# Markov processes

University of Bonn, summer term 2008

Author: **Prof. Dr. Andreas Eberle**
Edited by: Sebastian Riedel

Latest Version: January 22, 2009

# Contents

# Chapter 1

# Continuous-time Markov chains

Additional reference for this chapter:

- Asmussen [4]
- Stroock [20]
- Norris [16]
- Kipnis, Landim [13]

.

## 1.1 Markov properties in discrete and continuous time

Let $(S, \mathscr{S})$ be a measurable space, $p_n(x, dy)$ transition probabilities (Markov kernels) on $(S, \mathscr{S})$, $(\Omega, \mathcal{A}, P)$ a probability space and $\mathcal{F}_n(n = 0, 1, 2, ...)$ a filtration on $(\Omega, \mathcal{A})$.

**Definition 1.1.** *A stochastic process $(X_n)_{n=0,1,2,...}$ on $(\Omega, \mathcal{A}, P)$ is called **$(\mathcal{F}_n)$-Markov chain with transition probabilities $p_n$** if and only if*

*(i)* $X_n$ *is $\mathcal{F}_n$-measurable* $\forall\, n \geq 0$

*(ii)* $P[X_{n+1} \in B|\mathcal{F}_n] = p_{n+1}(X_n, B)$ *P-a.s.* $\forall n \geq 0, B \in \mathscr{S}$.

**Example** (***Non-linear state space model***)**.** *Consider*

$$\begin{aligned} &X_{n+1} = F_{n+1}(X_n, U_{n+1}),\\ &U_n \colon \Omega \to \tilde{S}_n \text{ *independent random variable,* noise}\\ &F_n \colon S \times \tilde{S}_n \to S \text{ *measurable,* rule for dynamics} \end{aligned}$$

*$(X_n)_{n\geq 0}$ is a Markov chain with respect to $\mathcal{F}_n = \sigma(X_0, U_1, U_2, \dots, U_n)$ with transition probabilities $p_n(x, \cdot) =$ distribution of $F_n(x, U_n)$.*

**Remark .** 1. ***Longer transitions:***

$$P[X_n \in B|\mathcal{F}_m] = (\underbrace{p_n \cdots p_{m+2}p_{m+1}}_{=:p_{m,n}})(X_n, B)$$

*where*

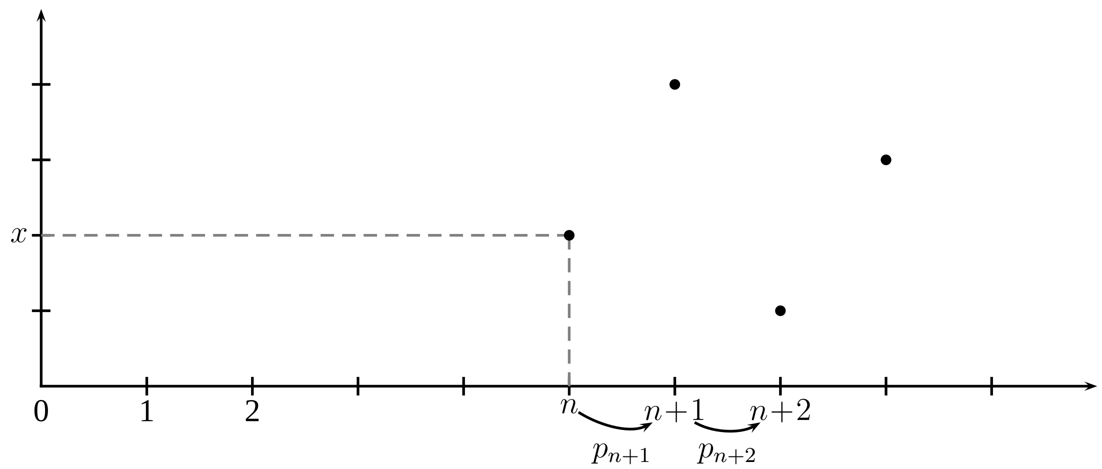
$$(pq)(x,B) := \int_S p(x,dy)q(y,B)$$

*Time-homogeneous case:* $p_n = p \; \forall n$, $p_{m,n} = p^{n-m}$.

2. ***Reduction to time-homogeneous case:***
*$X_n$ time-inhomogeneous Markov chain, $\hat{X}_n := (n, X_n)$ space-time process is time-homogeneous Markov chain on $\hat{S} = \{0,1,2,\dots\} \times S$ with transition probabilities*

$$\hat{p}((n,x), \{m\} \times B) = \delta_{m,n+1} \cdot p_{n+1}(x,B).$$

3. ***Canonical model for space-time chain with start in $(n,x)$:***
*There is an unique probability measure $P_{(n,x)}$ on $\Omega := S^{\{0,1,2,\dots\}}$ so that under $P_{(n,x)}$, $X_k(\omega) = \omega(k)$ is a Markov chain with respect to $\mathcal{F}_n = \sigma(X_0, \dots X_n)$ and with transition kernels $p_{n+1}, p_{n+2}, \dots$ and initial condition $X_0 = x$ $P_{(n,x)}$-a.s.*




Let $T\colon \Omega \to \{0,1,\dots\} \cup \{\infty\}$ be a $(\mathcal{F}_n)$-stopping time, i.e.

$$\{T = n\} \in \mathcal{F}_n \forall n \geq 0,$$
$$\mathcal{F}_T = \{A \in \mathcal{A} \mid A \cap \{T = n\} \in \mathcal{F}_n \; \forall n \geq 0\} \quad \textit{events observable up to time T.}$$

**Theorem 1.2** (Strong Markov property)**.**

$$\mathbb{E}[F(X_T, X_{T+1}, \ldots) \cdot I_{\{T<\infty\}}|\mathcal{F}_T] = \mathbb{E}_{(T,X_T)}[F(X_0, X_1, \ldots)] \quad P\text{-a.s. on } \{T < \infty\}$$

*Proof.* Exercise (Consider first $T = n$). □

## 1.2 From discrete to continuous time:

Let $t \in \mathbb{R}^+$, S Polish space (complete, separable metric space), $\mathscr{S} = \mathcal{B}(S)$ Borel $\sigma$-algebra, $p_{s,t}(x, dy)$ transition probabilities on $(S, \mathscr{S}),\ 0 \leq s \leq t < \infty,\ (\mathcal{F}_t)_{t\geq 0}$ filtration on $(\Omega, \mathcal{A}, P)$.

**Definition 1.3.** *A stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{A}, P)$ is called a **$(\mathcal{F}_t)$-Markov process with transition functions $p_{s,t}$** if and only if*

*(i)* $X_t$ *is* $\mathcal{F}_t$*-measurable* $\quad \forall\, t \geq 0$

*(ii)* $P[X_t \in B|\mathcal{F}_s] = p_{s,t}(X_s, B)$ *P-a.s.* $\quad \forall\, 0 \leq s \leq t,\ B \in \mathscr{S}$

**Lemma 1.4.** *The transition functions of a Markov process satisfy*

1. $p_{s,s}f = f$
2. $p_{s,t}p_{t,u}f = p_{s,u}f$ ***Chapman-Kolmogorov equation***

$P \circ X^{-1}$*-a.s. for every* $f : S \to \mathbb{R}$ *and* $0 \leq s \leq t \leq u$*.*

*Proof.* 1. $(p_{s,s}f)(X_s) = \mathbb{E}[f(X_s)|\mathcal{F}_s] = f(X_s) \quad P$-a.s.

2. $(p_{s,u}f)(X_s) = \mathbb{E}[f(X_u)|\mathcal{F}_s] = \mathbb{E}[\overbrace{\mathbb{E}[f(X_n)|\mathcal{F}_t]}^{(p_{t,u}f)(X_t)}|\mathcal{F}_s] = (p_{s,t}p_{t,u}f)(X_s) \quad P$-a.s.

□

**Remark .** 1. *Time-homogeneous case:* $p_{s,t} = p_{t-s}$
*Chapman-Kolmogorov:* $p_s p_t f = p_{s+t} f$ *(semigroup property).*

2. *Kolmogorov existence theorem: Let $p_{s,t}$ be transition probabilities on $(S, \mathscr{S})$ such that*

   *(i)* $p_{t,t}(x, \cdot) = \delta_x \quad \forall x \in S,\ t \geq 0$
   *(ii)* $p_{s,t}p_{t,u} = p_{s,u} \quad \forall\, 0 \leq s \leq t \leq u$

   *Then there is an unique canonical Markov process $(X_t, P_{(s,x)})$ on $S^{[0,\infty)}$ with transition functions $p_{s,t}$.*

**Problems:**

- regularity of paths $t \mapsto X_t$. One can show: If $S$ is locally compact and $p_{s,t}$ Feller, then $X_t$ has càdlàg modification (cf. Revuz, Yor [17]).
- in applications, $p_{s,t}$ is usually not known explicitly.

We take a more constructive approach instead.

Let $(X_t, P)$ be an $(\mathcal{F}_t)$-Markov process with transition functions $p_{s,t}$.

**Definition 1.5.** *$(X_t, P)$ has the **strong Markov property** w.r.t. a stopping time $T\colon \Omega \to [0,\infty]$ if and only if*

$$\mathbb{E}\left[f(X_{T+s}) I_{\{T<\infty\}} \mid \mathcal{F}_T\right] = (p_{T,T+s} f)(X_T)$$

*$P$-a.s. on $\{T < \infty\}$ for all measurable $f\colon S \to \mathbb{R}^+$.*
*In the time-homogeneous case, we have*

$$\mathbb{E}\left[f(X_{T+s}) I_{\{T<\infty\}} \mid \mathcal{F}_T\right] = (p_s f)(X_T)$$

**Lemma 1.6.** *If the strong Markov property holds then*

$$\mathbb{E}\left[F(X_{T+\cdot}) I_{\{T<\infty\}} \mid \mathcal{F}_T\right] = \mathbb{E}_{(T,X_T)}\left[F(X_{T+\cdot})\right]$$

*$P$-a.s. on $\{T < \infty\}$ for all measurable functions $F\colon S^{[0,\infty)} \to \mathbb{R}^+$.*
*In the time-homogeneous case, the right term is equal to*

$$\mathbb{E}_{X_T}[F(X)]$$

*Proof.* Exercise. □

**Definition 1.7.**

$$\mathrm{PC}(\mathbb{R}^+, S) := \{x\colon [0,\infty) \to S \mid \forall t \geq 0\ \exists \varepsilon > 0 : x \textit{ constant on } [t, t+\varepsilon)\}$$

**Definition 1.8.** *A Markov process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{A}, P)$ is called a **pure jump process** or **continuous time Markov chain** if and only if*

$$(t \mapsto X_t) \in \mathrm{PC}(\mathbb{R}^+, S)$$

*$P$-a.s.*

Let $q_t \colon S \times \mathscr{S} \to [0,\infty]$ be a ***kernel of positive measure***, i.e. $x \mapsto q_t(x,A)$ is measurable and $A \mapsto q_t(x,A)$ is a positive measure.

**Aim:** Construct a pure jump process with instantaneous jump rates $q_t(x,dy)$, i.e.

$$P_{t,t+h}(x,B) = q_t(x,B) \cdot h + o(h) \quad \forall\, t \geq 0,\ x \in S,\ B \subseteq S \setminus \{x\} \text{ measurable}$$

$(X_t)_{t\geq 0} \leftrightarrow (Y_n, J_n)_{n\geq 0} \leftrightarrow (Y_n, S_n)$ with $J_n$ ***holding times***, $S_n$ ***jumping times*** of $X_t$.

$J_n = \sum_{i=1}^{n} S_i \in (0,\infty]$ with jump time $\{J_n : n \in \mathbb{N}\}$ point process on $\mathbb{R}^+$, $\zeta = \sup J_n$ ***explosion time***.

## Construction of a process with initial distribution $\mu \in M_1(S)$:

$\lambda_t(x) := q_t(x, S \setminus \{x\})$ intensity, *total rate of jumping away from x.*

***Assumption:*** $\lambda_t(x) < \infty \quad \forall x \in S$ *, no instantaneous jumps.*

$\pi_t(x,A) := \frac{q_t(x,A)}{\lambda_t(x)}$ transition probability *, where jumps from $x$ at time $t$ go to.*

a) ***Time-homogeneous case:*** $q_t(x,dy) \equiv q(x,dy)$ independent of $t$, $\lambda_t(x) \equiv \lambda(x)$, $\pi_t(x,dy) \equiv \pi(x,dy)$.

$Y_n$ $(n = 0,1,2,\ldots)$ Markov chain with transition kernel $\pi(x,dy)$ and initial distribution $\mu$

$S_n := \dfrac{E_n}{\lambda(Y_{n-1})}$, $E_n \sim \mathrm{Exp}(1)$ independent and identically distributed random variables, independent of $Y_n$, i.e.

$S_n | (Y_0, \ldots Y_{n-1}, E_1, \ldots E_{n-1}) \sim \mathrm{Exp}(\lambda(Y_{n-1}))$,

$$J_n = \sum_{i=1}^{n} S_i$$

$$X_t := \begin{cases} Y_n & \text{for} \quad t \in [J_n, J_{n+1}),\ n \geq 0 \\ \Delta & \text{for} \quad t \geq \zeta = \sup J_n \end{cases}$$

where $\Delta$ is an extra point, called the *cemetery*.

**Example .** *1)* ***Poisson process with intensity $\lambda > 0$***

$S = \{0, 1, 2, \ldots\}, \quad q(x, y) = \lambda \cdot \delta_{x+1,y}, \quad \lambda(x) = \lambda \ \forall x, \quad \pi(x, x+1) = 1$
$S_i \sim \mathrm{Exp}(\lambda)$ *independent and identically distributed random variables*, $Y_n = n$
$N_t = n \ :\Leftrightarrow \ J_n \leq t \leq J_{n+1}$,
$N_t = \#\{i \geq 1 \ : \ J_i \leq t\}$ counting process of point process $\{J_n \mid n \in \mathbb{N}\}$.

*Distribution at time t:*

$$P[N_t \geq n] = P[J_n \leq t] \overset{J_n = \sum_{i=1}^{n} S_i \sim \Gamma(\lambda, n)}{=} \int_0^t \frac{(\lambda s)^{n-1}}{(n-1)!} \lambda e^{-\lambda s} \, ds \overset{\text{differentiate r.h.s.}}{=} e^{-t\lambda} \sum_{k=n}^{\infty} \frac{(t\lambda)^k}{k!},$$

$N_t \sim \mathrm{Poisson}(\lambda t)$

*2)* ***Continuization of discrete time chain***

*Let $(Y_n)_{n \geq 0}$ be a time-homogeneous Markov chain on $S$ with transition functions $p(x, dy)$,*

$X_t = Y_{N_t}, \quad N_t$ *Poisson(1)-process independent of* $(Y_n)$,
$q(x, dy) = \pi(x, dy), \quad \lambda(x) = 1$

*e.g.* ***compound Poisson process*** *(continuous time random walk):*

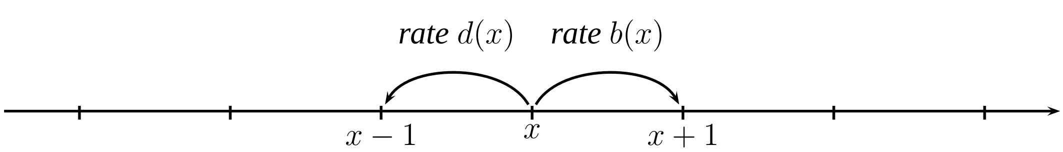
$$X_t = \sum_{i=1}^{N_t} Z_i,$$

$Z_i \colon \Omega \to \mathbb{R}^d$ *independent and identically distributed random variables, independent of* $(N_t)$.

*3)* ***Birth and death chains***

$S = \{0, 1, 2, \ldots\}$.

$$q(x, y) = \begin{cases} b(x) & \text{if } y = x+1 \quad \text{"birth rate"} \\ d(x) & \text{if } y = x-1 \quad \text{"death rate"} \\ 0 & \text{if } |y - x| \geq 2 \end{cases}$$



b) ***Time-inhomogeneous case:***

**Remark** (Survival times). *Suppose an event occurs in time interval $[t, t+h]$ with probability $\lambda_t \cdot h + o(h)$ provided it has not occurred before:*

$$\begin{aligned}
& P[T \le t+h | T > t] = \lambda_t \cdot h + o(h) \\
\Leftrightarrow \quad & \underbrace{\frac{P[T > t+h]}{P[T > t]}}_{\text{survival rate}} = P[T > t+h | T > t] = 1 - \lambda_t h + o(h) \\
\Leftrightarrow \quad & \frac{\log P[T > t+h] - \log P[T > t]}{h} = -\lambda_t + o(h) \\
\Leftrightarrow \quad & \frac{d}{dt} \log P[T > t] = -\lambda_t \\
\Leftrightarrow \quad & P[T > t] = \exp\left( - \int_0^t \lambda_s \, ds \, t \right)
\end{aligned}$$

*where the integral is the* accumulated hazard rate up to time $t$,

$$f_T(t) = \lambda_t \exp\left( - \int_0^t \lambda_s \, ds \right) \cdot I_{(0,\infty)}(t) \quad \textit{the } \text{survival distribution with hazard rate } \lambda_s$$

**Simulation of *T*:**

$E \sim \mathrm{Exp}(1), \; T := \inf\{t \ge 0 \; : \; \int_0^t \lambda_s \, ds \ge E\}$

$$\Rightarrow \quad P[T > t] = P\left[ \int_0^t \lambda_s \, ds < E \right] = e^{-\int_0^t \lambda_s \, ds}$$

**Construction of time-inhomogeneous jump process:**

Fix $t_0 \ge 0$ and $\mu \in M_1(S)$ (the initial distribution at time $t_0$).
Suppose that with respect to $P_{(t_0,\mu)}$,

$$J_0 := t_0, \qquad Y \sim \mu$$

and

$$P_{(t_0,\mu)}[J_1 > t \mid Y_0] = e^{-\int_{t_0}^{t \vee t_0} \lambda_s(Y_0) \, ds}$$

for all $t \ge t_0$, and $(Y_{n-1}, J_n)_{n \in \mathbb{N}}$ is a *time-homogeneous* Markov chain on $S \times [0, \infty)$ with transition law

$$P_{(t_0,\mu)}[Y_n \in dy,\ J_{n+1} > t \mid Y_0, J_1, \dots, Y_{n-1}, J_n] = \pi_{J_n}(Y_{n-1}, dy) \cdot e^{-\int_{J_n}^{t \vee J_n} \lambda_s(y)\, ds}$$

i.e.

$$P_{(t_0,\mu)}[Y_n \in A,\ J_{n+1} > t \mid Y_0, J_1, \dots, Y_{n-1}, J_n] = \int_A \pi_{J_n}(Y_{n-1}, dy) \cdot e^{-\int_{J_n}^{t \vee J_n} \lambda_s(y)\, ds}$$

$P$-a.s. for all $A \in \mathscr{S},\ t \geq 0$

**Explicit (algorithmic) construction:**

- $J_0 := t_0,\ Y_0 \sim \mu$

For $n = 1, 2, \dots$ do

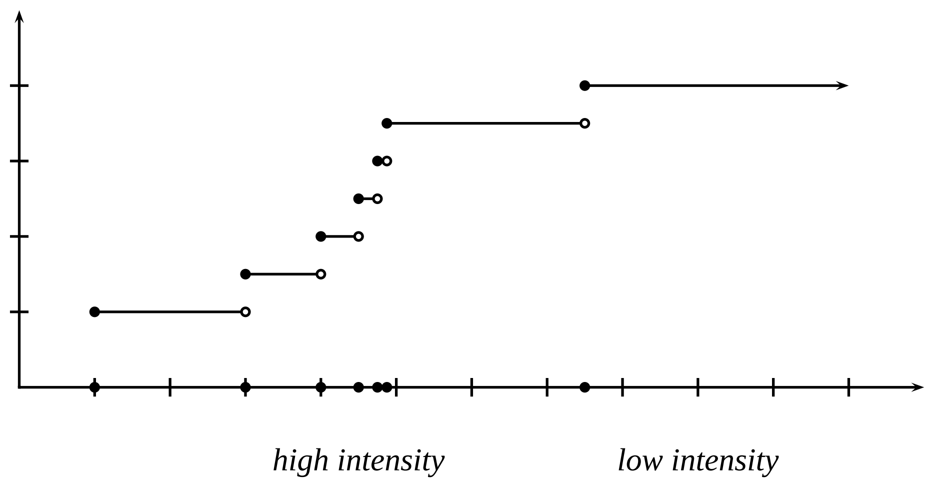
- $E_n \sim \mathrm{Exp}(1)$ independent of $Y_0, \dots, Y_{n-1}, E_1, \dots, E_{n-1}$
- $J_n := \inf \left\{ t \geq 0 \ : \ \int_{J_{n-1}}^t \lambda_s(Y_{n-1})\, ds \geq E_n \right\}$
- $Y_n | (Y_0, \dots, Y_{n-1}, E_0, \dots, E_n) \sim \pi_{J_n}(Y_{n-1}, \cdot)$ where $\pi_\infty(x, \cdot) = \delta_x$ (or other arbitrary definition)

**Example** (Non-homogeneous Poisson process on $\mathbb{R}^+$).

$S = \{0, 1, 2, \dots\}$, $\quad q_t(x, y) = \lambda_t \cdot \delta_{x+1,y}$,

$Y_n = n$, $\quad J_{n+1} | J_n \sim \lambda_t \cdot e^{-\int_{J_n}^t \lambda_s\, ds}\, dt$,

$N_t := \#\{n \geq 1 \ : \ J_n \leq t\}$ *the* associated counting process




***Claim:***

1. $f_{J_n}(t) = \frac{1}{(n-1)!}\left(\int_0^t \lambda_s\,ds\right)^{n-1}\lambda_t e^{-\int_0^t \lambda_s\,ds}$

2. $N_t \sim \mathrm{Poisson}\left(\int_0^t \lambda_s\,ds\right)$

*Proof.* 1. By induction:

$$f_{J_{n+1}}(t) = \int_0^t f_{J_{n-1}|J_n}(t|r) f_{J_n}(r)\,dr$$
$$= \int_0^t \lambda_t e^{-\int_r^t \lambda_s\,ds}\frac{1}{(n-1)!}\left(\int_0^r \lambda_s\,ds\right)^{n-1}\lambda_r e^{-\int_0^r \lambda_s\,ds}\,dr \;= \dots$$

2. $P[N_t \geq n] = P[J_n \leq t]$

□

**Lemma 1.9.** *(Exercise)*

1. $(Y_{n-1}, J_n)$ *is a Markov chain with respect to* $\mathcal{G}_n = \sigma(Y_0,\dots,Y_{n-1},E_1,\dots,E_n)$ *with transition functions*

$$p\left((x,s),dydt\right) = \pi_s(x,dy)\lambda_t(y)\cdot e^{-\int_s^t \lambda_r(y)\,dr} I_{(s,\infty)}(t)\,dt$$

2. $(J_n, Y_n)$ *is a Markov chain with respect to* $\tilde{\mathcal{G}} = \sigma(Y_0,\dots,Y_n,E_1,\dots,E_n)$ *with transition functions*

$$\tilde{p}\left((x,s),dtdy\right) = \lambda_t(x)\cdot e^{-\int_s^t \lambda_r(x)\,dr} I_{(s,\infty)}(t)\pi_t(x,dy)$$

**Remark .** *1.* $J_n$ *strictly increasing.*

2. $J_n = \infty\;\forall\, n,m$ *is possible* $\rightsquigarrow$ $X_t$ *absorbed in state* $Y_{n-1}$.

3. $\sup J_n < \infty$ $\rightsquigarrow$ *explosion in finite time*

4. $\{s < \zeta\} = \{X_s \neq \Delta\} \in \mathcal{F}_s \rightsquigarrow$ *no explosion before time* $s$.

$K_s := \min\{n \;:\; J_n > s\}$ first jump after time $s$. Stopping time with respect to $\mathcal{G}_n = \sigma\left(E_1,\dots,E_n,Y_0,\dots,Y_{n-1}\right)$,

$$\{K_s < \infty\} = \{s < \zeta\}$$

**Lemma 1.10** (Memoryless property)**.** *Let $s \geq t_0$. Then for all $t \geq s$,*

$$P_{(t_0,\mu)}\left[\{J_{K_s} > t\} \cap \{s < \zeta\} \mid \mathcal{F}_s\right] = e^{-\int_s^t \lambda_r(X_s)\,dr} \quad P\text{-a.s. on } \{s < \zeta\}$$

*i.e.*

$$P_{(t_0,\mu)}\left[\{J_{K_s} > t\} \cap \{s < \zeta\} \cap A\right] = \mathbb{E}_{(t_0,\mu)}\left[e^{-\int_s^t \lambda_r(X_s)\,dr}\ ;\ A \cap \{s < \zeta\}\right] \quad \forall A \in \mathcal{F}_s$$

**Remark .** *The assertion is a restricted form of the Markov property in continuous time: The conditional distribution with respect to $P_{(t_0,\mu)}$ of $J_{K_s}$ given $\mathcal{F}_s$ coincides with the distribution of $J_1$ with respect to $P_{(s,X_s)}$.*

*Proof.*

$$A \in \mathcal{F}_s \overset{\text{(Ex.)}}{\Rightarrow} A \cap \{K_s = n\} \in \sigma\left(J_0, Y_0, \dots, J_{n-1}, Y_{n-1}\right) = \tilde{\mathcal{G}}_{n-1}$$

$$\Rightarrow \quad P\left[\{J_{K_s} > t\} \cap A \cap \{K_s = n\}\right] = \mathbb{E}\left[P[J_n > t \mid \tilde{\mathcal{G}}_{n-1}]\ ;\ A \cap \{K_s = n\}\right]$$

where

$$P[J_n > t \mid \tilde{\mathcal{G}}_{n-1}] = \exp\left(-\int_{J_{n-1}}^{t} \lambda_r(Y_{n-1})\,dr\right) = \exp\left(-\int_s^t \lambda_r(\underbrace{Y_{n-1}}_{=X_s})\,dr\right) \cdot P[J_n > s \mid \tilde{\mathcal{G}}_{n-1}],$$

hence we get

$$P[J_n > t \mid \tilde{\mathcal{G}}_{n-1}] = \mathbb{E}\left[e^{-\int_s^t \lambda_r(X_s)\,dr}\ ;\ A \cap \{K_s = n\} \cap \{J_n > s\}\right] \quad \forall n \in \mathbb{N}$$

where $A \cap \{K_s = n\} \cap \{J_n > s\} = A \cap \{K_s = n\}$.
Summing over $n$ gives the assertion since

$$\{s < \zeta\} = \dot{\bigcup_{n \in \mathbb{N}}} \{K_s = n\}.$$

□

For $y_n \in S$, $t_n \in [0, \infty]$ strictly increasing define

$$x := \Phi\left((t_n, y_n)_{n=0,1,2,\dots}\right) \in \mathrm{PC}\left([t_0, \infty),\ S \dot{\cup} \{\Delta\}\right)$$

by

$$x_t := \begin{cases} Y_n & \text{for} \quad t_n \leq t < t_{n+1}\ ,\ n \geq 0 \\ \Delta & \text{for} \quad t \geq \sup t_n \end{cases}$$

Let

$$\begin{aligned}(X_t)_{t\geq t_0} &:= \Phi\left((J_n, Y_n)_{n\geq 0}\right)\\ \mathcal{F}_t^X &:= \sigma\left(X_s \mid s\in[t_0,t]\right), \quad t\geq t_0\end{aligned}$$

**Theorem 1.11** (Markov property)**.** *Let $s\geq t_0$, $X_{s:\infty} := (X_t)_{t\geq s}$. Then*

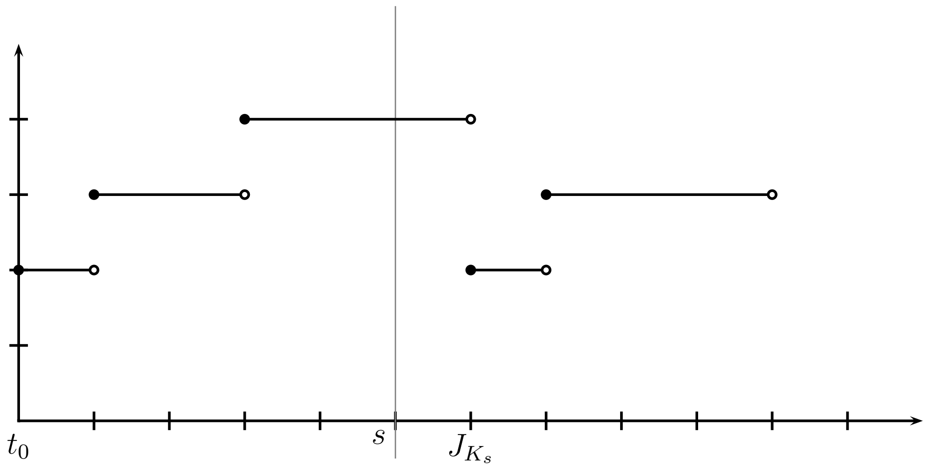
$$\mathbb{E}_{(t_0,\mu)}\left[F(X_{s:\infty})\cdot I_{\{s<\zeta\}} \mid \mathcal{F}_s^X\right](\omega) = \mathbb{E}_{(s,X_s(\omega))}\left[F(X_{s:\infty})\right] \quad P\text{-a.s. } \{s<\zeta\}$$

*for all*

$$F\colon \operatorname{PC}([s,\infty),\ S\cup\{\Delta\})\to\mathbb{R}^+$$

*measurable with respect to $\sigma\left(x\mapsto x_t \mid t\geq s\right)$.*

*Proof.* $X_{s:\infty} = \Phi(s, Y_{K_s-1}, J_{K_s}, Y_{K_s}, J_{K_s+1},\ldots)$ on $\{s<\zeta\}=\{K_s<\infty\}$




i.e. the process after time $s$ is constructed in the same way from $s, Y_{K_s-1}, J_{K_s},\ldots$ as the original process is constructed from $t_0, Y_0, J_1,\ldots$. By the Strong Markov property for the chain $(Y_{n-1}, J_n)$,

$$\begin{aligned}&\mathbb{E}_{(t_0,\mu)}\left[F(X_{s:\infty})\cdot I_{\{s<\zeta\}} \mid \mathcal{G}_{K_s}\right]\\ =&\mathbb{E}_{(t_0,\mu)}\left[F\circ\Phi(s, Y_{K_s-1}, J_{K_s},\ldots)\cdot I_{\{K_s<\infty\}} \mid \mathcal{G}_{K_s}\right]\\ =&\mathbb{E}^{\text{Markov chain}}_{(Y_{K_s-1},J_{K_s})}\left[F\circ\Phi(s,(Y_0,J_1),(Y_1,J_2),\ldots)\right] \quad \text{a.s. on } \{K_s<\infty\}=\{s<\zeta\}.\end{aligned}$$

Since $\mathcal{F}_s\subseteq\mathcal{G}_{K_s}$, we obtain by the projectivity of the conditional expectation,

$$\begin{aligned}&\mathbb{E}_{(t_0,\mu)}\left[F(X_{s:\infty})\cdot I_{\{s<\zeta\}} \mid \mathcal{F}_s\right]\\ =&\mathbb{E}_{(t_0,\mu)}\left[\mathbb{E}^{\text{Markov chain}}_{(X_s,J_{K_s})}\left[F\circ\Phi(s,(Y_0,J_1),\ldots)\cdot I_{\{s<\zeta\}} \mid \mathcal{F}_s\right]\right]\end{aligned}$$

taking into account that the conditional distribution given $\mathcal{G}_{K_s}$ is 0 on $\{s \geq \zeta\}$ and that $Y_{K_s-1} = X_s$.
Here the conditional distribution of $J_{K_s}$ ist $k(X_s, \cdot)$, by Lemma 1.10

$$k(x, dt) = \lambda_t(x) \cdot e^{-\int_s^t \lambda_r(x)\, dr} \cdot I_{(s,\infty)}(t)\, dt$$

hence

$$\mathbb{E}_{(t_0,\mu)}\left[F(X_{s:\infty}) \cdot I_{\{s<\zeta\}} \mid \mathcal{F}_s\right] = \mathbb{E}^{\text{Markov chain}}_{(X_s,k(X_s,\cdot))}\left[F \circ \Phi(\ldots)\right] \quad \text{a.s. on } \{s < \zeta\}$$

Here $k(X_s, \cdot)$ is the distribution of $J_1$ with respect to $P_{s,X_s}$, hence we obtain

$$\begin{aligned}\mathbb{E}_{(t_0,\mu)}\left[F(X_{s:\infty}) \cdot I_{\{s<\zeta\}} \mid \mathcal{F}_s\right] &= \mathbb{E}_{(s,X_s)}\left[F(\Phi(s, Y_0, J_1, \ldots))\right] \\ &= \mathbb{E}_{(s,X_s)}\left[F(X_{s:\infty})\right]\end{aligned}$$

□

**Corollary 1.12.** *A non-homogeneous Poisson process $(N_t)_{t\geq 0}$ with intensity $\lambda_t$ has independent increments with distribution*

$$N_t - N_s \sim \text{Poisson}\left(\int_s^t \lambda_r\, dr\right)$$

*Proof.*

$$\begin{aligned}P\left[N_t - N_s \geq k \mid \mathcal{F}_s^N\right] &\overset{\text{MP}}{=} P_{(s,N_s)}\left[N_t - N_s \geq k\right] = P_{(s,N_s)}\left[J_k \leq t\right] \\ &\overset{\text{as above}}{=} \text{Poisson}\left(\int_s^t \lambda_r\, dr\right)(\{k, k+1, \ldots\})\,.\end{aligned}$$

Hence $N_t - N_s$ independent of $\mathcal{F}_s^N$ and Poisson $\left(\int_s^t \lambda_r\, dr\right)$ distributed. □

Recall that the ***total variation norm*** of a signed measure $\mu$ on $(S, \mathscr{S})$ is given by

$$\|\mu\|_{\text{TV}} = \mu^+(S) + \mu^-(S) = \sup_{|f|\leq 1} \int f\, dr$$

**Theorem 1.13.** *1. Under $P_{(t_0,\mu)}$, $(X_t)_{t\geq t_0}$ is a Markov jump process with initial distribution $X_{t_0} \sim \mu$ and transition probabilities*

$$p_{s,t}(x, B) = P_{(s,x)}\left[X_t \in B\right] \quad (0 \leq s \leq t,\ x \in S,\ B \in \mathscr{S})$$

*satisfying the Chapman-Kolmogorov equations $p_{s,t}p_{t,u} = p_{s,u} \quad \forall\, 0 \leq s \leq t \leq u$.*

2. *The **integrated backward equation***

$$p_{s,t}(x,B) = e^{-\int_s^t \lambda_r(x)\,dr}\delta_x(B) + \int_s^t e^{-\int_s^r \lambda_u(x)\,du}(q_r p_{r,t})(x,B)\,dr \tag{1.1}$$

   *holds for all* $0 \leq s \leq t$ , $x \in S$ *and* $B \in \mathscr{S}$.

