Lecture 23. Infinitesimal operators.

Example 23.1. For a bounded linear operator A in a Banach space $\mathbf{E}$ we can define the exponential function of it, e^A, by

$$e^A = I + \frac{1}{1!}\,A + \frac{1}{2!}\,A^2 + ... + \frac{1}{n!}\,A^n + ..., \tag{23.1}$$

or, showing how this operator acts on $f \in \mathbf{E}$:

$$e^A f = f + \frac{1}{1!}\,Af + \frac{1}{2!}\,A^2 f + ... + \frac{1}{n!}\,A^n f + ... \tag{23.2}$$

The series converges because the series of norms $\|f\| + \frac{1}{1!}\,\|Af\| + \frac{1}{2!}\,\|A^2 f\| + ... + \frac{1}{n!}\,\|A^n f\|$ $+ ... \le \sum_{k=0}^{\infty} \|A\|^k \|f\|/k!$ converges (its sum being $e^{\|A\|}\|f\|$; in a Banach space, convergence of the series of norms is sufficient for convergence of the series of elements of the space).

The main property of the exponential function for numbers is $e^{a+b} = e^a \cdot e^b$; is it true for operator exponents: $e^{A+B} = e^A \cdot e^B$? It is easily proved that this is true if A and B commute: $A \cdot B = B \cdot A$ (we open the parentheses in the series $\sum_{k=0}^{\infty} \frac{1}{k!}\,(A+B)^k$, and instead of, say, $(A+B)^2 = A^2 + A\cdot B + B\cdot A + B^2$ take $A^2 + 2A\cdot B + B^2$).

The family of operators $P^t = e^{tA}$, $t \ge 0$, clearly forms a semigroup (in fact, the same formula with $t \in (-\infty, \infty)$ defines a whole *group* of operators; but we don't need it now); its infinitesimal operator has the domain being the whole space $\mathbf{E}$, and it is equal to A: for every $f \in \mathbf{E}$

$$\|h^{-1}(P^h f - f) - Af\| = \Big\|\sum_{k=2}^{\infty} \frac{h^{k-1}}{k!}\,A^k f\Big\| \le \sum_{k=2}^{\infty} \frac{h^{k-1}}{k!}\,\|A\|^k \|f\| \to 0 \quad (h \to 0^+). \tag{23.3}$$

This example is not one about Markov processes yet: we don't have the properties $P^t\mathbf{1} = \mathbf{1}$, $f \ge 0 \;\Rightarrow\; P^t f \ge 0$ yet (and even not the property $\|P^t\| \le 1$). But in an *arbitrary* Banach space $\mathbf{E}$ there may be no element $\mathbf{1}$ (which is our notation for the function being identically equal to 1), and we may not be able to speak of nonnegative elements or inequalities. Let us go to the space $\mathbf{B} = \mathbf{B}(X, \mathcal{X})$ of bounded measurable functions on a measurable space $(X, \mathcal{X})$.

Let $A(x, C)$, $x \in X$, $C \in \mathcal{X}$, be a number-valued function such that

- for every $x \in X$, $A(x, \bullet)$ is a countably additive function (a signed measure) on the σ-algebra $\mathcal{X}$ such that

$$A(x, C) \ge 0 \text{ for } C \subseteq X \setminus \{x\}, \quad A(x, \{x\}) = -v(x) \le 0, \quad A(x, X) = 0, \tag{23.4}$$

$$A(x, X \setminus \{x\}) \le a < \infty \text{ for all } x \in X; \tag{23.5}$$

- for every $C \in \mathcal{X}$ the function $A(\bullet,\, C)$ is $\mathcal{X}$-measurable in x.

Define the linear operator $A\colon \mathbf{B} \mapsto \mathbf{B}$ by

$$Af(x) = \int_X f(y)\; A(x,\, dy). \tag{23.6}$$

The integral with respect to a signed measure can be defined as one over the set where the signed measure is nonnegative, plus one over the set where it is nonpositive:

$$Af(x) = \int_{X \setminus \{x\}} f(y)\; A(x,\, dy) - f(x) \cdot v(x); \tag{23.7}$$

or, because $v(x) = A(x,\, X \setminus \{x\})$:

$$Af(x) = \int [f(y) - f(x)]\; A(x,\, dy), \tag{23.8}$$

where the integral can be taken either over the whole X, or only over the set $X \setminus \{x\}$ on which $A(x,\, \bullet)$ is a *measure.*

Formula (23.6) is quite similar to the matrix-vector formula $A\boldsymbol{f}$ that we had in the case of continuous-time Markov chains.

