Probabilistic Interpretation of the Heat Equation

In this lecture we give a probabilistic interpretation of the heat equation.

1. Brownian motion.

Consider a particle moving around in a set S. For two times $t \leqslant s$ and $x \in S$, $E \subset S$, we denote by $P(t, x; s, E)$ the probability of this particle ends up in E at time s while it starts from x at time t. In particular, we have

- $P(t, x; s, S) = 1$ for any $t \leqslant s$, $x \in S$, since the particle is always contained in S;
- $P(t, x; s, \phi) = 0$ where ϕ is the empty set.

We consider a "memoryless" particle, that is, if a particle is at point y at time τ, then its movement afterwards is independent of how it gets there. In this case, we have the following Chapman-Kolmogorov equation.

$$P(t, x; s, E) = \int_S P(\tau, y; s, E)\, P(t, x; \tau, y)\, \mathrm{d}y \tag{1}$$

where τ is any time between t and s.

Next we further simplify the situation by assuming that the movement of the particle is independent of time, or more specifically,

$$P(t, x; s, E) = P(t + t', x; s + t', E) \tag{2}$$

or equivalently, $P(t, x; s, E)$ only depends on $t - s$, x and E. In this case we can simplify the notation by denoting

$$P(t, x, E) \equiv P(\tau, x; \tau + t, E) \tag{3}$$

In other words, $P(t, x, E)$ is the probability of a particle staring at x (the initial time doesn't matter now) ending up in E after time t.

The Chapman-Kolmogorov equation becomes

$$P(t + \tau, x, E) = \int_S P(\tau, y, E)\, P(t, x, y)\, \mathrm{d}y \tag{4}$$

for any $t, \tau \geqslant 0$.

Remark 1. Such a family of probability distribution functions $\{P(t, x, E)\}$ is called a *Markov process*.

Remark 2. In fact here it may be better to use $p(t, x, y)$ for $P(t, x, y)$ as it is really a probability density. Thus we should write

$$P(\tau + t, x, E) = \int_S p(t, x, y)\, P(\tau, y, E)\, \mathrm{d}y. \tag{5}$$

If we write $u(t, x) \equiv P(\tau + t, x, E)$, then we have

$$u(t, x) = \int_S p(t, x, y)\, u_0(y)\, \mathrm{d}y. \tag{6}$$

We clearly see that there is a possible relation to an evolution equation. For example,

$$p(t, x, y) = \frac{1}{(4\pi t)^{n/2}}\, e^{-\frac{|x - y|^2}{4t}} \tag{7}$$

in the case of the heat equation. Note that $p(t, x, y)$ is indeed a probability density.

In the next two sections, we will try to find out what this equation is.

In the following we consider the case $S = \mathbb{R}^n$. One can add further symmetry (besides time-translation invariance) to the motion of the particle.

Definition 3. (Brownian motion) *A Markov process $P(t,x;E)$ is a* Brownian motion *if*

a) it is spatially homogeneous:

$$P(t,x+z,E+z)=P(t,x,E), \tag{8}$$

b) for all $\rho>0$ and all $x\in\mathbb{R}^n$,

$$\lim_{t\searrow 0}\frac{1}{t}\int_{|x-y|>\rho} p(t,x,y)\,\mathrm{d}y=0. \tag{9}$$

Thus for Brownian motion, the density can be further simplifed to $p(t,x-y)$.

2. Semigroup properties.

For any bounded function f we can define a family of operators by

$$(T_t f)(x)\equiv\int_{\mathbb{R}^n} p(t,x-y)\,f(y)\,\mathrm{d}y. \tag{10}$$

The goal of this section is to show the following theorem.

Theorem 4. *Let B be the Banach space of bounded and uniformly continuous functions on $\mathbb{R}^n$, with norm*

$$\|u\|\equiv\sup_{x\in\mathbb{R}^n}|u(x)|. \tag{11}$$

Let $P(t,x,E)$ be a Brownian motion. Define

$$(T_t f)(x)\equiv\int_{\mathbb{R}^n} p(t,x-y)\,f(y)\,\mathrm{d}y \quad t>0; \qquad T_0 f=f. \tag{12}$$

Then $\{T_t\}_{t\geqslant 0}$ constitutes a contracting semigroup on B.

Proof.