3. *If* $t \mapsto \lambda_t(x)$ *is continuous for all* $x \in S$, *then*

$$(p_{s,s+h}f)(x) = (1 - \lambda_s(x)\cdot h)f(x) + h\cdot(q_s f)(x) + o(h) \tag{1.2}$$

   *holds for all* $s \geq 0$ , $x \in S$ *and bounded functions* $f\colon \ \to \mathbb{R}$ *such that* $t \mapsto (q_t f)(x)$ *is continuous.*

**Remark .** 1. *(1.2) shows that* $(X_t)$ *is the continuous time Markov chain with intensities* $\lambda_t(x)$ *and transition rates* $q_t(x,dy)$.

2. *If* $\zeta = \sup J_n$ *is finite with strictly positive probability, then there are other possible continuations of* $X_t$ *after the explosion time* $\zeta$.

   $\rightsquigarrow$ *non-uniqueness.*

   *The constructed process is called the **minimal chain** for the given jump rates, since its transition probabilities* $p_t(x,B)$ , $B \in \mathscr{S}$ *are minimal for all continuations, cf. below.*

3. *The integrated backward equation extends to bounded functions* $f\colon S \to \mathbb{R}$

$$(p_{s,t}f)(x) = e^{-\int_s^t \lambda_r(x)\,dr} f(x) + \int_s^t e^{-\int_s^r \lambda_u(x)\,du}(q_r p_{r,t} f)(x)\,dr \tag{1.3}$$

*Proof.* 1. By the Markov property,

$$P_{(t_0,\mu)}\left[X_t \in B \middle| \mathcal{F}_s^X\right] = P_{(s,X_s)}\left[X_t \in B\right] = p_{s,t}(X_s,B) \quad \text{a.s.}$$

since $\{X_t \in B\} \subseteq \{t < \zeta\} \subseteq \{s < \zeta\}$ for all $B \in \mathscr{S}$ and $0 \leq s \leq t$.

Thus $\left((X_t)_{t \geq t_0}\ ,\ P_{(t_0,\mu)}\right)$ is a Markov jump process with transition kernels $p_{s,t}$. Since this holds for any initial condition, the Chapman-Kolmogorov equations

$$(p_{s,t}p_{t,u}f)(x) = (p_{s,u}f)(x)$$

are satisfied for all $x \in S$ , $0 \leq s \leq t \leq u$ and $f\colon S \to \mathbb{R}$.

2. ***First step analysis***: Condition on $\tilde{G}_1 = \sigma(J_0, Y_0, J_1, Y_1)$:

   Since $X_t = \Phi_t(J_0, Y_0, J_1, Y_1, J_2, Y_2, \ldots)$, the Markov property of $(J_n, Y_n)$ implies

   $$P_{(s,x)}\left[X_t \in B | \tilde{G}_1\right](\omega) = P_{(J_1(\omega), Y_1(\omega))}\left[\Phi_t(s, x, J_0, Y_0, J_1, Y_1, \ldots) \in B\right]$$

   On the right side, we see that

   $$\Phi_t(s, x, J_0, Y_0, J_1, Y_1, \ldots) = \begin{cases} x & \text{if } t < J_1(\omega) \\ \Phi_t(J_0, Y_0, J_1, Y_1, \ldots) & \text{if } t \geq J_1(\omega) \end{cases}$$

   and hence

   $$P_{(s,x)}\left[X_t \in B | \tilde{G}_1\right](\omega) = \delta_x(B) \cdot I_{\{t < J_1\}}(\omega) + P_{(J_1(\omega), Y_1(\omega))}[X_t \in B] \cdot I_{\{t \geq J_1\}}(\omega)$$

   $P_{(s,x)}$-a.s. We conclude

   $$\begin{aligned} p_{s,t}(x, B) &= P_{(s,x)}[X_t \in B] \\ &= \delta_x(B) P_{(s,x)}[J_1 > t] + \mathbb{E}_{(s,x)}[p_{J_1,t}(Y_1, B); t \geq J_1] \\ &= \delta_x(B) \cdot e^{-\int_s^t \lambda_r(x)\,dr} + \int_s^t \lambda_r(x) e^{-\int_s^r \lambda_u(x)\,du} \int \underbrace{\pi_r(x, dy) p_{r,t}(y, B)}_{=(\pi_r p_{r,t})(x,B)}\,dr \\ &= \delta_x(B) \cdot e^{-\int_s^t \lambda_r(x)\,dr} + \int_s^t e^{-\int_s^r \lambda_u(x)\,du} (q_r p_{r,t})(x, B)\,dr \end{aligned}$$

3. This is a direct consequence of (1.1).

   Fix a bounded function $f\colon S \to \mathbb{R}$. Note that

   $$0 \leq (q_r p_{r,t} f)(x) = \lambda_r(x)(\pi_r p_{r,t} f)(x) \leq \lambda_r(x) \sup |f|$$

   for all $0 \leq r \leq t$ and $x \in S$. Hence if $r \mapsto \lambda_r(x)$ is continuous (and locally bounded) for all $x \in S$, then

   $$(p_{r,t} f)(x) \longrightarrow f(x) \tag{1.4}$$

   as $r, t \downarrow s$ for all $x \in S$.

   Thus by dominated convergence,

   $$\begin{aligned} &(q_r p_{r,t} f)(x) - (q_s f)(x) \\ &= \int q_r(x, dy)(p_{r,t} f(y) - f(y)) + (q_r f)(x) - (q_s f)(x) \longrightarrow 0 \end{aligned}$$

   as $r, t \downarrow s$ provided $r \mapsto (q_r f)(x)$ is continuous. The assertion now follows from (1.3).

□

**Theorem 1.14** (A first non-explosion criterion). *If* $\bar{\lambda} := \sup_{\substack{t \geq 0 \\ x \in S}} \lambda_t(x) < \infty$*, then*

$$\zeta = \infty \quad P_{(t_0,\mu)}\text{-a.s.} \; \forall \, t_0, \; \mu$$

*Proof.*

$$\begin{aligned} J_n &= \inf \left\{ t \geq 0 \; : \; \int_{J_{n-1}}^{t} \overbrace{\lambda_s(Y_{n-1})}^{\leq \bar{\lambda}} \, ds \geq E_n \right\} \\ &\geq J_{n-1} + \bar{\lambda}^{-1} E_n \\ \implies \quad \zeta &= \sup J_n \geq \bar{\lambda}^{-1} \sum_{n-1}^{\infty} E_n = \infty \quad \text{a.s.} \end{aligned}$$

□

**Remark .** *In the time-homogeneous case,*

$$J_n = \sum_{k=1}^{n} \frac{E_k}{\lambda(Y_{n-1})}$$

*is a sum of conditionally independent exponentially distributed random variables given* $\{Y_k \mid k \geq 0\}$.
*From this one can conclude that the events*

$$\{\zeta < \infty\} = \left\{ \sum_{k=1}^{\infty} \frac{E_k}{\lambda(Y_{k-1})} < \infty \right\} \text{ and } \left\{ \sum_{k=0}^{\infty} \frac{1}{\lambda(Y_k)} < \infty \right\}$$

*coincide almost sure (apply Kolmogorov's 3-series Theorem).*

## 1.3 Forward and Backward Equations

**Definition 1.15.** *The* ***infinitesimal generator*** *(or* ***intensity matrix***, ***kernel****) of a Markov jump process at time t is defined by*

$$\mathscr{L}_t(x, dy) = q_t(x, dy) - \lambda_t(x)\delta_x(dy)$$

*i.e.*

$$\begin{aligned} (\mathscr{L}_t f)(x) &= (q_t f)(x) - \lambda_t(x) f(x) \\ &= \int q_t(x, dy) \cdot (f(y) - f(x)) \end{aligned}$$

*for all bounded and measurable* $f \colon S \to \mathbb{R}$.

**Remark .** 1. *$\mathscr{L}_t$ is a linear operator on functions $f\colon S \to \mathbb{R}$.*

2. *If $S$ is discrete, $\mathscr{L}_t$ is a matrix, $\mathscr{L}_t(x,y) = q_t(x,y) - \lambda_t(x)\delta(x,y)$. This matrix is called **Q-Matrix**.*

**Example .** *Random walk on $\mathbb{Z}^d$:*

$$q_t(x,y) = \begin{cases} \frac{1}{2d} & \text{if } |x-y| = 1 \\ 0 & \text{else} \end{cases}$$

*then*

$$(\mathscr{L}_t f)(x) = \frac{1}{2d}\sum_{k=1}^{d} (f(x+e_i) + f(x-e_i) - 2f(x)) = \frac{1}{2d}\left(\Delta_{\mathbb{Z}^d} f\right)(x)$$

**Theorem 1.16** (Kolmogorov's backward equation)**.** *If $t \mapsto q_t(x,\cdot)$ is continuous in total variation norm for all $x \in S$, then the transition kernels $p_{s,t}$ of the Markov jump process constructed above are the minimal solutions of the **backward equation** (BWE)*

$$-\frac{\partial}{\partial s}\left(p_{s,t} f\right)(x) = -\left(\mathscr{L}_s p_{s,t} f\right)(x) \quad \textit{for all bounded } f\colon S \to \mathbb{R},\ 0 \le s \le t \tag{1.5}$$

*with terminal condition $(p_{t,t} f)(x) = f(x)$.*

*Proof.* 1. **$p_{s,t} f$ solves (1.5):**

Informally, the assertion follows by differentiating the integrated backward equation (1.3) with respect to $s$. To make this rigorous, we proceed similarly as in the proof of (1.2). First, one shows similarly to the derivation of (1.4) that

$$(p_{r,t} f)(x) \longrightarrow (p_{s,t} f)(x) \quad \text{as } r \to s$$

for all $0 \le s \le t$, $x \in S$, and bounded functions $f\colon S \to \mathbb{R}$. This combined with the assumption implies that also

$$\begin{aligned} &|(q_r p_{r,t} f)(x) - (q_s p_{s,t} f)(x)| \\ \le\ &\|q_r(x,\cdot) - q_s(x,\cdot)\|_{\mathrm{TV}} \cdot \sup|p_{s,t} f| + \int q_s(x,dy)\left(p_{r,t} f(y) - p_{s,t} f(y)\right) \\ &\longrightarrow\ 0 \quad \text{as } r \to s \text{ for all } x \in S \end{aligned}$$

by dominated convergence, because $\sup|p_{r,t} f| \le \sup|f|$. Hence the integrand in (1.3) is continuous in $r$ at $s$, and so there exists

$$-\frac{\partial}{\partial s}\left(p_{s,t} f\right)(x) = -\lambda_s(x)(p_{s,t} f)(x) + (q_s p_{s,t} f)(x) = \left(\mathscr{L}_s p_{s,t} f\right)(x)$$

2. **Minimality:**

   Let $(x,B) \mapsto \tilde{p}_{s,t}(x,B)$ be an arbitrary non-negative solution of (1.5). Then

$$\begin{aligned} &-\frac{\partial}{\partial r}\tilde{p}_{r,t}(x,B) = (q_r\tilde{p}_{r,t})(x,B) - \lambda_r(x)\tilde{p}_{r,t}(x,B) \\ &\Rightarrow -\frac{\partial}{\partial r}\left(e^{\int_r^s \lambda_u(x)\,du}\tilde{p}_{r,t}(x,B)\right) = e^{-\int_s^r \lambda_u(x)\,du}(q_r\tilde{p}_r,t)(x,B) \\ &\overset{\text{integrate}}{\Rightarrow} \tilde{p}_{s,t}(x,B) - e^{-\int_s^t \lambda_u(x)\,du}\delta_x(B) = \int_s^t e^{-\int_s^r \lambda_u(x)\,du}(q_r\tilde{p}_{r,t})(x,B)\,dr \end{aligned}$$

   integrated backward-equation.

   **Claim:**

$$\tilde{p}_{s,t}(x,B) \geq p_{s,t}(x,B) = P_{(s,x)}[X_t \in B] \quad \forall\, x \in S,\ B \in \mathscr{S}$$

   This is OK if

$$P_{(s,x)}[X_t \in B, t < J_n] \leq \tilde{p}_{s,t}(x,B) \quad \forall\, n \in \mathbb{N}$$

   $\boldsymbol{n = 0}$ : ✓

   $\boldsymbol{n \to n+1}$: by First step analysis:

$$\begin{aligned} &P_{(s,x)}[X_t \in B, t < J_{n+1} \mid J_1, Y_1] \\ &\overset{\text{MP}}{=} \delta_x(B) \cdot I_{\{t<J_1\}} + P_{(J_1,Y_1)}[X_t \in B, t < J_n] \cdot I_{\{t \geq J_1\}} \end{aligned}$$

   where by induction

$$P_{(J_1,Y_1)}[X_t \in B, t < J_n] \leq \tilde{p}_{J_n,t}(Y_1,B)$$

   Hence we conclude with the integrated backward equation

$$\begin{aligned} &P_{(s,x)}[X_t \in B, t < J_{n+1}] \leq \mathbb{E}_{(s,x)}[\ldots] \\ &= \delta_x(B)e^{-\int_s^t \lambda_r(x)\,dr} + \int_s^t e^{-\int_r^s \lambda_u(x)\,du}(q_r\tilde{p}_{r,t})(x,B)\,dr \\ &\leq \tilde{p}_{s,t}(B) \end{aligned}$$

□

**Remark .** *1. (1.5) describes the backward evolution of the expectation values $\mathbb{E}_{(s,x)}[f(X_t)]$ respectively the probabilities $P_{(s,x)}[X_t \in B]$ when varying the starting times $s$.*

*2. In a **discrete state space**, (1.5) reduces to*

$$-\frac{\partial}{\partial s}p_{s,t}(x,z) = \sum_{y \in S} \mathscr{L}_s(x,y)p_{s,t}(y,z), \qquad p_{t,t}(x,z) = \delta_{xz}$$

*a system of ordinary differential equations.*

*For $S$ being finite,*

$$p_{s,t} = \exp\left(\int_s^t \mathscr{L}_r \, dr\right) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\int_s^t \mathscr{L}_r \, dr\right)^n$$

*is the unique solution.*

*If $S$ is infinite, the solution is not necessarily unique (hence the process is not unique).*

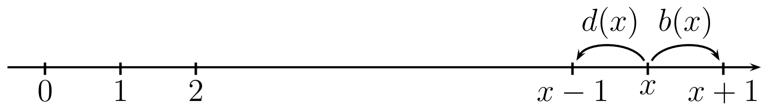
3. ***Time-homogeneous case:*** $p_{s,t} = p_{t-s}, \mathscr{L}_s = \mathscr{L}$. *The backward-equation (1.5) then becomes*

$$\frac{d}{dt}(p_t f)(x) = (\mathscr{L} p_t f)(x), \qquad p_0 f = f$$

*If $S$ is infite, $p_t = e^{t\mathscr{L}}$. In particular, if $\mathscr{L}$is diagonalizable with left/right eigenvectors $u_i, v_j$ s.t. $u_i^T v_j = \delta_{ij}$ and eigenvalues $\lambda_i$, then*

$$\mathscr{L} = \sum_{i=1}^{n} \lambda_i v_i \otimes u_i^T, \qquad p_t = \sum_{i=1}^{n} e^{t\lambda_i} v_i \otimes u_i^T$$

**Example .** 1. ***General birth-and-death process***




$$\begin{aligned}\frac{d}{dt} p_t(x,z) &= \sum_{|x-y|=1} q(x,y)\left(p_t(y,z) - p_t(x,z)\right) \\ &= b(x)\left(p_t(x+1,z) - p_t(x,z)\right) + d(x)\left(p_t(x-1,z) - p_t(x,z)\right) \\ p_0(x,z) &= \delta_{xz}\end{aligned}$$

2. ***Continuous-time branching process***

*The particles in a population die with rate $d > 0$ and divide into two particles with rate $b > 0$, independently from each other.*

$$X_t = \text{total number of particles at time } t$$

*is a birth-death process on $S = \{0, 1, 2, \ldots\}$ with total birth/death rates*

$$b(n) = q(n, n+1) = n \cdot b, \qquad d(n) = q(n, n-1) = n \cdot d, \lambda(n) = n \cdot (b + d)$$

*Let*

$$\eta(t) := P_1[X_t = 0] = p_t(1,0)$$

*the extinction probability. Equation (1.5) gives*

$$\begin{aligned}\eta'(t) &= dp_t(0,0) - (b+d)p_t(1,0) + b_t(2,0)\\ &= d - (b+d)\eta(t) + b\eta(t)^2,\\ \eta(0) &= 0\end{aligned}$$

*Hence we get*

$$P_1[X_t \neq 0] = 1 - \eta(t) = \begin{cases} \frac{1}{1+bt} & \text{if} \quad b = d \\ \frac{b-d}{b-d\cdot e^{t(d-b)}} & \text{if} \quad b \neq d \end{cases}$$

*i.e.*

- *exponentially decay if $d > b$*
- *polynomial decay if $d = b$ (critical case)*
- *strictly positive survival probability if $d < b$*

**Theorem 1.17** (Kolmogorov's forward equation)**.** *Suppose that*

$$\bar{\lambda}_t = \sup_{0 \le s \le t} \sup_{x \in S} \lambda_s(x) < \infty$$

*for all $t > 0$. Then the **forward equation***

$$\frac{d}{dt}(p_{s,t}f)(x) = (p_{s,t}\mathscr{L}_t f)(x), \qquad\qquad (p_{s,s}f)(x) = f(x) \tag{1.6}$$

*holds for all $0 \le s \le t,\ x \in S$ and all bounded functions $f\colon S \to \mathbb{R}$ such that $t \mapsto (q_t f)(x)$ and $t \mapsto \lambda_t(x)$ are continuous for all $x$.*

*Proof.* 1. ***Strong continuity:*** Fix $t_0 > 0$. Note that $\|q_r g\|_{\sup} \le \bar{\lambda}_r \|f\|_{\sup}$ for all $0 \le r \le t_0$. Hence by the assumption and the integrated backward equation (1.3),

$$\begin{aligned}\|p_{s,t}f - p_{s,r}f\|_{\sup} &= \|p_{s,r}(p_{r,t}f - f)\|_{\sup}\\ \le&\|p_{r,t}f - f\|_{\sup} \le \varepsilon(t-r)\cdot\|f\|_{\sup}\end{aligned}$$

for all $0 \le s \le r \le t \le t_0$ and some function $\varepsilon\colon \mathbb{R}^+ \to \mathbb{R}^+$ with $\lim_{h \downarrow 0} \varepsilon(h) = 0$.

2. ***Differentiability:*** By 1.) and the assumption,

$$(r, u, x) \mapsto (q_r p_{r,u} f)(x)$$

is uniformly bounded for $0 \le r \le u \le t_0$ and $x \in S$, and

$$q_r p_{r,u} f = \underbrace{q_r(p_{r,u}f - f)}_{\longrightarrow 0 \text{ uniformly}} + q_r f \longrightarrow q_t f$$

pointwise as $r, u \longrightarrow t$. Hence by the integrated backward equation (1.3) and the continuity of $t \mapsto \lambda_t(x)$,

$$\frac{p_{t,t+h}f(x) - f(x)}{h} \xrightarrow{h\downarrow 0} -\lambda_t(x)f(x) + q_t f(x) = \mathscr{L}_t f(x)$$

for all $x \in S$, and the difference quotients are uniformly bounded.

Dominated convergence now implies

$$\frac{p_{s,t+h}f - p_{s,t}f}{h} = p_{s,t}\frac{p_{t,t+h}f - f}{h} \longrightarrow p_{s,t}\mathscr{L}_t f$$

pointwise as $h \downarrow 0$. A similar argument shows that also

$$\frac{p_{s,t}f - p_{s,t-h}f}{h} = p_{s,t-h}\frac{p_{t-h,t}f - f}{h} \longrightarrow p_{s,t}\mathscr{L}_t f$$

pointwise.

□

**Remark .** *1. The assumption implies that the operators $\mathscr{L}_s, 0 \leq s \leq t_0$, are uniformly bounded with respect to the supremum norm:*

$$\|\mathscr{L}_s f\|_{\sup} \leq \lambda_t \cdot \|f\|_{\sup} \quad \forall\, 0 \leq s \leq t.$$

*2. Integrating (1.5) yields*

$$p_{s,t}f = f + \int_s^t p_{s,r}\mathscr{L}_r f\, dr \tag{1.7}$$

*In particular, the difference quotients $\frac{p_{s,t+h}f - p_{s,t}f}{h}$ converge uniformly for $f$ as in the assertion.*

**Notation:**

$$< \mu, f >:= \mu(f) = \int f\, d\mu$$

$\mu \in M_1(S),\ s \geq 0,\ \mu_t := \mu p_{s,t} = P_{(s,\mu)} \circ X_t^{-1}$ *mass distribution at time $t$*

**Corollary 1.18** (Fokker-Planck equation)**.** *Under the assumptions in the theorem,*

$$\frac{d}{dt} < \mu_t, f >=< \mu_t, \mathscr{L}_t f >$$

*for all $t \geq s$ and bounded functions $f\colon S \to \mathbb{R}$ such that $t \mapsto q_t f$ and $t \mapsto \lambda_t$ are pointwise continuous. Abusing notation, one sometimes writes*

$$\frac{d}{dt}\mu_t = \mathscr{L}_t^* \mu_t$$

*Proof.*

$$< \mu_t, f >=< \mu p_{s,t}, f >= \int \mu(dx) \int p_{s,t}(x, dy) f(y) =< \mu, p_{s,t} f >$$

hence we get

$$\frac{< \mu_{t+h}, f > - < \mu_t, f >}{h} =< \mu p_{s,t}, \frac{p_{t,t+h} f - f}{h} > \longrightarrow < \mu_t, \mathscr{L}_t f >$$

as $h \downarrow 0$ by dominated convergence. □

**Remark .** ***(Important!)***

$$\begin{aligned} &P_{(s,\mu)}[\zeta < \infty] > 0 \\ &\Rightarrow < \mu_t, 1 >= \mu_t(S) < 1 \quad \textit{for large } t \end{aligned}$$

*hence the Fokker-Planck equation does* not *hold for* $f \equiv 1$*:*

$$\underbrace{< \mu_t, 1 >}_{<1} < \underbrace{< \mu, 1 >}_{=1} + \int_0^t < \mu_s, \mathscr{L}_s 1 > \, ds$$

*where* $\mathscr{L}_s 1 = 0$.

**Example .** *Birth process on* $S = \{0, 1, 2, \ldots\}$

$$\begin{aligned} &q(i, j) = \begin{cases} b(i) & \text{if } j = i + 1 \\ 0 & \text{else} \end{cases} \\ &\pi(i, j) = \delta_{i+1,j}, \\ &Y_n = n, \\ &S_n = J_n - J_{n-1} \sim \text{Exp}(b(n-1)) \quad \textit{independent,} \\ &\zeta = \sup J_n = \sum_{n=1}^{\infty} S_n < \infty \iff \sum_{n=1}^{\infty} b(n)^{n-1} < \infty \end{aligned}$$

*In this case, Fokker-Planck does not hold.*

The question whether one can extend the forward equation to unbounded jump rates leads to the *martingale problem*.

## 1.4 The martingale problem

**Definition 1.19.** *A Markov process $(X_t, P_{(s,x)} \mid 0 \leq s \leq t, x \in S)$ is called* ***non-explosive*** *(or* ***conservative****) if and only if $\zeta = \infty \quad P_{(s,x)}$-a.s. for all $s, x$.*

Now we consider again the minimal jump process $(X_t, P_{(t_0,\mu)})$ constructed above. A function

$$\begin{aligned} f\colon [0,\infty) \times S &\to \mathbb{R} \\ (t,x) &\mapsto f_t(x) \end{aligned}$$

is called ***locally bounded*** if and only if there exists an increasing sequence of open subsets $B_n \subseteq S$ such that $S = \bigcup B_n$, and

$$\sup_{\substack{x \in B_n \\ 0 \leq s \leq t}} |f_s(x)| < \infty$$

for all $t > 0,\ n \in \mathbb{N}$.
The following theorem gives a probabilistic form of Kolmogorov's forward equation:

**Theorem 1.20** (Time-dependent martingale problem)**.** *Suppose that $t \mapsto \lambda_t(x)$ is continuous for all $x$. Then:*

1. *The process*

$$M_t^f := f_t(X_t) - \int\limits_{t_0}^{t} \left( \frac{\partial}{\partial r} + \mathscr{L}_r \right) f_r(X_r)\, dr, \quad t \geq t_0$$

*is a local $(\mathcal{F}_t^X)$-martingale up to $\zeta$ with respect to $P_{(t_0,\mu)}$ for any locally bounded function $f\colon \mathbb{R}^+ \times S \to \mathbb{R}$ such that $t \mapsto f_t(x)$ is $C^1$ for all $x$, $(t,x) \mapsto \frac{\partial}{\partial t} f_t(x)$ is locally bounded, and $r \mapsto (q_{r,t} f_t)(x)$ is continuous at $r = t$ for all $t, x$.*

2. *If $\bar{\lambda}_t < \infty$ and $f$ and $\frac{\partial}{\partial t} f$ are bounded functions, then $M^f$ is a global martingale.*

3. *More generally, if the process is non-explosive then $M^f$ is a global martingale provided*

$$\sup_{\substack{x \in S \\ t_0 \leq s \leq t}} \left( |f_s(x)| + \left| \frac{\partial}{\partial s} f_s(x) \right| + |(\mathscr{L}_s f_s)(x)| \right) < \infty \tag{1.8}$$

*for all $t > t_0$.*

**Corollary 1.21.** *If the process is conservative then the forward equation*

$$p_{s,t} f_t = f_s + \int\limits_{s}^{t} p_{r,t} \left( \frac{\partial}{\partial r} + \mathscr{L}_r \right) f_r \, dr, \quad t_0 \leq s \leq t \tag{1.9}$$

*holds for functions $f$ satisfying (1.8).*

*Proof of corollary.* $M^f$ being a martingale, we have

$$(p_{s,t}f_r)(x) = \mathbb{E}_{(s,x)}[f_t(X_t)] = \mathbb{E}_{(s,x)}\left[f_s(X_s) + \int_s^t \left(\frac{\partial}{\partial r} + \mathscr{L}_r\right) f_r(X_r)\,dr\right]$$
$$= f_s(x) + \int_s^t p_{s,r}\left(\frac{\partial}{\partial r} + \mathscr{L}_r\right) f_r(x)\,dr$$

for all $x \in S$. □

**Remark .** *The theorem yields the Doob-Meyer decomposition*

$$f_t(X_t) = \textit{local martingale} + \textit{bounded variation process}$$

**Remark .** 1. *Time-homogeneous case:*

*If $h$ is an harmonic function, i.e. $\mathscr{L}h = 0$, then $h(X_t)$ is a martingale*

2. *In general:*

   *If $h_t$ is* space-time harmonic, *i.e. $\frac{\partial}{\partial t}h_t + \mathscr{L}_t h_t = 0$, then $h(X_t)$ is a martingale. In particular, $(p_{s,t}f)(X_t),\ (t \geq s)$ is a martingale for all bounded functions $f$.*

3. *If $h_t$ is superharmonic (or* excessive*), i.e. $\frac{\partial}{\partial t}h_t + \mathscr{L}_t h_t \leq 0$, then $h_t(X_t)$ is a supermartingale. In particular, $\mathbb{E}[h_t(X_t)]$ is decreasing*

   $\rightsquigarrow$ *stochastic Lyapunov function, stability criteria*

   *e.g.*

$$h_t(x) = e^{-tc}h(tc), \qquad \mathscr{L}_t h \leq ch$$

*Proof of theorem.* 2. Similarly to the derivation of the forward equation, one shows that the assumption implies

$$\frac{\partial}{\partial t}(p_{s,t}f_t)(x) = (p_{s,t}\mathscr{L}_t f_t)(x) + \left(p_{s,t}\frac{\partial}{\partial t}f_t\right)(x) \qquad \forall\, x \in S,$$

or, in a integrated form,

$$p_{s,t}f_t = f_s + \int_s^t p_{s,r}\left(\frac{\partial}{\partial r} + \mathscr{L}_r\right) f_r\,dr$$

for all $0 \leq s \leq t$. Hence by the Markov property, for $t_0 \leq s \leq t$,

$$\begin{aligned}
&\mathbb{E}_{(t_0,\mu)}[f_t(X_t) - f_s(X_s) \mid \mathcal{F}_s^X] \\
=&\mathbb{E}_{(s,X_s)}[f_t(X_t) - f_s(X_s)] = (p_{s,t}f_t)(X_s) - f_s(X_s) \\
=&\int_s^t \left(p_{s,r}\left(\frac{\partial}{\partial r} + \mathscr{L}_r\right) f_r\right)(X_s)\, dr \\
=&\mathbb{E}_{(t_0,\mu)}\left[\int_s^t \left(\frac{\partial}{\partial r} + \mathscr{L}_r\right) f_r(X_r)\, dr \mid \mathcal{F}_r^X\right],
\end{aligned}$$

because all the integrands are uniformly bounded.

1. For $k \in \mathbb{N}$ let

$$q_t^{(k)}(x, B) := (\lambda_t(x) \wedge k) \cdot \pi_t(x, B)$$

denote the jump rates for the process $X_t^{(k)}$ with the same transition probabilities as $X_t$ and jump rates cut off at $k$. By the construction above, the process $X_t^{(k)},\ k \in \mathbb{N}$, and $X_t$ can be realized on the same probability space in such a way that

$$X_t^{(k)} = X_t \quad \text{a.s. on } \{t < T_k\}$$

where

$$T_k := \inf\{t \geq 0 \ :\ \lambda_t(X_t) \geq k,\ X_t \notin B_k\}$$

for an increasing sequence $B_k$ of open subsets of $S$ such that $f$ and $\frac{\partial}{\partial t} f$ are bounded on $[0, t] \times B_k$ for all $t, k$ and $S = \bigcup B_k$. Since $t \mapsto \lambda_t(X_t)$ is piecewise continuous and the jump rates do not accumulate before $\zeta$, the function is locally bounded on $[0, \zeta)$. Hence

$$T_k \nearrow \zeta \quad \text{a.s. as } k \to \infty$$

By the theorem above,

$$M_t^{f,k} = f_t(X_t^{(k)}) - \int_{t_0}^t \left(\frac{\partial}{\partial r} + \mathscr{L}_r^{(k)}\right) f_r(X_r^{(k)})\, dr, \qquad t \geq t_0,$$

is a martingale with respect to $P_{(t_0,\mu)}$, which coincides a.s. with $M_t^f$ for $t < T_k$. Hence $M_t^f$ is a local martingale up to $\zeta = \sup T_k$.

3. If $\zeta = \sup T_k = \infty$ a.s. and $f$ satisfies (1.8), then $(M_t^f)_{t \geq 0}$ is a bounded local martingale, and hence, by dominated convergence, a martingale.

