**Lecture 18. Proof of Theorem 17.6. Other forms of the strong Markov property.**

**Proof** of Theorem 17.6. It's enough to prove that for every $f \in \mathbf{C}$ and for every $(\mathcal{F}_{\le t^+})$-stopping time $\tau$

$$E\big(f(\xi_{\tau+s})\big|\mathcal{F}_{\le s^+}\big) = P^s f(\xi_\tau), \tag{18.1}$$

where we replace $f(\xi_\infty)$, $P^t f(\xi_\infty)$ with 0.

This means two things: the simpler one, that the right-hand side is $\mathcal{F}_{\le\tau^+}$-measurable; the more complicated one, that for every $A \in \mathcal{F}_{\le\tau^+}$

$$E\big(I_A \cdot f(\xi_{\tau+s})\big) = E\big(I_A \cdot P^s f(\xi_\tau)\big). \tag{18.2}$$

Let us introduce new stopping times, discrete ones: for every $h > 0$ take

$$\tau^h = ih \quad \text{for } (i-1)h \le \tau < ih, \qquad \tau^h = \infty \quad \text{if } \tau = \infty. \tag{18.3}$$

We have:

$$\{\tau^h \le t\} = \bigcup_{i:\, ih \le t} \{(i-1)h \le \tau < ih\} = \bigcup_{i:\, ih\le t} [\{\tau < ih\} \setminus \{\tau < (i-1)h\}]. \tag{18.4}$$

Since $\tau$ is a $(\mathcal{F}_{\le t^+})$-stopping time, the event $\{\tau < ih\}$ in the right-hand side belongs to $\mathcal{F}_{\le ih}$, and $\{\tau < (i-1)h\}$ belongs to $\mathcal{F}_{\le (i+1)h}$. All these events belong to $\mathcal{F}_{\le t}$, and $\{\tau^h \le t\} \in \mathcal{F}_{\le t}$. So the random variables $\tau^h$ are stopping times *with respect to the non-extended* $\sigma$-algebras $\mathcal{F}_{\le t}$.

Our next move: the random variable $\xi_{\tau^h}$ is $\mathcal{F}_{\le\tau^h}$-measurable. This means that for every $C \in \mathcal{X} = \mathcal{B}_X$ and every $t \in [0, \infty)$

$$\{\xi_{\tau^h} \in C\} \cap \{\tau^h \le t\} \in \mathcal{F}_{\le t}. \tag{18.5}$$

We have:

$$\begin{aligned} \{\xi_{\tau^h} \in C\} \cap \{\tau^h \le t\} &= \bigcup_{i:\, ih\le t} \{\xi_{\tau^h} \in C\} \cap \{\tau^h = ih\} \\ &= \bigcup_{i:\, ih \le t} \{\xi_{ih} \in C\} \cap [\{\tau < ih\} \setminus \{\tau < (i-1)h\}]. \end{aligned} \tag{18.6}$$

All events here belong to $\mathcal{F}_{\le t}$, which proves (18.5).

And the next move: $\tau^h \to \tau$ as $h \to 0^+$, $\tau^h > \tau$ if $\tau < \infty$, so by the right continuity $\xi_\tau = \lim_{h\to 0^+} \xi_{\tau^h}$ (for $\tau < \infty$). For every $\delta > 0$ the random variables $\tau^h$, $0 < h \le \delta$, are measurable with respect to $\mathcal{F}_{\le\tau+\delta}$, so their limit is also measurable with respect to this

$\sigma$-algebra. Since $\delta$ is an arbitrary positive number, we get that the random variable $\xi_\tau$ is $\mathcal{F}_{\le\tau+}$-measurable.

Now to (18.2). For $f \in \mathbf{C}$, $P^s f \in \mathbf{C}$ we have: $f(\xi_{\tau+s}) = \lim_{h\to 0+} f(\xi_{\tau^h+s})$, $P^s f(\xi_{\tau+s}) = \lim_{h\to 0+} P^s f(\xi_{\tau^h+s})$; so it's enough to prove

$$E\big(I_A \cdot f(\xi_{\tau^h+s})\big) = E\big(I_A \cdot P^s f(\xi_{\tau^h})\big). \tag{18.7}$$

But this follows from Theorem 17.5.

In fact, we repeated, with some additions, the proof of Theorem 14.4.

A particular case of Theorem 17.6: all our continuous-time Markov chains are strong Markov (the space $\mathbf{C}$ consists of *all* bounded functions, so $P^t\mathbf{C} \subseteq \mathbf{C}$).