- T_t maps B into B. That is we first verify that $T_t f$ remains bounded and uniformly continuous.
  - Bounded. We have

$$\sup_{x\in\mathbb{R}^n}|(T_t f)(x)|=\sup_{x\in\mathbb{R}^n}\left|\int p(t,x,y)\,f(y)\,\mathrm{d}y\right|\leqslant\sup_{x\in\mathbb{R}^n}|f(x)|\int p(t,x,y)\,\mathrm{d}y\leqslant\sup_{x\in\mathbb{R}^n}|f(x)|. \tag{13}$$

  - Uniformly continuous. We estimate

$$\begin{aligned}|(T_t f)(x)-(T_t f)(x+z)| &= \left|\int p(t,x-y)\,f(y)\,\mathrm{d}y-\int p(t,x+z-y)\,f(y)\,\mathrm{d}y\right|\\ &= \left|\int p(t,x-y)\,f(y)\,\mathrm{d}y-\int p(t,x-y')\,f(y'+z)\,\mathrm{d}y'\right|\\ &\quad \text{(We set } y'=y-z)\\ &= \left|\int p(t,x-y)\,[f(y)-f(y+z)]\,\mathrm{d}y\right|\\ &\leqslant \int p(t,x-y)\,|f(y)-f(y+z)|\,\mathrm{d}y\\ &\leqslant \sup_y|f(y)-f(y+z)|\int p(t,x-y)\,\mathrm{d}y\\ &= \sup_y|f(y)-f(y+z)|\,. \end{aligned} \tag{14}$$

    The uniform continuity of $T_t f$ now follows from the uniform continuity of f.

- $\{T_t\}_{t\geqslant 0}$ is a continuous semigroup.

  i. $T_0=\mathrm{Id}$. This is exactly how T_0 is defined.

ii. $T_{t+s} = T_t \circ T_s$. We need to show

$$\int p(t+s, x-y)\, f(y)\, \mathrm{d}y = \int p(t, x-z) \left[\int p(s, z-y)\, f(y)\, \mathrm{d}y \right] \mathrm{d}z. \tag{15}$$

This is equivalent to

$$p(t+s, x-y) = \int p(t, x-z)\, p(s, z-y)\, \mathrm{d}z. \tag{16}$$

Recall the Chapman-Kolmogorov equation

$$P(t+s, x, E) = \int_S P(s, z, E)\, p(t, x-z)\, \mathrm{d}z, \tag{17}$$

taking the “density” on both sides, we get the desired relation.

iii. $\lim_{t \to t_0} T_t f = T_{t_0} f$ for all $t_0 \geqslant 0$ and $f \in B$.

We estimate $|T_t f - T_s f|$ for $t \geqslant s$. Denote $\tau = t - s$.

$$\begin{aligned} |T_t f(x) - T_s f(x)| &= |T_\tau (T_s f)(x) - (T_s f)(x)| \\ &= \left| \int_{\mathbb{R}^n} p(\tau, x-y)\, [(T_s f)(y) - (T_s f)(x)]\, \mathrm{d}y \right| \\ &\leqslant \left| \int_{|x-y| \leqslant \delta} p(\tau, x-y)\, [(T_s f)(y) - (T_s f)(x)]\, \mathrm{d}y \right| \\ &\quad + \left| \int_{|x-y| > \delta} p(\tau, x-y)\, [(T_s f)(y) - (T_s f)(x)]\, \mathrm{d}y \right| \\ &\leqslant \sup_{|y-z| \leqslant \delta} |(T_s f)(y) - (T_s f)(z)| \int_{|x-y| \leqslant \varepsilon} p(\tau, x-y)\, \mathrm{d}y \\ &\quad + 2 \sup_y |(T_s f)(y)| \int_{|x-y| > \delta} p(\tau, x-y)\, \mathrm{d}y \\ &\leqslant \sup_{|y-z| \leqslant \delta} |(T_s f)(y) - (T_s f)(z)| \\ &\quad + 2 \sup_y |(T_s f)(y)| \int_{|x-y| > \delta} p(\tau, x-y)\, \mathrm{d}y. \end{aligned} \tag{18}$$

Now for any $\varepsilon > 0$, we take δ such that

$$\sup_{|y-z| \leqslant \delta} |(T_s f)(y) - (T_s f)(z)| < \frac{\varepsilon}{2}, \tag{19}$$

and then take $0 < \tau_0 < \varepsilon/2$ such that

$$\frac{1}{\tau_0} \int_{|x-y| > \delta} p(\tau, x-y)\, \mathrm{d}y < \frac{1}{2 \sup_y |(T_s f)(y)|}. \tag{20}$$

Then for all $\tau < \tau_0$ we have

$$|T_t f(x) - T_s f(x)| < \varepsilon. \tag{21}$$

- “Contracting”: We need to show

$$\sup_{x \in \mathbb{R}^n} |(T_t f)(x)| \leqslant \sup_{x \in \mathbb{R}^n} |f(x)| \qquad \forall f \in B, t \geqslant 0. \tag{22}$$

But this is already done when showing T_t is bounded. □

3. From Brownian motion to the heat equation.

The goal of this section is to prove the following theorem.