□

**Theorem 1.22** (Lyapunov condition for non-explosion)**.** *Let $B_n,\ n \in \mathbb{N}$ be an increasing sequence of open subsets of $S$ such that $S = \bigcup B_n$ and the intensities $(s,x) \mapsto \lambda_s(x)$ are bounded on $[0,t] \times B_n$ for all $t \geq 0$ and $n \in \mathbb{N}$. Suppose that there exists a function $\varphi\colon \mathbb{R}^+ \times S \to \mathbb{R}$ satisfying the assumption in Part 1) of the theorem above such that for all $t \geq 0$ and $x \in S$,*

$$(i) \quad \varphi_t(x) \geq 0 \qquad \text{non-negative}$$

$$(ii) \quad \inf_{\substack{0 \leq s \leq t \\ x \in B_n^c}} \varphi_s(x) \longrightarrow 0 \quad \text{as } n \to \infty \qquad \text{tends to infinity}$$

$$(iii) \quad \frac{\partial}{\partial t}\varphi_t(x) + \mathscr{L}_t \varphi_t(x) \leq 0 \qquad \text{superharmonic}$$

*Then the minimal Markov jump process constructed above is non-explosive.*

*Proof.* ***Claim:*** $P[\zeta > t] = 1$ for all $t \geq 0$ and any initial condition.
First note that since the sets $B_n^c$ are closed, the hitting times

$$T_n = \inf\{t \geq 0 \mid X_t \in B_n^c\}$$

are $(\mathcal{F}_t^X)$-stopping times and $X_{T_n} \in B_n^c$ whenever $T_n < \infty$. As the intensities are uniformly bounded on $[0,t] \times B_n$ for all $t, n$, the process can not explode before time $t$ without hitting $B_n^c$ (Exercise),i.e. $\zeta \geq T_n$ almost sure for all $n \in \mathbb{N}$. By (iii), the process $\varphi_t(X_t)$ is a local supermartingale up to $\zeta$. Since $\varphi_t \geq 0$, the stopped process

$$\varphi_{t \wedge T_n}(X_{t \wedge T_n})$$

is a supermartingale by Fatou. Therefore, for $0 \leq s \leq t$ and $x \in S$,

$$\begin{aligned}
\varphi_s(x) &\geq \mathbb{E}_{(s,x)}\left[\varphi_{t \wedge T_n}(X_{t \wedge T_n})\right] \\
&\geq \mathbb{E}_{(s,x)}\left[\varphi_{T_n}(X_{T_n})\ ;\ T_n \leq t\right] \\
&\geq P_{(s,x)}\left[T_n \leq t\right] \cdot \inf_{\substack{s \leq r \leq t \\ x \in B_n^c}} \varphi_r(x)
\end{aligned}$$

Hence by (ii), we obtain

$$P_{(s,x)}\left[\zeta \leq t\right] \leq \liminf_{n \to \infty} P_{(s,x)}[T_n \leq t] = 0$$

for all $t \geq 0$. □

**Remark .** 1. *If there exists a function $\psi\colon S \to \mathbb{R}$ and $\alpha > 0$ such that*

$$(i) \quad \psi \geq 0$$

$$(ii) \quad \inf_{x \in B_n^c} \psi(x) \longrightarrow \infty \quad \text{as } n \to \infty$$

$$(iii) \quad \mathscr{L}_t \psi \leq \alpha \psi \quad \forall\, t \geq 0$$

*then the theorem applies with $\varphi_t(x) = e^{-\alpha t}\psi(x)$:*

$$\frac{\partial}{\partial t}\varphi_t + \mathscr{L}_t \varphi_t \leq -\alpha \varphi_t + \alpha \varphi_t \leq 0$$

*This is a standard criterion in the time-homogeneous case!*

2. *If $S$ is a locally compact connected metric space and the intensities $\lambda_t(x)$ depend continuously on $s$ and $x$, then we can choose the sets*

$$B_n = \{x \in S \mid d(x_0, x) < n\}$$

*as the balls around a fixed point $x_0 \in S$, and condition (ii) above then means that*

$$\lim_{d(x,x_0)\to\infty} \psi(x) = \infty$$

**Example .** ***Time-dependent branching***

*Suppose a population consists initially $(t = 0)$ of one particle, and particles die with time-dependent rates $d_t > 0$ and divide into two with rates $b_t > 0$ where $d, b\colon \mathbb{R}^+ \to \mathbb{R}^+$ are continuous functions, and $b$ is bounded. Then the total number $X_t$ of particles at time $t$ is a birth-death process with rates*

$$q_t(n,m) = \begin{cases} n\cdot b_t & \text{if} \quad m = n+1 \\ n\cdot d_t & \text{if} \quad m = n-1 \\ 0 & \text{else} \end{cases}, \qquad\qquad \lambda_t(n) = n\cdot(b_t + d_t)$$

*The generator is*

$$\mathscr{L}_t = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ d_t & -(d_t+b_t) & b_t & 0 & 0 & 0 & \cdots \\ 0 & 2d_t & -2(d_t+b_t) & 2b_t & 0 & 0 & \cdots \\ 0 & 0 & 3d_t & -3(d_t+b_t) & 3b_t & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

*Since the rates are unbounded, we have to test for explosion. choose $\psi(n) = n$ as Lyapunov function. Then*

$$(\mathscr{L}_t\psi)(n) = n\cdot b_t\cdot(n+1-n) + n\cdot d_t\cdot(n-1-n) = n\cdot(b_t - d_t) \le n\sup_{t\ge 0} b_t$$

*Since the individual birth rates $b_t$, $t \ge 0$, are bounded, the process is non-explosive. To study long-time survival of the population, we consider the generating functions*

$$G_t(s) = \mathbb{E}\left[s^{X_t}\right] = \sum_{n=0}^{\infty} s^n P[X_t = n], \quad 0 < s \le 1$$

*of the population size. For $f_s(n) = s^n$ we have*

$$\begin{aligned} (\mathscr{L}_t f_s)(n) &= n b_t s^{n+1} - n(b_t+d_t)s^n + n d_t s^{n-1} \\ &= \left(b_t s^2 - (b_t+d_t)s + d_t\right)\cdot \frac{\partial}{\partial s} f_s(n) \end{aligned}$$

*Since the process is non-explosive and $f_s$ and $\mathscr{L}_t f_s$ are bounded on finite time-intervals, the forward equation holds. We obtain*

$$\begin{aligned}\frac{\partial}{\partial t} G_t(s) &= \frac{\partial}{\partial t}\mathbb{E}\left[f_s(X_t)\right] = \mathbb{E}\left[(\mathscr{L}_t f_s)(X_t)\right] \\ &= (b_t s^2 - (b_t + d_t)s + d_t) \cdot \mathbb{E}\left[\frac{\partial}{\partial s} s^{X_t}\right] \\ &= (b_t s - d_t)(s-1) \cdot \frac{\partial}{\partial s} G_t(s), \\ G_0(s) &= \mathbb{E}\left[s^{X_0}\right] = s\end{aligned}$$

*The solution of this first order partial differential equation for $s < 1$ is*

$$G_t(s) = 1 - \left(\frac{e^{\varrho_t}}{1-s} + \int_0^t b_n e^{\varrho_u}\, du\right)^{-1}$$

*where*

$$\varrho_t := \int_0^t (d_u - b_u)\, du$$

*is the accumulated death rate. In particular, we obtain an explicit formula for the extinction probability:*

$$\begin{aligned}P[X_t = 0] = \lim_{s \downarrow 0} G_t(s) &= \left(e^{\varrho_t} + \int_0^t b_n e^{\varrho_u}\, du\right)^{-1} \\ &= 1 - \left(1 + \int_0^t d_u e^{\varrho_u}\, du\right)^{-1}\end{aligned}$$

*since $b = d - \varrho'$. Thus we have shown:*

**Theorem 1.23.**

$$P[X_t = 0 \text{ eventually}] = 1 \iff \int_0^\infty d_u e^{\varrho_u}\, du = \infty$$

**Remark .** *Informally, the mean and the variance of $X_t$ can be computed by differentiating $G_t$ at $s = 1$ :*

$$\begin{aligned}\frac{d}{ds}\mathbb{E}\left[s^{X_t}\right]\Big|_{s=1} &= \mathbb{E}\left[X_t s^{X_t - 1}\right]\Big|_{s=1} = \mathbb{E}[X_t] \\ \frac{d^2}{ds^2}\mathbb{E}\left[s^{X_t}\right]\Big|_{s=1} &= \mathbb{E}\left[X_t(X_t-1)s^{X_t-2}\right]\Big|_{s=1} = \mathrm{Var}(X_t)\end{aligned}$$

## Explosion in time-homogeneous case

***Distribution of explosion time:*** By Kolmogorov's backward equation,

$$F(t,x) := P_x[\zeta \le t] = P_x[X_t = \Delta] = 1 - \mathbb{E}_x[I_s(X_t)]$$

where $\mathbb{E}_x[I_s(X_t)]$ is the minimal non-negative solution of

$$\frac{\partial u}{\partial t} = \mathscr{L}u, \qquad u(0,x) = 1,$$

hence $F(t,x)$ is the maximal solution $\le 1$ of

$$\frac{\partial u}{\partial t} = \mathscr{L}u, \qquad u(0,x) = 0$$

*Laplace transformation*: $\alpha > 0$,

$$F_\alpha(x) := \mathbb{E}_x\left[e^{-\alpha\varrho}\right] = \frac{1}{\alpha}\int_0^\infty e^{-\alpha t} F(t,x)\,dt$$

since

$$e^{-\alpha\varrho} = \frac{1}{\alpha}\int_\varrho^\infty e^{-\alpha t}\,dt = \frac{1}{\alpha}\int_0^\infty e^{-\alpha t}\cdot I_{\{t\ge\varrho\}}\,dt$$

Informally by integrating by parts,

$$(\mathscr{L}F_\alpha)(x) = \frac{1}{\alpha}\int_0^\infty e^{-\alpha t}\underbrace{\mathscr{L}F(t,x)}_{=\frac{\partial F}{\partial t}}\,dt = \int_0^\infty e^{-\alpha t}F(t,x)\,dt = (\alpha F_\alpha)(x)$$

**Theorem 1.24.** (Necessary and sufficient condition for non-explosion, ***Reuter's criterion***)

1. *$F_\alpha$ is the maximal solution of*

$$\mathscr{L}g = \alpha g \tag{1.10}$$

   *satisfying* $0 \le y \le 1$.

2. *The minimal Markov jump process is non-explosive if and only if (1.10) has only the trivial solution satisfying* $0 \le y \le 1$.

*Proof.* 1. by first step analysis (Exercise)

2. $\zeta = \infty \quad P_x$-a.s. if and only if $F_\alpha(x) = 0$.

□

## 1.5 Asymptotics of Time-homogeneous jump processes

Let $(X_t, P_x)$ be a minimal Markov jump process with jump rates $q(x, dy)$. The generator is given by

$$(\mathscr{L}f)(x) = (qf)(x) - \lambda(x)f(x) = \lambda(x) \cdot ((\pi f)(x) - f(x))$$

i.e.

$$\mathscr{L} = \lambda \cdot (\pi - I) \tag{1.11}$$

where $\pi - I$ is the generator of the jump chain $(Y_n)$. Let

$$T := \inf \{t \geq 0 \mid X_t \in D^c\}, \qquad D \subseteq S \text{ open}$$

**Theorem 1.25** (Dirichlet and Poisson problem)**.** *For any measurable functions $c\colon D \to \mathbb{R}^+$ and $f\colon D^c \to \mathbb{R}^+$,*

$$u(x) := \mathbb{E}_x \left[ \int_0^T c(X_t)\, dt + f(X_T) \cdot I_{\{T<\zeta\}} \right]$$

*is the minimal non-negative solution of the* ***Poisson equation***

$$\begin{aligned} -\mathscr{L}u &= c \qquad \text{on } D \\ u &= f \qquad \text{on } D^c \end{aligned} \tag{1.12}$$

*Proof.* 1. For $c \equiv 0$ this follows from the corresponding result in discrete time. In fact, the exit points from $D$ of $X_t$ and $Y_t$ coincide, and hence

$$f(X_T) \cdot I_{\{T<\zeta\}} = f(Y_\tau) \cdot I_{\{\tau<\infty\}}$$

where $\tau = \inf\{n \geq 0 \mid Y_n \notin D\}$. Therefore $u$ is the minimal non-negative solution of

$$\begin{aligned} \pi u &= u \qquad \text{on } D \\ u &= f \qquad \text{on } D^c \end{aligned}$$

which is equivalent to (1.12) by (1.11).

2. In the general case the assertion can be proven by first step analysis (Exercise).

□

**Example .** 1. ***Hitting probability of $\mathbf{D^c}$:***

$u(x) = P_x[T < \zeta]$ *solves* $\mathscr{L}u = 0$ *on* $D$*,* $u = 1$ *on* $D^c$*.*

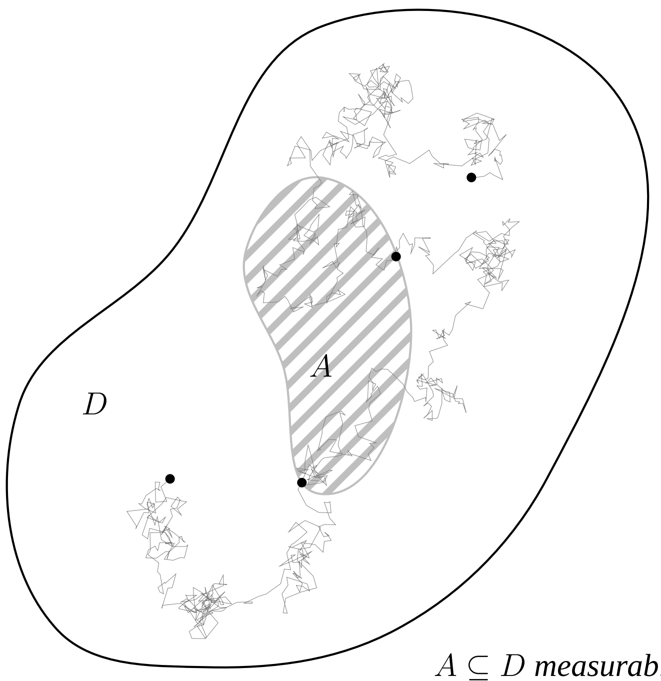
2. ***Distribution of $X_T$:***

   *$u(x) = \mathbb{E}_x[f(X_T)\ ;\ T < \zeta]$ solves $\mathscr{L}u = 0$ on $D$, $u = f$ on $D^c$.*

3. ***Mean exit time:***

   *$u(x) = \mathbb{E}_x[T]$ solves $-\mathscr{L}u = 1$ on $D$, $u = 0$ on $D^c$.*

4. ***Mean occupation time of $A$ before exit from $D$:***




*$A \subseteq D$ measurable*

$$u(x) = \mathbb{E}_x\left[\int_0^T I_A(X_t)\,dt\right] = \int_0^\infty P_x[X_t \in A,\ t < T]\,dt$$

*solves $-\mathscr{L}u = I_A$ on $D$, $u = 0$ on $D^c$. $u(x) = G_D(x, A)$ is called **Green function** of a Markov process in a domain $D$.*

**Assumption** (from now on): $S$ is countable.

**Definition 1.26.** *1. A state $x \in S$ is called **recurrent** if and only if*

$$P_x\left[\{t \geq 0\ :\ X_t = x\} \text{ is unbounded}\right] = 1,$$

*and **transient** if and only if*

$$P_x\left[\{t \geq 0\ :\ X_t = x\} \text{ is unbounded}\right] = 0$$

2. *A state $y \in S$ is called* ***accessible*** *from $x$ ($x \rightsquigarrow y$) if and only if*

$$P_x[X_t = y \text{ for some } t \geq 0] > 0.$$

*$x$ and $y$* ***communicate*** *($x \leftrightsquigarrow y$) if and only if $x \rightsquigarrow y$ and $y \rightsquigarrow x$. The Markov chain $(X_t, P_x)$ is called* ***irreducible*** *if and only if all states communicate. It is called* ***recurrent*** *respectively* ***transient*** *if and only if all states are recurrent respectively transient.*

**Lemma 1.27.** *Suppose that $\lambda(x) > 0$ for all $x \in S$. Then:*

1. *For $x, y \in S$ the following assertions are equivalent:*
   - *(i)* $x \rightsquigarrow y$
   - *(ii)* $x \rightsquigarrow y$ *for the jump chain $(Y_n)$*
   - *(iii) There is a $k \in \mathbb{N}$ and $x_1, \ldots, x_k \in S$ such that*

     $$q(x, x_1)q(x_1, x_2) \cdots q(x_{k-1}, x_k)q(x_k, y) > 0$$

   - *(iv)* $p_t(x, y) > 0$ *for all $t > 0$*
   - *(v)* $p_t(x, y) > 0$ *for some $t > 0$.*
2. *A state $x \in S$ is recurrent (respectively transient) if and only if it is recurrent (respectively transient) for the jump chain.*

*Proof.* 1. $(i) \Leftrightarrow (ii)$ since $Y_n$ visits the same states as $X_t$

$(ii) \Rightarrow \exists x_1, \ldots x_k$ such that $\pi(x, x_1)\pi(x_1, x_2) \cdots \pi(x_k, y) > 0 \Rightarrow (iii)$ since $\lambda > 0$.

$(iii) \Rightarrow (iv):\ q(a, b) > 0$ and hence

$$p_t(a, b) \geq P_a[J_1 \leq t < J_2,\ Y_1 = b] > 0$$

for all $t > 0$. Hence with $(iii)$ and the independence of the states,

$$p_t(x, y) \geq p_{\frac{t}{k+1}}(x, x_1)p_{\frac{t}{k+1}}(x_1, x_2) \cdots p_{\frac{t}{k+1}}(x_k, y) > 0$$

$(iv) \Rightarrow (v) \Rightarrow (i)$ is obvious.

2. If $x$ is recurrent for $Y_n$ then

$$P_x[\lambda(Y_n) = \lambda(x) \text{ infinitely often}] = 1,$$

and hence

$$\zeta = \sum_1^{\infty} (J_i - J_{i-1}) = \infty \qquad P_x\text{-a.s.}$$

since $(J_1 - J_{i-1})$ are conditional independent given $(Y_n)$, $\mathrm{Exp}(\lambda(x))$ distributed infinitely often.

Since $\lambda > 0$, the process $(X_t)_{t\geq 0}$ does not get stuck, and hence visits the same states as the jump chain $(Y_n)_{n\geq 0}$. Thus $X_t$ is recurrent. Similarly, the converse implication holds.

If $x$ is transient for $(X_t)$ then it is transient for $(Y_n)$ since otherwise it would be recurrent by the dichotomy in discrete time. Finally, if $x$ is transient for $(Y_n)$ then it is transient for $X_n$ since the process spends only finite time in each state.

□

Let

$$T_x := \inf\{t \geq J_1 \;:\; X_t = x\}$$

denote the ***first passage time*** of $x$.

**Theorem 1.28** (Recurrence and transience). *1. Every $x \in S$ is either recurrent or transient.*

*2. If $x$ is recurrent and $x \rightsquigarrow y$, then $y$ is recurrent.*

*3. $x$ recurrent $\Leftrightarrow$ $\lambda(x) = 0$ or $P_x[T_x < \infty] = 1 \Leftrightarrow$ $G(x,x) = \int_0^\infty p_t(x,x)\,dt = \infty$*

*Proof.* Under the assumption $\lambda > 0$, the first two assertions follow from the corresponding result in discrete time. If $\lambda(x) = 0$ for some $x$, we can apply the same arguments if we construct the process $(X_t)$ form a jump chain which is absorbed at $x \in S$ with $\lambda(x) = 0$.
3. If $\lambda(x) > 0$ then by the discrete time result $x$ is recurrent if and only if

$$P_x[Y_n = x \text{ for some } n \geq 1] = 1,$$

i.e., if and only if

$$P_x[T_x < \infty]$$

Moreover, the Green function of $(X_t)$ can be computed from the Green function of the jump chain $(Y_n)$:

$$\begin{aligned}
\int_0^\infty p_t(x,x)\,dt &= \mathbb{E}_x\left[\int_0^\infty I_{\{x\}}(X_t)\,dt\right] \\
&= \mathbb{E}_x\left[\sum_{n=0}^\infty (J_{n+1} - J_n) I_{\{x\}}(Y_n)\right] \\
&= \sum_{n=0}^\infty \mathbb{E}[\underbrace{J_{n+1} - J_n}_{\sim \mathrm{Exp}(\lambda(x))} \mid Y_n = x] \cdot P_x[Y_n = x] \\
&= \frac{1}{\lambda(x)} \underbrace{\sum_{n=0}^\infty \pi^n(x,x)}_{\text{discrete-time Green function}}
\end{aligned}$$

Hence

$$G(x,x)=\infty \ \Leftrightarrow\ \lambda(x)=0 \text{ or } x \text{ recurrent for } Y_n \ \Leftrightarrow\ x \text{ recurrent for } X_t$$

□

**Remark** (Strong Markov property)**.**

$$\mathbb{E}_\nu\left[F(X_{T+\bullet})\cdot I_{\{T<\zeta\}}\mid \mathcal{F}_T\right]=\mathbb{E}_{X_T}[F(X)]$$

*$P_\nu$-a.s. on $\{T<\zeta\}$ for any $\mathcal{F}_t^X$ stopping time T, and any measurable function $F\colon S^{\mathbb{R}^+}\to\mathbb{R}^+$, and any initial distribution $\nu\in M_1(S)$.*

*Proof.* Either directly from the strong Markov property for the jump chain (Exercise). A more general proof that applies to other continuous time Markov processes as well will be given in the next chapter. □

**Definition 1.29.** *A positive measure $\mu$ on $S$ is called* ***stationary*** *(or* ***invariant****) with respect to $(p_r)_{t\geq 0}$ if and only if*

$$\mu p_t=\mu$$

*for all $t\geq 0$.*

**Theorem 1.30** (***Existence and uniqueness of stationary measure***)**.** *Suppose that $x\in S$ is recurrent. Then:*

1. $$\mu(B):=\mathbb{E}_x\left[\int_0^{T_x} I_B(X_t)\,dt\right],\qquad B\subseteq S,$$

   *is a stationary measure. If $x$ is* ***positive recurrent*** *(i.e. $\mathbb{E}_x[T_x]<\infty$) then*

   $$\bar{\mu}(B)=\frac{\mu(B)}{\mu(S)}$$

   *is a stationary probability distribution.*

2. *If $(X_t,P_x)$ is irreducible then any stationary measure is a multiple of $\mu$.*

**Remark .** 1. *$\mu(B)=$ expected time spent in $B$ before returning to $x$.*

2. *Relation to stationary measure of the jump chain:*

$$\begin{aligned}
\tau_x &:= \inf\{n \geq 1 \ : \ Y_n = x\} \\
\mu(y) &= \mathbb{E}_x\left[\int_0^{T_x} I_{\{y\}}(X_t)\,dt\right] \\
&= \mathbb{E}_x\left[\sum_{n=0}^{\tau_x-1} (J_{n+1}-J_n) I_{\{y\}}(Y_n)\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_x\left[\mathbb{E}_x[\underbrace{J_{n+1}-J_n}_{\sim \mathrm{Exp}(\lambda(y))} \mid (Y_k)] \cdot I_{\{n<\tau_x\}} \cdot I_{\{y\}}(Y_n)\right] \\
&= \frac{1}{\lambda(y)} \qquad \underbrace{\mathbb{E}_x\left[\sum_{n=0}^{\tau_x-1} I_{\{y\}}(Y_n)\right]}_{=:\nu(y) \text{ stationary measure for jump chain}}
\end{aligned}$$

*Proof of theorem.* 1. Fix $B \subseteq S$ and $s \geq 0$.

$$\begin{aligned}
\mu(B) &= \mathbb{E}_x\left[\int_0^{s} I_B(X_t)\,dt\right] + \mathbb{E}_x\left[\int_s^{T_x} I_B(X_t)\,dt\right] \\
&= \mathbb{E}_x\left[\int_{T_x}^{T_x+s} I_B(X_t)\,dt\right] + \mathbb{E}_x\left[\int_s^{T_x} I_B(X_t)\,dt\right] \\
&= \mathbb{E}_x\left[\int_s^{T_x+s} I_B(X_t)\,dt\right] = \mathbb{E}_x\left[\int_0^{T_x} I_B(X_{s+t})\,dt\right] \\
&= \int_0^{\infty} \mathbb{E}_x[\underbrace{P_x[X_{s+t} \in B \mid \mathcal{F}_t]}_{=p_s(X_t,B)};\ T_x > t]\,dt \\
&= \sum_{y \in S} \underbrace{\mathbb{E}_x\left[\int_0^{T_x} I_{\{y\}}(X_t)\,dt\right]}_{=\mu(y)} \cdot p_s(y,B) = (\mu p_s)(B)
\end{aligned}$$

Here we have used in the second step that

$$\mathbb{E}_x\left[\int_{T_x}^{T_x+s} I_B(X_t)\,dt \mid \mathcal{F}_{T_x}\right] = \mathbb{E}_x\left[\int_0^{s} I_B(X_t)\,dt\right]$$

by the strong Markov property for $(X_t)$.

We have shown that $\mu$ is a stationary measure. If $x$ is positive recurrent then $\mu(S)$ is finite, and hence $\mu$ can be normalized to a stationary probability distribution.

2. If $(X_t)$ is irreducible then the skeleton chain $(X_n)_{n=0,1,2,\ldots}$ is a discrete-time Markov chain with transition kernel

$$p_1(x, y) > 0 \qquad \forall\, x, y \in S$$

Hence $(X_n)$ is irreducible. If we can show that $(X_n)$ is recurrent, than by the discrete-time theory, $(X_n)$ has at most one invariant measure (up to a multiplicative factor), and thus the same holds for $(X_t)$. Since $x$ is recurrent for $(X_t)$, the jump chain $(Y_n)$ visits $x$ infinitely often with probability 1. Let $K_1 < K_2 < \cdots$ denote the successive visit times. Then

$$X_{J_{K_i}} = Y_{K_i} = x \tag{1.13}$$

for all $i$. We claim that also

$$X_{\lceil J_{K_i} \rceil} = x \quad \text{infinitely often} \tag{1.14}$$

In fact, the holding times $J_{K_{i+1}} - J_{K_i},\ i \in \mathbb{N}$, are conditionally independent given $(Y_n)$ with distribution $\mathrm{Exp}(\lambda(x))$. Hence

$$P_x[J_{K_{i+1}} - J_{K_i} > 1 \text{ infinitely often}] = 1,$$

which implies (1.14) by (1.13). The recurrence of $(X_n)$ follows from (1.14) by irreducibility.

□

If $S$ is finite, by the Kolmogorov backward equation

$$\mu p_t = \mu \,\forall\, t \geq 0 \quad \Leftrightarrow \quad \mu\mathscr{L} = 0$$

In the general case this infinitesimal characterization of stationary measures does not always hold, cf. the example below. However, as a consequence of the theorem we obtain:

**Corollary 1.31** (Infinitesimal characterization of stationary distribution). *Suppose that $(X_t, P_x)$ is irreducible, and $\mu \in M_1(S)$. Then:*

1. *If $\mu$ is a stationary distribution of $(p_t)_{t\geq 0}$ then all states are positive recurrent, and*

$$(\mu\mathscr{L})(y) = 0 \quad \forall\, y \in S \tag{1.15}$$

2. *Conversely, if (1.15) holds then $\mu$ is stationary provided $(X_t, P_x)$ is recurrent. This is for example the case if $\sum_{x\in S} \lambda(x)\mu(x) < \infty$.*

**Remark .** *Condition on (1.15) means that*

$$< \mu\mathscr{L}, f >=< \mu, \mathscr{L} f >= 0$$

*for all* finitely supported *functions* $f : S \to \mathbb{R}$*. Note that if* $\sum \lambda(x)\mu(x) = \infty$ *then* $\mu\mathscr{L} = \mu q - \lambda\mu$ *is not a signed measure with finite variation. In particular,* $< \mu\mathscr{L}, 1 >$ *is not defined!*

*Proof.* A stationary distribution $\mu$ of $(X_t)$ is also stationary for the skeleton chain $(X_n)_{n=0,1,2,\ldots}$, which is irreducible as noted above. Therefore, the skeleton chain and thus $(X_t)$ are positive recurrent. Now the theorem and the remark above imply that in the recurrent case, a measure $\mu$ is stationary if and only if $\lambda \cdot \mu$ is stationary for the jump chain $(Y_n)$, i.e.

$$(\mu q)(y) = \sum_{x \in S} \lambda(x)\mu(x)\pi(x, y) = \lambda(y)\mu(y) \tag{1.16}$$

which is equivalent to $(\mu\mathscr{L})(y) = 0$.
In particular, if $\sum \lambda(x)\mu(x) < \infty$ and $\mu\mathscr{L} = 0$ then $\lambda\mu$ is a stationary distribution for $(Y_n)$, where $(Y_n)$ and thus $(X_t)$ are positive recurrent. □

**Example .** *We consider the minimal Markov jump process with jump chain* $Y_n = Y_0 + n$ *and intensities* $\lambda(x) = 1 + x^2$*. Since* $\nu(y) \equiv 1$ *is a stationary measure for* $(Y_n)$*, i.e.* $\nu\pi = \nu$*, we see that*

$$\mu(y) := \frac{\nu(y)}{\lambda(y)} = \frac{1}{1 + y^2}$$

*is a finite measure with* $(\mu\mathscr{L})(y) = 0$ *for all* $y$*. However,* $X_t$ *is not recurrent (since* $Y_n$ *is transient), and hence* $\mu$ *is not stationary for* $X_t$*!*

Actually, in the example above, $X_t$ is explosive. In fact, one can show:

**Theorem 1.32.** *If* $(X_t)$ *is irreducible and non-explosive, and* $\mu \in M_1(S)$ *satisfies (1.16), then* $\mu$ *is a stationary distribution.*

*Proof.* Omitted, cf. Asmussen [4], Theorem 4.3 in chapter II. □

**Remark .** ***Detailed balance:***

*Condition (1.16) is satisfied provided the* ***detailed balance condition***

$$\mu(x)q(x, y) = \mu(y)q(y, x) \tag{1.17}$$

*holds for all* $x, y \in S$*. In fact, (1.17) implies*

$$(\mu q)(y) = \sum_{x \in S} \mu(x)q(x, y) = \mu(y) \sum_{x \in S} q(y, x) = \lambda(y)\mu(y)$$

*for all* $y \in S$*.*

**Example .** ***Stationary distributions and mean hitting times of birth-death process:***

*For a birth-death process on* $\{0, 1, 2, \dots\}$ *with strictly positive birth rates* $b(x)$ *and death rates* $d(x)$ *the detailed balance condition is*

$$\mu(x)b(x) = \mu(x+1)d(x+1) \tag{1.18}$$

*for all* $x \geq 0$*. Hence detailed balance holds if and only id* $\mu$ *is a multiple of*

$$\nu(x) := \prod_{i=1}^{x} \frac{d(i)}{b(i-1)}$$

*Suppose that*

$$\sum_{n=0}^{\infty} \nu(n) < \infty \tag{1.19}$$

*Then*

$$\mu(x) := \frac{\nu(x)}{\sum_{y=0}^{\infty} \nu(y)}$$

*is a probability distribution satisfying (1.17), and hence (1.16). By irreducibility,* $\mu$ *is the unique stationary probability distribution provided the process is non-explosive. The example above shows that explosion may occur even when (1.19) holds.*

**Theorem 1.33.** *Suppose (1.19) holds and*

$$\sum_{n=0}^{\infty} \frac{1}{b(n)} = \infty. \tag{1.20}$$

*Then:*

1. *The minimal birth-death process in non explosive, and* $\mu$ *is the unique stationary probability distribution.*

2. *The* ***mean hitting times*** *are given by*

   *(a)*
   $$\mathbb{E}_x[T_y] = \sum_{n=x}^{y-1} \frac{\mu(\{0, 1, \dots, n\})}{\mu(n) \cdot b(n)} \quad \textit{for all } 0 \leq x \leq y \textit{ and}$$

   *(b)*
   $$\mathbb{E}_x[T_y] = \sum_{n=y+1}^{x} \frac{\mu(\{n, n+1, \dots\})}{\mu(n) \cdot d(n)} \quad \textit{for all } 0 \leq y \leq x \textit{ respectively.}$$

   *In particular, the* ***mean commute time*** *between* $x$ *and* $y$ *is given by*

*(c)*

$$\mathbb{E}_x[T_y] + \mathbb{E}_y[T_x] = \sum_{n=x}^{y-1} \frac{1}{\mu(n) \cdot b(n)} \quad \text{for all } 0 \leq x < y.$$

*Proof.* 1. Reuter's criterion implies that the process is non-explosive of and only if

$$\sum_{n=0}^{\infty} \frac{\nu(\{0, \dots, n\})}{b(n)} = \infty$$

(Exercise, cf. Brémaud [7], Chapter 8, Theorem 4.5).

If (1.19) holds then this condition is equivalent to (1.20).

2. Fix $y \in \mathbb{N}$. The function

$$u(x) = \mathbb{E}_x[T_y] = \mathbb{E}_x\left[\int_0^{T_y} 1\, dt\right], \quad 0 \leq x \leq y$$

is the minimal non-negative solution of the Poisson equation

$$-\mathscr{L}u = 1 \text{ on } \{0, 1, \dots, y-1\}, \qquad u(y) = 0$$

Hence $u'(n) := u(n+1) - u(n)$ solves the *ordinary difference equation*

$$\begin{aligned} &b(0)u'(0) = 1, \\ &b(n)u'(n) + d(n)u'(n-1) = 1 \end{aligned}$$

for all $1 \leq n < y$. By the detailed balance condition (1.18) the unique solution of this equation is given by

$$u'(n) = -\sum_{k=0}^{n} \frac{1}{b(k)} \prod_{l=k+1}^{n} \frac{d(l)}{b(l)} \overset{(1.18)}{=} -\sum_{k=0}^{n} \frac{\mu(k)}{\mu(n)b(n)} \quad \forall\, 0 \leq n \leq y$$

Assertion $(a)$ now follows by summing over $n$ and taking into account the boundary condition $u(y) = 0$. The proof of $(b)$ is similar and $(c)$ follows from $(a)$ and $(b)$ since

$$\mu(n)d(n) = \mu(n-1)b(n-1)$$

by (1.18).

□

**Remark .** *Since $\mu(n) \cdot b(n)$ is the flow through the edge $\{n, n+1\}$, the right hand side of $(c)$ can be interpreted as the **effective resistance** between $x$ and $y$ of the corresponding electrical network. With this interpretation, the formula carries over to Markov chains on general graphs and the corresponding electrical networks, cf. Aldous, Fill [1].*

**Theorem 1.34** (Ergodic Theorem)**.** *Suppose that $(X_t, P_x)$ is irreducible and has stationary probability distribution $\bar{\mu}$. Then*

$$\frac{1}{t}\int_0^t f(X_s)\,ds \longrightarrow \int f\,d\bar{\mu}$$

*$P_\nu$-a.s. as $t \to \infty$ for any non-negative function $f\colon S \to \mathbb{R}$ and any initial distribution $\nu \in M_1(S)$.*

**Remark .** 1. *In particular,*

$$\bar{\mu}(B) = \lim_{t\to\infty}\frac{1}{t}\int_0^t I_B(X_s)\,ds \quad P_\nu\text{-a.s.}$$

stationary probability $=$ mean occupation time.

2. *More generally: If $(X_t, P_x)$ is irreducible and recurrent then*

$$\frac{1}{L_t^x}\int_0^t f(X_s)\,ds \longrightarrow \int f\,d\mu_x \quad P_\nu\text{-a.s. for all } x \in S$$

*where $L_t^x = \int_0^t I_{\{x\}}(X_s)\,ds$ is the occupation time of $x$, and*

$$\mu_x(B) = \mathbb{E}_x\left[\int_0^{T_x} I_B(X_s)\,ds\right]$$

*is a stationary measure.*

*Proof.* Similar to the discrete time case. Fix $x \in S$ and define recursively the successive leaving and visit times of the state $x$:

$$\begin{aligned}
\tilde{T}^0 &= \inf\{t \geq 0 \;:\; X_t \neq x\} \\
T^n &= \inf\left\{t \geq \tilde{T}^{n-1} \;:\; X_t = x\right\} \quad \text{visit times of } x \\
\tilde{T}^n &= \inf\{t \geq T^n \;:\; X_t \neq x\} \quad \text{leaving times of } x
\end{aligned}$$

We have

$$\int_{T^1}^{T^n} f(X_s)\,ds = \sum_{k=1}^{n-1} Y_k$$

where

$$Y_k := \int_{T^k}^{T^{k+1}} f(X_s)\,ds = \int_0^{T^{k+1}-T^k} f(X_{s+T_k})\,ds$$

Note that $T^{k+1}(X) = T^k(X) + T^1(X_{T^k+\bullet})$. Hence by the strong Markov property the random variables are independent and identically distributed with expectation

$$\mathbb{E}_\nu[Y_k] = \mathbb{E}_\nu[\mathbb{E}_\nu[Y_k \mid \mathcal{F}_{T^k}]] = \mathbb{E}_x\left[\int_0^{T_x} f(X_s)\, ds\right] = \int f\, d\mu$$

The law of large numbers now implies

$$\frac{1}{n}\int_0^{T^n} f(X_s)\, ds \longrightarrow \int f\, d\mu_x \quad P_\nu\text{-a.s. as } n\to\infty \tag{1.21}$$

In particular,

$$\frac{T_n}{n} \longrightarrow \mu_x(S) \quad P_\nu\text{-a.s.}$$

By irreducibility, the stationary measure is unique up to a multiplicative factor. Hence $\mu_x(S) < \infty$ and $\bar{\mu} = \frac{\mu_x}{\mu_x(S)}$. Thus we obtain

$$\begin{aligned}\int f\, d\bar{\mu} = \frac{\int f\, d\mu}{\mu(S)} &= \lim_{n\to\infty} \frac{n}{T_{n+1}}\cdot\frac{1}{n}\int_0^{T_n} f(X_s)\, ds \leq \liminf_{t\to\infty}\frac{1}{t}\int_0^t f(X_s)\, ds \\ &\leq \limsup_{t\to\infty}\frac{1}{t}\int_0^t f(X_s)\, ds \leq \limsup_{n\to\infty}\frac{n}{T_n}\cdot\frac{1}{n}\int_0^{T_n} f(X_s)\, ds = \int f\, d\bar{\mu}_s\end{aligned}$$

i.e.

$$\frac{1}{t}\int_0^t f(X_s)\, ds \longrightarrow \int f\, d\bar{\mu} \quad P_\nu\text{-a.s.}$$

□

# Chapter 2

# Interacting particle systems

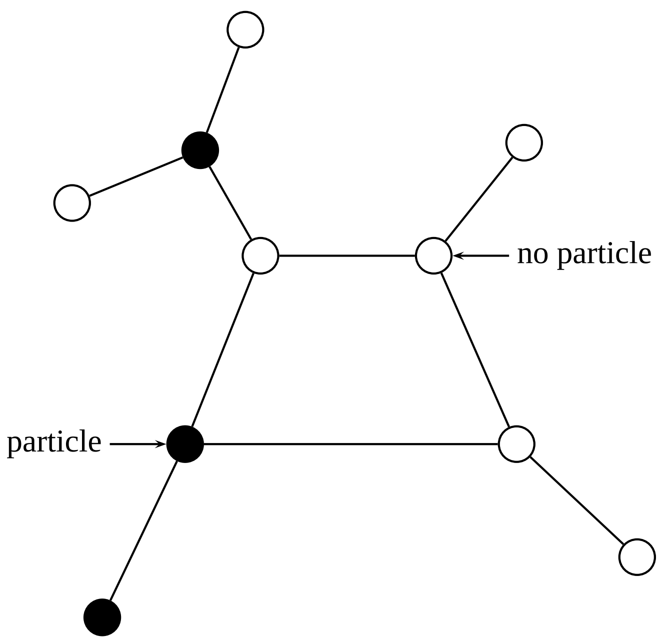
## 2.1 Interacting particle systems - a first look

Let $G = (V, E)$ be an (undirected) graph with $V$ the set of vertices and $E$ the set of edges. We write $x \sim y$ if and only if $\{x, y\} \in E$. We call

$$S = T^V = \{\eta \colon V \to T\}$$

the ***configuration space***. $T$ can be the space of types, states, spins etc.
E.g.

$$T = \{0, 1\}, \qquad \eta(x) = \begin{cases} 1 & \text{particle at } x \\ 0 & \text{no particle at } x \end{cases}$$




Markovian dynamics: $\eta(x)$ changes to state $i$ with rate

$$c_i(x, \eta) = g_i\left((\eta(x), (\eta(y))_{y \sim x}\right)$$

i.e.

$$q(\eta,\xi)=\begin{cases}c_i(x,\eta) & \text{if } \xi=\eta^{x,i}\\ 0 & \text{otherwise}\end{cases}$$

where

$$\eta^{x,i}(y)=\begin{cases}\eta(y) & \text{for } y\neq x\\ i & \text{for } y=x\end{cases}$$

**Example .** *1.* ***Contact process:*** *(Spread of plant species, infection,...)* $T=\{0,1\}$. *Each particle dies with rate $d>0$, produces descendent at any neighbor site with rate $b>0$ (if not occupied)*

$$\begin{aligned}c_0(x,\eta)&=d\\ c_1(x,\eta)&=b\cdot N_1(x,\eta); \qquad N_1(x,\eta):=|\{y\sim x\ :\ \eta(y)=1\}|\end{aligned}$$

Spatial branching process with exclusion rule (only one particle per site).

