)\,dW_s. \end{aligned} \tag{10.49}$$