The operator A is bounded, with $\|A\| \le 2a$; so we can define the semigroup $P^t = e^{tA}$. Let us check that $P^t\mathbf{1} = \mathbf{1}$ first: by formula (23.8) $A\mathbf{1} = 0$, $A^k\mathbf{1} = 0$ for $k \ge 1$, and

$$e^{tA}\mathbf{1} = \mathbf{1} + \sum_{k=1}^{\infty} \frac{1}{k!}\, A^k\mathbf{1} = \mathbf{1}. \tag{23.9}$$

Now let us prove $f \ge 0 \;\Rightarrow\; P^t f \ge 0$. This is not clear from formula (23.2); but we can write:

$$P^t f = e^{tA} f = e^{t(A+aI)} \cdot e^{-atI} f = (e^{-at} \cdot I) \cdot e^{t(A+aI)} f = e^{-at} \cdot e^{t(A+aI)} f. \tag{23.10}$$

The operator $A + aI$ can be written as

$$(A + aI)f(x) = \int_X f(y)\; \tilde{A}(x,\, dy), \tag{23.11}$$

where

$$\tilde{A}(x,\, C) = A(x,\, C) + a \cdot \delta_x(C) \tag{23.12}$$

is a *measure* (not just a *signed* measure): its value at the one-point set $\{x\}$ is equal to $-v(x) + a \ge 0$.

The operator $A + aI$ takes nonnegative functions to nonnegative, so do all its powers, and so does the operator P^t being defined as an infinite sum of these powers with nonnegative coefficients.

Finally, since $\tilde{A}(x,\, X) = a$, the norm $\|A + a \cdot I\| = a$, and $\|P^t\| = e^{-at} \cdot \|e^{t(A+aI)}\| \le e^{-at} \cdot \sum_{k=0}^{\infty} t^k \|A + a \cdot I\|^k / k! = e^{-at} \cdot e^{at} = 1$ (in fact, because of $P^t\mathbf{1} = \mathbf{1}$ this norm is *equal* to 1).

Does there exist a Markov process corresponding to the semigroup P^t? We have to handle transition probabilities first: is $P^t f(x) = \int_X f(y)\ P(t,\, x,\, dy)$?

Let us define, successively: $\tilde{A}^{(0)}(x,\, C) = \delta_x(C)$,

$$\tilde{A}^{(k)}(x,\, C) = \int_X \tilde{A}^{(k-1)}(x,\, dy)\ \tilde{A}(y,\, C) \tag{23.13}$$

($\tilde{A}^{(1)}(x,\, C) = \tilde{A}(x,\, C)$). Clearly

$$(A + aI)^k f(x) = \int_X \tilde{A}^{(k)}(x,\, dy)\ f(y). \tag{23.14}$$

By formulas (23.10), (23.12) we have:

$$P^t f(x) = e^{-at} \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_X f(y)\ \tilde{A}^{(k)}(x,\, dy) = \int_X f(y)\ P(t,\, x,\, dy), \tag{23.15}$$

where

$$P(t,\, x,\, C) = e^{-at} \sum_{k=0}^{\infty} \frac{t^k}{k!}\, \tilde{A}^{(k)}(x,\, C). \tag{23.16}$$

Here we are with a transition function: it is a measure as a function of C, measurable in x, it is a *probability* measure because of $P^t \mathbf{1} = \mathbf{1}$, we have $P(0,\, x,\, C) = \delta_x(C)$, and the Chapman – Kolmogorov equation is satisfied because $P(t,\, x,\, C)$ corresponds to a *semigroup*.

It turns out that the transition function $P(t,\, x,\, C)$ satisfies the condition of uniform continuity in probability – with the metric $\operatorname{dist}(x,\, y) = 1$ for $y = x$ (the metric corresponding to the discrete topology). Indeed, by formula (23.16)

$$P(h,\, x,\, X \setminus \{x\}) \le e^{-ah} \cdot \sum_{k=1}^{\infty} h^k a^k / k! = e^{-ah} \cdot (e^{ah} - 1) = 1 - e^{-ah} \to 0 \qquad (h \to 0^+). \tag{23.17}$$

By Theorem 21.1, there exists a Markov process with the transition function (23.16) (and with infinitesimal operator A) having trajectories that are right-continuous in the discrete topology and has left limits at all time points. This means that the trajectories are as follows: the process starts at a point ξ_0, spends there some time, and jumps to another point (at the time of the jump it is already at the new point); spends some time there and jumps to another point; etc. In every finite time interval we have only a finite number of jumps.

Such Markov processes are called *pure-jump* processes; their infinitesimal operators are integral operators.

It can be proved that the time spent at a state $\xi_0 = x$ is exponential with parameter $v(x)$; and all the rest of what we proved for continuous-time-discrete-space Markov chains.