Why is the process $\tilde{\xi}_t$ of Example 17.2 not strong Markov? Its trajectories are continuous, so the only thing is that the transition probabilities (17.15) don't satisfy the Feller property: indeed, $\tilde{P}(t, x, \bullet)$ does not depend on $x$ in a weakly continuous way: the normal distributions $\tilde{P}(t, x, \bullet)$ do not converge to $\tilde{P}(t, 0, \bullet) = \delta_0(\bullet)$ as $x \to 0$, but rather to the normal distribution with parameters $(0, t)$.

Other forms of the strong Markov property.

Just as from the ordinary Markov property we deduced

$$P(\theta_t^{-1} A \| \mathcal{F}_t) = P_{\xi_t}(A) \tag{18.8}$$

first for $A = \{(\xi_{s_1}, ..., \xi_{s_n}) \in C_n\}$, and then for general $A \in \mathcal{F}_{\ge 0}$, we can deduce other forms of the strong Markov property. But first we need to consider the shift operators $\theta_\tau$ and $\theta_\tau^{-1}$.

The operator $\theta_\tau$ is defined as taking an $\omega \in \Omega$ to $\theta_{\tau(\omega)}\omega$. But there is no such thing as a shift by an infinite space $\theta_\infty \omega$; so $\theta_{\tau(\omega)}\omega$ is *not defined* for $\omega$'s with $\tau(\omega) = \infty$. So the operator $\theta_\tau$ is defined on the set $\{\tau < \infty\}$ only: $\theta_\tau : \{\tau < \infty\} \mapsto \Omega$ (not necessarily *onto* $\Omega$). The inverse image $\theta_\tau^{-1} A$ of every subset $A \subset \Omega$ is a subset of the event $\{\tau < \infty\}$.

So we formulate the new form of the strong Markov property:

$$P(\theta_\tau^{-1} A \| \mathcal{F}_\tau) = P_{\xi_\tau}(A), \tag{18.9}$$

where $P_{\xi_\tau}(A)$ is replaced with 0 if $\tau = \infty$ (and $\xi_\tau$ makes no sense). It is proved in exactly the same way as the similar form of the ordinary Markov property, for $A = \{(\xi_{s_1}, ..., \xi_{s_n}) \in C_n\}$ first, and for arbitrary $A \in \mathcal{F}_{\ge 0}$ after that.

**Example 18.1.** For continuous-time Markov chains, taking $\tau = \tau_1$ and $A = \{\tau_1 \in C\}$, $C \in \mathcal{B}_{(0,\infty)}$, we have: $\theta_{\tau_1}^{-1} A = \{\tau_2 - \tau_1 \in C\}$; and the strong Markov property yields

$$P\{\tau_2 - \tau_1 \in C \| \mathcal{F}_{\tau_1}\} = P_{\xi_{\tau_1}}\{\tau_1 \in C\} = \int_C v_{\xi_{\tau_1}} e^{-v_{\xi_{\tau_1}} s}\, ds: \tag{18.10}$$

the conditional distribution is exponential with parameter $v_{\xi_{\tau_1}}$. In particular,

$$\begin{aligned} P\{\tau_2 - \tau_1 \in C \| \tau_1, \xi_{\tau_1}\} &= E\big(P\{\tau_2 - \tau_1 \in C \| \mathcal{F}_{\le\tau_1}\} \big\| \tau_1, \xi_{\tau_1}\big) \\ &= E(P_{\xi_{\tau_1}}\{\tau_1 \in C\} \| \tau_1, \xi_{\tau_1}) = P_{\xi_{\tau_1}}\{\tau_1 \in C\}. \end{aligned} \tag{18.11}$$

We have already got it in Theorem 9.2; in fact its proof and that of Theorem 9.1 was using the strong Markov property for the discrete Markov chain $\xi_0$, $\xi_{1/2^n}$, $\xi_{2/2^n}$, ..., and then the limit passage to go from the discrete stopping times $\tau_1^n$ to the continuous $\tau_1$ – just the same as in the proof of Theorem 17.6.

**Example 18.2.** Not only the Wiener process starting from 0 at time 0 changes its sign infinitely many times in any small interval of time from 0 to $\delta$, but this also occurs almost surely after the first time $\tau_0 = \min\{t\colon \xi_t = 0\}$ of reaching 0.

Another form of the strong Markov property:

Let $u \in T$; let $\tau$ be a stopping time. Then for $C \in \mathcal{X}$ almost surely on the set $\{\tau \le u\}$

$$P\{\xi_u \in C \| \mathcal{F}_\tau\} = P(u - \tau,\, \xi_\tau,\, C) \tag{18.12}$$

(on the complement $\{\tau > u\}$ of this set the left-hand-side conditional probability is obviously equal to the random variable $I_{\{\xi_u \in C\}} = I_C(\xi_u)$).