Theorem 5. *Let $P(t, x, E)$ be a Brownian motion that is invariant under all isometries of Euclidean space.*[1] *Then the infinitesimal generator of the contracting semigroup defined by this process is*

$$A = c \triangle \tag{23}$$

where $c > 0$ is a constant. Furthermore the density is simply

$$p(t, x - y) = \frac{1}{(4\pi c t)^{n/2}} e^{-\frac{|x-y|^2}{4t}}. \tag{24}$$

Proof. First note that, since $P(t, x, E)$ is invariant under all isometries of Euclidean space, the density $p(t, x - y)$ can only depend on the distance between x, y, that is

$$p(t, x - y) = p(t, |x - y|). \tag{25}$$

- Main idea.

  We would like to find out

  $$\lim_{t \searrow 0} \frac{1}{t} (T_t f - f)(x) = \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) [f(y) - f(x)] \, \mathrm{d}y. \tag{26}$$

  Now we expand f at x by Taylor expansion:

  $$f(y) - f(x) = \sum_i (\partial_i f)(x) (y_i - x_i) + \frac{1}{2} \sum_{i,j} (\partial_{ij} f)(x) (y_i - x_i) (y_j - x_j) + R(x, y), \tag{27}$$

  where $R(x, y) = o\left(|x - y|^2\right)$ as $|x - y| \searrow 0$.

  Plugging this expansion into the limit we obtain

  $$\begin{aligned} A f \equiv \lim_{t \searrow 0} \frac{1}{t} (T_t f - f)(x) &= \sum_i \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) (\partial_i f)(x) (y_i - x_i) \\ &\quad + \frac{1}{2} \sum_{i,j} \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) (\partial_{ij} f)(x) (y_i - x_i) (y_j - x_j) \\ &\quad + \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) R(x, y) \, \mathrm{d}y \\ &= \sum_i \left[\lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) (y_i - x_i) \right] (\partial_i f)(x) \\ &\quad + \sum_{i,j} \left[\frac{1}{2} \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) (y_i - x_i) (y_j - x_j) \right] (\partial_{ij} f)(x) \\ &\quad + \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) R(x, y) \, \mathrm{d}y. \end{aligned} \tag{28}$$

  We see that, if the last term vanishes, then formally we have shown that the infinitesimal generator

  $$A = \sum_{i,j} a_{ij} \partial_{ij} + \sum_i b_i \partial_i \tag{29}$$

  where

  $$a_{ij} = \frac{1}{2} \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) (y_i - x_i) (y_j - x_j); \qquad b_i = \lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|) (y_i - x_i). \tag{30}$$

  Now we formally argue that $b_i = 0$ and $a_{ij} = c \, \delta_{ij}$.

  - $b_i = 0$.

    We try to show that

    $$\int p(t, |x - y|) (x_i - y_i) \, \mathrm{d}y = 0 \tag{31}$$

    or equivalently

    $$\int p(t, |x|) x_i \, \mathrm{d}x = 0. \tag{32}$$

---

1. That is, translation (respected by any Brownian motion), rotation, and reflection. To see this, first note that any isometry must also preserve angles. From this it is easy to show that it must be linear after setting the image of the origin to be the new origin. Thus it can be represented by a matrix which must be orthogonal. And we know that any orthogonal matrix is either a rotation or a composition of a rotation and a reflection.

But this is obvious as the integrand is odd in x_i.

- $a_{ij} = 0$ when $i \neq j$.

  We need to show

$$\int p(t, |x|)\, x_i\, x_j\, \mathrm{d}x = 0. \tag{33}$$

  The integrand is odd in both x_i and x_j.

- $a_{ii} = a_{jj}$.

  That is

$$\int p(t, |x|)\, x_i^2\, \mathrm{d}x = \int p(t, |x|)\, x_j^2\, \mathrm{d}x. \tag{34}$$

  This follows from a change of variables.