2. ***Voter model:*** *$\eta(x)$ opinion of voter at $x$,*

$$c_i(x,y)=N_i(x,y):=|\{y\sim x\ :\ \eta(y)=i\}|$$

*changes opinion to $i$ with rate equal to number of neighbors with opinion $i$.*

3. ***Ising model with Glauber (spin flip) dynamics:*** $T=\{-1,1\}$, $\beta>0$ *inverse temperature.*

   *(a)* Metropolis dynamics:

$$\begin{aligned}\Delta(x,\eta)&:=\sum_{y\sim x}\eta(y)=N_1(x,\eta)-N_{-1}(x,\eta)\quad \textit{total magnetization}\\ c_1(x,\eta)&:=\min\left(e^{2\beta\cdot\Delta(x,\eta)},1\right)\\ c_0(x,\eta)&:=\min\left(e^{-2\beta\cdot\Delta(x,\eta)},1\right)\end{aligned}$$

   *(b)* Heath bath dynamics / Gibbs sampler:

$$\begin{aligned}c_1(x,\eta)&=\frac{e^{\beta\Delta(x,\eta)}}{e^{\beta\Delta(x,\eta)}+e^{-\beta\Delta(x,\eta)}}\\ c_0(x,\eta)&=\frac{e^{-\beta\Delta(x,\eta)}}{e^{\beta\Delta(x,\eta)}+e^{-\beta\Delta(x,\eta)}}\end{aligned}$$

   $\boldsymbol{\beta=0}$*: (infinite temperature)* $c_1\equiv c_0\equiv\frac{1}{2}$, *random walk on $\{0,1\}^V$ (hypercube)*
   $\boldsymbol{\beta\to\infty}$*: (zero temperature)*

$$c_1(x,\eta)=\begin{cases}1 & \textit{if } \Delta(x,y)>0\\ \frac{1}{2} & \textit{if } \Delta(x,y)=0\\ 0 & \textit{if } \Delta(x,y)<0\end{cases}, \qquad c_0(x,\eta)=\begin{cases}1 & \textit{if } \Delta(x,y)<0\\ \frac{1}{2} & \textit{if } \Delta(x,y)=0\\ 0 & \textit{if } \Delta(x,y)>0\end{cases}$$

   *Voter model with majority vote.*

In the rest of this section we will assume that the vertex set $V$ is finite. In this case, the configuration space $S = T^V$ is finite-dimensional. If, moreover, the type space $T$ is also finite then $S$ itself is a finite graph with respect to the ***Hamming distance***

$$d(\eta, \xi) = |\{x \in V \; ; \; \eta(x) \neq \xi(x)\}|$$

Hence a continuous-time Markov chain $(\eta_t, P_x)$ can be constructed as above from the jump rates $q_t(\eta, \xi)$. The process is non-explosive, and the asymptotic results from the last section apply. In particular, if irreducibility holds the there exists a unique stationary probability distribution, and the ergodic theorem applies.

**Example .** 1. ***Ising Model:*** *The Boltzman distribution*

$$\mu_\beta(\eta) = \frac{1}{Z_\beta} e^{-\beta H(\eta)}, \qquad\qquad Z_\beta = \sum_\eta e^{-\beta H(\eta)},$$

*with Hamiltonian*

$$H(\eta) = \frac{1}{2} \sum_{\{x,y\} \in E} (\eta(x) - \eta(y))^2 = \sum_{\{x,y\} \in E} \eta(x)\eta(y) + |E|$$

*is stationary, since it satisfies the detailed balance condition*

$$\mu_\beta(\eta) q(\eta, \xi) = \mu_\beta(\xi) q(\xi, \eta) \quad \forall\, \xi, \eta \in S.$$

*Moreover, irreducibility holds - so the stationary distribution is unique, and the ergodic theorem applies (Exercise).*

2. ***Voter model:*** *The constant configurations* $\underline{i}(x) \equiv i, \; i \in T$, *are* absorbing states, *i.e.* $c_j(x, \underline{i}) = 0$ *for all* $j \neq i, x$. *Any other state is transient, so*

$$P\left[\bigcup_{i \in T} \{\eta_t = \underline{i} \textit{ eventually}\}\right] = 1.$$

*Moreover,*

$$N_i(\eta_t) := |\{x \in V \;:\; \eta_t(x) = i\}|$$

*is a martingale (Exercise), so*

$$N_i(\eta) = \mathbb{E}_\eta[N_i(\eta_t)] \xrightarrow{t \to \infty} \mathbb{E}_\eta[N_i(\eta_\infty)] = N \cdot P[\eta_t = \underline{i} \textit{ eventually}]$$

*i.e.*

$$P[\eta_t = \underline{i} \textit{ eventually}] = \frac{N_i(\eta)}{N}$$

*The stationary distributions are the Dirac measures* $\delta_{\underline{i}}, \; i \in T$, *and their convex combinations.*

3. ***Contact process:*** *The configuration $\underline{0}$ is absorbing, all other states are transient. Hence $\delta_{\underline{0}}$ is the unique invariant measure and ergodicity holds.*

We see that on finite graphs the situation is rather simple as long as we are only interested in existence and uniqueness of invariant measures, and ergodicity. Below, we will show that on infinite graphs the situation is completely different, and phase transitions occur. On finite subgraphs on an infinite graph these phase transitions effect the rate of convergence to the stationary distribution and the variances of ergodic averages but not the ergodicity properties themselves.

**Mean field models:**

Suppose that $G$ is the complete graph with $n$ vertices, i.e.

$$V = \{1, \dots, n\} \quad \text{and} \quad E = \{\{x, y\} \, : \, x, y \in V\}$$

Let

$$L_n(\eta) = \frac{1}{n} \sum_{x=1}^{n} \delta_{\eta(x)}$$

denote the ***empirical distribution*** of a configuration $\eta \colon \{1, \dots, n\} \to T$, the *mean field*. In a ***mean-field model*** the rates

$$c_i(x, \eta) = f_i(L_n(\eta))$$

are *independent of* $x$, and *depend on* $\eta$ *only through the* mean field $L_n(\eta)$.

**Example .** ***Multinomial resampling*** *(e.g. population genetics), mean field voter model.*
*With rate 1 replace each type $\eta(x)$, $x \in V$, by a type that is randomly selected from $L_n(\eta)$:*

$$c_i(x, \eta) = L_n(\eta)(i) = \frac{1}{n} \left| \{x \in \eta \, : \, \eta(x) = i\} \right|$$

As a special case we now consider mean-field models with type space $T = \{0, 1\}$ or $T = \{-1, 1\}$. In this case the empirical distribution is completely determined by the frequence of type 1 in a configuration:

$$L_n(\eta) \longleftrightarrow N_1(\eta) = \left| \{x \, : \, \eta(x) = 1\} \right|$$
$$c_i(x, y) = \tilde{f}(N_1(\eta))$$

If $(\eta_t, P_x)$ is the corresponding mean field particle system, then (Exercise) $X_t = N_1(\eta)$ is a birth-death process on $\{0, 1, \dots, n\}$ with birth/death rates

$$b(k) = (n - k) \cdot \tilde{f}_1(k), \qquad d(k) = k \cdot \tilde{f}_0(k)$$

where $(n - k)$ is the number of particles with state 0 and $\tilde{f}_1(k)$ is the birth rate per particle.
$\rightsquigarrow$ Explicit computation of hitting times, stationary distributions etc.!

**Example .** 1. ***Binomial resampling:*** *For multinomial resampling with $T = \{0,1\}$ we obtain*

$$b(k) = d(k) = \frac{k \cdot (n-k)}{n}$$

2. ***Mean-field Ising model:*** *For the Ising model on the complete graph with inverse temperature $\beta$ and interaction strength $\frac{1}{n}$ the stationary distribution is*

$$\mu_\beta(\eta) \propto e^{-\frac{\beta}{4n}\sum_{x,y}(\eta(x)-\eta(y))^2} \propto e^{\frac{\beta}{2n}\sum_x \eta(x)\cdot\sum_y \eta(y)} = e^{\frac{\beta}{2n}m(\eta)^2}$$

*where*

$$m(\eta) = \sum_{x=1}^{n} \eta(x) = N_1(\eta) - N_{-1}(\eta) = 2N_1(\eta) - n$$

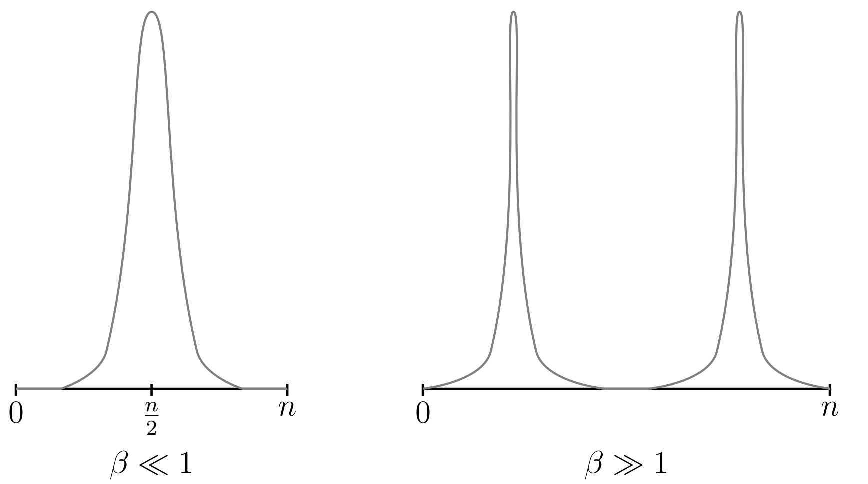
*is the **total magnetization**. Note that each $\eta(x)$ is interacting with the mean field $\frac{1}{n}\sum \eta(y)$, which explains the choice of interacting strength of order $\frac{1}{n}$. The birth-death chain $N_1(\eta_t)$ corresponding to the heat bath dynamics has birth and death rates*

$$b(k) = (n-k) \cdot \frac{e^{\beta\frac{k}{n}}}{e^{\beta\frac{k}{n}} + e^{\beta\frac{n-k}{n}}}, \qquad d(k) = k \cdot \frac{e^{\beta\frac{n-k}{n}}}{e^{\beta\frac{k}{n}} + e^{\beta\frac{n-k}{n}}}$$

*and stationary distribution*

$$\bar{\mu}_\beta(k) = \sum_{\eta\,:\,N_1(\eta)=k} \mu_\beta(\eta) \propto \binom{n}{k} 2^{-n} e^{\frac{2\beta}{n}\left(k-\frac{n}{2}\right)^2}, \quad 0 \le k \le n$$

*The binomial distribution $\mathrm{Bin}(n,\frac{1}{2})$ has a maximum at its mean value $\frac{n}{2}$, and standard deviation $\frac{\sqrt{n}}{2}$. Hence for large n, the measure $\bar{\mu}_\beta$ has one sharp mode of standard deviation $O(\sqrt{n})$ if $\beta$ is small, and two modes if $\beta$ is large:*



*The transition from uni- to multimodality occurs at an inverse temperature $\beta_n$ with*

$$\lim_{n\to\infty} \beta_n = 1 \qquad \textit{(Exercise)}$$

*The asymptotics of the stationary distribution as $n \to \infty$ can be described more accurately using large deviation results, cf. below.*

Now consider the heat bath dynamics with an initial configuration $\eta_0$ with $N_1(\eta_0) \leq \frac{n}{2}$, $n$ even, and let

$$T := \inf\left\{t \geq 0 \ : \ N_1(\eta_t) > \frac{n}{2}\right\}.$$

By the formula for mean hitting times for a birth-and-death process,

$$\mathbb{E}[T] \geq \frac{\bar{\mu}_\beta\left(\{0,1,\dots,\frac{n}{2}\}\right)}{\bar{\mu}_\beta\left(\frac{n}{2}\right)\cdot b\left(\frac{n}{2}\right)} \geq \frac{\frac{1}{2}}{\bar{\mu}_\beta\left(\frac{n}{2}\right)\cdot\frac{n}{2}} \geq \frac{e^{\beta\frac{n}{2}}}{n2^n}$$

since

$$\bar{\mu}_\beta\left(\frac{n}{2}\right) = \binom{n}{\frac{n}{2}} \cdot e^{-\frac{\beta n}{2}}\bar{\mu}_\beta(0) \leq 2^n e^{-\frac{\beta n}{2}}.$$

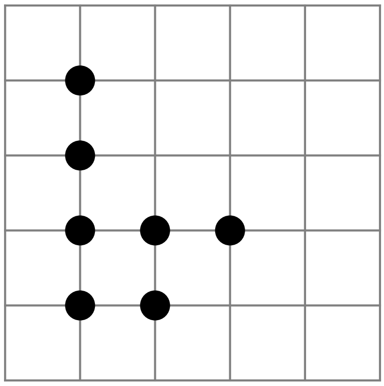
Hence the average time needed to go from configurations with negative magnetization to states with positive magnetization is increasing exponentially in $n$ for $\beta > 2\log 2$. Thus although ergodicity holds, for large $n$ the process gets stuck for a very large time in configurations with negative resp. positive magnetization.
$\rightsquigarrow$ *Metastable behaviour*.
More precisely, one can show using large deviation techniques that metastability occurs for any inverse temperature $\beta > 1$, cf. below.

## 2.2 Particle systems on $\mathbb{Z}^d$

Reference:

- Durett [9]
- Liggett [15]

| | |
|---|---|
| $V = \mathbb{Z}^d$, | $T$ finite |
| $E = \{(x,y) \ : \ \lvert x-y\rvert_{l^1} = 1\}$ | $S = T^{\mathbb{Z}^d}$ with product topology, compact |
| $\mu_n \to \mu \ \Leftrightarrow \ \mu_n(x) \to \mu(x) \,\forall\, x \in \mathbb{Z}^d$ | |



**Assumptions:**

$$(i)\quad \bar{\mu} := \sup_{\substack{i\in T\\ x\in\mathbb{Z}^d}} c_i(x,y) < \infty$$

$$(ii)\quad c_i(x,y) = g_i\left(\eta(x), (\eta(y))_{y\sim x}\right) \qquad \text{translation invariant and nearest neighbor}$$

## Graphical construction of associated particle systems:

Hier fehlt ein Bild!

$N_t^{x,i}$ independent Poisson process with rate $\bar{\lambda}$ (alarm clock for transition at $x$ to $i$)

$T_n^{x,i}$ $n$-th. arrival time of $N_t^{x,i}$

$U_n^{x,i}$ independent random variables uniformly distributed on $[0,1]$

**Recipe:** At time $T_n^{x,i}$, change $\eta(x)$ to $i$ provided

$$U_n^{x,i} \leq \frac{c_i(x,y)}{\bar{\lambda}} \qquad \left(\text{i.e. with probability } \frac{c_i(x,y)}{\bar{\lambda}}\right)$$

**Problem:** Infinitely many Poisson processes, hence transitions in arbitrary small time, no first transition.
How can we consistently define a process from the jump times? For a finite subset $A \subset \mathbb{Z}^d$ and $\xi \in S$, the restricted configuration space

$$S_{\xi,A} := \{\eta \in S \mid \eta = \xi \text{ on } A^c\}$$

is finite. Hence for all $s \geq 0$ there exists a unique Markov jump process $\left(\eta_t^{(s,\xi,A)}\right)_{t\geq s}$ on $S_{\xi,A}$ with initial condition $\eta_s^{(s,\xi,A)} = \xi$ and transitions $t \geq s,\ \eta \to \eta^{x,i}$ at times $T_n^{x,i}$ whenever $U_n^{x,i} \leq \frac{c_i(x,y)}{\bar{\lambda}}$, $x \in A$. The idea is now to define a Markov process $\eta_t^{(s,\xi)}$ on $S$ for $t-s$ small by

$$\eta_t^{(s,\xi)} := \eta_t^{(s,\xi,A)}$$

where $A$ is an appropriately chosen finite neighborhood of $x$. The neighborhood should be chosen in such a way that during the considered time interval, $\eta_t^{(s,\xi)}(x)$ has only been effected by previous values on $A$ of the configuration restricted to $A$. That this is possible is guaranteed by the following observation:
For $0 \leq s \leq t$ we define a random subgraph $(\mathbb{Z}^d, E_{s,t}(\omega))$ of $(V, E)$ by:

$$E_{s,t}(\omega) = \big\{\{x,y\} \;:\; T_n^{x,i} \in (s,t] \text{ or } T_n^{y,i} \in (s,t] \text{ for some } n \in \mathbb{N} \text{ and } i \in T\big\}$$

If $x$ effects $y$ in the time interval $(s,t]$ or vice versa then $\{x,y\} \in E_{s,t}$.

**Lemma 2.1.** *If*

$$t - s \leq \frac{1}{8 \cdot d^2 \cdot |T| \cdot \bar{\lambda}} =: \delta$$

*then*

$$P\left[\textit{all connected components of } (\mathbb{Z}^d, E_{s,t}) \textit{ are finite}\right] = 1.$$

**Consequence:** For small time intervals $[s,t]$ we can construct the configuration at time $t$ form the configuration at time $s$ independently for each component by the standard construction for jump processes with finite state space.

*Proof.* By translation invariance it suffices to show

$$P[|C_0| < \infty] = 1$$

where $C_0$ is the component of $(\mathbb{Z}^d, E_{s,t})$ containing $0$. If $x$ is in $C_0$ then there exists a self-avoiding path in $(\mathbb{Z}^d, E_{s,t})$ starting at $0$ with length $d_{l^1}(x,0)$. Hence

$$\begin{aligned}
&P[\exists\, x \in C_0 \;:\; d_{l^1}(x,0) \geq 2n-1] \\
&\leq P[\exists \text{ self-avoiding path } z_1 = 0, z_2, \ldots, z_{2n-1} \text{ s.t. } (z_i, z_{i+1}) \in E_{s,t}\ \forall\, i] \\
&\leq (2d)^{2n-1} \cdot \prod_{i=0}^{n-1} P[(z_{2i}, z_{2i+1}) \in E_{s,t}]
\end{aligned}$$

where $(2d)^{2n-1}$ is a bound for the number of self-avoiding paths starting at 0 and independent events $\{(z_{2i}, z_{2i+1}) \in E_{s,t}\}$.
Hence

$$\begin{aligned}
P[\exists\, x \in C_0 \;:\; d_{l^1}(x,0) \geq 2n-1] \;&\leq\; \left(4d^2 \cdot \left(1 - e^{-2|T|\bar{\lambda}(t-s)}\right)\right)^n \\
&\leq (8d^2 \cdot |T|\bar{\lambda} \cdot (t-s))^n \;\longrightarrow\; 0
\end{aligned}$$

as $n \to \infty$, where $e^{-2|T|\bar{\lambda}(t-s)}$ is the probability for no arrival in $[s,t]$ in a $2|T|$ $\text{Poisson}(\bar{\lambda})$ process and $1 - e^{-2|T|\bar{\lambda}(t-s)} \leq 2|T|\bar{\lambda} \cdot (t-s)$. □

By the lemma, $P$-almost sure for all $s > 0$ and $\xi \in T^{\mathbb{Z}^d}$, there is an unique function $t \mapsto \eta_t^{(s,\xi)}$, $t \geq s$, such that

(i) $\eta_s^{(s,\xi)} = \xi$

(ii) For $s \leq t,\ h \leq \delta$, and each connected component $C$ of $(\mathbb{Z}^d, E_{t,t+h})$, $\eta_{t+h}^{(s,\xi)}\Big|_C$ is obtained from $\eta_t^{(s,\xi)}\Big|_C$ by subsequently taking into account the finite number of transitions in $C$ during $[t, t+h]$.

We set

$$\eta_t^\xi := \eta_t^{0,\xi}.$$

By construction,

$$\eta_t^\xi = \eta_t^{(s,\eta_s^\xi)} \quad \forall\, 0 \leq s \leq t \tag{2.1}$$

**Corollary 2.2.** *(i)* ***Time-homogeneity:***

$$\left(\eta_{s+t}^{(s,\xi)}\right)_{t\geq 0} \sim \left(\eta_t^\xi\right)_{t\geq 0}$$

*(ii)* $(\eta_t^\xi, P)$ *is a Markov process with transition semigroup*

$$(p_t f)(\xi) = \mathbb{E}[f(\eta_t^\xi)]$$

*(iii)* ***Feller property:***

$$f \in C_b(S) \implies p_t f \in C_b(S)\ \forall\, t \geq 0$$

*Or, equivalently, $p_t f$ is continuous whenever $f$ is continuous with respect to the product topology. Since $S$ is compact, any continuous function is automatically bounded.*

*(iv)* ***Translation invariance:*** *Let $\xi \colon \Omega \to S$ be a random variable, independent of all $N_t^{x,i}$ and translation invariant, i.e. $\xi(x + \bullet) \sim \xi$ for all $x \in \mathbb{Z}^d$. Then $\eta_t^\xi$ is translation invariant for all $t \geq 0$ $P$-a.s.*

*Sketch of proof:* (i) by the time homogeneity of the Poisson arrivals.

(ii)

$$\mathbb{E}\left[f\left(\eta_t^\xi\right) \mid \mathcal{F}_s\right](\omega) \overset{2.1}{=} \mathbb{E}\left[f\left(\eta_t^{(s,\eta_s^\xi)}\right) \mid \mathcal{F}_s\right](\omega)$$

taking into account the $\mathcal{F}_s$-measurability of $\eta_s^\xi$ and $\eta_t^{(s,\xi)}$ being independent of $\mathcal{F}_s$ for fixed $\xi$, we conclude with $(i)$

$$\begin{aligned}\mathbb{E}\left[f\left(\eta_t^\xi\right) \mid \mathcal{F}_s\right](\omega) &= \mathbb{E}\left[f\left(\eta_t^{(s,\eta_s^\xi(\omega))}\right)\right] \\ &= \mathbb{E}\left[f\left(\eta_{t-s}^{\eta_s^\xi(\omega)}\right)\right] \\ &= (p_{t-s} f)\left(\eta_s^\xi(\omega)\right)\end{aligned}$$

(iii)

$$\xi_n \to \xi \;\Rightarrow\; \xi_n(x) \to \xi(x)\,\forall\, x \in \mathbb{Z}^d$$

Hence $\xi_n = \xi$ eventually on each finite set $C \subset \mathbb{Z}^d$, and hence on each component of $\left(\mathbb{Z}^d, E_{0,\delta}\right)$. By the componentwise construction,

$$\eta_t^{\xi_n} = \eta_t^{\xi} \quad \forall\, t \le \delta$$

eventually on each component. Hence

$$\eta_t^{\xi_n} \to \eta_t^{\xi} \quad \text{(pointwise)}\,\forall\, t \le \delta$$

and for $f \in C_b(S)$,

$$f\left(\eta_t^{\xi_n}\right) \to f\left(\eta_t^{\xi}\right)$$

for all $t \le \delta$. With Lebesgue we conclude

$$p_t f(\xi_n) = \mathbb{E}\left[f\left(\eta_t^{\xi_n}\right)\right] \longrightarrow p_t f(\xi)\ \forall t \le \delta$$

Hence OK for $t \le \delta$. General case by semigroup property:

$$p_t = p_{t-\lfloor \frac{t}{\delta}\rfloor \cdot \delta}\, p_\delta^{\lfloor \frac{t}{\delta}\rfloor} : C_b(S) \to C_b(S)$$

(iv) The $c_i(x,y)$ are translation invariant by assumption,

$$\left(\left(N_n^{x,i}\right)_{t,i}, \left(U_n^{x,i}\right)_{n,i}\right)$$

are identically distributed. This gives the claim.

□

**Theorem 2.3** (Forward equation)**.** *For any cylinder function*

$$f(\eta) = \varphi\left(\eta(x_1), \dots, \eta(x_n)\right), \quad n \in \mathbb{N}\ \varphi\colon T^n \to \mathbb{R}$$

*the forward equation*

$$\frac{d}{dt}\left(p_t f\right)(\xi) = \left(p_t \mathscr{L} f\right)(\xi)$$

*holds for all* $\xi \in S$ *where*

$$(\mathscr{L} f)(\xi) = \sum_{\substack{x \in \mathbb{Z}^d \\ i \in T}} c_i(x,\xi) \cdot \left(f(\xi^{\xi,i}) - f(\xi)\right)$$

**Remark .** *Since $f$ is a cylinder function, the sum in the formula for the generator has only finitely many non-zero summands.*

*Proof.*

$$P\left[\sum_{k=1,\dots,n i\in T} N_t^{x_k,i} > 1\right] \leq \text{const.} \cdot t^2$$

where $\{N_t^{x_k,i} > 1\}$ means that there is more than one transition in the time interval $[0,t]$ among $\{x_1,\dots,x_n\}$ and const. is a global constant.

$$P\left[N_t^{x_k,i} = 1\right] = \bar{\lambda} \cdot t + O(t^2)$$

and hence

$$\begin{aligned}
(p_t f)(\xi) &= \mathbb{E}[f(\eta_t^\xi)] \\
&= f(\xi) \cdot P\left[N_t^{x_k,i} = 0 \quad \forall\, 1 \leq k \leq n,\ i \in T\right] \\
&+ \sum_{i,k} f(\xi^{x_k,i}) \cdot P\left[N_t^{x_k,i} = 1,\ U_1^{x_k,i} \leq \frac{c_i(x,\xi)}{\bar{\lambda}}\right] + O(t^2) \\
&= f(\xi) + \sum_{i,k} t \cdot \frac{\bar{\lambda} c_i(x_k,\xi)}{\bar{\lambda}} \cdot \big(f(\xi^{x_k,i}) - f(\xi)\big) + O(t^2) \\
&= f(\xi) + t \cdot (\mathscr{L} f)(\xi) + O(t^2)
\end{aligned}$$

where the constants $O(t^2)$ do not depend on $\xi$. Hence

$$p_{t+h} f = p_t p_h f = p_t f + h p_t \mathscr{L} f + O(h^2)$$

□

## 2.3 Stationary distributions and phase transitions

The reference for this chapter is Liggett [15].

From now on we assume $T = \{0,1\}$. We define a ***partial order*** on configurations $\eta, \tilde{\eta} \in S = \{0,1\}^{\mathbb{Z}^d}$ by

$$\eta \leq \tilde{\eta} \ :\Leftrightarrow\ \eta(x) \leq \tilde{\eta}(x)\ \forall\, x \in \mathbb{Z}^d$$

A function $f\colon S \to \mathbb{R}$ is called ***increasing*** if and only if

$$f(\eta) \leq f(\tilde{\eta}) \text{ whenever } \eta \leq \tilde{\eta}.$$

**Definition 2.4** (Stochastic dominance). *For probability measures $\mu, \nu \in M_1(S)$ we set*

$$\mu \preccurlyeq \nu \;:\Leftrightarrow\; \int f\, d\mu \leq \int f\, d\nu \text{ for any increasing bounded function } f\colon S \to \mathbb{R}$$

**Example .** *For $\mu, \nu \in M_1(\mathbb{R})$,*

$$\mu \preccurlyeq \nu \;\Leftrightarrow\; F_\mu(c) = \mu\left((-\infty, c]\right) \geq F_\nu(c) \quad \forall\, c \in \mathbb{R}$$

Now consider again the stochastic dynamics constructed above.

$$\begin{array}{ll} c_1(x,\eta) & \text{birth rates} \\ c_0(x,\eta) & \text{death rates} \end{array}$$

**Definition 2.5.** *The Markov process $\left(\eta_t^\xi, P\right)$ is called **attractive** if and only if for all $x \in \mathbb{Z}^d$,*

$$\eta \leq \tilde{\eta},\; \eta(x) = \tilde{\eta}(x) \quad \Rightarrow \quad \begin{cases} c_1(x,\eta) \leq c_1(x,\tilde{\eta}) & \text{and} \\ c_0(x,\eta) \geq c_0(x,\tilde{\eta}) \end{cases}$$

**Example .** *Contact process, voter model, as well as the Metropolis and heat-bath dynamics for the (ferromagnetic) Ising model are attractive*

**Theorem 2.6.** *If the dynamics is attractive then:*

1. *If $\xi \leq \tilde{\xi}$ then $\eta_t^\xi \leq \eta_t^{\tilde{\xi}}$ for all $t \geq 0$ $P$-a.s.*
2. *If $f\colon S \to \mathbb{R}$ is increasing then $p_t f$ is increasing for all $t \geq 0$.*
3. *If $\mu \preccurlyeq \nu$ then $\mu p_t \preccurlyeq \nu p_t$ for all $t \geq 0$ (*Monotonicity*).*

*Proof.* 1. The dynamics is attractive and $\xi \leq \tilde{\xi}$, hence every single transition preserves order. Hence

$$\begin{array}{lll} & \eta_t^{(s,\xi,A)} \leq \eta_t^{(s,\tilde{\xi},A)} & \forall\, 0 \leq s \leq t,\; A \subset \mathbb{Z}^d \text{ finite} \\ \Rightarrow & \eta_t^{(s,\xi)} \leq \eta_t^{(s,\tilde{\xi})} & \forall\, s \geq 0,\; t \in [s, s+\delta] \end{array}$$

and by induction

$$\eta_t^{(s,\xi)} \leq \eta_t^{(s,\tilde{\xi})} \quad \forall\, t \geq s \geq 0$$

since $\eta_t^{(s,\xi)} = \eta_t^{\left(s+\delta,\eta_{s+\delta}^{(s,\xi)}\right)}$.

(If, for example, before a possible transition at time $T_n^{x,1}$, $\eta \leq \tilde{\eta}$ and $\eta(x) = \tilde{\eta}(x) = 0$, then after the transition, $\eta(x) = 1$ if $U_n^{x,1} \leq \frac{c_1(x,\eta)}{\lambda}$, but in this case also $\tilde{\eta}(x) = 1$ since $c_1(x,\eta) \leq c_1(x,\tilde{\eta})$ by attractiveness. The other cases are checked similarly.)

2. Since $f$ is increasing and $\xi \leq \tilde{\xi}$,

$$(p_t f)(\xi) = \mathbb{E}\left[f(\eta_t^\xi)\right] \leq \mathbb{E}\left[f(\eta_t^{\tilde{\xi}})\right] = (p_t f)(\tilde{\xi})$$

3. If $f$ is increasing, $p_t f$ is increasing as well and hence by Fubini

$$\int f\, d(\mu p_t) = \int p_t f\, d\mu \overset{\mu \preccurlyeq \nu}{\leq} \int p_t f\, d\nu = \int f\, d(\nu p_t)$$

□

Let $0, 1 \in S$ denote the constant configurations and $\delta_0, \delta_1$ the minimal respectively maximal element in $M_1(S)$.

**Theorem 2.7.** *For an attractive particle system on $\{0,1\}^{\mathbb{Z}^d}$ we have*

1. *The functions $t \mapsto \delta_0 p_t$ and $t \mapsto \delta_1 p_t$ are decreasing respectively increasing with respect to $\preccurlyeq$.*
2. *The limits $\underline{\mu} := \lim_{t\to\infty} \delta_0 p_t$ and $\bar{\mu} := \lim_{t\to\infty} \delta_1 p_t$ exist with respect to weak convergence in $M_1(S)$*
3. *$\underline{\mu}$ and $\bar{\mu}$ are stationary distributions for $p_t$*
4. *Any stationary distribution $\pi$ satisfies*

$$\underline{\mu} \preccurlyeq \pi \preccurlyeq \bar{\mu}.$$

*Proof.* 1.

$$0 \leq s \leq t \quad \Rightarrow \quad \delta_0 \preccurlyeq \delta_0 p_{t-s}$$

and hence by monotonicity

$$\delta_0 p_s \preccurlyeq \delta_0 p_{t-s} p_s = \delta_0 p_t$$

2. By monotonicity and compactness, since $S = \{0,1\}^{\mathbb{Z}^d}$ is compact with respect to the product topology, $M_1(S)$ is compact with respect to weak convergence. Thus it suffices to show that any two subsequential limits $\mu_1$ and $\mu_2$ of $\delta_0 p_t$ coincide. Now by 1),

$$\int f\, d(\delta_0 p_t)$$

is increasing in $t$, and hence

$$\int f\,d\mu_1 = \lim_{t\uparrow\infty} \int f\,d(\delta_0 p_t) = \int f\,d\mu_2$$

for any continuous increasing function $f\colon S \to \mathbb{R}$, which implies $\mu_1 = \mu_2$.

