Lecture 31. Stochastic integrals of step random functions.

Now let us go to stochastic equations. In what sense should we understand the stochastic differential equation (30.6)?

An ordinary differential equation

$$d\boldsymbol{\xi}_t = \boldsymbol{b}(\boldsymbol{\xi}_t)\,dt \tag{31.1}$$

is, by definition, a statement about the derivative $\dfrac{d\boldsymbol{\xi}_t}{dt}$. From the equation (31.1) can be *deduced* the integral equation

$$\boldsymbol{\xi}_t = \boldsymbol{\xi}_{t_0} + \int_{t_0}^{t} \boldsymbol{b}(\boldsymbol{\xi}_s)\ ds. \tag{31.2}$$

The equation (31.2) turns out to be equivalent to the equation (31.1) together with the initial condition at the time point t_0.

So: the main, basic thing is the differential equation; and the integral equation is just one of its possible consequences.

In the theory of stochastic equations the situation is opposite: by definition, a random function (stochastic process) $\boldsymbol{\xi}_t$ is the solution of the stochastic differential equation

$$d\boldsymbol{\xi}_t = \boldsymbol{b}(\boldsymbol{\xi}_t)\,dt + \sigma(\boldsymbol{\xi}_t)\,d\boldsymbol{W}_t \tag{31.3}$$

if it is a solution of the stochastic *integral* equation

$$\boldsymbol{\xi}_t = \boldsymbol{\xi}_{t_0} + \int_{t_0}^{t} \boldsymbol{b}(\boldsymbol{\xi}_s)\ ds + \int_{t_0}^{t} \sigma(\boldsymbol{\xi}_s)\ d\boldsymbol{W}_s, \tag{31.4}$$

the second integral being a *stochastic* integral (about what stochastic integrals are we'll speak a little later).

So the basic thing in the theory of stochastic equations is stochastic *integral* equations; and stochastic *differential* equations are just a shorter form to express integral equations (we can see that indeed formula (31.3) is shorter than (31.4)). In addition to being shorter, this form also may be in some ways more intuitive – which is a good thing.

So now we'll follow this way: we define what a stochastic integral is; we consider examples and study properties of stochastic integrals – not forgetting to give their expressions in the form with "stochastic differentials"; then we go to stochastic integral equations – which are, by definition, the thing that is meant by stochastic differential equations.

The theory of stochastic integrals, and of stochastic integral equations (which underlie stochastic differential equations) was developed independently in approximately 1947 (published in 1948) by Kiyoshi Itô in Japan and Joseph Gikhman in the Soviet Union.

Their theories coincided in very many details, and were simple enough to gain wide popularity. But Gikhman's theory contained also some things that were deeper and more complicated; in addition, the isolation of the Soviet Union lead to the theory becoming popular under Itô's name: we speak of Itô integrals, of Itô's formula (even if Gikhman's work contained exactly the the same formula), etc. Besides that, Professor Itô came to the United States and worked at Cornell University, which also added him popularity as the author of this theory.

I think this is unfair.

So how is the stochastic integral (with respect to the one-dimensional Wiener process to begin with)

$$\int_a^b f(t,\,\omega)\;dW_t \tag{31.5}$$

of a random function $f(t,\,\omega)$ defined? (Note that here I chose not to call the random function $f(t,\,\omega)$ a *stochastic process*. Of course, this is only a question of verbal expression, and does not matter much; but it is more natural to speak of integrating a *function* than doing this with a *process*.)

Let us remember how the usual Riemann integral $\int_a^b f(t)\;dt$ of a function $f(t)$ is defined.

We take a partition $\mathfrak{T}$ of the interval $[a,\,b]$:

$$a = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = b; \tag{31.6}$$

in each of the small intervals $[t_{i-1},\,t_i]$ we choose a point t_i^*, and form a Riemann sum:

$$\sum_{i=1}^{n} f(t_i^*)\cdot(t_i - t_{i-1}). \tag{31.7}$$

Then we say that a function f is integrable if there exists a (finite) limit (which is called its Riemann integral)

$$\int_a^b f(t)\;dt = \lim_{\max_{1\le i\le n}(t_i - t_{i-1})\to 0} \sum_{i=1}^{n} f(t_i^*)\cdot(t_i - t_{i-1}). \tag{31.8}$$

Not every function $f(t)$, $t\in[a,\,b]$, is integrable; but it is proved that every continuous function is. Also every function taking finitely many values, being constant on some subintervals of the interval $[a,\,b]$, is integrable, and its integral is equal to the sum of its values multiplied by the lengths of the intervals on which they are taken.