- Furthermore we obtain

$$a_{ii} = \frac{1}{n} \int p(t, |x|)\, |x|^2\, \mathrm{d}x \tag{35}$$

  therefore

$$a_{ij} = c\, \delta_{ij} \tag{36}$$

  with

$$c = \frac{1}{n} \int p(t, |x|)\, |x|^2\, \mathrm{d}x. \tag{37}$$

Thus formally we have shown that

$$A = c\, \triangle. \tag{38}$$

In the following we will try to make the argument rigorous. We need to do the following

1. Justify the Taylor expansion. Note that a priori we do not know whether $f \in D(A)$ is differentiable or not.
2. Show that

$$\lim_{t \searrow 0} \frac{1}{t} \int p(t, |x - y|)\, R(x, y)\, \mathrm{d}y = 0; \tag{39}$$

3. Justify the computations for a_{ij}, b_i.

- Rigorous argument.

  Before we start, first note that since $P(x, t, E)$ is a Brownian motion,

$$\lim_{t \searrow 0} \frac{1}{t} \int_{|x - y| > \rho} p(t, x, y)\, \mathrm{d}y = 0 \tag{40}$$

for any $\rho > 0$. Thus for any fixed ρ, we have

$$\lim_{t \searrow 0} \frac{1}{t} (T_t f - f)(x) = \lim_{t \searrow 0} \frac{1}{t} \int_{|x - y| < \rho} p(t, |x - y|)\, [f(y) - f(x)]\, \mathrm{d}y \tag{41}$$

as any $f \in B$ is bounded.

1. First we justify the Taylor expansion of f by showing that $D(A) \cap C^\infty$ is dense in B. We do this by showing that whenever $f \in D(A)$, $f_\varepsilon \equiv h_\varepsilon * f$ is also contained in $D(A)$, where $h_\varepsilon \equiv \varepsilon^{-n}\, h(x/\varepsilon)$ is the usual (rescaled) mollifier.

   Thus it suffices to show that the limit

$$\lim_{t \searrow 0} \frac{1}{t} [(T_t\, f_\varepsilon) - f_\varepsilon] \tag{42}$$

   exists.

   First write

$$f_\varepsilon(x) = \int \varepsilon^{-n}\, h\left(\frac{z}{\varepsilon}\right) f(x - z)\, \mathrm{d}z = \int h(w)\, f(x - \varepsilon\, w)\, \mathrm{d}w. \tag{43}$$

Then compute

$$\begin{aligned}
\frac{1}{t}[(T_t f_\varepsilon) - f_\varepsilon] &= \frac{1}{t}\left\{ \int \left[p(t,|x-y|) \int h(w)\, f(y-\varepsilon\, w)\,\mathrm{d}w - \int h(w)\, f(x-\varepsilon\, w)\,\mathrm{d}w \right] \mathrm{d}y \right\} \\
&= \frac{1}{t}\left\{ \int p(t,|x-y|) \left[\int h(w)\, [f(y-\varepsilon\, w) - f(x-\varepsilon\, w)]\,\mathrm{d}w \right] \mathrm{d}y \right\} \\
&= \int h(w) \left\{ \frac{1}{t}\left[\int p(t,|x-y|)\, f(y-\varepsilon\, w)\,\mathrm{d}y - f(x-\varepsilon\, w) \right] \right\} \mathrm{d}w. \qquad (44)
\end{aligned}$$

Now because of the translation invariance, $f \in D(A)$ implies $f(\,\cdot - \varepsilon\, w) \in D(A)$ too, and furthermore the convergence (in the space B, that is, uniform convergence)

$$\frac{1}{t}\left[\int p(t,|x-y|)\, f(y-\varepsilon\, w)\,\mathrm{d}y - f(x-\varepsilon\, w) \right] \longrightarrow (A f)(\,\cdot - \varepsilon\, w) \qquad (45)$$

is uniform with respect to ε and w. Thus in particular,

$$\frac{1}{t}\left[\int p(t,|x-y|)\, f(y-\varepsilon\, w)\,\mathrm{d}y - f(x-\varepsilon\, w) \right]$$

as functions in B are uniformly bounded in t. Application of Lebesgue's dominated convergence theorem gives

$$\lim_{t\searrow 0} \int h(w) \left\{ \frac{1}{t}\left[\int p(t,\,|x-y|)\, f(y-\varepsilon\, w)\,\mathrm{d}y - f(x-\varepsilon\, w) \right] \right\} \mathrm{d}w = \int h(w)\, (A f)(x - \varepsilon\, w)\,\mathrm{d}w. \qquad (46)$$

Therefore $f_\varepsilon \in D(A)$, and we can safely work on smooth functions only.

2. Next we show that

$$\lim_{t\searrow 0} \frac{1}{t} \int p(t,|x|)\, R(x)\,\mathrm{d}y = 0 \qquad (47)$$

when $R(x,y) = o\left(|x|^2\right)$.