3. Since $p_t$ is Feller,

$$\begin{aligned}\int f\,d(\underline{\mu} p_t) &= \int p_t f\,d\underline{\mu} = \lim_{s\to\infty} \int p_t f\,d(\delta_0 p_s) = \lim_{s\to\infty} \int f\,d(\delta_0 p_s p_t)\\ &= \lim_{s\to\infty} \int f\,d(\delta_0 p_s) = \int f\,d\underline{\mu}\end{aligned}$$

for all $f \in C_b(S)$.

4. Since $\pi$ is stationary,

$$\delta_0 p_t \preccurlyeq \pi p_t = \pi \preccurlyeq \delta_1 p_t$$

for all $t \geq 0$ and hence for $t \to \infty$,

$$\underline{\mu} \preccurlyeq \pi \preccurlyeq \bar{\mu}.$$

□

**Corollary 2.8.** *For an attractive particle system, the following statements are equivalent:*

1. $\underline{\mu} = \bar{\mu}$.
2. *There is an unique stationary distribution.*
3. ***Ergodicity** holds:*

$$\exists\, \mu \in M_1(S): \quad \nu p_t \longrightarrow \mu \quad \forall\, \nu \in M_1(S).$$

*Proof.* 1. $\Leftrightarrow$ 2. : by the theorem.

1. $\Rightarrow$ 3. : Since $\delta_0 \preccurlyeq \nu \preccurlyeq \delta_1$,

$$\delta_0 p_t \preccurlyeq \nu p_t \preccurlyeq \delta_1 p_t$$

and since $\delta_0 p_t \to \underline{\mu}$ and $\delta_1 p_t \to \bar{\mu}$ for $t \to \infty$,

$$\nu p_t \to \underline{\mu} = \bar{\mu}$$

3. $\Rightarrow$ 1.: obvious.

□

**Example 1: Contact process on $\mathbb{Z}^d$**

For the contact process, $c_0(x, \eta) = \delta$ and $c_1(x, \eta) = b \cdot N_1(x, \eta)$ where the birth rate $b$ and the death rate $\delta$ are positive constants. Since the 0 configuration is an absorbing state, $\underline{\mu} = \delta_0$ is the minimal stationary distribution. The question now is if there is another (non-trivial) stationary distribution, i.e. if $\bar{\mu} \neq \underline{\mu}$.

**Theorem 2.9.** *If $2db < \delta$ then $\delta_0$ is the only stationary distribution, and ergodicity holds.*

*Proof.* By the forward equation and translation invariance,

$$\begin{aligned}\frac{d}{dt} P\left[\eta_t^1(x) = 1\right] &= -\delta P\left[\eta_t^1(x) = 1\right] + \sum_{y\,:\,|x-y|=1} b \cdot P\left[\eta_t^1(x) = 0,\ \eta_t^1(y) = 1\right] \\ &\leq (-\delta + 2db) \cdot P\left[\eta_t^1(x) = 1\right]\end{aligned}$$

for all $x \in \mathbb{Z}^d$. Hence if $2db < \delta$ then

$$\begin{aligned}\bar{\mu}\left(\{\eta \,:\, \eta(x) = 1\}\right) &= \lim_{t\to\infty} (\delta_1 p_t)(\{\eta \,:\, \eta(x) = 1\}) \\ &= \lim_{t\to\infty} P\left[\eta_t^1(x) = 1\right] \\ &= 0\end{aligned}$$

for all $x \in \mathbb{Z}^d$ and thus $\bar{\mu} = \delta_0$. □

Conversely, one can show that for $b$ sufficiently small (or $\delta$ sufficiently large), there is nontrivial stationary distribution. The proof is more involved, cf. Liggett [15]. Thus a phase transition from ergodicity to non-ergodicity occurs as $b$ increases.

**Example 2: Ising model on $\mathbb{Z}^d$**

We consider the heat bath or Metropolis dynamics with inverse temperature $\beta > 0$ on $S = \{-1, +1\}^{\mathbb{Z}^d}$.

a) ***Finite volume:*** Let $A \subseteq \mathbb{Z}^d$ be finite,

$$\begin{aligned}S_{+,A} &:= \{\eta \in S \mid \eta = +1 \text{ on } A^c\} \qquad \text{(finite!)} \\ S_{-,A} &:= \{\eta \in S \mid \eta = -1 \text{ on } A^c\}\,.\end{aligned}$$

For $\xi \in S_{+,A}$ resp. $\xi \in S_{-,A}$, $\eta_t^{\xi,A} = \eta_t^{(0,\xi,A)}$, the dynamics taking into account only transitions in $A$.

$\left(\eta_t^{\xi,A}, P\right)$ is a Markov chain on $S_{+,A}$ resp. $S_{-,A}$ with generator

$$(\mathscr{L}f)(\eta) = \sum_{\substack{x \in A \\ i \in \{-1,+1\}}} c_i(x,\eta) \cdot \big(f(\eta^{x,i}) - f(\eta)\big)$$

Let

$$H(\eta) = \frac{1}{4} \sum_{\substack{x,y \in \mathbb{Z}^d \\ |x-y|=1}} (\eta(x) - \eta(y))^2$$

denote the ***Ising Hamiltonian***. Note that for $\eta \in S_{+,A}$ or $\eta \in S_{-,A}$ only finitely many summands do not vanish, so $H(\eta)$ is finite. The probability measure

$$\mu_\beta^{+,A}(\eta) = \frac{1}{Z_\beta^{+,A}} e^{-\beta H(\eta)}, \quad \eta \in S_{+,A}$$

where

$$Z_\beta^{+,A} = \sum_{\eta \in S_{+,A}} e^{-\beta H(\eta)}$$

on $S_{+,A}$ and $\mu_\beta^{-,A}$ on $S_{-,A}$ defined correspondingly satisfy the detailed balance conditions

$$\mu_\beta^{+,A}(\xi)\mathscr{L}(\xi,\eta) = \mu_\beta^{+,A}(\eta)\mathscr{L}(\eta,\xi) \quad \forall\, \xi,\eta \in S_{+,A}$$

respectively

$$\mu_\beta^{-,A}(\xi)\mathscr{L}(\xi,\eta) = \mu_\beta^{-,A}(\eta)\mathscr{L}(\eta,\xi) \quad \forall\, \xi,\eta \in S_{-,A}.$$

Since $S_{+,A}$ and $S_{-,A}$ are finite and irreducible this implies that $\mu_\beta^{+,A}$ respectively $\mu_\beta^{-,A}$ is the unique stationary distribution of $\left(\mu_t^{\xi,A}, P\right)$ for $\xi \in S_{+,A}, S_{-,A}$ respectively. Thus in finite volume there are several processes corresponding to different boundary conditions (which effect the Hamiltonian) but each of them has a unique stationary distribution. Conversely, in infinite volume there is only one process, but it may have several stationary distributions:

b) ***Infinite volume:*** To identify the stationary distributions for the process on $\mathbb{Z}^d$, we use an approximation by the dynamics in finite volume. For $n \in \mathbb{N}$ let

$$A_n := [-n,n]^d \cap \mathbb{Z}^d,$$

$$\xi_n(x) := \begin{cases} \xi(x) & \text{for } x \in A_n \\ +1 & \text{for } x \in \mathbb{Z}^d \setminus A_n \end{cases}$$

The sequences $\mu_\beta^{+,A_n}$ and $\mu_\beta^{-,A_n}$, $n \in \mathbb{N}$, are decreasing respectively increasing with respect to stochastic dominance. Hence my compactness of $\{-1,+1\}^{\mathbb{Z}^d}$ there exist

$$\mu_\beta^+ := \lim_{n \uparrow \infty} \mu_\beta^{+,A_n} \quad \text{and} \quad \mu_\beta^- := \lim_{n \uparrow \infty} \mu_\beta^{-,A_n}$$

**Remark** (Gibbs measures)**.** *A probability measure $\mu$ on $S$ is called **Gibbs measure** for the Ising Hamiltonian on $\mathbb{Z}^d$ and inverse temperature $\beta > 0$ if and only if for all finite $A \subseteq \mathbb{Z}^d$ and $\xi \in S$,*

$$\mu_\beta^{\xi,A}(\eta) := \frac{1}{Z_\beta^{\xi,A}} e^{-\beta H(\eta)}, \quad \eta \in S_{\xi,A} := \{\eta \in S \mid \eta = \xi \text{ on } A^c\},$$

*is a version of the conditional distribution of $\mu_\beta$ given $\eta(x) = \xi(x)$ for all $x \in A^c$. One can show that $\mu_\beta^+$ and $\mu_\beta^-$ are the extremal Gibbs measures for the Ising model with respect to stochastic dominance, cf. e.g. [Milos] ???.*

**Definition 2.10.** *We say that a **phase transition** occurs for $\beta > 0$ if and only if $\mu_\beta^+ \neq \mu_\beta^-$*

For $\xi \in S$ define $\xi_n \in S_{+,A_n}$ by

$$\xi_n(x) := \begin{cases} \xi(x) & \text{for } x \in A_n \\ +1 & \text{for } x \in \mathbb{Z}^d \setminus A_n \end{cases}$$

**Lemma 2.11.** *For all $x \in \mathbb{Z}^d$ and $f \in [0, \delta]$,*

$$P\left[\eta_t^\xi(x) \neq \eta_t^{\xi_n,A_n}(x) \quad \text{for some } \xi \in S\right] \longrightarrow 0 \tag{2.2}$$

*as $n \to \infty$.*

*Proof.* Let $C_x$ denote the component containing $x$ in the random graph $\left(\mathbb{Z}^d, E_{0,\delta}\right)$. If $C_x \subseteq A_n$ then the modifications in the initial condition and the transition mechanism outside $A_n$ do not effect the value at $x$ before time $\delta$. Hence the probability in (2.2) can be estimated by

$$P\left[C_x \cap A_n^c \neq \varnothing\right]$$

which goes to $0$ as $n \to \infty$ by Lemma (2.1) above. □

Let $p_t$ denote the transition semigroup on $\{-1, 1\}^{\mathbb{Z}^d}$. Since the dynamics is attractive,

$$\bar{\mu}_\beta = \lim_{t\to\infty} \delta_{+1} p_t \quad \text{and} \quad \underline{\mu}_\beta = \lim_{t\to\infty} \delta_{-1} p_t$$

are extremal stationary distributions with respect to stochastic dominance. The following theorem identifies $\bar{\mu}$ and $\underline{\mu}$ as the extremal Gibbs measures for the Ising Hamiltonian on $\mathbb{Z}^d$:

**Theorem 2.12.** *The upper and lower invariant measures are*

$$\bar{\mu}_\beta = \mu_\beta^+ \quad \text{and} \quad \underline{\mu}_\beta = \mu_\beta^-.$$

*In particular, ergodicity holds if and only if there is no phase transition (i.e. iff $\mu_\beta^+ = \mu_\beta^-$).*

*Proof.* We show:

1. $\bar{\mu}_\beta \preccurlyeq \mu_\beta^+$
2. $\mu_\beta^+$ is a stationary distribution with respect to $p_t$.

This implies $\bar{\mu}_\beta = \mu_\beta^+$, since by 2. and the corollary above, $\mu_\beta^+ \preccurlyeq \bar{\mu}_\beta$, and thus $\mu_\beta^+ = \bar{\mu}_\beta$ by 1. $\mu_\beta^- = \underline{\mu}_\beta$ follows similarly.

1. It can be shown similarly as above that, the attractiveness of the dynamics implies

$$\mu_t^1 \leq \mu_t^{1,A_n}$$

$P$-a.s. for all $n \in \mathbb{N}$ and $t \geq 0$. As $t \to \infty$,

$$\mu_t^1 \xrightarrow{\mathscr{D}} \bar{\mu}_\beta \quad \text{and} \quad \eta_t^{1,A_n} \xrightarrow{\mathscr{D}} \mu_\beta^{+,A_n},$$

hence

$$\bar{\mu}_\beta \preccurlyeq \mu_\beta^{+,A_n}$$

for all $n \in \mathbb{N}$. The assertion follows as $n \to \infty$.

2. It is enough to show

$$\mu_\beta^+ p_t = \mu_\beta^+ \quad \text{for } t \leq \delta, \tag{2.3}$$

then the assertion follows by the semigroup property of $(p_t)_{t\geq 0}$. Let

$$(p_t^n f)(\xi) := \mathbb{E}\left[f(\eta_t^{\xi_n,A_n})\right]$$

denote the transition semigroup on $S_{\xi_n,A_n}$. We know:

$$\mu_\beta^{+,n} p_t^n = \mu_\beta^{+,n} \tag{2.4}$$

To pass to the limit $n \to \infty$ let $f(\eta) = \varphi(\eta(x_1), \dots, \eta(x_k))$ be a cylinder function on $S$. Then

$$\int p_t f \, d\mu_\beta^{+,n} = \int p_t^n f \, d\mu_\beta^{+,n} + \int (p_t^n f - p_t f) \, d\mu_\beta^{+,n} \tag{2.5}$$

and by (2.4) this is equal to

$$\int f \, d\mu_\beta^{+,n} + \int (p_t^n f - p_t f) \, d\mu_\beta^{+,n}$$

But by the lemma above, for $t \leq \delta$,

$$\begin{aligned} |(p_t^n f)(\xi) - (p_t f)(\xi)| &\leq \mathbb{E}\left[\left|f\left(\eta_t^{\xi_n,A_n}\right) - f\left(\eta_t^\xi\right)\right|\right] \\ &\leq 2 \cdot \sup|f| \cdot P\left[\eta_t^{\xi_n,A_n}(x_i) \neq \eta_t^\xi(x_i) \text{ for some } i\right] \longrightarrow 0 \end{aligned}$$

uniformly in $\xi$.

Since $\mu_\beta^{+,n} \overset{w}{to} \mu_\beta^+$, and $f$ and $p_t f$ are continuous by the Feller property, taking the limit in (2.5) as $n \to \infty$ yields

$$\int f \, d\left(\mu_\beta^+ p_t\right) = \int p_t f \, d\mu_\beta^+ = \int f \, d\mu_\beta^+$$

for all cylinder functions $f$, which implies (2.3).

□

The question now is: when does a phase transition occur?
For $\beta = 0$, there is no interaction between $\eta(x)$ and $\eta(y)$ for $x \neq y$. Hence $\eta_\beta^{+,n}$ and $\eta_\beta^{-,n}$ are the uniform distributions on $S_{+,A_n}$ and $S_{-,A_n}$, and

$$\mu_\beta^+ = \mu_\beta^- = \bigotimes_{z \in \mathbb{Z}^d} \nu, \quad \text{where } \nu\left(\pm 1\right) = \frac{1}{2}$$

On the other hand, phase transition occur for $d \geq 2$ and large values of $\beta$:

**Theorem 2.13** (PEIERL). *For $d = 2$ there exists $\beta_c \in (0, \infty)$ such that for $\beta > \beta_c$,*

$$\mu_\beta^+\left(\{\eta \,:\, \eta(0) = -1\}\right) < \frac{1}{2} < \mu_\beta^-\left(\{\eta \,:\, \eta(0) = -1\}\right),$$

*and thus $\mu_\beta^+ \neq \mu_\beta^-$.*

*Proof.* Let $C_0(\eta)$ denote the connected component of $0$ in $\{x \in \mathbb{Z}^d \mid \eta(x) = -1\}$, and set $C_0(\eta) = \varnothing$ if $\eta(0) = +1$. Let $A \subseteq \mathbb{Z}^d$ be finite and non-empty. For $\eta \in S$ with $C_0 = A$ let $\tilde{\eta}$ denote the configuration obtained by reversing all spins in $A$. Then

$$H(\tilde{\eta}) = H(\eta) - 2|\partial A|,$$

and hence

$$\begin{aligned}\mu_\beta^{+,n}(C_0 = A) &= \sum_{\eta \,:\, C_0(\eta) = A} \mu_\beta^{+,n}(\eta) \\ &\leq e^{-2\beta|\partial A|} \underbrace{\sum_{\eta \,:\, C_0(\eta) = A} \mu_\beta^{+,n}(\tilde{\eta})}_{\leq 1} \leq e^{-2\beta|\partial A|}\end{aligned}$$

Thus

$$\begin{aligned}\mu_\beta^{+,n}\left(\{\eta\ :\ \eta(0)=-1\}\right) &= \sum_{\substack{A\subset\mathbb{Z}^d\\ A\neq\varnothing}} \mu_\beta^{+,n}(C_0=A)\\ &\le \sum_{L=1}^{\infty} e^{-2\beta L}\left|\{A\subset\mathbb{Z}^d\ :\ |\partial A|=L\}\right|\\ &\le \sum_{L=4}^{\infty} e^{-2\beta L}\cdot 4\cdot 3^{L-1}\cdot L^2\\ &\le \frac{1}{2}\quad \text{for } \beta>\beta_c\end{aligned}$$

where $\partial A$ is a self-avoiding path in $\mathbb{Z}^2$ by length $L$, starting in $\left(-\frac{L}{2},\frac{L}{2}\right)^2$. Hence for $n\to\infty$,

$$\mu_\beta^{+}\left(\{\eta\ :\ \eta(0)=-1\}\right)<\frac{1}{2}$$

and by symmetry

$$\mu_\beta^{-}\left(\{\eta\ :\ \eta(0)=-1\}\right)=\mu_\beta^{+}\left(\{\eta\ :\ \eta(0)=1\}\right)>\frac{1}{2}$$

for $\beta>\beta_c$. □

## 2.4 Poisson point process

Let $S$ be a polish space (e.g. $\mathbb{R}^d$) and $\nu$ a $\sigma$-finite measure on the Borel $\sigma$-algebra $\mathcal{S}$.

**Definition 2.14.** *A collection of random variables $N(B),\ B\in\mathcal{S}$, on a probability space $(\Omega,\mathcal{A},P)$ is called a* ***Poisson random measure (Poisson random field, spatial Poisson process) of intensity $\nu$****, if and only if*

*(i)* $B\mapsto N(B)(\omega)$ *is a positive measure for all* $\omega\in\Omega$.

*(ii) If* $B_1,\dots,B_n\in\mathcal{S}$ *are disjoint, then the random variables* $N(B_1),\dots,N(B_n)$ *are independent.*

*(iii)* $N(B)$ *is* Poisson$(\nu(B))$*-distributed for all* $B\in\mathcal{S}$ *with* $\nu(B)<\infty$.

**Example .** *If $N_t$ is a standard Poisson process with intensity $\lambda>0$ the number*

$$N(B):=\left|\{t\in B\mid N_{t-}\neq N_t\}\right|,\quad B\in\mathcal{B}(\mathbb{R}^+)$$

*of arrivals in a time set $B$ is a Poisson random measure on $\mathbb{R}^+$ of intensity $\nu=\lambda\,dx$, and*

$$N_t-N_s=N([s,t]),\quad \forall\, 0\le s\le t$$

**Construction of Poisson random measures:**

a) $\boldsymbol{\nu(S) < \infty}$ **:** Define $\lambda := \nu(S)$. Let $X_1, X_2, \ldots$ be independent and identically distributed random variables, $\lambda^{-1}\nu$-distributed. Let $K$ be a $\text{Poisson}(\lambda)$ distributed random variable, independent of $X_i$. Then

$$N := \sum_{k=1}^{K} \delta_{X_i}$$

is a Poisson random measure of intensity $\nu$.

b) $\boldsymbol{\nu}$ $\boldsymbol{\sigma}$**-finite:** Let $S = \dot{\bigcup}_{i\in\mathbb{N}} S_i$ with $\nu(S_i) < \infty$. Let $N_i$ be independent Poisson random measures with intensity $I_{S_i} \cdot \nu$. Then

$$N := \sum_{i=1}^{\infty} N_i$$

is a Poisson random measure with intensity $\nu = \sum_{i=1}^{\infty} I_{S_i} \cdot \nu$.

**Definition 2.15.** *A collection $N_t(B),\ t \geq 0,\ B \in \mathcal{S}$, of random variables on a probability space $(\Omega, \mathcal{A}, P)$ is called a* ***Poisson point process of intensity $\nu$*** *if and only if*

*(i)* $B \mapsto N_t(B)(\omega)$ *is a positive measure for all* $t \geq 0,\ \omega \in \Omega$.

*(ii) If* $B_1, \ldots, B_n \in \mathcal{S}$ *are disjoint, then* $(N_t(B_1))_{t\geq 0}, \ldots, (N_t(B_n))_{t\geq 0}$ *are independent.*

*(iii)* $(N_t(B))_{t\geq 0}$ *is a Poisson process of intensity* $\nu(B)$ *for all* $B \in \mathcal{S}$ *with* $\nu(B) < \infty$.

**Remark .** *A Poisson random measure (respectively a Poisson point process) is a random variable (respectively a stochastic process) with values in the space*

$$M_c^+(S) = \left\{ \sum_{x\in A} \delta_x \mid A \subseteq S \text{ countable subset} \right\} \subseteq M^+(S)$$

*of all counting measures on $S$. The distribution of a Poisson random measure and a Poisson point process of given intensity is determined uniquely by the definition.*

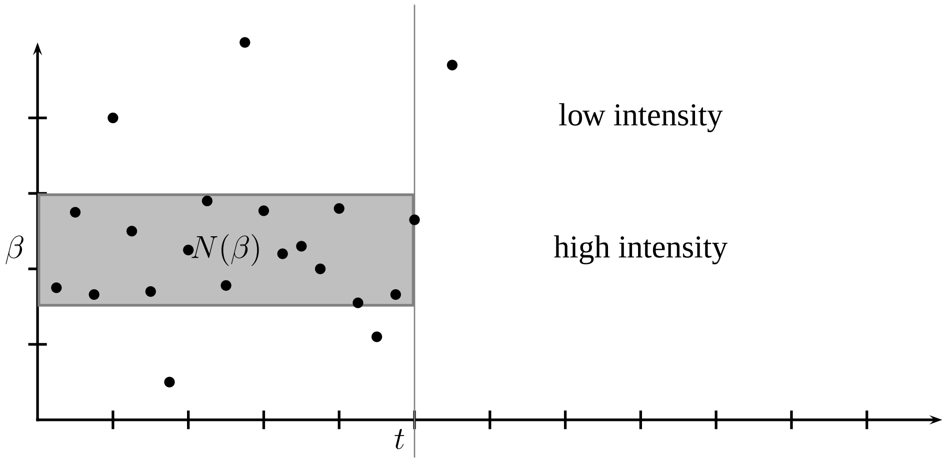
**Theorem 2.16** (Construction of Poisson point processes). *1. If $N$ is a Poisson random measure on $\mathbb{R}^+ \times S$ of intensity $dt \otimes \nu$ then*

$$N_t(B) := N((0,t] \times B), \quad t \geq 0,\ B \in \mathcal{S},$$

*is a Poisson point process on intensity $\nu$.*

2. *Suppose $\lambda := \nu(S) < \infty$. Then*
$$N_t = \sum_{i=1}^{K_t} \delta_{Z_i}$$
*is a Poisson point process of intensity $\nu$ provided the random variables $Z_i$ are independent with distribution $\lambda^{-1}\nu$, and $(K_t)_{t\geq 0}$ is an independent Poisson process of intensity $\lambda$.*

*Proof.* Exercise.




□

**Corollary 2.17.** *If $\nu(S) < \infty$ then a Poisson point process of intensity $\nu$ is a Markov jump process on $M_c^+(S)$ with finite jump measure*
$$q(\pi, \bullet) = \int (\pi + \delta_y)\,\nu(dy), \quad \pi \in M_c^+(S)$$
*and generator*
$$(\mathscr{L}F)(\pi) = \int \left(F(\pi + \delta_y) - F(\pi)\right)\,\nu(dy), \tag{2.6}$$
*$F\colon M_c^+(S) \to \mathbb{R}$ bounded. If $\nu(S) = \infty$, (2.6) is not defined for all bounded functions $F$.*

# Chapter 3

# Markov semigroups and Lévy processes

## 3.1 Semigroups and generators

Suppose that $p_t,\ t \geq 0$, are the transition kernels of a time-homogeneous Markov process on a Polish space $S$ with Borel $\sigma$-algebra $\mathcal{S}$.

**Properties:**

1. *Semigroup*:
$$p_s p_t = p_{s+t} \quad \forall\, s,t \geq 0$$

2. *(sub-)Markov*:
$$(i) \qquad f \geq 0 \quad \Rightarrow \quad p_t f \geq 0 \qquad \textit{positivity preserving}$$
$$(ii) \qquad p_t 1 = 1 \quad \text{(respectively } p_t 1 \leq 1 \text{ if } \zeta \neq \infty)$$

3. *Contraction with respect to sup-norm*:
$$\|p_t f\| \leq \|f\|_{\sup}$$

4. $L^p$*-Contraction*: If $\mu$ is a stationary distribution, then
$$\|p_t f\|_{\mathscr{L}^p(S,\mu)} \leq \|f\|_{\mathscr{L}^p(S,\mu)}$$
for all $1 \leq p \leq \infty$.

*Proof of 4.* If $p = \infty$, the claim follows by 3. If $1 \leq p < \infty$, we conclude with the Jensen inequality
$$|p_t f|^p \leq (p_t|f|)^p \leq p_t|f|^p$$
and since $\mu$ is a stationary distribution,
$$\int |p_t f|^p \, d\mu \leq \int p_t|f|^p \, d\mu = \int |f|^p \, d(\mu p_t) = \int |f|^p \, d\mu$$

□

**Consequence:**
$(p_t)_{t\geq 0}$ induces a semigroup of linear contractions $(P_t)_{t\geq 0}$ on the following Banach spaces:

1. $\mathcal{F}_b(S)$ which is the space of all bounded measurable functions $f\colon S \to \mathbb{R}$ endowed with the sup-norm.
2. $C_b(S)$ provided $p_t$ is Feller.
3. $L^p(S,\mu)$ provided $\mu$ is a stationary distribution.

Now let $(P_t)_{t\geq 0}$ be a general semigroup of linear contractions on a Banach space $B$.

**Definition 3.1.** *1. The linear operator*

$$Lf := \lim_{t\downarrow 0} \frac{P_t f - f}{t}$$

*with limit in $B$ an domain*

$$\mathrm{Dom}(L) = \left\{ f \in B \mid \textit{the limit } \lim_{t\downarrow 0} \frac{P_t f - f}{t} \textit{ exists} \right\}$$

*is called the **generator** of $(P_t)_{t\geq 0}$.*

2. *$(P_t)_{t\geq 0}$ is called **strongly continuous ($C_0$ semigroup)** if and only if*

$$\lim_{t\downarrow 0} P_t f = f$$

*for all $f \in B$.*

**Remark .** *1. $P_t f \to f$ as $t \downarrow 0$ for all $f \in \mathrm{Dom}(L)$, and hence, by contractivity, for all $f \in \overline{\mathrm{Dom}(L)}$:*

$$\begin{aligned}\|p_t f - f\| &\leq \|p_t f - p_t g\| + \|p_t g - g\| + \|g - f\|, \qquad g \in \mathrm{Dom}(L)\\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}\end{aligned}$$

*i.e. $P_t$ is a $C_0$ semigroup on $B$ if $\mathrm{Dom}(L)$ is dense in $B$. In general, $P_t$ is a $C_0$ semigroup on $\mathrm{Dom}(\overline{L})$.*

2. *Conversely, if $P_t$ is strongly continuous on $B$ then $\mathrm{Dom}(L)$ is dense in $B$ (for the proof see script* Stochastic analysis II*).*

3. *The transition semigroup $(p_t)_{t\geq 0}$ of a right-continuous Markov process induces a $C_0$ semigroup $(P_t)_{t\geq 0}$ on*

   *(i) $B = L^p(S,\mu)$ provided $\mu$ is stationary.*

*(ii)* $B = C_\infty(S) :=$ *all continuous functions vanishing at infinity provided* $S$ *is locally compact and* $p_t(C_\infty(S)) \subseteq C_\infty(S)$ *(proof omitted, see script* Stochastic analysis II*).*

**Theorem 3.2** (Maximum principle)**.** *The generator* $L$ *of a Markov semigroup* $(p_t)_{t\geq 0}$ *on* $\mathcal{F}_b(S)$ *satisfies*

1. ***Maximum principle:*** *If* $f \in \mathrm{Dom}(L),\ x_0 \in S$ *with* $f(x_0) = \sup_{s\in S} f(x)$*, then*

$$(Lf)(x_0) = \lim_{t\downarrow 0} \frac{p_t f(x_0) - f(x_0)}{t} \leq 0$$

2. $1 \in \mathrm{Dom}(L)$ *and* $L1 = 0$*.*

*Proof.* 1. $f \leq f(x_0)$, hence

$$p_t f \leq f(x_0) \cdot p_t 1 \leq f(x_0)$$

for all $t \geq 0$ and hence

$$(Lf)(x_0) = \lim_{t\downarrow 0} \frac{p_t f(x_0) - f(x_0)}{t} \leq 0$$

2. $P_t 1 = 1$ for all $t \geq 0$.

□

**Theorem 3.3** (Kolmogorov equations)**.** *If* $(P_t)_{t\geq 0}$ *is a* $C_0$ *semigroup with generator* $L$ *then* $t \mapsto P_t f$ *is continuous for all* $f \in B$*. Moreover, if* $f \in \mathrm{Dom}(L)$ *then* $P_t f \in \mathrm{Dom}(L)$ *for all* $t \geq 0$*, and*

$$\frac{d}{dt} P_t f = P_t L f = L P_t f$$

*Proof.* 1. For $h > 0$,

$$\|P_{t+h} f - P_t f\| = \|P_t(P_h f - f)\| \leq \|P_h f - f\| \to 0$$

and

$$\|P_{t-h} f - P_t f\| = \|P_{t-h}(f - P_h f)\| \leq \|f - P_h f\| \to 0$$

as $h \downarrow 0$.

2. For $f \in \mathrm{Dom}(L)$ and $h > 0$ we have

$$\frac{1}{h}\left(P_{t+h} - P_t f\right) = P_t \frac{P_h f - f}{h} \to P_t L f$$

as $h \downarrow 0$, because the operators $P_t$ are contractions. On the other hand,

$$\frac{1}{-h}\left(P_{t-h} f - P_t f\right) = P_{t-h} \frac{P_h f - f}{h} \to P_t L f$$

as $h \downarrow 0$ by 1.) and the contractivity.

3. We use 2.) so conclude that

$$\frac{1}{h}\left(P_h P_t f - P_t f\right) = \frac{1}{h}\left(P_{t+h} f - P_t f\right) \to P_t L f$$

as $h \downarrow 0$. Hence by 1.), $P_t f \in \mathrm{Dom}(L)$ and $L P_t f = P_t L f$.

□

Application to the martingale problem:

**Corollary 3.4.** *Suppose $(X_t, P)$ is a right-continuous $(\mathcal{F}_t)$-Markov process with transition semigroup $(p_t)_{t \geq 0}$.*

1. *Suppose $(p_t)$ induces a $C_0$ semigroup with generator L on a closed subspace of $\mathcal{F}_b(S)$ (e.g. on $C_b(S)$ or $C_\infty(S)$). Then $(X_t, P)$ is a solution of the martingale problem for $(L, \mathrm{Dom}(L))$ (independently of the initial distribution).*

2. *Suppose $(X_t, P)$ is stationary with initial distribution $\mu$, and $L^{(p)}$ is the generator of the corresponding $C_0$ semigroup on $L^p(S, \mu)$ for some $p \in [1, \infty)$. Then for $f \in \mathrm{Dom}(L^{(p)})$,*

$$M_t^f = f(X_t) - \int_0^t \left(L^{(p)} f\right)(X_s)\, ds, \quad t \geq 0,$$

*is $P$-almost sure independent of the chosen version $L^{(p)} f$ and $(X_t, P)$ solves the martingale problem for $(L^{(p)}, \mathrm{Dom}(L^{(p)}))$.*

*Proof.* 1. Since $f \in \mathrm{Dom}(L)$, $f$ and $Lf$ are bounded and

$$M_t^f = f(X_t) - \int_0^t (Lf)(X_s)\, ds \in \mathscr{L}^1(P),$$

and

$$\begin{aligned}\mathbb{E}[M_t^f - M_s^f \mid \mathcal{F}_s] &= \mathbb{E}\left[f(X_t) - f(X_s) - \int\limits_s^t Lf(X_u)\,du \mid \mathcal{F}_s\right] \\ &= \mathbb{E}[f(X_t) \mid \mathcal{F}_s] - f(X_s) - \int\limits_s^t \mathbb{E}[Lf(X_u) \mid ]\,du \\ &= p_{t-s}f(X_s) - f(X_s) - \int\limits_s^t p_{u-s}Lf(X_s)\,du = 0\end{aligned}$$

$P$-almost sure by Kolomogorov's forward equation.

2. Exercise.

□

## 3.2 Lévy processes

Additional reference for this chapter: Applebaum [3] and Bertoin [5].

**Definition 3.5.** *An $\mathbb{R}^d$-valued stochastic process $((X_t)_{t\geq 0}, P)$ with càdlàg paths is called a* ***Lévy process*** *if and only if it has stationary independent increments, i.e.,*

*(i)* $X_{s+t} - X_s \coprod \mathcal{F}_s = \sigma(X_r \mid r \leq s)$ *for all* $s, t \geq 0$.

*(ii)* $X_{s+t} - X_s \sim X_t - X_0$ *for all* $s, t \geq p$.

**Example .** 1. *Diffusion processes with constant coefficients:*

$$X_t = \sigma B_t + bt$$

*with $B_t$ a Brownian motion on $\mathbb{R}^n$, $\sigma \in \mathbb{R}^{d\times n}$, $b \in \mathbb{R}^d$.*

2. *Compound Poisson process:*

$$X_t = \sum_{i=1}^{N_t} Z_i$$

*with $Z_i$ independent identically distributed random variables, $N_t$ Poisson process, independent of $Z_i$.*

More interesting examples will be considered below.