There is also a concept that is similar to that of Riemann integral: that of the *Stieltjes integral* $\int_a^b f(t)\;dg(t)$ of a function $f(t)$, $t\in[a,\,b]$, with respect to another function $g(t)$, $a\le t\le b$:

$$\int_a^b f(t)\;dg(t) = \lim_{\max_{1\le i\le n}(t_i - t_{i-1})\to 0} \sum_{i=1}^{n} f(t_i^*)\cdot\big(g(t_i) - g(t_{i-1})\big) \tag{31.9}$$

(in the case of $g(t) = t$, it reduces to the Riemann integral).

Just as in the case of Riemann integrals, not every function $f(t)$ is integrable with respect to every other function $g(t)$. It is proved (just the same way as in the case of Riemann integrals) that every *continuous* function $f(t)$ is integrable with respect to every function $g(t)$ *of bounded variation* on the interval $[a, b]$.

We know that the trajectories of the Wiener process are not functions of bounded variation, so in all probability we cannot define the stochastic integral $\int_a^b f(t, \omega)\, dW_t(\omega)$ by taking, for every (or *almost* every) $\omega \in \Omega$, the Stieltjes integral. ("In all probability" rather than "definitely", because the condition of $f(t)$ being continuous and $g(t)$ having bounded variation is not *necessary* for the Stieltjes integral making sense, but only *sufficient*. But we'll see later that indeed this Stieltjes-integral approach is impossible for stochastic integrals with respect to the Wiener process in the general case.)

We can represent the definitions (31.8), (31.9) in another form.

First we consider *step functions* $f(t)$ having the form

$$f(t) = \sum_{i=1}^{n} c_i \cdot I_{[t_{i-1},\, t_i)}(t), \tag{31.10}$$

where, for a set A, I_A is the notation for its indicator function. We included into the interval from t_{i-1} to t_i its left end but not the right: if we had included both, the sum (31.10) would have taken the value $c_i + c_{i+1}$ at the point t_i, and no one would want this; we had to decide on something, so we opted for the left end (anyway, we know that for the Riemann integral the values of the integrand at one point, or finitely many points, don't have any significance). We could also have taken the last interval containing its right end – the function (31.10) could take a non-zero value at the point t_n; but again, this is not important for the same reason that one point, or finitely many points, don't matter.

The function (31.10) can be rewritten as

$$f(t) = c_i, \qquad t_{i-1} \le t < t_i. \tag{31.11}$$

This is a step function, taking finitely many values on finitely many subintervals of the interval from a to b (the word "step" refers to the image that you have before your eyes if you make a picture of the graph of such a function).

We can divide the definitions (31.8), (31.9) into two steps. The first step is defining the integral for step functions of the form (31.10) by

$$\int_a^b f(t)\, dg(t) = \sum_{i=1}^{n} c_i \cdot [g(t_i) - g(t_{i-1})] \tag{31.12}$$

(I am not writing this for the integral with respect to dt: it is just a particular case with $g(t) = t$). The first step is, in the case of Riemann integrals, just evaluating the areas of unions of rectangles.

The second step is approximating an arbitrary function $f(t)$ by step functions and performing a limit passage. Namely, for a partition $\mathfrak{T}$ given by (31.6), and a choice of points t_i^* in each interval $[t_{i-1},\, t_i]$ we take

$$f_{\mathfrak{T},\, t_1^*, \ldots, t_n^*}(t) = \sum_{i=1}^{n} f(t_i^*) \cdot I_{[t_{i-1},\, t_i)}(t). \tag{31.13}$$

(that is, we define the approximating function by formula (31.10) with $c_i = f(t_i^*)$).

