Lecture note 5. Definition of stochastic integrals.

The size of a lecture note of mine fluctuates because it is determined by the number of things I forgot to mention, or stress enough, in the corresponding lecture, or in the previous one.

So I am introducing the thing that was mentioned only shortly in my lectures, and was not written in Lecture note 4. The stochastic integral (of step random functions: we don't have stochastic integrals of any random functions of a different kind) is *linear* with respect to the integrand: for a constant c

$$\int_a^b c \cdot f(t, \omega)\, dW_t = c \cdot \int_a^b f(t, \omega)\, dW_t; \tag{5.1}$$

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

$$\int_a^b [f(t, \omega) + g(t, \omega)]\, dW_t = \int_a^b f(t, \omega)\, dW_t + \int_a^b g(t, \omega)\, dW_t. \tag{5.2}$$

As for (5.1), it is very simple: the integral in the left-hand side is just the sum (4.41) 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 (5.1).

The statement (5.2) is not this simple: the partition $t_0 = a < t_1 < t_2 < ... < t_n = b$ used in the expression (4.41) 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 (4.41) 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 $f + g$): $a = 0$, $b = 4$,

$$f(t, \omega) = f(t) = \begin{cases} 1, & 0 \le t \le 2, \\ 2, & 2 < t \le 3, \\ 3, & 3 < t \le 4, \end{cases} \tag{5.3}$$

$$g(t) = \begin{cases} 1, & 0 \le t \le 1, \\ 2, & 1 < t \le 3, \\ 1, & 3 < t \le 4; \end{cases} \tag{5.4}$$

$$f(t) + g(t) = \begin{cases} 2, & 0 \le t \le 1, \\ 3, & 1 < t \le 2, \\ 4, & 2 < t \le 4. \end{cases} \tag{5.5}$$

We see that the function $f(t)$ is represented in the form (4.40) 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 (4.41) do not change (see the previous lecture note), 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.

Now let us go to the statements I finished the previous lecture with.

I forgot to mention that the equalities (4.55), (4.56) hold if the expectations $E(Y_i^2)$ are finite. I am correcting this now.

We have to prove that, if $E(Y_i^2) < \infty$,

$$E\big(\sum_{i=1}^{\infty} Y_i(W_s,\, s \le t_{i-1}) \cdot (W_{t_i} - W_{t_{i-1}})\big) = 0, \tag{5.6}$$

$$E\big(\big(\sum_{i=1}^{\infty} Y_i(W_s,\, s \le t_{i-1}) \cdot (W_{t_i} - W_{t_{i-1}})\big)^2\big) = \sum_{i=1}^{n} E(Y_i^2) \cdot (t_i - t_{i-1}) = E\big(\sum_{i=1}^{n} Y_i^2 \cdot (t_i - t_{i-1})\big). \tag{5.7}$$

For every i, $1 \le i \le n$, the random variable $Y_i = Y_i(W_s,\, s \le t_{i-1})$ is independent from the increment $W_{t_i} - W_{t_{i-1}}$ of the Wiener process. This is because the collection of random variables W_s, $t_0 = a \le s \le t_{i-1}$, (which are, of course, dependent between them) is independent with the random variable $W_{t_i} - W_{t_{i-1}}$.

What does independence of random variables mean if it is about *infinitely many* random variables? By definition, it means that the independence holds for every *finite subcollection* of them. That is, in our case, we have to prove that for every finite collection of time moments $s_0 = t_0 = a < s_1 < s_2 < \dots < s_m = t_{i-1}$ (of course, we are not speaking of $i = 1$, for which $Y_1 = Y_1(W_{t_0}) = Y_1(x_0)$ is just a constant, and the independence holds in the trivial way) the collection of random variables $W_{s_0} = W_{t_0} = x_0$, W_{s_1}, W_{s_2}, ..., W_{s_m}, on one hand, and the increment $W_{t_i} - W_{t_{i-1}}$, on the other, are independent.