**Remark .** *If $(X_t, P)$ is a Lévy process then the increments $X_{s+t} - X_s$ are* ***infinitely divisible random variables****, i.e. for any $n \in \mathbb{N}$, there exist independent identically distributed random variables $Y_1^{(n)}, \dots, Y_n^{(n)}$ such that*

$$X_{s+t} - X_s \sim \sum_{i=1}^{n} Y_i^{(n)} \quad \left(\text{e.g. } Y_i^{(n)} := X_{\frac{it}{n}} - X_{\frac{(i-1)t}{n}}\right)$$

*The Lévy-Khinchin formula gives a classification of the distributions of all infinitely divisible random variables on $\mathbb{R}^d$ in terms of their characteristic functions.*

**Theorem 3.6.** *A Lévy process is a time-homogeneous Markov process with translation invariant transition functions*

$$p_t(x, B) = \mu_t(B - x) = p_t(a + x, a + B) \quad \forall\, a \in \mathbb{R}^d \tag{3.1}$$

*where $\mu_t = P \circ (X_t - X_0)^{-1}$.*

*Proof.*

$$\begin{aligned} P[X_{s+t} \in B \mid \mathcal{F}_s](\omega) &= P[X_s + (X_{s+t} - X_s) \in B \mid \mathcal{F}_s](\omega) \\ &= P[X_{s+t} - X_s \in B - X_s(\omega)] \\ &= P[X_t - X_0 \in B - X_s(\omega)] \\ &= \mu(B - X_s(\omega)), \end{aligned}$$

so $(X_t, P)$ is Markov with transition function $p_t(x, B) = \mu_t(B - x)$ which is clearly translation invariant. □

**Remark .** 1. *In particular, the transition semigroup of a Lévy process is* Feller*: If $f \in C_b(\mathbb{R}^d)$, then*

$$(p_t f)(x) = \int f(x + y)\, \mu_t(dy)$$

*is continuous by dominated convergence. If $f \in C_\infty(\mathbb{R}^d)$, then $p_t f \in C_\infty(\mathbb{R}^d)$.*

2. *$p_t$ defined by (3.1) is a semigroup if and only if $\mu_t$ is a* ***convolution semigroup****, i.e.,*

$$\mu_t * \mu_s = \mu_{t+s} \quad \forall\, t, s \geq 0$$

*E.g.*

$$\mu_t * \mu_s(B) = \int \mu_t(dy)\mu_s(B - y) = \int p_t(0, dy) p_s(y, B) = p_{t+s}(0, B) = \mu_{t+s}(B)$$

*if $p_t$ is a semigroup. The inverse implication follows similarly.*

From now on we assume w.l.o.g. $X_0 = 0$, and set

$$\mu := \mu_1 = P \circ X_1^{-1}.$$

**Definition 3.7.** *A continuous function $\psi\colon \mathbb{R}^d \to \mathbb{C}$ is called* ***characteristic exponent*** *of the measure $\mu$ or the Lévy process $(X_t, P)$ if and only if $\psi(0) = 0$ and*

$$\mathbb{E}\left[e^{ip\cdot X_1}\right] = \varphi_\mu(p) = e^{-\psi(p)}$$

*One easily verifies that for any Lévy process there exists a unique characteristic exponent.*

**Theorem 3.8.** *1. $\mathbb{E}\left[e^{ip\cdot X_t}\right] = e^{-t\psi(p)}$ for all $t \geq 0, p \in \mathbb{R}^n$.*

*2. $M_t^p := e^{ip\cdot X_t + t\psi(p)}$ is a martingale for any $p \in \mathbb{R}^n$.*

*Proof.* 1. a) For $t \in \mathbb{N}$, define

$$X_t = \sum_{i=1}^{t} (X_i - X_{i-1})$$

Since $(X_i - X_{i-1})$ are independent identically distributed random variables with the same distribution as $X_1$,

$$\varphi_{X_t}(p) = \varphi_{X_1}(p)^t = e^{-t\psi(p)}$$

b) Let $t = \frac{m}{n} \in \mathbb{Q}$ and

$$X_m = \sum_{i=1}^{n} \left( X_{\frac{im}{n}} - X_{\frac{(i-1)m}{n}} \right)$$

Hence

$$\varphi_{X_m} = \varphi_{X_t}^n$$

and since $\varphi_{X_m} = e^{-m\psi}$,

$$\varphi_{X_t} = e^{-\frac{m}{n}\psi} = e^{-t\psi}$$

c) Let $t_n \in \mathbb{Q},\ t_n \downarrow t$. Since $X_t$ is right-continuous, we conclude with Lebesque

$$\mathbb{E}\left[e^{ip\cdot X_t}\right] = \lim_{t_n \to t} \mathbb{E}\left[e^{ip\cdot X_{t_n}}\right] = \lim_{t_n \to t} e^{-t_n\psi(p)} = e^{-t\psi(p)}$$

2. Exercise.

□

Since $X_{s+t}-X_s \sim X_t$, independent of $\mathcal{F}_s$, the marginal distributions of a Lévy process $((X_t)_{t\geq 0}, P)$ are completely determined by the distributions of $X_t$, and hence by $\psi$! In particular:

**Corollary 3.9** (Semigroup and generator of a Lévy process). *1. For all $f \in \mathscr{S}(\mathbb{R}^d)$ and $t \geq 0$,*

$$p_t f = \left(e^{-t\psi}\hat{f}\right)^{\check{}}$$

*where*

$$\hat{f}(p) = (2\pi)^{-\frac{d}{2}} \int e^{ip\cdot x} f(x)\,dx, \quad \textit{and}$$
$$\check{g}(x) = (2\pi)^{-\frac{d}{2}} \int e^{ip\cdot x} g(p)\,dp$$

*denote the* Fourier transform *and the* inverse Fourier transform *of functions $f, g \in \mathscr{L}^1(\mathbb{R}^d)$.*

2. *$\mathscr{S}(\mathbb{R}^d)$ is contained in the domain of the generator $L$ of the semigroup induced by $(p_t)_{t\geq 0}$ on $C_\infty(\mathbb{R}^d)$, and*

$$Lf = (-\psi\hat{f})^{\check{}} \qquad \textbf{\textit{(Pseudo-differential operator)}}. \tag{3.2}$$

*In particular, $p_t$ is strongly continuous on $C_\infty(\mathbb{R}^d)$.*

*Here $\mathscr{S}(\mathbb{R}^d)$ denotes the Schwartz space of rapidly decreasing smooth functions on $\mathbb{R}^d$. Recall that the Fourier transform maps $\mathscr{S}(\mathbb{R}^d)$ one-to-one onto $\mathscr{S}(\mathbb{R}^d)$.*

*Proof.* 1. Since $(p_t f)(x) = \mathbb{E}[f(X_t + x)]$, we conclude with Fubini

$$\begin{aligned}(\widehat{p_t f})(p) &= (2\pi)^{-\frac{d}{2}} \int e^{-ip\cdot x}(p_t f)(x)\,dx \\ &= (2\pi)^{-\frac{d}{2}} \cdot \mathbb{E}\left[\int e^{-ip\cdot x} f(X_t + x)\,dx\right] \\ &= \mathbb{E}\left[e^{ip\cdot X_t}\right] \cdot \hat{f}(p) \\ &= e^{-t\psi(p)}\hat{f}(p)\end{aligned}$$

for all $p \in \mathbb{R}^d$. The claim follows by the Fourier inversion theorem, noting that $\left|e^{-t\psi}\right| \leq 1$.

2. For $f \in \mathscr{S}(\mathbb{R}^d)$, $\hat{f}$ is in $\mathscr{S}(\mathbb{R}^d)$ as well. The Lévy-Khinchin formula that we will state below gives an explicit representation of all possible Lévy exponents which shows in particular that $\psi(p)$ is growing at most polynomial as $|p| \to \infty$. Hence

$$\left|\frac{e^{-t\psi}\hat{f} - \hat{f}}{t} + \psi\hat{f}\right| = \left|\frac{e^{-t\psi} - 1}{t} + \psi\right| \cdot |\hat{f}|$$

and

$$\frac{e^{-t\psi}-1}{t}+\psi=-\frac{1}{t}\int_0^t \psi\left(e^{-s\psi}-1\right)\,ds=\frac{1}{t}\int_0^t\int_0^s \psi^2 e^{-r\psi}\,dr\,ds$$

hence

$$\left|\frac{e^{-t\psi}\hat{f}-\hat{f}}{t}+\psi\hat{f}\right|\leq t\cdot|\psi^2|\cdot|\hat{f}|\;\in\mathscr{L}^1(\mathbb{R}^d),$$

and therefore:

$$\frac{p_t f-f}{t}-(-\psi\hat{f})\check{}=(2\pi)^{-\frac{d}{2}}\int e^{ip\cdot x}\cdot\cdots\,dp\to 0$$

as $t\downarrow 0$ uniformly in $x$. This shows $f\in\mathrm{Dom}(L)$ and $Lf=(-\psi\hat{f})\check{}$. In particular, $p_t$ is strongly continuous on $\mathscr{S}(\mathbb{R}^d)$. Since $\mathscr{S}(\mathbb{R}^d)$ is dense in $C_\infty(\mathbb{R}^d)$ and $p_t$ is contractive this implies strong continuity on $C_\infty(\mathbb{R}^d)$.

□

**Remark .** *$p_t$ is not necessarily strongly continuous on $C_b(\mathbb{R}^d)$. Consider e.g. the deterministic process*

$$X_t=X_0+t$$

*on $\mathbb{R}^1$. Then*

$$(p_t f)(x)=f(x+t),$$

*and one easily verifies that there exists $f\in C_b(\mathbb{R})$ such that $p_t f\nrightarrow f$ uniformly.*

**Corollary 3.10.** *$(X_t,P)$ solves the martingale problem for the operator $(L,\mathscr{S}(\mathbb{R}^d))$ defined by (3.2).*

**Example .** *1. **Translation invariant diffusions:***

$$X_t=\sigma B_t+bt,\qquad \sigma\in\mathbb{R}^{d\times n}, b\in\mathbb{R}^d, B_t \text{ Brownian motion on } \mathbb{R}^n$$

*We have*

$$\begin{aligned}\mathbb{E}\left[e^{ip\cdot X_t}\right]&=\mathbb{E}\left[e^{i(\sigma^T p)\cdot B_t}\right]e^{ip\cdot bt}\\&=e^{-\psi(p)\cdot t}\end{aligned}$$

*where*

$$\psi(p)=\frac{1}{2}(\sigma^T p)^2-ib\cdot p=\frac{1}{2}p\cdot ap-ib\cdot p,\quad a:=\sigma\sigma^T$$

*and*

$$Lf=-(\psi\hat{f})\check{}=\frac{1}{2}\operatorname{div}(a\nabla f)-b\cdot\nabla f,\quad f\in\mathscr{S}(\mathbb{R}^n)$$

2. ***Compound Poisson process:***

$$X_t = \sum_{i=1}^{N_t} Z_i,$$

*$Z_i$ independent identically distributed random variables on $\mathbb{R}^d$ with distribution $\pi$, $N_t$ Poisson process of intensity $\lambda$, independent of $Z_i$.*

$$\begin{aligned}\mathbb{E}\left[e^{ip\cdot X_t}\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{ip\cdot X_t} \mid N_t = n\right] \cdot P[N_t = n] \\ &= \sum_{n=0}^{\infty} \varphi_\pi(p)^n \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\ &= e^{-\lambda\cdot(1-\varphi_\pi(p))\cdot t}\end{aligned}$$

*and hence*

$$\psi(p) = \lambda \cdot (1 - \varphi_\pi(p)) = \int \left(1 - e^{ip\cdot y}\right) \lambda\pi(dy)$$

*and*

$$(Lf)(x) = (-\psi \hat{f})\check{}(x) = \int (f(x+y) - f(x))\lambda\pi(dy), \quad f \in \mathscr{S}(\mathbb{R}^n).$$

*The jump intensity measure $\nu := \lambda\pi$ is called the **Lévy measure** of the compound Poisson process.*

3. ***Compensated Poisson process:*** *$X_t$ as above, assume $Z_i \in \mathscr{L}^1$.*

$$M_t := X_t - \mathbb{E}[X_t] = \sum_{i=1}^{N_t} Z_i - \lambda\mathbb{E}[Z_1] \cdot t$$

*is a Lévy process with generator*

$$\begin{aligned}(L^{comp} f)(x) &= Lf(x) - \lambda \cdot \int y\pi(dy) \cdot \nabla f(x) \\ &= \int \left(f(x+y) - f(x) - y \cdot \nabla f(x)\right) \lambda\pi(dy)\end{aligned}$$

**Remark/Exercise** (Martingales of compound Poisson process). *The following processes are martingales:*

*(a)* $M_t = x_t - b \cdot t$, *where* $b := \lambda \cdot \mathbb{E}[Z_1] = \int y\,\nu(dy)$ *provided* $Z_1 \in \mathscr{L}^1$.

*(b)* $|M_t|^2 - a \cdot t$, *where* $a := \lambda \cdot \mathbb{E}[|Z_1|^2] = \int |y|^2\,\nu(dy)$ *provided* $Z_1 \in \mathscr{L}^2$.

*(c)* $\exp(ip \cdot X_t + \psi(p) \cdot t),\ p \in \mathbb{R}^n$.

*Proof.* e.g.

$$\begin{aligned}
\mathbb{E}[M_{s+t}^2 - M_s^2 \mid \mathcal{F}_s] &= \mathbb{E}[(M_{s+t} - M_s)^2 \mid \mathcal{F}_s] = \mathbb{E}[(M_{s+t} - M_s)^2] \\
&= \mathbb{E}[|M_t|^2] = \mathrm{Var}(X_t) = \mathrm{Var}\left(\sum_{i=1}^{N_t} Z_i\right) \\
&= \mathbb{E}\left[\mathrm{Var}\left(\sum_{i=1}^{N_t} Z_i \,\middle|\, N_t\right)\right] + \mathrm{Var}\left(\mathbb{E}\left[\sum_{i=1}^{N_t} Z_i \,\middle|\, N_t\right]\right)
\end{aligned}$$

and since $\mathrm{Var}\left(\sum_{i=1}^{N_t} Z_i \mid \right) = N_t \cdot \mathrm{Var}(Z_1)$ and $\mathbb{E}\left[\sum_{i=1}^{N_t} Z_i \mid \right] = N_t \cdot \mathbb{E}[Z_1]$,

$$\mathbb{E}[M_{s+t}^2 - M_s^2 \mid \mathcal{F}_s] = \mathbb{E}[N_t] \cdot \mathrm{Var}(Z_1) + \mathrm{Var}(N_t) \cdot |\mathbb{E}[Z_1]|^2 = \lambda t \mathbb{E}[|Z_1|^2].$$

□

4. ***Symmetric stable processes:*** *Stable processes appear as continuous-time scaling limits of random walks. By Donsker's invariance principle, if*

$$S_n = \sum_{i=1}^{n} Z_i$$

*is a random walk with independent identically distributed increments $Z_i \in \mathscr{L}^2$ then the rescaled processes*

$$X_t^{(k)} := k^{-\frac{1}{2}} S_{\lfloor kt \rfloor}$$

*converge in distribution to a Brownian motion. This functional central limit theorem fails (as does the classical central limit theorem) if the increments are not square integrable, i.e., if their distribution has **heavy tails**. In this case, one looks more generally for scaling limits of rescaled processes of type*

$$X_t^{(k)} := k^{-\frac{1}{\alpha}} S_{\lfloor kt \rfloor}$$

*for some $\alpha > 0$. If $(X_t^{(k)})_{t \geq 0}$ converges in distribution then the limit should be a **scale-invariant Lévy process**, i.e.*

$$k^{-\frac{1}{\alpha}} X_{kt} \sim X_t \quad \text{for all } k > 0 \tag{3.3}$$

*This motivates looking for Lévy processes that satisfy the scaling relation (3.3). Clearly, (3.3) is equivalent to*

$$e^{-t\psi(p)} = \mathbb{E}\left[e^{ip\cdot cX_t}\right] = \mathbb{E}\left[e^{ip\cdot X_{c^\alpha t}}\right] = e^{-c^\alpha t\psi(p)} \quad \forall\, c > 0$$

*i.e.*

$$\psi(cp) = c^\alpha \psi(p) \qquad \forall\, c > 0$$

*The simplest choice of a Lévy exponent satisfying (3.3) is*

$$\psi(p) = \frac{\alpha}{2} \cdot |p|^\alpha$$

*for some $\sigma > 0$. In this case, the generator of a corresponding Lévy process would be the fractional power*

$$Lf = -(\psi \hat{f})\check{} = \frac{\sigma}{2} \Delta^{\frac{\alpha}{2}} f$$

*of the Laplacien.*

*For $\alpha = 2$ and $\sigma = 1$ the corresponding Lévy process is a Brownian motion, the scaling limit in the classical central limit theorem. For $\alpha > 2$, $L$ does not satisfy the maximum principle, hence the corresponding semigroup is not a transition semigroup of a Markov process.*

Now suppose $\alpha \in (0, 2)$.

**Lemma 3.11.** *For $\alpha \in (0, 2)$,*

$$\psi(p) = \textit{const.} \cdot \lim_{\varepsilon \downarrow 0} \psi_\varepsilon(p)$$

*where*

$$\psi_\varepsilon(p) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(0,\varepsilon)} \left(1 - e^{ip\cdot y}\right) |y|^{-\alpha-1} \, dy$$

*Proof.* By substitution $x = |p|y$ and $\nu := \frac{p}{|p|}$,

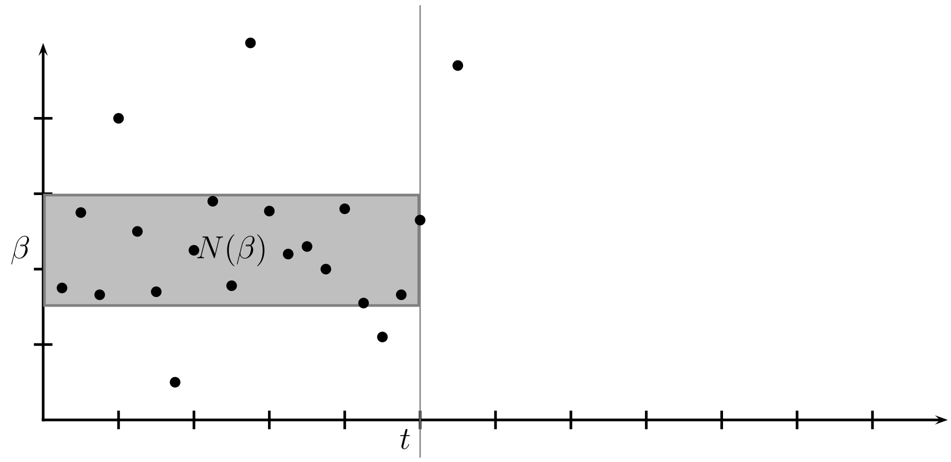
$$\int_{\mathbb{R}^d \setminus B(0,\varepsilon)} \left(1 - e^{ip\cdot y}\right) \cdot |y|^{-\alpha-1} \, dy = \int_{\mathbb{R}^d \setminus B(0,\varepsilon\cdot|p|)} \left(1 - e^{i\nu x}\right) |x|^{-\alpha-1} \, dx \cdot |p|^\alpha \to \text{const.} |p|^\alpha$$

as $\varepsilon \downarrow 0$ since $(1 - e^{i\nu x}) = i\nu x + O(|x|^2)$ by Taylor expansion. □

Note that $\psi_\varepsilon$ is the symbol of a compound Poisson process with Lévy measure proportional to $|y|^{-\alpha-1} \cdot I_{\{|y|>\varepsilon\}} \, dy$. Hence we could expect that $\psi$ is a symbol of a similar process with Lévy measure proportional to $|y|^{-\alpha-1} \, dy$. Since this measure is infinite, a corresponding process should have infinitely many jumps in any non-empty time interval. To make this heuristics rigorous we now give a construction of Lévy processes from Poisson point process:

## 3.3 Construction of Lévy processes from Poisson point processes:

**Idea:** Jumps of given size for Lévy processes $\leftrightarrow$ Points of a Poisson point process on $\mathbb{R}^d$.




Position of a Lévy process after time $t$:

$$X_t = \sum_y y N_t(\{y\}) \qquad \text{if } \operatorname{supp}(N_t) \text{ is countable}$$

$$X_t = \int y\, N_t(dy) \qquad \text{in general}$$

### a) Finite intensity

**Theorem 3.12.** *Suppose $\nu$ is a* finite *measure on $\mathbb{R}^d$. If $(N_t)_{t\geq 0}$ is a Poisson point process of intensity $\nu$ then*

$$X_t := \int y\, N_t(dy)$$

*is a compound Poisson process with Lévy measure $\nu$ (i.e. total intensity $\lambda = \nu(\mathbb{R}^d)$ and jump distribution $\pi = \frac{\nu}{\nu(\mathbb{R}^d)}$ ).*

*Proof.* By the theorem in Section 1.9 above and the uniqueness of a Poisson point process of intensity $\nu$ we may assume

$$N_t = \sum_{i=1}^{K_t} \delta_{Z_i}$$

where $Z_i$ are independent random variables of distribution $\lambda^{-1}\nu$, and $(K_t)_{t\geq 0}$ is an independent Poisson process of intensity $\lambda$.

Hence

$$X_t = \int y\, N_t(dy) = \sum_{k=1}^{K_t} Z_i$$

is a compound Poisson process. □

## b) Infinite symmetric intensity

Now assume that

**(A1)** $\nu$ is a ***symmetric measure*** on $\mathbb{R}^d \setminus \{0\}$, i.e.

$$\nu(B) = \nu(-B) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$$

and

**(A2)** $\int (1 \wedge |y|^2)\ \nu(dy) < \infty$.
(i.e. $\nu\,(|y| \geq \varepsilon) < \infty$ and $\int_{|y|<\varepsilon} |y|^2\, \nu(dy) < \infty \quad \forall\, \varepsilon > 0$ )

For example, we could choose $\nu(dy) = |y|^{-\alpha-1},\ \alpha \in (0,2)$ which is our candidate for the Lévy measure on an $\alpha$-stable process. Let $(N_t)_{t\geq 0}$ be a Poisson point process with intensity $\nu$. Our aim is to prove the existence of a corresponding Lévy process by an approximation argument. For $\varepsilon > 0$,

$$N_t^\varepsilon(dy) := I_{\{|y|>\varepsilon\}} \cdot N_t(dy)$$

is a Poisson point process with *finite* intensity $\nu^\varepsilon(dy) = I_{\{|y|>\varepsilon\}} \cdot \nu(dy)$, and hence

$$X_t^\varepsilon := \int\limits_{|y|>\varepsilon} y\, N_t(dy) = \int y\, N_t^\varepsilon(dy)$$

is a compound Poisson process with Lévy measure $\nu^\varepsilon$.

**Lemma 3.13.** *If (A1) holds, then for all* $0 < \delta \leq \varepsilon$ *and* $t \geq 0$*,*

$$\mathbb{E}\left[\sup_{s\leq t} \left|X_s^\delta - X_s^\varepsilon\right|^2\right] \leq 4t \cdot \int\limits_{\delta<|y|\leq\varepsilon} |y|^2\, \nu(dy)$$

*Proof.*

$$X_t^\delta - X_t^\varepsilon = \int\limits_{\delta<|y|\leq\varepsilon} y\, N_t(dy) = \int y\, N_t^{\delta,\varepsilon}(dy)$$

where

$$N_t^{\delta,\varepsilon}(dy) := I_{\{\delta<|y|\leq\varepsilon\}} \cdot N_t(dy)$$

is a Poisson point process of intensity $\nu^{\delta,\varepsilon}(dy) = I_{\{\delta<|y|\leq\varepsilon\}} \cdot \nu(dy)$. Hence $X_t^\delta - X_t^\varepsilon$ is a compound Poisson process with finite Lévy measure $\nu^{\delta,\varepsilon}$. In particular,

$$M_t := X_t^\delta - X_t^\varepsilon - t \cdot \int y\nu^{\delta,\varepsilon}(dy) = X_t^\delta - X_t^\varepsilon - t \cdot \int\limits_{\delta<|y|\leq\varepsilon} y\,\nu(dy)$$

and

$$|M_t|^2 - t \cdot \int y^2\,\nu^{\delta,\varepsilon}(dy) = |M_t|^2 - t \cdot \int\limits_{\delta<|y|\leq\varepsilon} |y|^2\,\nu(dy)$$

are right-continuous martingales. Since $\nu$ is symmetric, $M_t = X_t^\delta - X_t^\varepsilon$. Hence by Doob's maximal inequality,

$$\mathbb{E}\left[\sup_{s\leq t} \left|X_s^\delta - X_s^\varepsilon\right|^2\right] = \mathbb{E}\left[\sup_{s\leq t} |M_s|^2\right] \leq \left(\frac{2}{2-1}\right)^2 \cdot \mathbb{E}[|M_t|^2] = 4t \cdot \int\limits_{\delta<|y|\leq\varepsilon} |y|^2\,\nu(dy).$$

□

**Theorem 3.14.** *Let $t \geq 0$. If (A1) and (A2) hold then the process $X^\varepsilon, \varepsilon > 0$, form a Cauchy sequence with respect to the norm*

$$\|X^\delta - X^\varepsilon| := \mathbb{E}\left[\sup_{s\leq t} \left|X_s^\delta - X_s^\varepsilon\right|^2\right].$$

*The limit process*

$$X_s = \lim_{\varepsilon\downarrow 0} X_s^\varepsilon = \lim_{\varepsilon\downarrow 0} \int\limits_{|y|>\varepsilon} y\,N_s(dy), \quad 0 \leq s \leq t,$$

*is a Lévy process with symbol*

$$\psi(p) = \lim_{\varepsilon\downarrow 0} \int\limits_{|y|>\varepsilon} \left(1 - e^{ip\cdot y}\right)\,\nu(dy).$$

**Remark .** 1. *Representation of Lévy process with symbol $\psi$ as jump process with infinite jump intensity.*

2. *For $\nu(dy) = |y|^{-\alpha-1}$, $\alpha \in (0,2)$, we obtain an $\alpha$-stable process.*

*Proof.* Lemma and (A2) yields that $(X^\varepsilon)_{\varepsilon>0}$ is a Cauchy sequence with respect to $\|\cdot\|$. Since the processes $X_s^\varepsilon$ are right-continuous and the convergence is uniform, the limit process $X_s$ is right-continuous as well. Similarly, it has independent increments, since the approximating processes have independent increments, and by dominated convergence

$$\mathbb{E}\left[e^{ip\cdot(X_{s+t}-X_s)}\right] = \lim_{\varepsilon\downarrow 0} \mathbb{E}\left[e^{ip\cdot(X_{s+t}^\varepsilon - X_s^\varepsilon)}\right] = \lim_{\varepsilon\downarrow 0} e^{-t\psi_\varepsilon(p)}$$

where

$$\psi_\varepsilon(p) = \int \left(1 - e^{ip\cdot y}\right) \nu_\varepsilon(dy) = \int\limits_{|y|>\varepsilon} \left(1 - e^{ip\cdot y}\right) \nu(dy)$$

is the symbol of the approximating compound Poisson process. □

## General Lévy processes

**Theorem 3.15** (LÉVY-KHINCHIN). *For $\psi\colon \mathbb{R}^d \to \mathbb{C}$ the following statements are equivalent:*

*(i) $\psi$ is the characteristic exponent of a Lévy process.*

*(ii) $e^{-\psi}$ is the characteristic function of an infinitely divisible random variable.*

*(iii)*
$$\psi(p) = \frac{1}{2} p \cdot ap - ib + \int_{\mathbb{R}^d} \left(1 - e^{ip\cdot y} + ip \cdot y I_{\{|y|\le 1\}}\right) \nu(dy)$$

*where $a \in \mathbb{R}^{d^2}$ is a non-negative definite matrix, $b \in \mathbb{R}^d$, and $\nu$ is a positive measure on $\mathbb{R}^d \setminus \{0\}$ satisfying (A2).*

*Sketch of proof:* (i)⇒(ii): If $X_t$ is a Lévy process with characteristic exponent $\psi$ then $X_1$ is an infinitely divisible random variable with the same characteristic exponent.

(ii)⇒(iii): This is the classical Lévy-Khinchin theorem which is proven in several textbooks on probability theory, cf. e.g. Feller [10] and Varadhan [22].

(iii)⇒(i): The idea for the construction of a Lévy process with symbol $\psi$ is to define

$$X_t = X_t^{(1)} + X_t^{(2)} + X_t^{(3)}$$

where $X^{(1)}$, $X^{(2)}$ and $X^{(3)}$ are independent Lévy processes,

$X_t^{(1)} = \sqrt{a} B_t + b$ (diffusion part),
$X_t^{(2)}$ compound Poisson process with Lévy measure $I_{\{|y|\ge 1\}} \cdot \nu(dy)$ (big jumps)
$X_t^{(3)} = \lim_{\varepsilon \downarrow 0} X_t^{(3,\varepsilon)}$ (small jumps),
$X_t^{(3,\varepsilon)}$ compensated Poisson process with Lévy measure $I_{\{\varepsilon<|y|\le 1\}} \cdot \nu(dy)$

Since $X_t^{(3,\varepsilon)}$ is a martingale for all $\varepsilon > 0$, the existence of the limit as $\varepsilon \downarrow 0$ can be established as above via the maximal inequality. One then verifies as above that $X^{(1)}$, $X^{(2)}$

and $X^{(3)}$ are Lévy processes with symbols

$$\psi^{(1)}(p) = \frac{1}{2}p \cdot ap - ib,$$
$$\psi^{(2)}(p) = \int\limits_{|y|>1} \left(1 - e^{ip\cdot y}\right) \nu(dy), \text{ and}$$
$$\psi^{(3)}(p) = \int\limits_{|y|\leq 1} \left(1 - e^{ip\cdot y} + ip \cdot y\right) \nu(dy).$$

Thus by independence, $X = X^{(1)} + X^{(2)} + X^{(3)}$ is a Lévy process with symbol $\psi = \psi^{(1)} + \psi^{(2)} + \psi^{(3)}$, cf. Bertoin [5] or Applebaum [3] for details.

□

**Remark .** *1.* ***Lévy-Itô representation:***

$$X_t = X_0 + \underbrace{\sqrt{a}B_t + bt}_{\text{diffusion part}} + \underbrace{\int\limits_{|y|\geq 1} y\, N_t(dy)}_{\text{big jumps}} + \underbrace{\int\limits_{0<|y|<1} y\ (N_t(dy) - t\nu(dy))}_{\text{small jumps compensated by drift}}$$

2. *The compensation for small jumps ensures that the infinite intensity limit exists by martingale arguments. If $\int |y|\, \nu(dy) = \infty$ then the uncompensated compound Poisson process do not converge!*

3. *In the construction of the $\alpha$-stable process above, a compensation was not required because for a symmetric Lévy measure the approximating processes are already martingales.*

Extension to non-translation invariant case:

**Theorem 3.16** (Classification of Feller semigroups in $\mathbb{R}^d$)**.** [DYNKIN, COURRÈGE, KUNITA, ROTH] *Suppose $(P_t)_{t\geq 0}$ is a $C_0$ contraction semigroup on $C_\infty(\mathbb{R}^d)$, such that $C_0^\infty(\mathbb{R}^d)$ is contained in the domain of the generator $L$. Then:*

*1.*

$$\begin{aligned} Lf(x) = \sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b(x)\nabla f(x) + c(x)\cdot f(x) \\ + \int\limits_{\mathbb{R}^d\setminus\{x\}} \left(f(y) - f(x) - I_{\{|y-x|<1\}}\cdot(y-x)\cdot\nabla f(x)\right) \nu(x, dy) \end{aligned} \tag{3.4}$$

*for all $f \in C_0^\infty(\mathbb{R}^d)$, where $a_{ij}, b, c \in C(\mathbb{R}^d)$, $a(x)$ non-negative definit and $c(x) \leq 0$ for all $x$, and $\nu(x, \cdot)$ is a kernel of positive (Radon) measures.*

2. *If $P_t$ is the transition semigroup of a non-explosive Markov process then $c \equiv 0$.*

3. *If $P_t$ is the transition semigroup of a diffusion (i.e. a Markov process with continuous paths) then $L$ is a local operator, and a representation of type (3.4) holds with $\nu \equiv 0$.*

**Remark .** *Corresponding Markov processes can be constructed as solutions of stochastic differential equations with combined Gaussian and Poisson noise.*

We will not prove assertion 1. The proof of 2. is left as an exercise. We now sketch an independent proof of 3., for a detailed proof we refer to volume one of Rogers / Williams [18]:

If $f \equiv 1$ on a neighborhood of $x$ and $0 \leq f \leq 1$, then by the maximum principle,

$$0 = (Lf)(x) = o(x) + \int\limits_{\mathbb{R}^d \setminus \{x\}} (f(y) - f(x))\ \nu(dy)$$

this is only possible for all $f$ as above if $\nu(x, \cdot) = 0$.

*Proof of 3.:* a) **Locality:** If $x \in \mathbb{R}^d$ and $f, g \in \mathrm{Dom}(L)$ with $f = g$ in a neighborhood of $x$, then $Lf(x) = Lg(x)$. Since $f \in \mathrm{Dom}(L)$ and $x \in \mathbb{R}^d$,

$$f(X_t) - \int\limits_0^t Lf(X_s)\, ds$$

is a martingale and hence

$$\mathbb{E}_x[f(X_T)] = f(x) + \mathbb{E}_x\left[\int\limits_0^T (Lf)(X_s)\, ds\right]$$

for all bounded stopping times $T$ (Dynkin's formula). Hence

$$\begin{aligned}(Lf)(x) &= \lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}_x\left[\int_0^{T_\varepsilon} Lf(X_s)\, ds\right]}{\mathbb{E}_x[T_\varepsilon]} \\ &= \lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}_x[f(X_{T_\varepsilon})] - f(x)}{\mathbb{E}_x[T_\varepsilon]} \qquad \text{(Dynkin's characteristic operator)}\end{aligned}$$

where

$$T_\varepsilon := \inf\{t \geq 0 \ :\ X_t \notin B(x, \varepsilon)\} \wedge 1$$

since in the equation above, $Lf(X_s) = Lf(x) + O(1)$ by right continuity.

If the paths are continuous then $X_{T_\varepsilon} \in \overline{B(x, \varepsilon)}$. Hence for $f, g \in \mathrm{Dom}(L)$ with $f = g$ in a neighborhood of $x$,

$$f(X_{T_\varepsilon}) = g(X_{T_\varepsilon})$$

for small $\varepsilon > 0$, and thus

$$Lf(x) = Lg(x).$$

b) **Local maximum principle:** Locality and the maximum principle imply:

$$f \in \mathrm{Dom}(L) \text{ with local maximum at } x \quad \Longrightarrow \quad Lf(x) \leq 0$$

c) **Taylor expansion:** Fix $x \in \mathbb{R}^d$ and $f \in C_0^\infty(\mathbb{R}^d)$. Let $\varphi, \varphi_i \in C_0^\infty(\mathbb{R}^d)$ such that $\varphi(y) = 1$ and $\varphi_i(y) = y_i - x_i$ for all $y$ in a neighborhood of $x$. Then in a neighborhood $U$ of $x$,

$$f(y) = f(x) \cdot \varphi(y) + \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(x)\varphi_i(y) + \frac{1}{2} \sum_{i,j=1} d \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\varphi_i(y)\varphi_j(y) + R(y)$$

where $R$ is a function in $C_0^\infty(\mathbb{R}^d)$ with $R(y) = o(|y-x|^2)$. Hence

$$(Lf)(x) = c \cdot f(x) + b\nabla f(x) + \frac{1}{2} \sum a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + (LR)(x)$$

where $c := L\varphi(x),\ b_i := L\varphi_i(x)$ and $a_{ij} := L(\varphi_i\varphi_j)(x)$. In order to show $(LR)(x) = 0$ we apply the local maximum principle. For $\varepsilon \in \mathbb{R}$ choose $R_\varepsilon \in C_0^\infty(\mathbb{R}^d)$ such that

$$R_\varepsilon(y) = R(y) - \varepsilon|y-x|^2$$

on $U$. Then for $\varepsilon > 0$, $R_\varepsilon$ has a local maximum at $x$, and hence $LR_\varepsilon \leq 0$. For $\varepsilon \downarrow 0$ we obtain $LR(x) \leq 0$. Similarly, for $\varepsilon < 0$, $-R_\varepsilon$ has a local maximum at $x$ and hence $LR_\varepsilon \geq 0$. For $\varepsilon \uparrow 0$ we obtain $LR(x) \geq 0$, and thus $LR(x) = 0$.