And we define the Stieltjes integral by

$$\int_a^b f(t)\; dg(t) = \lim_{\max_{1 \le i \le n}(t_i - t_{i-1}) \to 0} \int_a^b f_{\mathfrak{T},\, t_1^*, \ldots, t_n^*}(t)\; dg(t). \tag{31.14}$$

For a continuous function $f(t)$, $t \in [a,\, b]$, the existence of the limit (31.14) is proved using the fact that

$$\lim_{\max_{1 \le i \le n}(t_i - t_{i-1}) \to 0} f_{\mathfrak{T},\, t_1^*, \ldots, t_n^*}(t) = f(t), \tag{31.15}$$

uniformly for t between a and b; that is, that

$$\sup_{a \le t \le b} |f(t) - f_{\mathfrak{T},\, t_1^*, \ldots, t_n^*}(t)| \to 0 \tag{31.16}$$

as $\max_{1 \le i \le n}(t_i - t_{i-1}) \to 0$.

We are going to implement this two-step plan for stochastic integrals of random functions.

The first part is very simple: for a step random function $f(t,\, \omega)$ that can be represented in the form

$$f(t,\, \omega) = \sum_{i=1}^{n} \eta_i(\omega) \cdot I_{[t_{i-1},\, t_i)}(t), \tag{31.17}$$

that is,

$$f(t,\, \omega) = \eta_i, \qquad t_{i-1} \le t < t_i, \tag{31.18}$$

where η_1, η_2, ..., η_n are some random variables, we take

$$\int_a^b f(t,\, \omega)\; dW_t = \sum_{i=1}^{n} \eta_i \cdot (W_{t_i} - W_{t_{i-1}}). \tag{31.19}$$

As a matter of fact, here we should stop a little and consider the possibility of the same function $f(t,\, \omega)$ being represented in the form (31.17), (31.18) in two different ways (this problem also arises when we define the Stieltjes or Riemann integral, but I did not stop at it when introducing formula (31.12) because we seem to know almost everything about Riemann integrals, and can guess at almost everything for Stieltjes integrals). Draw a picture of a step function with pretty large small intervals $[t_{i-1},\, t_i)$; these intervals can be subdivided into smaller ones for which a representation (31.17) will also take place. But putting smaller intervals together, and using the fact that

$$(W_{t_{i-1}} - W_{t_{i-2}}) + (W_{t_i} - W_{t_{i-1}}) = W_{t_i} - W_{t_{i-2}}, \tag{31.20}$$

we get that for every representation the value (31.19) of the integral is the same.

Theorem 31.1. *The stochastic integral of step random functions is* linear *with respect to the integrand*: *for a constant* c

$$\int_a^b c \cdot f(t,\,\omega)\;dW_t \;=\; c \cdot \int_a^b f(t,\,\omega)\;dW_t; \tag{31.21}$$

and for any two step random functions $f(t,\,\omega)$, $g(t,\,\omega)$

$$\int_a^b [f(t,\,\omega) + g(t,\,\omega)]\;dW_t = \int_a^b f(t,\,\omega)\;dW_t + \int_a^b g(t,\,\omega)\;dW_t. \tag{31.22}$$

Proof. As for (31.21), it is very simple: the integral in the left-hand side is just the sum (31.19) with $f(t,\,\omega)$ multiplied by a constant factor c; the points of partition $t_0 = a$, t_1, t_2, ..., $t_n = b$ are the same for the random function $c \cdot f(t,\,\omega)$ as for $f(t,\,\omega)$, and taking the factor c outside the sum, we get (31.21).

The statement (31.22) is not this simple: the partition $t_0 = a < t_1 < t_2 < ... < t_n = b$ used in the expression (31.19) for the random function $f(t,\omega)$ may be different from the partition $t'_0 = a < t'_1 < t'_2 < ... < t'_{n'} = b$ for the random function $g(t,\,\omega)$, and the partition used in the representation (31.19) for the random function $f(t,\,\omega) + g(t,\,\omega)$ may be different from both of them. Consider the following example (and draw the graphs of f, g, and $f+g$): $a = 0$, $b = 4$,

$$f(t,\,\omega) = f(t) = \begin{cases} 1, & 0 \le t < 2, \\ 2, & 2 \le t < 3, \\ 3, & 3 \le t \le 4, \end{cases} \tag{31.23}$$

$$g(t) = \begin{cases} 1, & 0 \le t < 1, \\ 2, & 1 \le t < 3, \\ 1, & 3 \le t \le 4; \end{cases} \tag{31.24}$$

$$f(t) + g(t) = \begin{cases} 2, & 0 \le t < 1, \\ 3, & 1 \le t < 2, \\ 4, & 2 \le t \le 4. \end{cases} \tag{31.25}$$