But the random variables W_{s_j}, $0 \le j \le m$, can be expressed as functions of the increments $W_{s_1} - W_{s_0}$, $W_{s_2} - W_{s_1}$, ..., $W_{s_m} - W_{s_{m-1}} = W_{t_{i-1}} - W_{s_{m-1}}$:

$$W_{s_0} \equiv x_0, \quad W_{s_1} = x_0 + (W_{s_1} - W_{s_0}), \quad W_{s_2} = x_0 + (W_{s_1} - W_{s_0}) + (W_{s_1} - W_{s_0}), \quad \dots,$$
$$W_{s_m} = x_0 + \sum_{j=1}^{m} (W_{s_j} - W_{s_{j-1}}). \tag{5.8}$$

The increments $W_{s_1} - W_{s_0}$, $W_{s_2} - W_{s_1}$, ..., $W_{s_m} - W_{s_{m-1}} = W_{t_{i-1}} - W_{s_{m-1}}$, $W_{t_i} - W_{t_{i-1}}$ are mutually independent by the requirement 1) in the definition of the Wiener process. The next step: the collection of random variables $W_{s_1} - W_{s_0}$, $W_{s_2} - W_{s_1}$, ..., $W_{s_m} - W_{s_{m-1}} = W_{t_{i-1}} - W_{s_{m-1}}$, on one hand, is independent from one random variable $W_{t_i} - W_{t_{i-1}}$, on the other. And from this we get our statement about the collection of random variables $W_{s_0} = W_{t_0} = x_0$, W_{s_1}, W_{s_2}, ..., W_{s_m}, expressed as the functions (5.8) of the increments $W_{s_j} - W_{s_{j-1}}$, being independent from $W_{t_i} - W_{t_{i-1}}$; and from this, the statement about infinitely many random variables W_s, $t_0 \le s \le t_{i-1}$, and one increment $W_{t_i} - W_{t_{i-1}}$ after the time moment t_{i-1}. Therefore any function of this collection W_s, $s \le t_{i-1}$, of random variables – which function we call a *functional* – is independent with $W_{t_i} - W_{t_{i-1}}$. So $Y_i = Y_i(W_s,\ s \le t_{i-1})$ is independent with $W_{t_i} - W_{t_{i-1}}$.

Now, the expectation of the product of independent random variables is equal to the product of their expectations (provided that their expectations exist, i. e., are finite); so we have:

$$E\big(Y_i \cdot (W_{t_i} - W_{t_{i-1}})\big) = E(Y_i) \cdot E(W_{t_i} - W_{t_{i-1}}) = E(Y_i) \cdot 0 = 0, \tag{5.9}$$

and the expectation (5.6) is just the sum of expectations (5.9): zero.

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

$$\begin{aligned}&\Big(\sum_{i=1}^{n} Y_i \cdot (W_{t_i} - W_{t_{i-1}})\Big)^2 \\ &\quad = \sum_{i=1}^{n} Y_i^2 \cdot (W_{t_i} - W_{t_{i-1}})^2 + 2 \sum_{1 \le j < i \le n} Y_j \cdot (W_{t_j} - W_{t_{j-1}}) \cdot Y_i \cdot (W_{t_i} - W_{t_{i-1}}).\end{aligned} \tag{5.10}$$

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

$$E(Y_i^2) \cdot E\big((W_{t_i} - W_{t_{i-1}})^2\big) = E(Y_i^2) \cdot (t_i - t_{i-1}). \tag{5.11}$$

As for the expectations of the summands in the second sum, they are at least finite. It seems that I did not have the opportunity to mention Schwarz's inequality: *If X and Y are two random variables with $E(X^2)$, $E(Y^2) < \infty$, then the product XY has an expectation, and*

$$|E(XY)| \le \sqrt{E(X^2) \cdot E(Y^2)}. \tag{5.12}$$

Since the expectations of the i-th and j-th summand in the first sum in (5.10) are finite, so is the expectation of the (j, i)-th summand in the second sum. Let us handle it.

Just as before, we see that the random variables $Y_j(W_s, s \le t_{j-1})$, $W_{t_j} - W_{t_{j-1}}$, $Y_i(W_s, s \le t_{i-1})$, all of which are determined by the increments of the Wiener process *up to* time t_{i-1}, are independent with the random variable $W_{t_i} - W_{t_{i-1}}$ being its increment *after* time t_{i-1}. So we have:

$$E\big(Y_j \cdot (W_{t_j} - W_{t_{j-1}}) \cdot Y_i \cdot (W_{t_i} - W_{t_{i-1}})\big) = E\big(Y_j \cdot (W_{t_j} - W_{t_{j-1}}) \cdot Y_i\big) \cdot E\big((W_{t_i} - W_{t_{i-1}})\big) = 0. \tag{5.13}$$

So the expectation of the second sum in (5.10) is equal to 0, and formula (5.11) gives us the first equality in (5.7). (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 (4.56) in the left-hand side we have the *expectation of the square* of an integral: a *stochastic integral*; whereas in its right-hand side we have a usual *Riemann integral* of the *expectation of the square* (of the integrand). So almost everything is reversed.