□

# Chapter 4

# Convergence to equilibrium

Our goal in the following sections is to relate the long time asymptotics $(t \uparrow \infty)$ of a time-homogeneous Markov process (respectively its transition semigroup) to its infinitesimal characteristics which describe the short-time behavior $(t \downarrow 0)$:

| Asymptotic properties | $\leftrightarrow$ | Infinitesimal behavior, generator |
|---|---|---|
| $t \uparrow \infty$ | | $t \downarrow 0$ |

Although this is usually limited to the time-homogeneous case, some of the results can be applied to time-inhomogeneous Markov processes by considering the space-time process $(t, X_t)$, which is always time-homogeneous. On the other hand, we would like to take into account processes that jump instantaneously (as e.g. interacting particle systems on $\mathbb{Z}^d$) or have continuous trajectories (diffusion-processes). In this case it is not straightforward to describe the process completely in terms of infinitesimal characteristics, as we did for jump processes. A convenient general setup that can be applied to all these types of Markov processes is the martingale problem of Stroock and Varadhan.

## 4.1 Setup and examples

In this section, we introduce the setup for the rest of the chapter IV. Let $S$ be a Polish space endowed with its Borel $\sigma$-algebra $\mathcal{S}$. By $\mathcal{F}_b(S)$ we denote the linear space of all bounded measurable functions $f\colon S \to \mathbb{R}$. Suppose that $\mathscr{A}$ is a linear subspace of $\mathcal{F}_b(S)$ such that

**(A0)** If $\mu$ is a signed measure on $S$ with finite variation and

$$\int f\, d\mu = 0 \quad \forall\, f \in \mathscr{A},$$

then $\mu = 0$

Let

$$\mathscr{L}\colon \mathscr{A} \subseteq \mathcal{F}_b(S) \to \mathcal{F}_b(S)$$

be a linear operator.

**Definition 4.1.** *An adapted right-continuous stochastic process $((X_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, P)$ is called a* ***solution for the (local) martingale problem for the operator $(\mathscr{L}, \mathscr{A})$*** *if and only if*

$$M_t^f := f(X_t) - \int_0^t (\mathscr{L}f)(X_s)\, ds$$

*is an $(\mathcal{F}_t)$-martingale for all $f \in \mathscr{A}$.*

**Example .** 1. ***Jump processes:*** *A* minimal Markov jump process *solves the martingale problem for its generator*

$$(\mathscr{L}f)(x) = \int q(x, dy)\,(f(y) - f(x))$$

*with domain*

$$\mathscr{A} = \{f \in \mathcal{F}_b(S) \ : \ \mathscr{L}f \in \mathcal{F}_b(S)\},$$

*cf. above.*

2. ***Interacting particle systems:*** *An* interacting particle system *with configuration space $T^{\mathbb{Z}^d}$ as constructed in the last section solves the martingale problem for the operator*

$$(\mathscr{L}f)(\mu) = \sum_{x\in\mathbb{Z}^d}\sum_{i\in T} c_i(x,\mu)\cdot\left(f\left(\mu^{x,i}\right) - f(\mu)\right) \tag{4.1}$$

*with domain given by the bounded cylinder functions*

$$\mathscr{A} = \left\{f\colon S\to\mathbb{R} \ : \ f(\mu) = \varphi\left(\mu(x_1),\dots,\mu(x_k)\right), k\in\mathbb{N}, x_1,\dots,x_k\in\mathbb{Z}^d, \varphi\in\mathcal{F}_b(T^k)\right\}$$

*Note that for a cylinder function only finitely many summands in (4.1) do not vanish. Hence $\mathscr{L}f$ is well-defined.*

3. ***Diffusions:*** *Suppose $S = \mathbb{R}^n$. By Itô's formula, any (weak) solution $((X_t)_{t\geq 0}, P)$ of the stochastic differential equation*

$$dX_t = \sigma(X_t)\, dB_t + b(X_t)\, dt$$

*with an $\mathbb{R}^d$-valued Brownian motion $B_t$ and locally bounded measurable functions $\sigma\colon \mathbb{R}^n \to \mathbb{R}^{d\times n}$, $b\colon \mathbb{R}^n \to \mathbb{R}^n$, solves the martingale problem for the differential operator*

$$(\mathscr{L}f)(x) = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + b(x)\cdot\Delta f(x),$$
$$a(x) = \sigma(x)\sigma(x)^T,$$

*with domain $C^2(\mathbb{R}^n)$, and the martingale problem for the same operator with domain $\mathscr{A} = C_0^2(\mathbb{R}^n)$, provided there is no explosion in finite time. The case of explosion can be included by extending the state space to $\mathbb{R}^n\,\dot\cup\{\Delta\}$ and setting $f(\Delta) = 0$ for $f\in C_0^2(\mathbb{R}^n)$.*

4. **Lévy processes** *A Lévy process solves the martingale problem for its generator*

$$\mathscr{L}f = -(\psi\hat{f})\check{}$$

*with domain* $\mathscr{A} = \mathscr{S}(\mathbb{R}^n)$.

From now on we assume that we are given a right continuous time-homogeneous Markov process $((X_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, (P_x)_{x\in S})$ with transition semigroup $(p_t)_{t\geq 0}$ such that for any $x \in S$, $(X_t)_{t\geq 0}$ is under $P_x$ a solution of the martingale problem for $(\mathscr{L}, \mathscr{A})$ with $P_x\,[X_0 = x] = 1$.

**Remark** (Markov property of solutions of martingale problems)**.** *Suppose $P_x, x \in S$, are probability measures on*

$$\mathscr{D}(\mathbb{R}^+, S) := \text{ all càdlàg functions } \omega\colon \mathbb{R}^+ \to S$$

*such that with respect to $P_x$, the canonical process $X_t(\omega) = \omega(t)$ is a solution of the martingale problem for $(\mathscr{L}, \mathscr{A})$ satisfying $P_x[X_0 = x] = 1$.*
*If*

*(i) $\mathcal{A}$ is separable with respect to $\|f\|_{\mathscr{L}} := \|f\|_{\sup} + \|\mathscr{L}f\|_{\sup}$,*

*(ii) $x \mapsto P_x(B)$ is measurable for all $B \in \mathcal{S}$,*

*(iii) For any $x \in S$, $P_x$ is the* unique *probability measure on $\mathscr{D}(\mathbb{R}^+, S)$ solving the martingale problem,*

*then $(X_t, P_x)$ is a strong Markov process, cf. e.g. Rogers, Williams [18] Volume 1.*

Let $\bar{\mathscr{A}}$ denote the closure of $\mathscr{A}$ with respect to the supremum norm. For most results derived below, we will impose two additional assumptions:

**Assumptions:**

**(A1)** If $f \in \mathscr{A}$, then $\mathscr{L}f \in \bar{\mathscr{A}}$.

**(A2)** There exists a linear subspace $\mathscr{A}_0 \subseteq \mathscr{A}$ such that if $f \in \mathscr{A}_0$, then $p_t f \in \mathscr{A}$ for all $t \geq 0$, and $\mathscr{A}_0$ is dense in $\mathscr{A}$ with respect to the supremum norm.

**Example .** *1. For Lévy processes (A1) and (A2) hold with $\mathscr{A}_0 = \mathscr{A} = \mathscr{S}(\mathbb{R}^d)$, and $B = \bar{\mathscr{A}} = C_\infty(\mathbb{R}^d)$.*

*2. For a diffusion process in $\mathbb{R}^d$ with continuous non-degenerated coefficients satisfying an appropriate growth constraint at infinity, (A1) and (A2) hold with $\mathscr{A}_0 = C_0^\infty(\mathbb{R}^d)$, $\mathscr{A} = \mathscr{S}(\mathbb{R}^d) \cap C^2(\mathbb{R}^d)$ and $B = \bar{\mathscr{A}} = C_\infty(\mathbb{R}^d)$.*

3. *In general, it can be difficult to determine explicitly a space $\mathscr{A}_0$ such that (A2) holds. In this case, a common procedure is to approximate the Markov process and its transition semigroup by more regular processes (e.g. non-degenerate diffusions in $\mathbb{R}^d$), and to derive asymptotic properties from corresponding properties of the approximands.*

4. *For an interacting particle system on $T^{\mathbb{Z}^d}$ with bounded transition rates $c_i(x,\eta)$, the conditions (A1) and (A2) hold with*

$$\mathscr{A}_0 = \mathscr{A} = \left\{ f : T^{\mathbb{Z}^d} \to \mathbb{R} \;:\; |||f||| < \infty \right\}$$

*where*

$$|||f||| = \sum_{x\in\mathbb{Z}^d} \Delta_f(x), \qquad \Delta_f(x) = \sup_{i\in T} \left| f(\eta^{x,i}) - f(\eta) \right|,$$

*cf. Liggett [15].*

**Theorem 4.2** (From the martingale problem to the Kolmogorov equations)**.** *Suppose (A1) and (A2) hold. Then $(p_t)_{t\geq 0}$ induces a $C_0$ contraction semigroup $(P_t)_{t\geq 0}$ on the Banach space $B = \bar{\mathscr{A}} = \bar{\mathscr{A}_0}$, and the generator is an extension of $(\mathscr{L}, \mathscr{A})$. In particular, the forward and backward equations*

$$\frac{d}{dt} p_t f = p_t \mathscr{L} f \quad \forall\, f \in \mathscr{A}$$

*and*

$$\frac{d}{dt} p_t f = \mathscr{L} p_t f \quad \forall\, f \in \mathscr{A}_0$$

*hold.*

*Proof.* Since $M_t^f$ is a bounded martingale with respect to $P_x$, we obtain the integrated backward forward equation by Fubini:

$$\begin{aligned}(p_t f)(x) - f(x) = \mathbb{E}_x[f(X_t) - f(X_0)] &= \mathbb{E}_x\left[\int_0^t (\mathscr{L} f)(X_s)\, ds\right] \\ &= \int_0^t (p_s \mathscr{L} f)(x)\, ds\end{aligned} \tag{4.2}$$

for all $f \in \mathscr{A}$ and $x \in S$. In particular,

$$\|p_t f - f\|_{\sup} \leq \int_0^t \|p_s \mathscr{L} f\|_{\sup}\, ds \leq t \cdot \|\mathscr{L} f\|_{\sup} \to 0$$

as $t \downarrow 0$ for any $f \in \mathscr{A}$. This implies strong continuity on $B = \bar{\mathscr{A}}$ since each $p_t$ is a contraction with respect to the sup-norm. Hence by (A1) and (4.2),

$$\frac{p_t f - f}{t} - \mathscr{L} f = \frac{1}{t} \int_0^t (p_s \mathscr{L} f - \mathscr{L} f) \, ds \to 0$$

uniformly for all $f \in \mathscr{A}$, i.e. $\mathscr{A}$ is contained in the domain of the generator L of the semigroup $(P_t)_{t \geq 0}$ induced on $B$, and $Lf = \mathscr{L} f$ for all $f \in \mathscr{A}$. Now the forward and the backward equations follow from the corresponding equations for $(P_t)_{t \geq 0}$ and Assumption (A2). □

## 4.2 Stationary distributions and reversibility

**Theorem 4.3** (Infinitesimal characterization of stationary distributions)**.** *Suppose (A1) and (A2) hold. Then for $\mu \in M_1(S)$ the following assertions are equivalent:*

*(i) The process $(X_t, P_\mu)$ is stationary, i.e.*

$$(X_{s+t})_{t \geq 0} \sim (X_t)_{t \geq 0}$$

*with respect to $P_\mu$ for all $s \geq 0$.*

*(ii) $\mu$ is a stationary distribution for $(p_t)_{t \geq 0}$*

*(iii)*

$$\int \mathscr{L} f \, d\mu = 0 \quad \forall f \in \mathscr{A}$$

*(i.e. $\mu$ is infinitesimally invariant, $\mathscr{L}^* \mu = 0$).*

*Proof.* (i)⇒(ii) If (i) holds then in particular

$$\mu p_s = P_\mu \circ X_s^{-1} = P_\mu \circ X_0^{-1} = \mu$$

for all $s \geq 0$, i.e. $\mu$ is a stationary initial distribution.

(ii)⇒(i) By the Markov property, for any measurable subset $B \subseteq \mathscr{D}(\mathbb{R}^+, S)$,

$$P_\mu[(X_{s+t})_{t \geq 0} \in B \mid \mathcal{F}_s] = P_{X_s}[(X_t)_{t \geq 0} \in B]$$

$P_\mu$-a.s., and thus

$$P_\mu[(X_{s+t})_{t \geq 0} \in B] = \mathbb{E}_\mu[P_{X_s}((X_t)_{t \geq 0} \in B)] = P_{\mu p_s}[(X_t)_{t \geq 0} \in B] = P_\mu[X \in B]$$

(ii)⇒(iii) By the theorem above, for $f \in \mathscr{A}$,

$$\frac{p_t f - f}{t} \to \mathscr{L} f \text{ uniformly as } t \downarrow 0,$$

so

$$\int \mathscr{L} f \, d\mu = \lim_{t \downarrow 0} \frac{\int (p_t f - f) \, d\mu}{t} = \lim_{t \downarrow 0} \frac{\int f \, d(\mu p_t) - \int f \, d\mu}{t} = 0$$

provided $\mu$ is stationary with respect to $(p_t)_{t \geq 0}$.

(iii)⇒(ii) By the backward equation and (iii),

$$\frac{d}{dt} \int p_t f \, d\mu = \int \mathscr{L} p_t f \, d\mu = 0$$

since $p_t f \in \mathscr{A}$ for $f \in \mathscr{A}_0$ and hence

$$\int f \, d(\mu p_t) = \int p_t f \, d\mu = \int f \, d\mu \tag{4.3}$$

for all $f \in \mathscr{A}_0$ and $t \geq 0$. Since $\mathscr{A}_0$ is dense in $\mathscr{A}$ with respect to the supremum norm, (4.3) extends to all $f \in \mathscr{A}$. Hence $\mu p_t = \mu$ for all $t \geq 0$ by (A0).

□

**Remark .** *Assumption (A2) is required only for the implication (iii)⇒(ii).*

**Applicaton to Itô diffusions:**

Suppose that we are given non-explosive weak solutions $(X_t, P_x), x \in \mathbb{R}^d$, of the stochastic differential equation

$$dX_t = \sigma(X_t) \, dB_t + b(X_t) \, dt, \quad X_0 = x \quad P_x\text{-a.s.},$$

where $(B_t)_{t \geq 0}$ is a Brownian motion in $\mathbb{R}^d$, and the functions $\sigma \colon \mathbb{R}^n \to \mathbb{R}^{n \times d}$ and $b \colon \mathbb{R}^n \to \mathbb{R}^n$ are locally Lipschitz continuous. Then by Itô's formula $(X_t, P_x)$ solves the martingale problem for the operator

$$\mathscr{L} = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + b(x) \cdot \nabla, \quad a = \sigma \sigma^T,$$

with domain $\mathscr{A} = C_0^\infty(\mathbb{R}^n)$. Moreover, the local Lipschitz condition implies uniqueness of strong solutions, and hence, by the Theorem of Yamade-Watanabe, uniqueness in distribution of weak solutions and uniqueness of the martingale problem for $(\mathscr{L}, \mathscr{A})$, cf. e.g. Rogers/Williams [18]. Therefore by the remark above, $(X_t, P_x)$ is a Markov process.

**Theorem 4.4.** *Suppose $\mu$ is a stationary distribution of $(X_t, P_x)$ that has a smooth density $\varrho$ with respect to the Lebesgue measure. Then*

$$\mathscr{L}^*\varrho := \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}\varrho) - \operatorname{div}(b\varrho) = 0$$

*Proof.* Since $\mu$ is a stationary distribution,

$$0 = \int \mathscr{L}f\,d\mu = \int_{\mathbb{R}^n} \mathscr{L}f\,\varrho\,dx = \int_{\mathbb{R}^n} f\mathscr{L}^*\varrho\,dx \quad \forall\, f \in C_0^\infty(\mathbb{R}^n) \tag{4.4}$$

Here the last equation follows by integration by parts, because $f$ has compact support. □

**Remark .** *In general, $\mu$ is a distributional solution of $\mathscr{L}^*\mu = 0$.*

**Example** (One-dimensional diffusions)**.** *In the one-dimensional case,*

$$\mathscr{L}f = \frac{a}{2}f'' + bf',$$

*and*

$$\mathscr{L}^*\varrho = \frac{1}{2}(a\varrho)'' - (b\varrho)'$$

*where $a(x) = \sigma(x)^2$. Assume $a(x) > 0$ for all $x \in \mathbb{R}$.*

**a) Harmonic functions and recurrence:**

$$\begin{aligned}\mathscr{L}f = \frac{a}{2}f'' + bf' = 0 \quad &\Leftrightarrow \quad f' = C_1 \exp - \int_0^\bullet \frac{2b}{a}\,dx, \quad C_1 \in \mathbb{R}\\ &\Leftrightarrow \quad f = C_2 + C_1 \cdot s, \quad C_1, C_2 \in \mathbb{R}\end{aligned}$$

where

$$s := \int_0^\bullet e^{-\int_0^y \frac{2b(x)}{a(x)}\,dx}\,dy$$

is a strictly increasing harmonic function that is called the ***scale function*** or ***natural scale of the diffusion***. In particular, $s(X_t)$ is a martingale with respect to $P_x$. The stopping theorem implies

$$P_x[T_a < T_b] = \frac{s(b) - s(x)}{s(b) - s(a)} \quad \forall\, a < x < b$$

As a consequence,

(i) If $s(\infty) < \infty$ or $s(-\infty) > -\infty$ then $P_x[|X_t| \to \infty] = 1$ for all $x \in \mathbb{R}$, i.e., $(X_t, P_x)$ is *transient*.

(ii) If $s(\mathbb{R}) = \mathbb{R}$ then $P_x[T_a < \infty] = 1$ for all $x, a \in \mathbb{R}$, i.e., $(X_t, P_x)$ is irreducible and *recurrent*.

**b) Stationary distributions:**

(i) $\boldsymbol{s(\mathbb{R}) \neq \mathbb{R}}$: In this case, by the transience of $(X_t, P_x)$, a stationary distribution does not exist. In fact, if $\mu$ is a finite stationary measure, then for all $t, r > 0$,

$$\mu(\{x : |x| \leq r\}) = (\mu p_t)(\{x : |x| \leq r\}) = P_\mu[|X_t| \leq r].$$

Since $X_t$ is transient, the right hand side converges to $0$ as $t \uparrow \infty$. Hence $\mu(\{x : |x| \leq r\}) = 0$ for all $r > 0$, i.e., $\mu \equiv 0$.

(ii) $\boldsymbol{s(\mathbb{R}) = \mathbb{R}}$: We can solve the ordinary differential equation $\mathscr{L}^* \varrho = 0$ explicitly:

$$\begin{aligned}
& \mathscr{L}^* \varrho = \left( \frac{1}{2}(a\varrho)' - b\varrho \right)' = 0 & \\
\Leftrightarrow \quad & \frac{1}{2}(a\varrho)' - \frac{b}{a} a\varrho = C_1 & \text{with } C_1 \in \mathbb{R} \\
\Leftrightarrow \quad & \frac{1}{2}\left( e^{-\int_0^\bullet \frac{2b}{a}\,dx} a\varrho \right)' = C_1 \cdot e^{-\int_0^\bullet \frac{2b}{a}\,dx} & \\
\Leftrightarrow \quad & s' a\varrho = C_2 + 2C_1 \cdot s & \text{with } C_1, C_2 \in \mathbb{R} \\
\Leftrightarrow \quad & \varrho(y) = \frac{C_2}{a(y)s'(y)} = \frac{C_2}{a(y)} e^{\int_0^y \frac{2b}{a}\,dx} & \text{with } C_2 \geq 0
\end{aligned}$$

Here the last equivalence holds since $s'a\varrho \geq 0$ and $s(\mathbb{R}) = \mathbb{R}$ imply $C_2 = 0$. Hence a stationary distribution $\mu$ can only exist if the measure

$$m(dy) := \frac{1}{a(y)} e^{\int_0^y \frac{2b}{a}\,dx}\, dy$$

is finite, and in this case $\mu = \frac{m}{m(\mathbb{R})}$. The measure $m$ is called the ***speed measure*** of the diffusion.

**Concrete examples:**

1. **Brownian motion:** $a \equiv 1, b \equiv 0, s(y) = y$. Since $s(\mathbb{R}) = \mathbb{R}$, Brownian motion is transient and there is no stationary distribution. Lebesgue measure is an infinite stationary measure.

2. **Ornstein-Uhlenbeck process:**

$$\begin{aligned}
dX_t &= dB_t - \gamma X_t\,dt, && \gamma > 0,\\
\mathscr{L} &= \frac{1}{2}\frac{d^2}{dx^2} - \gamma x \frac{d}{dx}, && a \equiv 1,\\
b(x) &= -\gamma x, && s(y) = \int_0^y e^{\int_0^y 2\gamma x\,dx}\,dy = \int_0^y e^{\gamma y^2}\,dy \text{ recurrent},\\
m(dy) &= e^{-\gamma y^2}\,dy, && \mu = \frac{m}{m(\mathbb{R})} = N\left(0, \frac{2}{\gamma}\right) \text{ is the unique stationary distribution}
\end{aligned}$$

3.

$$dX_t = dB_t + b(X_t)\,dt, \qquad b \in C^2, \qquad b(x) = \frac{1}{x} \text{ for } |x| \geq 1$$

transient, two independent non-negative solutions of $\mathscr{L}^*\varrho = 0$ with $\int \varrho\,dx = \infty$.

(***Exercise:*** stationary distributions for $dX_t = dB_t - \frac{\gamma}{1+|X_t|}\,dt$)

**Example** (Deterministic diffusions)**.**

$$\begin{aligned}
dX_t &= b(X_t)\,dt, && b \in C^2(\mathbb{R}^n)\\
\mathscr{L}f &= b \cdot \nabla f\\
\mathscr{L}^*\varrho &= -\operatorname{div}(\varrho b) = -\varrho \operatorname{div} b - b \cdot \nabla \varrho, && \varrho \in C^1
\end{aligned}$$

**Proposition 4.5.**

$$\begin{aligned}
\mathscr{L}^*\varrho = 0 \quad &\Leftrightarrow \quad \operatorname{div}(\varrho b) = 0\\
&\Leftrightarrow \quad (\mathscr{L}, C_0^\infty(\mathbb{R}^n)) \textit{ anti-symmetric on } L^2(\mu)
\end{aligned}$$

*Proof.* First equivalence: cf. above

Second equivalence:

$$\begin{aligned}
\int f\mathscr{L}g\,d\mu &= \int fb \cdot \nabla g \varrho\,dx = -\int \operatorname{div}(fb\varrho)g\,dx\\
&= -\int \mathscr{L}f g\,d\mu - \int \operatorname{div}(\varrho b) f g\,dx \qquad \forall\, f, g \in C_0^\infty
\end{aligned}$$

Hence $\mathscr{L}$ is anti-symmetric if and only if $\operatorname{div}(\varrho b) = 0$

□

**Theorem 4.6.** *Suppose (A1) and (A2) hold. Then for $\mu \in M_1(S)$ the following assertions are equivalent:*

*(i) The process $(X_t, P_\mu)$ is* ***invariant with respect to time reversal****, i.e.,*

$$(X_s)_{0\le s\le t} \sim (X_{t-s})_{0\le s\le t} \quad \text{with respect to } P_\mu \; \forall\, t \ge 0$$

*(ii)*

$$\mu(dx)p_t(x,dy) = \mu(dy)p_t(y,dx) \quad \forall\, t \ge 0$$

*(iii) $p_t$ is $\mu$-symmetric, i.e.,*

$$\int f p_t g \, d\mu = \int p_t f g \, d\mu \quad \forall\, f, g \in \mathcal{F}_b(S)$$

*(iv) $(\mathscr{L}, \mathscr{A})$ is $\mu$-symmetric, i.e.,*

$$\int f \mathscr{L} g \, d\mu = \int \mathscr{L} f g \, d\mu \quad \forall\, f, g \in \mathscr{A}$$

**Remark .** *1. A reversible process $(X_t, P_\mu)$ is stationary, since for all $s, u \ge 0$,*

$$(X_{s+t})_{0\le t\le u} \sim (X_{u-t})_{0\le t\le u} \sim (X_t)_{0\le t\le u} \quad \text{with respect to } P_\mu$$

*2. Similarly (ii) implies that $\mu$ is a stationary distribution:*

$$\int \mu(dx)p_t(x,dy) = \int p_t(y,dx)\mu(dy) = \mu(dy)$$

*Proof of the Theorem.* (i)⇒(ii):

$$\mu(dx)p_t(x,dy) = P_\mu \circ (X_0, X_t)^{-1} = P_\mu \circ (X_t, X_0)^{-1} = \mu(dy)p_t(y,dx)$$

(ii)⇒(i): By induction, (ii) implies

$$\begin{aligned} &\mu(dx_0)p_{t_1-t_0}(x_0,dx_1)p_{t_2-t_1}(x_1,dx_2)\cdots p_{t_n-t_{n-1}}(x_{n-1},dx_n) \\ =&\mu(dx_n)p_{t_1-t_0}(x_n,dx_{n-1})\cdots p_{t_n-t_{n-1}}(x_1,dx_0) \end{aligned}$$

for $n \in \mathbb{N}$ and $0 = t_0 \le t_1 \le \cdots \le t_n = t$, and thus

$$\mathbb{E}_\mu[f(X_0, X_{t_1}, X_{t_2}, \ldots, X_{t_{n-1}}, X_t)] = \mathbb{E}_\mu[f(X_t, \ldots, X_{t_1}, X_0)]$$

for all measurable functions $f \ge 0$. Hence the time-reversed distribution coincides with the original one on cylinder sets, and thus everywhere.

(ii)⇔(iii): By Fubini,

$$\int f p_t g \, d\mu = \iint f(x)g(y)\mu(dx)p_t(x,dy)$$

is symmetric for all $f, g \in \mathcal{F}_b(S)$ if and only if $\mu \otimes p_t$ is a symmetric measure on $S \times S$.

(iii)⇔(iv): Exercise.

□

**Application to Itô diffusions in $\mathbb{R}^n$:**

$$\mathscr{L} = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + b\cdot\nabla, \qquad \mathscr{A} = C_0^\infty(\mathbb{R}^n)$$

$\mu$ probability measure on $\mathbb{R}^n$ (more generally locally finite positive measure)

*Question*: For which process is $\mu$ stationary?

**Theorem 4.7.** *Suppose $\mu = \varrho\, dx$ with $\varrho_i a_{ij} \in C^1, b \in C, \varrho > 0$. Then*

1. *We have*

$$\mathscr{L}g = \mathscr{L}_s g + \mathscr{L}_a g$$

   *for all $g \in C_0^\infty(\mathbb{R}^n)$ where*

$$\mathscr{L}_s g = \frac{1}{2}\sum_{i,j=1} n \frac{1}{\varrho}\frac{\partial}{\partial x_i}\left(\varrho a_{ij}\frac{\partial g}{\partial x_i}\right)$$

$$\mathscr{L}_a g = \beta\cdot\nabla g, \quad \beta_j = b_j - \sum_i \frac{1}{2\varrho}\frac{\partial}{\partial x_i}(\varrho a_{ij})$$

2. *The operator $(\mathscr{L}_s, C_0^\infty$ is symmetric with respect to $\mu$.*

3. *The following assertions are equivalent:*

   *(i)* $\mathscr{L}^*\mu = 0$ *(i.e. $\int \mathscr{L} f\, d\mu = 0$ for all $f \in C_0^\infty$).*

   *(ii)* $\mathscr{L}_a^*\mu = 0$

   *(iii)* $\operatorname{div}(\varrho\beta) = 0$

   *(iv)* $(\mathscr{L}_a, C_0^\infty)$ *is anti-symmetric with respect to $\mu$*

*Proof.* Let

$$\mathcal{E}(f,g) := -\int f\mathscr{L}g\, d\mu \qquad (f,g\in C_0^\infty)$$

denote the bilinear form of the operator $(\mathscr{L}, C_0^\infty(\mathbb{R}^n))$ on the Hilbert space $L^2(\mathbb{R}^n,\mu)$. We decompose $\mathcal{E}$ into a symmetric part and a remainder. An explicit computation based on the integration by parts formula in $\mathbb{R}^n$ shows that for $g \in C_0^\infty(\mathbb{R}^n)$ and $f \in C^\infty(\mathbb{R}^n)$:

$$\begin{aligned}
\mathcal{E}(f,g) &= -\int f\left(\frac{1}{2}\sum a_{ij}\frac{\partial^2 g}{\partial x_i\partial x_j} + b\cdot\nabla g\right)\varrho\, dt \\
&= \int \frac{1}{2}\sum_{i,j}\frac{\partial}{\partial x_i}(\varrho a_{ij} f)\frac{\partial g}{\partial x_j}\, dx - \int f b\cdot\nabla g \varrho\, dx \\
&= \int \frac{1}{2}\sum_{i,j} a_{i,j}\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}\varrho\, dx - \int f\beta\cdot\nabla g\varrho\, dx \qquad \forall f,g\in C_0^\infty
\end{aligned}$$

and set

$$\mathcal{E}_s(f,g) := \int \frac{1}{2} \sum_{i,j} a_{i,j} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} \varrho \, dx = - \int f \mathscr{L}_s g \, d\mu$$

$$\mathcal{E}_a(f,g) := \int f \beta \cdot \nabla g \varrho \, dx = - \int f \mathscr{L}_a g \, d\mu$$

This proves 1) and, since $\mathcal{E}_s$ is a symmetric bilinear form, also 2). Moreover, the assertions (i) and (ii) of 3) are equivalent, since

$$- \int \mathscr{L} g \, d\mu = \mathcal{E}(1,g) = \mathcal{E}_s(1,g) + \mathcal{E}_a(1,g) = - \int \mathscr{L}_a g \, d\mu$$

for all $g \in C_0^\infty(\mathbb{R}^n)$ since $\mathcal{E}_s(1,g) = 0$. Finally, the equivalence of (ii),(iii) and (iv) has been shown in the example above. □

**Example .** $\mathscr{L} = \frac{1}{2}\Delta + b \cdot \nabla, \; b \in C(\mathbb{R}^n, \mathbb{R}^n)$,

$$(\mathscr{L}, C_0^\infty) \; \mu\text{-symmetric} \quad \Leftrightarrow \quad \beta = b - \frac{1}{2\varrho} \nabla \varrho = 0$$

$$\Leftrightarrow \quad b = \frac{\nabla \varrho}{2\varrho} = \frac{1}{2} \nabla \log \varrho$$

*where* $\log \varrho = -H$ *if* $\mu = e^{-H} \, dx$.

$$\mathscr{L} \, symmetrizable \quad \Leftrightarrow \quad b \text{ is a gradient}$$

$$\mathscr{L}^* \mu = 0 \quad \Leftrightarrow \quad b = \frac{1}{2} \nabla \log \varrho + \beta$$

*when* $\operatorname{div}(\varrho\beta) = 0$.

**Remark .** *Probabilistic proof of reversibility for* $b := -\frac{1}{2}\nabla H, \; H \in C^1$*:*

$$X_t = x + B_t + \int_0^t b(X_s) \, ds, \quad \textit{non-explosive}, \quad b = -\frac{1}{2} \nabla h$$

*Hence* $P_\mu \circ X_{0:T}^{-1} \ll$ *Wiener measure with density*

$$\exp \left( -\frac{1}{2} H(B_0) - \frac{1}{2} H(B_T) - \int_0^T \left( \frac{1}{8} |\nabla H| - \frac{1}{2} \Delta H \right) (B_s) \, ds \right)$$

*which shows that* $(X_t, P_\mu)$ *is reversible.*

## 4.3 Dirichlet forms and convergence to equilibrium

Suppose now that $\mu$ is a stationary distribution for $(p_t)_{t\geq 0}$. Then $p_t$ is a contraction on $L^p(S,\mu)$ for all $p \in [1,\infty]$ since

$$\int |p_t f|^p \, d\mu \leq \int p_t |f|^p \, d\mu = \int |f|^p \, d\mu \quad \forall\, f \in \mathcal{F}_b(S)$$

by Jensen's inequality and the stationarity of $\mu$. As before, we assume that we are given a Markov process with transition semigroup $(p_t)_{t\geq 0}$ solving the martingale problem for the operator $(\mathscr{L}, \mathscr{A})$. The assumptions on $\mathscr{A}_0$ and $\mathscr{A}$ can be relaxed in the following way:

**(A0)** as above

**(A1')** $f,\ \mathscr{L}f \in \mathscr{L}^p(S,\mu)$ for all $1 \leq p < \infty$

**(A2')** $\mathscr{A}_0$ is dense in $\mathscr{A}$ with respect to the $L^p(S,\mu)$ norms, $1 \leq p < \infty$, and $p_t f \in \mathscr{A}$ for all $f \in \mathscr{A}_0$

In addition, we assume for simplicity

**(A3)** $1 \in \mathscr{A}$

**Remark .** *Condition (A0) implies that $\mathscr{A}$, and hence $\mathscr{A}_0$, is dense in $L^p(S,\mu)$ for all $p \in [1,\infty)$. In fact, if $g \in \mathscr{L}^q(S,\mu)$, $\frac{1}{q}+\frac{1}{q}=1$, with $\int fg\,d\mu = 0$ for all $f \in \mathscr{A}$, then $g\,d\mu = 0$ by (A0) and hence $g = 0$ $\mu$-a.e. Similarly as above, the conditions (A0), (A1') and (A2') imply that $(p_t)_{t\geq 0}$ induces a $C_0$ semigroup on $L^p(S,\mu)$ for all $p \in [1,\infty)$, and the generator $(L^{(p)}, \mathrm{Dom}(L^{(p)}))$ extends $(\mathscr{L}, \mathscr{A})$, i.e.,*

$$\mathscr{A} \subseteq \mathrm{Dom}(L^{(p)}) \quad \textit{and} \quad L^{(p)} f = \mathscr{L} f \quad \mu\textit{-a.e. for all } f \in \mathscr{A}$$

*In particular, the Kolmogorov forward equation*

$$\frac{d}{dt} p_t f = p_t \mathscr{L} f \quad \forall\, f \in \mathscr{A}$$

*and the backward equation*

$$\frac{d}{dt} p_t f = \mathscr{L} p_t f \quad \forall\, f \in \mathscr{A}_0$$

*hold with the derivative taken in the Banach space $L^p(S,\mu)$.*

We first restrict ourselves to the case $p = 2$. For $f, g \in \mathscr{L}^2(S,\mu)$ let

$$(f,g)_\mu = \int fg\,d\mu$$

denote the $L^2$ inner product.