The equality (18.12) can be deduced from (17.12), (18.9); but the proof could be rather long and complicated, so better I formulate the following couple of theorems:

**Theorem 18.1.** *Let $\xi_t$, $t \in T = \{0, h, 2h, ..., nh, ...\}$, be a Markov process; $\tau$ a stopping time for it, and $u \in T$. Then* (18.12) *is satisfied almost surely on the event $\{\tau \le u\}$.*

The **proof** is the standard one: we subdivide the probabilities and expectations involved into their parts on the events $\{\tau = ih\}$, and apply to each summand the ordinary Markov property with respect to the non-random time $ih$.

Before I formulate the next theorem, let me tell you this. If the process $\xi_t$, $t \ge 0$, has right-continuous trajectories, the function $P^t f(x)$, $f \in \mathbf{C}$, is right-continuous in $t$ for every $x$: $\lim_{t' \to t+} P^{t'} f(x) = \lim_{t' \to t+} E_x f(\xi_{t'}) = E_x \lim_{t' \to t+} f(\xi_{t'}) = E_x f(\xi_t) = P^t f(x)$. If $P^t \mathbf{C} \subseteq \mathbf{C}$, the function $P^t f(x)$ is continuous in $x$ for every $t$. But continuity in $x$ for every $t$, plus (right) continuity in $t$ for every $x$ is not the same as continuity in the pair $(t, x)$ (from the right for the argument $t$). I did not take this into account when I formulated Theorem 18.2 in the lecture. I am correcting this in the lecture note.

**Theorem 18.2.** *Let $\xi_t, t \ge 0$, be a Markov process on a metric space $X$ having right-continuous trajectories and such that $\lim_{s \to t+,\, y \to x} P^s f(y) = P^t f(x)$; let $\tau$ be a stopping time with respect to the $\sigma$-algebras $\mathcal{F}_{\le t+}$, and $u \in [0, \infty)$. Then* (18.12) *is satisfied almost surely on the event $\{\tau \le u\}$ with $\mathcal{F}_{\le \tau+}$ as $\mathcal{F}_\tau$.*

**Proof.** We reformulate the statement in terms of expectations:

$$E\big(f(\xi_u) \big\| \mathcal{F}_{\le \tau+}\big) = P^{u-\tau} f(\xi_\tau) \tag{18.13}$$

on $\{\tau \le u\}$ for every $f \in \mathbf{C}$. This means $\mathcal{F}_{\le \tau+}$-measurability of the right-hand side, plus

$$E\big(I_B \cdot f(\xi_u)\big) = E\big(I_B \cdot P^{u-\tau} f(\xi_\tau)\big) \tag{18.14}$$

for every event $B \in \mathcal{F}_{\le \tau+}$, $B \subseteq \{\tau \le u\}$.

Then we define $\tau^h$ by (18.3), and $u^h = (i+1)h$. For every $h > 0$ the random variable $\tau^h$ is a stopping time with respect to the non-extended $\sigma$-algebras $\mathcal{F}_{\leq t}$; and the event $B \in \mathcal{F}_{\leq \tau^h}$. By Theorem 18.1,

$$E\big(f(\xi_{u^h})\big\|\mathcal{F}_{\leq \tau^h}\big) = P^{\,u^h - \tau^h} f(\xi_{\tau^h}) \tag{18.15}$$

on the event $\{\tau^h \leq u^h\}$. The event $B \subseteq \{\tau \leq u\} \subseteq \{\tau^h \leq u^h\}$, so we have:

$$E\big(I_B \cdot f(\xi_{u^h})\big) = E\big(I_B \cdot P^{u^h - \tau^h} f(\xi_{\tau^h})\big). \tag{18.16}$$

Then we take $h \to 0+$. We have $\tau^h \to h$, $\tau^h > \tau$, $\xi_{\tau^h} \to \xi_\tau$, $u^h \to u$, $u^h > u$, $u^h - \tau^h \to u - \tau$, $u^h - \tau^h > u - \tau$ (this is why we added the additional $h$ defining $u^h$). We have $\lim_{h \to 0+} P^{u^h - \tau^h} f(\xi_{\tau^h}) = P^{u-\tau} f(\xi_\tau)$; and by the dominated-convergence theorem we get (18.14).

The semigroup $P^t$ associated with the Wiener process satisfies the condition of $(t, x)$-continuity of $P^t f(x)$ for every $f \in \mathbf{C}$, so this form of the strong Markov property can be applied to it. The example will be considered in the next lecture; in particular, the distribution of the first time $\tau_0$ of reaching 0 will be found.