We see that the function $f(t)$ is represented in the form (31.18) with the partition points 0 (common for all partitions of the interval $[0,\,4]$), 2, 3, and 4 (also common for all partitions); the function $g(t)$, with the partition $0 < 1 < 3 < 4$; and $f(t) + g(t)$ with 0, 1, 2, 4.

But in fact we have prepared everything for this: we go in all three cases to the partition containing all partition points for all three partitions; the sums (31.19) do not change (see (31.20)), and each summand in the sum for $f+g$ is the sum of the two summands with the same number in the sums for f and for g.

Theorem 31.2. *For a step function $f(t,\omega)$ defined for t between a and c, $a < b < c$, we have:* $\int_a^c f(t,\omega)\,dW_t = \int_a^b f(t,\omega)\,dW_t + \int_b^c f(t,\omega)\,dW_t$.

Proof: Obvious.

For a step function $f(t,\omega)$ defined on the interval from a to b we can consider the integral $\int_a^t f(s,\omega)\,dW_s$ with variable upper limit $t \in [a,b]$.

Theorem 31.3. *The integral $\int_a^t f(s,\omega)\,dW_s$ is for all $\omega \in \Omega$ continuous in t.*

Proof. This is because we have $\int_a^t f(s,\omega)\,dW_s = \eta_1 \cdot (W_{t_1} - W_{t_0}) + \dots + \eta_i \cdot (W_{t_i} - W_{t_{i-1}}) + \eta_{i+1} \cdot (W_t - W_{t_i})$ for $t_i \le t \le t_{i+1}$, including both ends, and the Wiener process W_t is continuous in t.

Up to now, we did not use the fact that W_t is a Wiener process (except the continuity of its trajectories): the integral of *step* functions can be defined without it. For more general functions we need a limit passage; and here the probabilities and everything associated with them become essential.

First of all, assume that a non-decreasing family of σ-algebras $\mathcal{F}_t$, $t \ge 0$, is given. And let W_t be a Wiener process with respect to this family of σ-algebras; that is, for $0 \le t < u$ and a Borel set $C \subseteq \mathbb{R}^1$

$$P\{W_u \in C \| \mathcal{F}_t\} = P(u-t, W_t, C) = \int_C \frac{1}{\sqrt{2\pi(u-t)}}\, e^{-(y-W_t)^2/2(u-t)}\, dy. \tag{31.26}$$

It follows from (31.26) that for $0 \le t \le u_1 \le u_2 \le \dots \le u_n$ the random variables $W_{u_2} - W_{u_1}$, ..., $W_{u_n} - W_{u_{n-1}}$ are independent between them, and all of them from the σ-algebra $\mathcal{F}_t$. As the family of the σ-algebras $\mathcal{F}_t$ with respect to which a Wiener process is still a Wiener process, we can take the family of σ-algebras $\mathcal{F}_{\le t}$ generated by this process; or the σ-algebras $\mathcal{F}_{\le t+}$ with "looking into the future for an infinitely small time"; or $\sigma(\eta;\, \xi_s,\, s \le t)$, where η is a random variable that is independent from ξ_t; or, if W_t is one coordinate of a multidimensional Wiener process $\boldsymbol{W}_t$, the σ-algebras $\sigma(\boldsymbol{W}_s,\, s \le t)$.

Now we'll reduce the class of random functions that we are taking stochastic integrals of by imposing some measurability conditions.