The next step in our definition of the stochastic integral is a *limit passage*. Just as in the definition of the Riemann or Stieltjes integral, we approximate the the non-step (random) integrand with step random functions, and take as the definition of the integral the limit of the integrals of the approximating step functions. The thing we have to do is proving that for a sufficiently large class of integrands, the limit exists.

But here we have to return to probability theory and speak about under what conditions limits of random variables *exist*. Perhaps this should be mentioned in Lecture 2, and included in Lecture note 1; but I wanted to get as early as possible to the Wiener process and stochastic integrals.

For sequences of *numbers*, we know the Cauchy principle: a finite $\lim_{n\to\infty} a_n$ exists if and only if

$$\lim_{\substack{n\to\infty \\ m\to\infty}} (a_n - a_m) = 0. \tag{5.14}$$

It turns out that the same principle is true for convergence of random variables – for all three of kinds of convergence that we have introduced: a finite $\lim\limits_{n\to\infty} (P) X_n$ exists if and only if

$$\lim_{\substack{n\to\infty \\ m\to\infty}} (P)\,(X_n - X_m) = 0; \tag{5.15}$$

the mean-square limit $\underset{n\to\infty}{\text{l.i.m.}}\, X_n$ exists if and only if

$$\underset{\substack{n\to\infty \\ m\to\infty}}{\text{l.i.m.}}\, (X_n - X_m) = 0; \tag{5.16}$$

and a finite limit $\lim\limits_{n\to\infty} X_n$ exists *almost surely* if and only if almost surely

$$\lim_{\substack{n\to\infty \\ m\to\infty}} (X_n - X_m) = 0. \tag{5.17}$$

The last statement is the easiest to prove. Indeed, (5.17) being satisfied almost surely means that for all ω except for those forming an event of probability 0 the double limit (5.14) for the numerical sequence $a_n = X_n(\omega)$ is equal to 0. Applying the classical Cauchy principle to this numerical sequence, we get that a finite limit of $X_n(\omega)$ exists almost surely (with exactly the same exceptional set).

The Cauchy principle in probability, with the double limit (5.15) is good, but we are not going to use it – so don't let us speak of it anymore. What we'll be using is the mean-square Cauchy principle.

I could have given it without any proof: we have already taken many things without proof. But we must prove at least something, so let us take it on.

That the existence of $\underset{n\to\infty}{\text{l.i.m.}}\, X_n = Z$ implies (5.16) is pretty simple:

$$\begin{aligned} E\big((X_n - X_m - 0)^2\big) &= E\Big(\big((X_n - Z) - (X_m - Z)\big)^2\Big) \\ &= E\big((X_n - Z)^2\big) + E\big((X_m - Z)^2\big) - 2E\big((X_n - Z)\cdot(X_m - Z)\big). \end{aligned} \tag{5.18}$$

The first two summands go to 0 as n, $m \to \infty$ because of $\operatorname*{l.i.m.}_{n\to\infty} X_n = Z$; and the subtrahend $2E\big((X_n - Z)\cdot(X_m - Z)\big)$ goes to 0 by the Schwarz inequality (5.12). The proof of the opposite implication is more complicated.

Let (5.16) hold; then the same thing is true *in probability*. We know that from a sequence of random variables converging in probability one can choose a subsequence converging to the same limit *almost surely*. The same is true for *double* sequences. So there exists a sequence n_k of natural numbers, $n_k \to \infty$ as $k \to \infty$, such that

$$\lim_{\substack{k\to\infty \\ l\to\infty}} (X_{n_k} - X_{n_l}) = 0. \tag{5.19}$$

By the Cauchy principle for almost-sure convergence, there exists a random variable Z such that

$$\lim_{k\to\infty} X_{n_k} = Z \tag{5.20}$$

almost surely.

Let us prove that the same random variable Z is the mean-square limit of the whole sequence X_n as $n \to \infty$ (but not necessarily almost surely).

We have to prove that for every positive ε there exists a natural N such that for $n \geq N$

$$E\big((X_n - Z)^2\big) \leq \varepsilon. \tag{5.21}$$

(Of course we are more accustomed to having here $< \varepsilon$; but the meaning of the statement with $\leq$ is just the same. Also if we have, say, $\leq 5\varepsilon$, this is also OK; but since I am writing now a lecture note and not giving a lecture, I think I'll be able to get $\leq \varepsilon$ rather than 5ε or something like that.