Let φ be any smooth function in $D(A)$. Then we have

$$\begin{aligned}
(A\varphi)(x) &= \lim_{t\searrow 0} \frac{1}{t} \int p(t,|x-y|)\, [\varphi(y) - \varphi(x)]\,\mathrm{d}y \\
&= \lim_{t\searrow 0} \frac{1}{t} \int p(t,|x-y|) \sum_i (\partial_i\varphi)(x)\,(y_i - x_i)\,\mathrm{d}y \\
&\quad + \lim_{t\searrow 0} \frac{1}{t} \int p(t,|x-y|) \sum_{i,j} (\partial_{ij}\varphi)(\xi)\,(y_i - x_i)\,(y_j - x_j)\,\mathrm{d}y \\
&= \lim_{t\searrow 0} \frac{1}{t} \int p(t,|x-y|) \sum_{i,j} (\partial_{ij}\varphi)(\xi)\,(y_i - x_i)\,(y_j - x_j)\,\mathrm{d}y. \qquad (48)
\end{aligned}$$

where ξ is a point between x and y obtained from mean value theorem.

Thus it suffices to find one particular $\varphi \in D(A) \cap C^\infty$ such that

$$\sum_{i,j} (\partial_{ij}\varphi)(\xi)\,(y_i - x_i)\,(y_j - x_j) \geqslant |x-y|^2. \qquad (49)$$

Since the integral domain can be taken as $|y| \leqslant \varepsilon$, by taking ε small enough, we further simplify the requirement to

$$\sum_{i,j} (\partial_{ij}\varphi)(x)\,(y_i - x_i)\,(y_j - x_j) \geqslant |x-y|^2 \qquad (50)$$

and by translation invariance, finally we only need to find $\varphi \in D(A) \cap C^\infty$ such that

$$(\partial_{ij}\varphi)(0)\, x_i\, x_j \geqslant |x|^2 \qquad \text{for } |x| \leqslant \varepsilon. \qquad (51)$$

Such a function can be obtained by cutting-off and then mollifying $2\,|x|^2$.

3. Finally we justify the computations for a_{ij} and b_i. Just note that now the integration domain is $|x-y| \leqslant \rho$. □

4. Particles in a lattice.

In this section we give a probabilistic interpretation of the heat equation from a discrete point of view. Consider the lattice $h\,\mathbb{Z}^n$, and a collection of particles moving on it. We denote by ρ_{ij} the number of particles at (i, j) ($(i\,h, j\,h)$ as a point in $\mathbb{R}^n$).

We consider the case that the particles only move at times $t_k = k\,\triangle t$, and further denote by ρ_{ij}^k the number of particles at (i, j) at time t_k. Further assume that at each step, a particle at (i, j) can only hop to the four adjacent grids $(i+1, j), (i-1, j), (i, j+1), (i, j-1)$ with equal probability (that is 1/4 each). Under this assumption we have

$$\rho_{ij}^{k+1} = \frac{1}{4}\left[\rho_{i+1,j}^k + \rho_{i-1,j}^k + \rho_{i,j+1}^k + \rho_{i,j-1}^k\right]. \tag{52}$$

Subtracting ρ_{ij}^k from both sides, we obtain

$$\rho_{ij}^{k+1} - \rho_{ij}^k = \frac{1}{4}\left[\rho_{i+1,j}^k + \rho_{i-1,j}^k + \rho_{i,j+1}^k + \rho_{i,j-1}^k - 4\,\rho_{ij}^k\right]. \tag{53}$$

Finally assume that $\Delta t = h^2/4$, we obtain

$$\frac{\rho_{ij}^{k+1} - \rho_{ij}^k}{\Delta t} = \frac{\rho_{i+1,j}^k + \rho_{i-1,j}^k + \rho_{i,j+1}^k + \rho_{i,j-1}^k - 4\,\rho_{ij}^k}{h^2}. \tag{54}$$

Now if

$$\rho_{ij}^k \approx \rho(k\,\Delta t, i\,h, j\,h) \tag{55}$$

for a smooth function $\rho(t, x_1, x_2)$, the above differencee equation approximates

$$\rho_t = \triangle \rho \tag{56}$$

which is the heat equation.

Exercises.

Exercise 1. Show formally that for general Brownian motions, that is without invariance under rotation and reflection, we will obtain

$$A = \sum_{i,j} a_{ij}(x)\,\partial_{ij}^2 + \sum_k b_k(x)\,\partial_k. \tag{57}$$

Write down the formulas for $a_{ij}(x)$ and $b_k(x)$, then prove (basing on the formulas) that for any fixed x, the matrix $(a_{ij}(x))$ is positive semidefinite.