A random function $f(t,\omega)$ defined for t in an interval I with left end a is called *progressively measurable* if for every $t \in I$ the restriction of this function $f(s,\omega)$, $a \le s \le t$, $\omega \in \Omega$, is measurable in (s,ω) with respect to the product σ-algebra $\mathcal{B}_{[a,\,t]} \times \mathcal{F}_t$.

We have already encountered this requirement occasionally; it can be formulated as measurability with respect to a certain σ-algebra in $I \times \Omega$.

By definition, $\mathcal{P}rog$ will be the class of subsets $A \subseteq I \times \Omega$ such that $A \cap ([a,\, t] \times \Omega) \in \mathcal{B}_{[a,\, t]} \times \mathcal{F}_t$ for every $t \in I$. It's easy to see that $\mathcal{P}rog$ is a σ-algebra, and a random function is progressively measurable if and only if it is measurable with respect to the σ-algebra $\mathcal{P}rog$.

Theorem 31.4. *If a random function has right-continuous sample functions* (or left-continuous) *and is adapted to our family of σ-algebras, it is progressively measurable.*

We have it proved, in fact (see Example 14.5), only at that time we hadn't introduced the term "progressively measurable".

Theorem 31.5. *A step random function $f(t,\, \omega)$ of the form* (31.17), (31.18) *is progressively measurable if and only if it is adapted to the family $(\mathcal{F}_t)$, and if and only if each random variable η_i is measurable with respect to $\mathcal{F}_{t_{i-1}}$.*

Proof. It's enough to mention that $\eta_i = f(t_{i-1},\, \omega)$ and that

$$\begin{aligned} &\{(s,\, \omega) \in [a,\, t] \times \Omega\colon\ f(s,\, \omega) \in C\} \\ &\quad = \big([t_0,\, t_1) \times \{\eta_1 \in C\}\big) \cup \ldots \cup \big([t_{i-1},\, t_i) \times \{\eta_i \in C\}\big) \cup \big([t_t,\, t] \times \{\eta_{i+1} \in C\}\big) \end{aligned} \tag{31.27}$$

for $t_i \le t < t_{i+1}$.

Note that for *non-step* random functions being adapted and being progressively measurable are, in general, different things: every progressively measurable function is adapted to our family of σ-algebras, but the reverse is not necessarily true.

Theorem 31.6. *For progressively measurable step random functions $f(t,\, \omega)$ of the form* (31.17), (31.18), *with $E\eta_i^2 < \infty$, we have, for $0 \le a < b$:*

$$E \int_a^b f(t,\, \omega)\; dW_t = 0, \tag{31.28}$$

$$E\Big(\int_a^b f(t,\, \omega)\; dW_t \Big\| \mathcal{F}_a\Big) = 0, \tag{31.29}$$

$$E\Big(\int_a^b f(t,\, \omega)\; dW_t\Big)^2 = \int_a^b E f(t,\, \omega)^2\, dt = E \int_a^b f(t,\, \omega)^2\; dt \tag{31.30}$$

(I am using the notation $f(t,\, \omega)^2$ instead of the more cumbersome $[f(t,\, \omega)]^2$ to avoid too many parentheses: my notation cannot be misunderstood in any way).

Proof. The random variables η_i are measurable with respect to $\mathcal{F}_{t_{i-1}}$. We have to prove that, if $E\eta_i^2 < \infty$,

$$E \sum_{i=1}^n \eta_i \cdot (W_{t_i} - W_{t_{i-1}}) = 0, \tag{31.31}$$

$$E\Big(\sum_{i=1}^{n} \eta_i \cdot (W_{t_i} - W_{t_{i-1}}) \Big\| \mathcal{F}_a\Big) = 0, \tag{31.32}$$

$$E\Big(\sum_{i=1}^{n} \eta_i(W_s,\, s \le t_{i-1}) \cdot (W_{t_i} - W_{t_{i-1}})\Big)^2 = \sum_{i=1}^{n} E\eta_i^2 \cdot (t_i - t_{i-1}) = E\sum_{i=1}^{n} \eta_i^2 \cdot (t_i - t_{i-1}). \tag{31.33}$$