Using (5.16), let us choose N so that for n, $m \geq N$

$$E\big((X_n - X_m)^2\big) \leq \varepsilon. \tag{5.22}$$

In particular, for $n \geq N$ and sufficiently large natural k we have:

$$E\big((X_n - X_{n_k})^2\big) \leq \varepsilon. \tag{5.23}$$

For fixed n, let us apply to the sequence of nonnegative random variables $(X_n - X_{n_k})^2$, $k = 1, 2, 3, \ldots$, Fatou's Lemma (Theorem 2.3):

$$E\big(\varliminf_{k\to\infty} (X_n - X_{n_k})^2\big) \leq \varliminf_{k\to\infty} E\big((X_n - X_{n_k})^2\big). \tag{5.24}$$

The expectation under the lower limit sign in the right-hand side here is $\leq \varepsilon$, so the lower limit is $\leq \varepsilon$. In the left-hand side, the *limit* of the sequence of random variables as $k \to \infty$ exists almost surely, and is equal to $(X_n - Z)^2$. Of course, for such ω's for which it is true (and this is for all ω's except for a set of them having probability equal to 0) the lower limit is equal to the limit, and we have:

$$E\big((X_n - Z)^2\big) \leq \varepsilon \tag{5.25}$$

for $n \geq N$.

So we have proved (5.21).

The mean-square Cauchy principle is true also for functions of continuous argument:

$$\underset{t \to t_0}{\text{l.i.m.}}\, X_t \text{ exists} \quad \Leftrightarrow \quad \underset{\substack{t \to t_0 \\ t' \to t_0}}{\text{l.i.m.}}\, (X_t - X_{t'}) = 0, \tag{5.26}$$

and for all other varieties of mean-square limit. In particular, if we have a random function $Y_{\mathfrak{T}}$ of a partition $\mathfrak{T}$ of the interval $[a, b]$, we have:

$$\underset{\max_{1 \leq t \leq n}(t_i - t_{i-1}) \to 0}{\text{l.i.m.}} Y_{\mathfrak{T}} \text{ exists} \quad \Leftrightarrow \quad \underset{\substack{\max_{1 \leq t \leq n}(t_i - t_{i-1}) \to 0 \\ \max_{1 \leq t \leq n'}(t'_i - t'_{i-1}) \to 0}}{\text{l.i.m.}} (Y_{\mathfrak{T}} - Y_{\mathfrak{T}'}) = 0. \tag{5.27}$$

Now let us go to stochastic integrals of non-step random functions $f(t, \omega)$ determined by the past (of our Wiener process). For a given random function $f(t, \omega)$, $a \leq t \leq b$, and for a partition $\mathfrak{T}$ with the partitioning points $a = t_0 < t_1 < ... < t_n = b$ of the interval $[a, b]$, let us introduce the step random function

$$f_{\mathfrak{T}}(t, \omega) = f(t_0, \omega) \cdot I_{[t_0,\, t_1]} + f(t_1, \omega) \cdot I_{(t_1,\, t_2]} + ... + f(t_{n-1}, \omega) \cdot I_{(t_{n-1},\, t_n]}, \tag{5.28}$$

or, which is the same,

$$f_{\mathfrak{T}}(t, \omega) = \begin{cases} f(t_0, \omega), & t_0 \leq t \leq t_1, \\ f(t_{i-1}, \omega), & t_{i-1} < t \leq t_i, \ \ 2 \leq i \leq n. \end{cases} \tag{5.29}$$

Theorem 5.1. *If* $f(t, \omega)$, $a \leq t \leq b$, *is a random function determined by the past, with finite expectation of the square* $E\big(f(t, \omega)^2\big) < \infty$ *and mean-square continuous in the interval* $[a, b]$:

$$\lim_{s \to t} E\Big(\big(f(s, \omega) - f(t, \omega)\big)^2\Big) = 0, \qquad t \in [a, b], \tag{5.30}$$

then the mean-square limit

$$\underset{\max_{1 \leq t \leq n}(t_i - t_{i-1}) \to 0}{\text{l.i.m.}} \int_a^b f_{\mathfrak{T}}(t, \omega)\, dW_t \tag{5.31}$$

exists.