Example 23.2. The space $X = (-\infty, \infty)$, the process ξ_t is uniform motion to the right at speed 1: $\xi_t = \xi_0 + t$. There is no randomness in this process (if we start from a non-random point); we call such a degenerate stochastic process *deterministic*.

But ξ_t *is* a Markov process with the transition function

$$P(t, x, C) = \delta_{x+t}(C). \tag{23.18}$$

Let us find its infinitesimal operator A.

The operators P^t on the space $\mathbf{B}$ are given by

$$P^t f(x) = f(x+t) \tag{23.19}$$

(the shift operators); it is easy to see that the space $\mathbf{B}_0$ on which the semigroup P^t is continuous is the space $\mathbf{C}_{\text{unif.}}$ of bounded uniformly continuous functions.

Let us find D_A and A. The norm convergence in the space $\mathbf{B}$ is the *uniform convergence*; so $f \in D_A$ if and only if the limit

$$\lim_{h \to 0+} h^{-1}\big(P^h f(x) - f(x)\big) = \lim_{h \to 0+} h^{-1}\big(f(x+h) - f(x)\big) \tag{23.20}$$

exists, uniformly in x. This limit is nothing but the right-hand derivative. By Theorem 22.5, this happens if and only if the function $f(x+t)$ as a function of t has a *two-sided* derivative that is continuous in t, uniformly in t and in x. So we have:

$$\begin{aligned} D_A = \{f\colon\ f \text{ is bounded and differentiable,} \\ \text{with a bounded and uniformly continuous derivative}\}, \end{aligned} \tag{23.21}$$

$$Af(x) = f'(x). \tag{23.22}$$

In contrast with example 23.1, this operator is not a bounded one: for $f_n(x) = \sin(nx)$, we have $\|f_n\| = 1$, while $\|Af_n\| = n$, which can ba arbitrarily large.

Example 23.3. Let $\boldsymbol{b}(\boldsymbol{x})$ be a bounded and Lipschitz-continuous function $\mathbb{R}^d \mapsto \mathbb{R}^d$. Let us consider the deterministic Markov process described by the differential equation $\dfrac{d\boldsymbol{\xi}_t}{dt} = \boldsymbol{b}(\boldsymbol{\xi}_t)$ with the initial condition $\boldsymbol{\xi}_0$. Can we find its infinitesimal operator A of this process?

Unfortunately, the continuity space $\mathbf{B}_0$ here is not the space $\mathbf{C}_{\text{unif.}}$ of bounded uniformly continuous functions, generally not even in the one-dimensional case: it depends on the concrete properties of the vector field $\boldsymbol{b}(\boldsymbol{x})$. But let us find some *restriction* of the infinitesimal operator.

Let $\mathbf{C}^1_{\text{unif.}} = \mathbf{C}^1_{\text{unif.}}(\mathbb{R}^d)$ be the space of all bounded differentiable functions whose partial derivatives are bounded and uniformly continuous. Let us check that $\mathbf{C}^1_{\text{unif.}} \subseteq D_A$, and find Af for $f \in \mathbf{C}^1_{\text{unif.}}$. We have:

$$\lim_{h \to 0+} h^{-1}\big(P^h f(\boldsymbol{x}) - f(\boldsymbol{x})\big) = \lim_{h \to 0+} h^{-1}\big(f(\boldsymbol{\xi}_t) - f(\boldsymbol{x})\big), \tag{23.23}$$

where $\boldsymbol{\xi}_t$ is the solution of our differential equation with the initial condition $\boldsymbol{x}$. We see that the limit exists, uniformly in $\boldsymbol{x}$, and is equal to

$$Af(\boldsymbol{x}) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\boldsymbol{x}) \cdot b_i(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x}) \cdot \nabla f(\boldsymbol{x}). \tag{23.24}$$

This is again a differential operator, but unlike the operator in the previous example with variable coefficients (and unlike the differential operators considered in Lecture 21, it is a *first-order* operator).

As a matter of fact, the relation between the first-order partial differential equation $\dfrac{\partial u}{\partial t}(t, \boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x}) \cdot \nabla u(t, \bullet)(\boldsymbol{x})$ and the ordinary differential equation $\dfrac{d\boldsymbol{\xi}_t}{dt} = \boldsymbol{b}(\boldsymbol{\xi}_t)$ is well known in the theory of partial differential equations: it is called "the method of characteristics for solving first-order partial differential equations" ($\boldsymbol{\xi}_t$ is what is called *characteristic*); the method is applied much wider, in particular to quasi-linear and nonlinear first-order equations.

We can consider *diffusion processes* as some kind of random characteristics for solving second-order partial differential equations.

Example 23.4. Let ξ_t be the one-dimensional Wiener process. Let us find its infinitesimal operator.