If we prove (31.32), the formula (31.31) will also be proved (the expectation is equal to the expectation of the conditional expectation with respect to $\mathcal{F}_a$). We have (using the fact that $\mathcal{F}_a \subseteq \mathcal{F}_{t_{i-1}}$):

$$\begin{aligned} E\big(\eta_i \cdot (W_{t_i} - W_{t_{i-1}}) \big\| \mathcal{F}_a\big) &= E\Big(E\big(\eta_i \cdot (W_{t_i} - W_{t_{i-1}}) \Big\| \mathcal{F}_{t_{i-1}}\big) \Big\| \mathcal{F}_a\Big) \\ &= E\Big(\eta_i \cdot E\big((W_{t_i} - W_{t_{i-1}}) \big\| \mathcal{F}_{t_{i-1}}\big) \Big\| \mathcal{F}_a\Big) = 0. \end{aligned} \tag{31.34}$$

Now to the expectation of the square of the sum (the square of the stochastic integral). We can write:

$$\begin{aligned} &\Big(\sum_{i=1}^{n} \eta_i \cdot (W_{t_i} - W_{t_{i-1}})\Big)^2 \\ &\quad = \sum_{i=1}^{n} \eta_i^2 \cdot (W_{t_i} - W_{t_{i-1}})^2 + 2 \sum_{1 \le j < i \le n} \eta_j \cdot (W_{t_j} - W_{t_{j-1}}) \cdot \eta_i \cdot (W_{t_i} - W_{t_{i-1}}). \end{aligned} \tag{31.35}$$

The expectation of the i-th summand in the first sum is equal to

$$E\eta_i^2 \cdot E(W_{t_i} - W_{t_{i-1}})^2 = E\eta_i^2 \cdot (t_i - t_{i-1}). \tag{31.36}$$

As for the expectations of the summands in the second sum, they are at least finite. Since the expectations of the i-th and j-th summand in the first sum in (31.35) are finite, by the Schwarz inequality so is the expectation of the $(j,\, i)$-th summand in the second sum. Let us handle it.

The random variables η_j, $W_{t_j} - W_{t_{j-1}}$, η_i are $\mathcal{F}_{t_{i-1}}$-measurable, and as such are independent with the random variable $W_{t_i} - W_{t_{i-1}}$ being the increment *after* the time t_{i-1}. So we have:

$$E\big(\eta_j \cdot (W_{t_j} - W_{t_{j-1}}) \cdot \eta_i \cdot (W_{t_i} - W_{t_{i-1}})\big) = E\big(\eta_j \cdot (W_{t_j} - W_{t_{j-1}}) \cdot \eta_i\big) \cdot E(W_{t_i} - W_{t_{i-1}}) = 0. \tag{31.37}$$

So the expectation of the second sum in (31.35) is equal to 0, and formula (31.45) gives us the first equality in (31.33). (The second equality in this formula is just the statement that the expectation of the sum is equal to the sum of expectations.)

Note that in the formula (31.30) in the left-hand side we have the *expectation of the square* of an integral: of a *stochastic integral*; whereas in its right-hand side we have a usual *Riemann*, or *Lebesgue integral* of the *expectation of the square* (of the integrand). So almost everything is reversed.

Theorem 31.7. *The stochastic integral* $\int_a^b f(t, \omega)\, dW_t$ *of a progressively measurable step random function* $f(t, \omega)$ *is a random variable that is measurable with respect to the* σ*-algebra* $\mathcal{F}_b$.

Proof: Obvious, because all summands in the sum (31.19) are $\mathcal{F}_b$-measurable.

Theorem 31.8. *Let* $f(t, \omega)$ *be a progressively measurable step random function with* $E\eta_i^2 < \infty$. *Then the stochastic integral* $\eta_t = \int_a^t f(s, \omega)\, dW_s$ *is a martingale with respect to the family of* σ*-algebras* $\mathcal{F}_t$.

Proof. That η_t is adapted to our family of σ-algebras follows from Theorem 31.7; what remains to prove is that for $a \le b < c$

$$E(\eta_c - \eta_b \| \mathcal{F}_b) = 0 \tag{31.38}$$

(almost surely, of course). But this difference is equal to $\int_b^c f(s, \omega)\, dW_s$, and the statement follows from (31.29) with a changed to b, and b to c.