And this limit is taken as the definition of the stochastic integral of the non-step random function f:

$$\int_a^b f(t, \omega)\, dW_t = \underset{\max_{1 \leq t \leq n}(t_i - t_{i-1}) \to 0}{\text{l.i.m.}} \int_a^b f_{\mathfrak{T}}(t, \omega)\, dW_t. \tag{5.32}$$

Proof. In order to prove the existence of this limit we have to check (5.27) for $Y_{\mathfrak{T}} = \displaystyle\int_a^b f_{\mathfrak{T}}(t, \omega)\, dW_t$:

$$\lim_{\substack{\max_{1 \leq t \leq n}(t_i - t_{i-1}) \to 0 \\ \max_{1 \leq t \leq n'}(t'_i - t'_{i-1}) \to 0}} E\bigg(\Big(\int_a^b f_{\mathfrak{T}}(t, \omega)\, dW_t - \int_a^b f_{\mathfrak{T}'}(t, \omega)\, dW_t\Big)^2\bigg) = 0. \tag{5.33}$$

We have prepared formulas (5.1), (5.2), by which the random variable under the expectation sign is equal to

$$\Big(\int_a^b [f_{\mathfrak{T}}(t,\omega) - f_{\mathfrak{T}'}(t,\omega)]\ dW_t\Big)^2; \tag{5.34}$$

by formula (4.56), which holds for step random functions determined by the past, we have:

$$E\Big(\big(\int_a^b f_{\mathfrak{T}}(t,\omega)\ dW_t - \int_a^b f_{\mathfrak{T}'}(t,\omega)\ dW_t\big)^2\Big) = \int_a^b E\big(\big(f_{\mathfrak{T}}(t,\omega) - f_{\mathfrak{T}'}(t,\omega)\big)^2\big)\ dt. \tag{5.35}$$

We have to prove that this integral converges to 0 as both partitions $\mathfrak{T}$ and $\mathfrak{T}'$ become infinitely small: as $\max_{1\le t\le n}(t_i - t_{i-1}) \to 0$, $\max_{1\le t\le n'}(t'_i - t'_{i-1}) \to 0$. This means that for every positive ε there exists a positive δ such that if for both partitions $\max_{1\le t\le n}(t_i - t_{i-1}) < \delta$, $\max_{1\le t\le n'}(t'_i - t'_{i-1}) < \delta$, we have

$$\int_a^b E\big(\big(f_{\mathfrak{T}}(t,\omega) - f_{\mathfrak{T}'}(t,\omega)\big)^2\big)\ dt < \varepsilon. \tag{5.36}$$

If we prove that this expectation is less than some 5ε, it would be also OK; but I'll try to make it less than ε.

We know that every continuous number-valued function $f(t)$, $a \le t \le b$, is automatically *uniformly* continuous on this interval: i. e., for every positive ε there exists a positive δ such that for t, $s \in [a,b]$

$$|t-s| < \delta \ \Rightarrow\ |f(t) - f(s)| < \varepsilon. \tag{5.37}$$

In exactly the same way it is proved that every mean-square continuous random function $f(t,\omega)$, $t \in [a,b]$, is automatically uniformly mean-square continuous on the same interval: for every positive ε there exists a positive δ such that for t, $s \in [a,b]$

$$|t-s| < \delta \ \Rightarrow\ E\big(\big(f(t,\omega) - f(s,\omega)\big)^2\big) < \varepsilon. \tag{5.38}$$

Instead of ε we can take here any other positive number; so let $\delta > 0$ be such that

$$|t-s| < \delta \ \Rightarrow\ E\big(\big(f(t,\omega) - f(s,\omega)\big)^2\big) < \frac{\varepsilon}{b-a}. \tag{5.39}$$

Let $\max_{1\le t\le n}(t_i - t_{i-1}) < \delta$, $\max_{1\le t\le n'}(t'_i - t'_{i-1}) < \delta$. For every point $t \in [a,b]$, let t_{i-1} be the left end of the small interval of partition $\mathfrak{T}$ that contains t, and $t'_{i'-1}$ the left end of the corresponding interval of partition $\mathfrak{T}'$. Both these points are to the left of the time point t (or they coincide with t if $t = a$); so $|t_{i-1} - t'_{i'-1}| < \delta$, and

$$E\big(\big(f_{\mathfrak{T}}(t,\omega) - f_{\mathfrak{T}'}(t,\omega)\big)^2\big) = E\big(\big(f(t_{i-1},\omega) - f(t'_{i'-1},\omega)\big)^2\big) < \frac{\varepsilon}{b-a}, \tag{5.40}$$

and

$$\int_a^b E\big(\big(f_{\mathfrak{T}}(t,\omega) - f_{\mathfrak{T}'}(t,\omega)\big)^2\big)\ dt < \frac{\varepsilon}{b-a}\cdot(b-a) = \varepsilon. \tag{5.41}$$

The inequality (5.36) for sufficiently small partitions $\mathfrak{T}$, $\mathfrak{T}'$ is proved, and with it, Theorem 5.1.