First, the continuity space $\mathbf{B}_0$. For every function $f \in \mathbf{B}$ and $h > 0$ the function

$$P^h f(x) = \int_{-\infty}^{\infty} f(y) \cdot \frac{1}{\sqrt{2\pi h}}\, e^{-(y-x)^2/2h}\, dy \tag{23.25}$$

is infinitely differentiable with bounded derivatives; so the function $P^h f$ belongs to $\mathbf{C}_{\text{unif.}}$. The uniform limit of uniformly continuous functions is again uniformly continuous, so we see that $\mathbf{B}_0 \subseteq \mathbf{C}_{\text{unif.}}$. It is easy to prove also that $\mathbf{C}_{\text{unif.}} \subseteq \mathbf{B}_0$, so $\mathbf{C}_{\text{unif.}} = \mathbf{B}_0$.

Now let a function f be bounded and twice differentiable with the first and the second derivative bounded and uniformly continuous (let us denote the space of all such functions $\mathbf{C}^2_{\text{unif.}}$). We have:

$$h^{-1}\big(P^h f(x) - f(x)\big) = h^{-1} \int_{-\infty}^{\infty} [f(y) - f(x)] \cdot \frac{1}{\sqrt{2\pi h}}\, e^{-(y-x)^2/2h}\, dy. \tag{23.26}$$

Let us write the Taylor formula for the function f:

$$f(y) - f(x) = f'(x) \cdot (y - x) + \tfrac{1}{2} f''(x) \cdot (y - x)^2 + \tfrac{1}{2} [f''(\tilde{y}) - f''(x)] \cdot (y - x)^2, \tag{23.27}$$

where $\tilde{y} = \tilde{y}(x, y)$ is some point between x and y; and let us put it in formula (23.26). The integral of $(y - x)$ multiplied by the rest of it is equal to 0; that with $\frac{1}{2} f''(x) \cdot (y - x)^2$ yields $h^{-1} \cdot f''(x)/2$ multiplied by the variance of the normal distribution with parameters (x, h), so it is $\frac{1}{2} f''(x)$. And we'll see that the integral with $\frac{1}{2} [f''(\tilde{y}) - f''(x)] \cdot (y - x)^2$ goes to 0 as $h \to 0^+$, uniformly in x.

We have to prove that for every $\gamma > 0$ there exists $h_0 > 0$ such that the integral is less than γ for all $h \le h_0$ and $x \in (-\infty, \infty)$. Let $\varepsilon > 0$ be such that

$|y-x| < \varepsilon \;\Rightarrow\; |f''(y) - f''(x)| < \gamma$ (uniform continuity). Let us break our integral into one over the set $\{y\colon |y-x| < \varepsilon\}$, and over the complement of this set. In the first integral we have $|\tilde{y} - x| \le |y - x| < \varepsilon$, $|f''(\tilde{y}) - f''(x)| < \gamma$,

$$\begin{aligned} h^{-1} \int_{x-\varepsilon}^{x+\varepsilon} \tfrac{1}{2} |f''(\tilde{y}) - f''(x)| \cdot (y-x)^2 \cdot \frac{1}{\sqrt{2\pi h}}\, e^{-(y-x)^2/2h}\, dy \\ < \frac{\gamma}{2}\, h^{-1} \int_{-\infty}^{\infty} (y-x)^2 \cdot \frac{1}{\sqrt{2\pi h}}\, e^{-(y-x)^2/2h}\, dy = \frac{\gamma}{2}; \end{aligned} \tag{23.27}$$

and

$$\begin{aligned} h^{-1} \int\limits_{\{y\colon |y-x| \ge \varepsilon\}} \tfrac{1}{2} |f''(\tilde{y}) - f''(x)| \cdot (y-x)^2 \cdot \frac{1}{\sqrt{2\pi h}}\, e^{-(y-x)^2/2h}\, dy \\ \le \|f''\| \cdot \int\limits_{\{z\colon |z| \ge \varepsilon/\sqrt{h}\}} \frac{z^2}{\sqrt{2\pi}}\, e^{-z^2/2}\, dz, \end{aligned} \tag{23.28}$$

which goes to 0 as $h \to 0^+$.

So $\mathbf{C}^2_{\text{unif.}} \subseteq D_A$, and for $f \in \mathbf{C}^2_{\text{unif.}}$ we have

$$Af(x) = \tfrac{1}{2}\, f''(x). \tag{23.29}$$

Is $\mathbf{C}^2_{\text{unif.}} = D_A$, or is it that $\mathbf{C}^2_{\text{unif.}} \subset D_A$? We'll see in the next lecture that it is $=$: $D_A = \mathbf{C}^2_{\text{unif.}}$.