**Definition 4.8.** *The bilinear form*

$$\mathscr{E}(f,g) := -(f,\mathscr{L}g)_\mu = -\frac{d}{dt}(f,p_t g)_\mu\Big|_{t=0},$$

*$f,g \in \mathscr{A}$, is called the* ***Dirichlet form*** *associated to $(\mathscr{L},\mathscr{A})$ on $L^2(\mu)$.*

$$\mathscr{E}_s(f,g) := \frac{1}{2}\left(\mathscr{E}(f,g)+\mathscr{E}(g,f)\right)$$

*is the* ***symmetrized Dirichlet form****.*

**Remark .** *More generally, $\mathscr{E}(f,g)$ is defined for all $f \in L^2(S,\mu)$ and $g \in \mathrm{Dom}(L^{(2)})$ by*

$$\mathscr{E}(f,g) = -(f,L^{(2)}g)_\mu = -\frac{d}{dt}(f,p_t g)_\mu\Big|_{t=0}$$

**Theorem 4.9.** *For all $f \in \mathscr{A}_0$ and $t \geq 0$*

$$\frac{d}{dt}\operatorname{Var}_\mu(p_t f) = \frac{d}{dt}\int (p_t f)^2\,d\mu = -2\mathscr{E}(p_t f, p_t f) = -2\mathscr{E}_s(p_t f, p_t f)$$

**Remark .**

1. *In particular,*

$$\mathscr{E}(f,f) = -\frac{1}{2}\int (p_t f)^2\,d\mu = -\frac{1}{2}\frac{d}{dt}\operatorname{Var}_\mu(p_t f)\Big|_{t=0},$$

infinitesimal change of variance

2. *The assertion extends to all $f \in \mathrm{Dom}(L^{(2)})$ if the Dirichlet form is defined with respect to the $L^2$ generator. In the symmetric case the assertion even holds for all $f \in L^2(S,\mu)$.*

*Proof.* By the backward equation,

$$\frac{d}{dt}\int (p_t f)^2\,d\mu = 2\int p_t \mathscr{L} p_t f\,d\mu = -2\mathscr{E}(p_t f, p_t f) = -2\mathscr{E}_s(p_t f, p_t f)$$

Moreover, since

$$\int p_t f\,d\mu = \int f\,d(\mu p_t) = \int f\,d\mu$$

is constant,

$$\frac{d}{dt}\operatorname{Var}_\mu(p_t) = \frac{d}{dt}\int (p_t f)^2\,d\mu$$

□

**Remark .** *1. In particular,*

$$\mathscr{E}(f,f) = -\frac{1}{2}\frac{d}{dt}\int (p_t f)^2\, d\mu\Big|_{t=0} = -\frac{1}{2}\frac{d}{dt}\operatorname{Var}_\mu(p_t f)$$

$$\mathscr{E}_s(f,g) = \frac{1}{4}\left(\mathscr{E}_s(f+g,f+g) + \mathscr{E}_s(f-g,f-g)\right) = -\frac{1}{2}\frac{d}{dt}\operatorname{Cov}_\mu(p_t f, p_t g)$$

*Dirichlet form = infinitesimal change of (co)variance.*

2. *Since $p_t$ is a contraction on $\mathscr{L}^2(\mu)$, the operator $(\mathscr{L},\mathscr{A})$ is negative-definite, and the bilinear form $(\mathscr{E},\mathscr{A})$ is positive definite:*

$$(-f,\mathscr{L}f)_\mu = \mathscr{E}(f,f) = -\frac{1}{2}\lim_{t\downarrow 0}\left(\int (p_t f)^2\, d\mu - \int f^2\, d\mu\right) \geq 0$$

**Corollary 4.10** (Decay of variance)**.** *For $\lambda > 0$ the following assertions are equivalent:*

*(i)* ***Poincaré inequality****:*

$$\operatorname{Var}_\mu(f) \leq \frac{1}{\lambda}\mathscr{E}_{(s)}(f,f) \quad \forall\, f \in \mathscr{A}$$

*(ii)* ***Exponential decay of variance****:*

$$\operatorname{Var}_\mu(p_t f) \leq e^{-2\lambda t}\operatorname{Var}_\mu(f) \quad \forall\, f \in L^2(S,\mu) \tag{4.5}$$

*(iii)* ***Spectral gap****:*

$$\operatorname{Re}\alpha \geq \lambda \quad \forall\, \alpha \in \operatorname{spec}\left(-L^{(2)}\Big|_{\operatorname{span}\{1\}^\perp}\right)$$

**Remark .** *Optimizing over $\lambda$, the corollary says that (4.5) holds with*

$$\lambda := \inf_{f\in\mathscr{A}} \frac{\mathscr{E}(f,f)}{\operatorname{Var}_\mu(f)} = \inf_{\substack{f\in\mathscr{A}\\ f\perp 1 \text{ in } L^2(\mu)}} \frac{(f,-\mathscr{L}f)_\mu}{(f,f)_\mu}$$

*Proof.* $(i) \Rightarrow (ii)$

$$\mathscr{E}(f,f) \geq \lambda\cdot\operatorname{Var}_\mu(f) \quad \forall\, f\in\mathscr{A}$$

By the theorem above,

$$\frac{d}{dt}\operatorname{Var}_\mu(p_t f) = -2\mathscr{E}(p_t f, p_t f) \leq -2\lambda\operatorname{Var}_\mu(p_t f)$$

for all $t \geq 0,\ f \in \mathscr{A}_0$. Hence

$$\operatorname{Var}_\mu(p_t f) \leq e^{-2\lambda t}\operatorname{Var}_\mu(p_0 f) = e^{-2\lambda t}\operatorname{Var}_\mu(f)$$

for all $f \in \mathscr{A}_0$. Since the right hand side is continuous with respect to the $L^2(\mu)$ norm, and $\mathscr{A}_0$ is dense in $L^2(\mu)$ by (A0) and (A2), the inequality extends to all $f \in L^2(\mu)$.

$(ii) \Rightarrow (iii)$ For $f \in \mathrm{Dom}(L^{(2)}$,

$$\frac{d}{dt} \operatorname{Var}_\mu(p_t f)\Big|_{t=0} = -2\mathscr{E}(f,f).$$

Hence if (4.5) holds then

$$\operatorname{Var}_\mu(p_t f) \leq e^{-2\lambda t} \operatorname{Var}_\mu(f) \quad \forall\, t \geq 0$$

which is equivalent to

$$\operatorname{Var}_\mu(f) - 2t\mathscr{E}(f,f) + o(t) \leq \operatorname{Var}_\mu(f) - 2\lambda t \operatorname{Var}_\mu(f) + o(t) \quad \forall\, t \geq 0$$

Hence

$$\mathscr{E}(f,f) \geq \lambda \operatorname{Var}_\mu(f)$$

and thus

$$-(L^{(2)} f, f)_\mu \geq \lambda \int f^2 \, d\mu \quad \text{for } f \perp 1$$

which is equivalent to (iii).

$(iii) \Rightarrow (i)$ Follows by the equivalence above.

□

**Remark .** *Since $(\mathscr{L}, \mathscr{A})$ is negative definite, $\lambda \geq 0$. In order to obtain exponentially decay, however, we need $\lambda > 0$, which is not always the case.*

**Example .** *1. **Finite state space:** Suppose $\mu(x) > 0$ for all $x \in S$.*

***Generator**:*

$$(\mathscr{L} f)(x) = \sum_y \mathscr{L}(x,y) f(y) = \sum_y \mathscr{L}(x,y)(f(y) - f(x))$$

***Adjoint**:*

$$\mathscr{L}^{*\mu}(y,x) = \frac{\mu(x)}{\mu(y)} \mathscr{L}(x,y)$$

*Proof.*

$$\begin{aligned}(\mathscr{L} f, g)_\mu &= \sum_{x,y} \mu(x) \mathscr{L}(x,y) f(y) g(x) \\ &= \sum \mu(y) f(y) \frac{\mu(x)}{\mu(y)} \mathscr{L}(x,y) g(x) \\ &= (f, \mathscr{L}^{*\mu} g)_\mu\end{aligned}$$

□

***Symmetric part:***

$$\mathscr{L}_s(x,y) = \frac{1}{2}\left(\mathscr{L}(x,y) + \mathscr{L}^{*\mu}(x,y)\right) = \frac{1}{2}\left(\mathscr{L}(x,y) + \frac{\mu(y)}{\mu(x)}\mathscr{L}(y,x)\right)$$
$$\mu(x)\mathscr{L}_s(x,y) = \frac{1}{2}\left(\mu(x)\mathscr{L}(x,y) + \mu(y)\mathscr{L}(y,x)\right)$$

***Dirichlet form:***

$$\begin{aligned}\mathscr{E}_s(f,g) = -(\mathscr{L}_s f, g) &= -\sum_{x,y}\mu(x)\mathscr{L}_s(x,y)\left(f(y)-f(x)\right)g(x)\\ &= -\sum_{x,y}\mu(y)\mathscr{L}_s(y,x)\left(f(x)-f(y)\right)g(y)\\ &= -\frac{1}{2}\sum \mu(x)\mathscr{L}_s(x,y)\left(f(y)-f(x)\right)\left(g(y)-g(x)\right)\end{aligned}$$

*Hence*

$$\mathscr{E}(f,f) = \mathscr{E}_s(f,f) = \frac{1}{2}\sum_{x,y} Q(x,y)\left(f(y)-f(x)\right)^2$$

*where*

$$Q(x,y) = \mu(x)\mathscr{L}_s(x,y) = \frac{1}{2}\left(\mu(x)\mathscr{L}(x,y) + \mu(y)\mathscr{L}(y,x)\right)$$

2. ***Diffusions in $\mathbb{R}^n$:*** *Let*

$$\mathscr{L} = \frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + b\cdot\nabla,$$

*and* $\mathscr{A} = C_0^\infty,\ \mu = \varrho\, dx,\ \varrho, a_{ij} \in C^1,\ b \in C\ \varrho \geq 0,$

$$\mathscr{E}_s(f,g) = \frac{1}{2}\int \sum_{i,j=1}^{n} a_{ij}\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}\, d\mu$$
$$\mathscr{E}(f,g) = \mathscr{E}_s(f,g) - (f, \beta\cdot\nabla g), \qquad \beta = b - \frac{1}{2\varrho}\operatorname{div}(\varrho a_{ij})$$

**Definition 4.11** ("Distances" of probability measures). *$\mu$, $\nu$ probability measures on $S$, $\mu - \nu$ signed measure.*

*(i)* ***Total variation distance:***

$$\|\nu - \mu\|_{TV} = \sup_{A\in\mathscr{S}} |\nu(A) - \mu(A)|$$

*(ii)* ***$\chi^2$-contrast:***

$$\chi^2(\mu|\nu) = \begin{cases} \int \left(\frac{d\mu}{d\nu} - 1\right)^2 \, d\mu = \int \left(\frac{d\nu}{d\mu}\right)^2 \, d\mu - 1 & \text{if } \nu \ll \mu \\ +\infty & \text{else} \end{cases}$$

*(iii)* ***Relative entropy**:*

$$H(\nu|\mu) = \begin{cases} \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} \, d\mu = \int \log \frac{d\nu}{d\mu} \, d\nu & \text{if } \nu \ll \mu \\ +\infty & \text{else} \end{cases}$$

*(where* $0 \log 0 := 0$).

**Remark .** *By Jensen's inequality,*

$$H(\nu|\mu) \geq \int \frac{d\nu}{d\mu} \, d\mu \log \int \frac{d\nu}{d\mu} \, d\mu = 0$$

**Lemma 4.12** (Variational characterizations)**.**

*(i)*

$$\|\nu - \mu\| = \frac{1}{2} \sup_{\substack{f \in \mathcal{F}_b(S) \\ |f| \leq 1}} \left( \int f \, d\nu - \int f \, d\mu \right)$$

*(ii)*

$$\lambda^2(\nu|\mu) = \sup_{\substack{f \in \mathcal{F}_b(S) \\ \int f^2 \, d\mu \leq 1}} \left( \int f \, d\nu - \int f \, d\mu \right)^2$$

*and by replacing $f$ by $f - \int f \, d\mu$,*

$$\lambda^2(\nu|\mu) = \sup_{\substack{f \in \mathcal{F}_b(S) \\ \int f^2 \, d\mu \leq 1 \\ \int f \, d\mu = 0}} \left( \int f \, d\nu \right)^2$$

*(iii)*

$$H(\nu|\mu) = \sup_{\substack{f \in \mathcal{F}_b(S) \\ \int e^f \, d\mu \leq 1}} \int f \, d\nu = \sup_{f \int \mathcal{F}_b(S)} \int f \, d\nu - \log \int e^f \, d\mu$$

**Remark .** *$\int e^f \, d\mu \le 1$, hence $\int f \, d\mu \le 0$ by Jensen and we also have*

$$\sup_{\int e^f \, d\mu \le 1} \left( \int f \, d\nu - \int f \, d\mu \right) \le H(\nu|\mu)$$

*Proof.* (i) " $\le$ "

$$\nu(A) - \mu(A) = \frac{1}{2} \left( \nu(A) - \mu(A) + \mu(A^c) - \nu(A^c) \right) = \frac{1}{2} \left( \int f \, d\nu - \int f \, d\mu \right)$$

and setting $f := I_A - I_{A^c}$ leads to

$$\|\nu - \mu\|_{\mathrm{TV}} = \sup_A \left( \nu(A) - \mu(A) \right) \le \frac{1}{2} \sup_{|f| \le 1} \left( \int f \, d\nu - \int f \, d\mu \right)$$

" $\ge$ " If $|f| \le 1$ then

$$\begin{aligned} \int f \, d(\nu - \mu) &= \int_{S_+} f \, d(\nu - \mu) + \int_{S_-} f \, d(\nu - \mu) \\ &\le (\nu - \mu)(S_+) - (\nu - \mu)(S_-) \\ &= 2(\nu - \mu)(S_+) \qquad \text{(since } (\nu - \mu)(S_+) + (\nu - \mu)(S_-) = (\nu - \mu)(S) = 0) \\ &\le 2\|\nu - \mu\|_{\mathrm{TV}} \end{aligned}$$

where $S = S_+ \dot{\bigcup} S_-$, $\nu - \mu \ge 0$ on $S_+$, $\nu - \mu \le 0$ on $S_-$ is the Hahn-Jordan decomposition of the measure $\nu - \mu$.

(ii) If $\nu \ll \mu$ with density $\varrho$ then

$$\chi^2(\nu|\mu)^{\frac{1}{2}} = \|\varrho - 1\|_{L^2(\mu)}^{\frac{1}{2}} = \sup_{\substack{f \in \mathscr{L}^2(\mu) \\ \|f\|_{L^2(\mu)} \le 1}} \int f(\varrho - 1) \, d\mu = \sup_{\substack{f \in \mathcal{F}_b(S) \\ \|f\|_{L^2(\mu)} \le 1}} \left( \int f \, d\nu - \int f \, d\mu \right)$$

by the Cauchy-Schwarz inequality and a density argument.

If $\nu \not\ll \mu$ then there exists $A \in \mathscr{S}$ with $\mu(A) = 0$ and $\nu(A) \neq 0$. Choosing $f = \lambda \cdot I_A$ with $\lambda \uparrow \infty$ we see that

$$\sup_{\substack{f \in \mathcal{F}_b(S) \\ \|f\|_{L^2(\mu)} \le 1}} \left( \int f \, d\nu - \int f \, d\mu \right)^2 = \infty = \chi^2(\nu|\mu).$$

This proves the first equation. The second equation follows by replacing $f$ by $f - \int f \, d\mu$.

(iii) First equation:

" $\geq$ " By Young's inequality,

$$uv \leq u \log u - u + e^v$$

for all $u \geq 0$ and $v \in \mathbb{R}$, and hence for $\nu \ll \mu$ with density $\varrho$,

$$\begin{aligned}
\int f \, d\nu &= \int f \varrho \, d\mu \\
&\leq \int \varrho \log \varrho \, d\mu - \int \varrho \, d\mu + \int e^f \, d\mu \\
&= H(\nu|\mu) - 1 + \int e^f \, d\mu \qquad \forall \, f \in \mathcal{F}_b(S) \\
&\leq H(\nu|\mu) \qquad \text{if} \int e^f \, d\mu \leq 1
\end{aligned}$$

" $\leq$ " $\nu \ll \mu$ with density $\varrho$:

a) $\varepsilon \leq \varrho \leq \frac{1}{\varepsilon}$ for some $\varepsilon > 0$: Choosing $f = \log \varrho$ we have

$$H(\nu|\mu) = \int \log \varrho \, d\nu = \int f \, d\nu$$

and

$$\int e^f \, d\mu = \int \varrho \, d\mu = 1$$

b) General case by an approximation argument.

Second equation: cf. Deuschel, Stroock [8].

□

**Remark .** *If $\nu \ll \mu$ with density $\varrho$ then*

$$\|\nu - \mu\|_{TV} = \frac{1}{2} \sup_{|f| \leq 1} \int f(\varrho - 1) \, d\mu = \frac{1}{2} \|\varrho - 1\|_{L^1(\mu)}$$

*However, $\|\nu - \mu\|_{TV}$ is finite even when $\nu \not\ll \mu$.*

**Corollary 4.13.** *The assertions $(i) - (iii)$ in the corollary above are also equivalent to*

*(iv)* ***Exponential decay of $\chi^2$ distance to equilibrium:***

$$\chi^2(\nu p_t|\mu) \leq e^{-2\lambda t} \chi^2(\nu|\mu) \quad \forall \, \nu \in M_1(S)$$

*Proof.* We show $(ii) \Leftrightarrow (iv)$.

" $\Rightarrow$ " Let $f \in \mathscr{L}^2(\mu)$ with $\int f\,d\mu = 0$. Then

$$\begin{aligned}\int f\,d(\nu p_t) - \int f\,d\mu = \int f\,d(\nu p_t) &= \int p_t f\,d\nu \\ &\le \|p_t f\|_{L^2(\mu)} \cdot \chi^2(\nu|\mu)^{\frac{1}{2}} \\ &\le e^{-\lambda t}\|f\|_{L^2(\mu)} \cdot \chi^2(\nu|\mu)^{\frac{1}{2}}\end{aligned}$$

where we have used that $\int p_t f\,d\mu = \int f\,d\mu = 0$. By taking the supremum over all $f$ with $\int f^2\,d\mu \le 1$ we obtain

$$\chi^2(\nu p_t|\mu)^{\frac{1}{2}} \le e^{-\lambda t}\chi^2(\nu|\mu)^{\frac{1}{2}}$$

" $\Leftarrow$ " For $f \in \mathscr{L}^2(\mu)$ with $\int f\,d\mu = 0$, (iv) implies

$$\begin{aligned}\int p_t f g\,d\mu \overset{\nu:=g\mu}{=} \int f\,d(\nu p_t) &\le \|f\|_{L^2(\mu)}\chi^2(\nu p_t|\mu)^{\frac{1}{2}} \\ &\le e^{-\lambda t}\|f\|_{L^2(\mu)}\chi^2(\nu|\mu)^{\frac{1}{2}} \\ &= e^{-\lambda t}\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}\end{aligned}$$

for all $g \in L^2(\mu), g \ge 0$. Hence

$$\|p_t f\|_{L^2(\mu)} \le e^{-\lambda t}\|f\|_{L^2(\mu)}$$

□

Example: $d = 1$!

**Example** (Gradient type of diffusions in $\mathbb{R}^n$).

$$dX_t = dB_t + b(X_t)\,dt, \qquad\qquad b \in C(\mathbb{R}^n, \mathbb{R}^n)$$

*Generator:*

$$\mathscr{L}f = \frac{1}{2}\Delta f + b\nabla f, \quad f \in C_0^\infty(\mathbb{R}^n)$$

*symmetric with respect to* $\mu = \varrho\,dx$, $\varrho \in C^1 \Leftrightarrow b = \frac{1}{2}\nabla\log\varrho$.
*Corresponding Dirichlet form on* $L^2(\varrho\,dx)$*:*

$$\mathscr{E}(f,g) = -\int \mathscr{L}f g\varrho\,dx = \frac{1}{2}\int \nabla f\nabla g\varrho\,dx$$

*Poincaré inequality:*

$$\mathrm{Var}_{\varrho\,dx}(f) \le \frac{1}{2\lambda}\cdot\int |\nabla f|^2\varrho\,dx$$

***The one-dimensional case:*** $n = 1, b = \frac{1}{2}(\log\varrho)'$ *and hence*

$$\varrho(x) = \textit{const.}\ e^{\int_0^x 2b(y)\,dy}$$

*e.g.* $b(x) = -\alpha x, \varrho(x) = \textit{const.}\ e^{-\alpha x^2}, \mu =$ *Gauss measure.*

**Theorem 4.14.** *???*

*Proof.* By the Cauchy-Schwarz inequality, for $x > 0$,

$$(f(x) - f(0))^2 = \left(\int_0^x f'\,dy\right)^2 \leq \int_0^x |f'|^2 g\varrho\,dy \int_0^x \frac{1}{g\varrho}\,dy$$

where $g\colon \mathbb{R}^+ \to \mathbb{R}^+$ is an arbitrary continuous strict positive function. Hence by Fubini

$$\begin{aligned}
&\int_0^\infty (f(x) - f(0))^2\,\varrho(x)\,dx \\
&\leq \int_0^\infty |f'(y)|^2 g(y)\varrho(y) \int_y^\infty \int_0^x \frac{1}{g(x)\varrho(z)}\,dz\,\varrho(x)\,dx \\
&\leq \int_0^\infty |f'(y)|^2 \varrho(y)\,dy \cdot \sup_{y>0} g(y) \cdot \int_y^\infty \int_0^x \frac{1}{g(x)\varrho(z)} dz\,\varrho(x)\,dx
\end{aligned}$$

Optimal choice for $g$:

$$g(y) = \left(\int_0^y \frac{1}{\varrho(x)}\,dx\right)^{\frac{1}{2}}$$

In this case:

$$\int_0^x \frac{1}{g\varrho}\,dz = 2\int_0^x g'\,dz = 2g(x),$$

so

$$\int_0^\infty (f - f(0))^2\,\varrho\,dx \leq \int_0^\infty |f'|^2 \varrho\,dy \cdot \sup_{y>0}$$

□

Bounds on the variation norm:

**Lemma 4.15.** *(i)*

$$\|\nu - \mu\|_{TV}^2 \leq \frac{1}{4}\chi^2(\nu|\mu)$$

*(ii) Pinsker's inequality:*

$$\|\nu - \mu\|_{TV}^2 \leq \frac{1}{2}H(\nu|\mu) \quad \forall\, \mu, \nu \in M_1(S)$$

*Proof.* If $\nu \not\ll \mu$, then $H(\nu|\mu) = \chi^2(\nu|\mu) = \infty$.
Now let $\nu \ll \mu$:

(i)

$$\|\nu - \mu\|_{\mathrm{TV}} = \frac{1}{2}\|\varrho - 1\|_{L^1(\mu)} \leq \frac{1}{2}\|\varrho - 1\|_{L^2(\mu)} = \frac{1}{2}\chi^2(\nu|\mu)^{\frac{1}{2}}$$

(ii) We have the inequality

$$3(x-1)^2 \leq (4+2x)(x\log x - x + 1) \quad \forall\, x \geq 0$$

and hence

$$\sqrt{3}|x-1| \leq (4+2x)^{\frac{1}{2}}(x\log x - x + 1)^{\frac{1}{2}}$$

and with the Cauchy Schwarz inequality

$$\begin{aligned}\sqrt{3}\int |\varrho - 1|\, d\mu &\leq \left(\int (4+2\varrho)\, d\mu\right)^{\frac{1}{2}} \left(\int (\varrho\log\varrho - \varrho + 1)\, d\mu\right)^{\frac{1}{2}} \\ &= \sqrt{6}\cdot H(\nu|\mu)^{\frac{1}{2}}\end{aligned}$$

□

**Remark .** *If $S$ is finite and $\mu(x) > 0$ for all $x \in S$ then conversely*

$$\begin{aligned}\chi^2(\nu|\mu) = \sum_{x\in S}\left(\frac{\nu(x)}{\mu(x)} - 1\right)^2 \mu(x) &\leq \frac{\left(\sum_{x\in S}\left|\frac{\nu(x)}{\mu(x)} - 1\right|\mu(x)\right)^2}{\min_{x\in S}\mu(x)} \\ &= \frac{4\|\nu-\mu\|_{TV}^2}{\min\mu}\end{aligned}$$

**Corollary 4.16.** *(i) If the Poincaré inequality*

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\lambda}\mathscr{E}(f,f) \quad \forall\, f \in \mathscr{A}$$

*holds then*

$$\|\nu p_t - \mu\|_{TV} \leq \frac{1}{2}e^{-\lambda t}\chi^2(\nu|\mu)^{\frac{1}{2}} \tag{4.6}$$

*(ii) In particular, if $S$ is finite then*

$$\|\nu p_t - \mu\|_{TV} \leq \frac{1}{\min_{x\in S}\mu(x)^{\frac{1}{2}}}e^{-\lambda t}\|\nu - \mu\|_{TV}$$

*where $\|\nu - \mu\|_{TV} \leq 1$. This leads to a bound for the **Dobrushin coefficient** (contraction coefficient with respect to $\|\cdot\|_{TV}$).*

*Proof.*

$$\|\nu p_t - \mu\|_{\text{TV}} \leq \frac{1}{2}\chi^2(\nu p_t|\mu)^{\frac{1}{2}} \leq \frac{1}{2}e^{-\lambda t}\chi^2(\nu|\mu)^{\frac{1}{2}} \leq \frac{2}{2}\frac{1}{\min \mu^{\frac{1}{2}}}e^{-\lambda t}\|\nu - \mu\|_{\text{TV}}$$

if $S$ is finite. □

***Consequence:*** Total variation mixing time: $\varepsilon \in (0,1)$,

$$\begin{aligned} T_{\text{mix}}(\varepsilon) &= \inf\{t \geq 0 \ : \ \|\nu p_t - \mu\|_{\text{TV}} \leq \varepsilon \text{ for all } \nu \in M_1(S)\} \\ &\leq \frac{1}{\lambda}\log\frac{1}{\varepsilon} + \frac{1}{2\lambda}\log\frac{1}{\min \mu(x)} \end{aligned}$$

where the first summand is the ***$L^2$ relaxation time*** and the second is called ***burn-in period***, i.e. the time needed to make up for a bad initial distribution.

**Remark .** *On high or infinite-dimensional state spaces the bound (4.6) is often problematic since $\chi^2(\nu|\mu)$ can be very large (whereas $\|\nu - \mu\|_{TV} \leq 1$). For example for product measures,*

$$\chi^2\left(\nu^n|\mu^n\right) = \int \left(\frac{d\nu^n}{d\mu^n}\right)^2 d\mu^n - 1 = \left(\int \left(\frac{d\nu}{d\mu}\right)^2 d\mu\right)^n - 1$$

*where $\int \left(\frac{d\nu}{d\mu}\right)^2 d\mu > 1$ grows exponentially in n.*

Are there improved estimates?

$$\int p_t f \, d\nu - \int f \, d\mu = \int p_t f \, d(\nu - \mu) \leq \|p_t f\|_{\sup} \cdot \|\nu - \mu\|_{\text{TV}}$$

Analysis: From the Sobolev inequality follows

$$\|p_t f\|_{\sup} \leq c \cdot \|f\|_{L^p}$$

However, Sobolev constants are dimension dependent! This leads to a replacement by the log Sobolev inequality.

## 4.4 Hypercontractivity

Additional reference for this chapter:

- Gross [11]
- Deuschel, Stroock [8]

- Ané [2]
- Royer [19]

We consider the setup from section 4.3. In addition, we now assume that $(\mathscr{L}, \mathscr{A})$ is symmetric on $L^2(S, \mu)$.

**Theorem 4.17.** *With assumptions (A0)-(A3) and $\alpha > 0$, the following statements are equivalent:*

*(i)* ***Logarithmic Sobolev inequality (LSI)***

$$\int_S f^2 \log \frac{f^2}{\|f\|^2_{L^2(\mu)}} \, d\mu \leq \alpha \mathscr{E}(f,f) \quad \forall\, f \in \mathscr{A}$$

*(ii)* ***Hypercontractivity*** *For $1 \leq p < q < \infty$,*

$$\|p_t f\|_{L^q(\mu)} \leq \|f\|_{L^p(\mu)} \quad \forall\, f \in L^p(\mu),\ t \geq \frac{\alpha}{4} \log \frac{q-1}{p-1}$$

*(iii) Assertion (ii) holds for $p = 2$.*

**Remark .** *Hypercontractivity and Spectral gap implies*

$$\|p_t f\|_{L^q(\mu)} = \|p_{t_0} p_{t-t_0} f\|_{L^q(\mu)} \leq \|p_{t-t_0} f\|_{L^2(\mu)} \leq e^{-\lambda(t-t_0)} \|f\|_{L^2(\mu)}$$

*for all $t \geq t_0(q) := \frac{\alpha}{4} \log(q-1)$.*

*Proof.* (i)⇒(ii) ***Idea:*** WLOG $f \in \mathscr{A}_0,\ f \geq \delta > 0$ (which implies that $p_t f \geq \delta\ \forall\, t \geq 0$).

Compute

$$\frac{d}{dt} \|p_t f\|_{L^{q(t)}(\mu)}, \quad q\colon \mathbb{R}^+ \to (1, \infty) \text{ smooth:}$$

1. Kolmogorov:

$$\frac{d}{dt} p_t f = \mathscr{L} p_t f \quad \text{derivation with respect to sup-norm}$$

implies that

$$\frac{d}{dt} \int (p_t f)^{q(t)} \, d\mu = q(t) \int (p_t f)^{q(t)-1} \mathscr{L} p_t f \, d\mu + q'(t) \int (p_t f)^{q(t)} \log p_t f \, d\mu$$

where

$$\int (p_t f)^{q(t)-1} \mathscr{L} p_t f \, d\mu = -\mathscr{E}\left((p_t f)^{q(t)-1}, p_t f\right)$$

2. Stroock estimate:

$$\mathscr{E}\left(f^{q-1}, f\right) \geq \frac{4(q-1)}{q^2} \mathscr{E}\left(f^{\frac{q}{2}}, f^{\frac{q}{2}}\right)$$

*Proof.*

$$\begin{aligned}
\mathscr{E}(f^{q-1}, f) &= -\left(f^{q-1}, \mathscr{L} f\right)_\mu = \lim_{t \downarrow 0} \frac{1}{t}\left(f^{q-1}, f - p_t f\right)_\mu \\
&= \lim_{t \downarrow 0} \frac{1}{2t} \iint \left(f^{q-1}(y) - f^{q-1}(x)\right)\left(f(y) - f(x)\right) p_t(x, dy)\, \mu(dx) \\
&\geq \frac{4(q-1)}{q^2} \lim_{t \downarrow 0} \frac{1}{2t} \iint \left(f^{\frac{q}{2}}(y) - f\frac{q}{2}(x)\right)^2 p_t(x, dy)\, \mu(dx) \\
&= \frac{4(q-1)}{q^2} \mathscr{E}\left(f^{\frac{q}{2}}, f^{\frac{q}{2}}\right)
\end{aligned}$$

where we have used that

$$\left(a^{\frac{q}{2}} - b\frac{q}{2}\right)^2 \leq \frac{q^2}{4(q-1)}\left(a^{q-1} - b^{q-1}\right)(a - b) \quad \forall\, a, b > 0,\ q \geq 1$$

□

**Remark .** – *The estimate justifies the use of functional inequalities with respect to $\mathscr{E}$ to bound $L^p$ norms.*

– *For generators of diffusions, equality holds, e.g.:*

$$\int \nabla f^{q-1} \nabla f \, d\mu = \frac{4(q-1)}{q^2} \int \left|\nabla f^{\frac{q}{2}}\right|^2 d\mu$$

*by the chain rule.*

3. Combining the estimates:

$$q(t) \cdot \|p_t f\|_{q(t)}^{q(t)-1} \frac{d}{dt}\|p_t f\|_{q(t)} = \frac{d}{dt}\int (p_t f)^{q(t)}\, d\mu - q'(t)\int (p_t f)^{q(t)} \log \|p_t f\|_{q(t)}\, d\mu$$

where

$$\int (p_t f)^{q(t)}\, d\mu = \|p_t f\|_{q(t)}^{q(t)}$$

This leads to the estimate

$$\begin{aligned}
&q(t) \cdot \|p_t f\|_{q(t)}^{q(t)-1} \frac{d}{dt}\|p_t f\|_{q(t)} \\
\leq &- \frac{4(q(t)-1)}{q(t)} \mathscr{E}\left((p_t f)^{\frac{q(t)}{2}}, (p_t f)^{\frac{q(t)}{2}}\right) + \frac{q'(t)}{q(t)} \cdot \int (p_t f)^{q(t)} \log \frac{(p_t f)^{q(t)}}{\int (p_t f)^{q(t)}\, d\mu}\, d\mu
\end{aligned}$$

4. Applying the logarithmic Sobolev inequality: Fix $p \in (1, \infty)$. Choose $q(t)$ such that

$$\alpha q'(t) = 4(q(t) - 1), \quad q(0) = p$$

i.e.

$$q(t) = 1 + (p-1)e^{\frac{4t}{\alpha}}$$

Then by the logarithmic Sobolev inequality, the right hand side in the estimate above is negative, and hence $\|p_t f\|_{q(t)}$ is decreasing. Thus

$$\|p_t f\|_{q(t)} \leq \|f\|_{q(0)} = \|f\|_p \quad \forall\, t \geq 0.$$

Other implication: Exercise. (Hint: consider $\frac{d}{dt}\|p_t f\|_{L^{q(t)}(\mu)}$). □

**Theorem 4.18** (ROTHAUS)**.** *A logarithmic Sobolev inequality with constant $\alpha$ implies a Poincaré inequality with constant $\alpha = \frac{2}{\alpha}$.*

*Proof.* $f \in L^2(\mu)$, $\int g\, d\mu = 0$, $f := 1 + \varepsilon g$, $f^2 = 1 + 2\varepsilon g + \varepsilon^2 g^2$,

$$\int f^2\, d\mu = 1 + \varepsilon^2 \int g^2\, d\mu, \qquad \mathscr{E}(f,f) = \mathscr{E}(1,1) + 2\mathscr{E}(1,g) + \varepsilon^2 \mathscr{E}(g,g)$$

and the Logarithmic Sobolev Inequality implies

$$\int (1+\varepsilon)^2 \log(1+\varepsilon g)^2\, d\mu \leq \alpha \mathscr{E}(f,f) + \int f^2\, d\mu \log \int f^2\, d\mu$$
$$\int f^2 \log f^2\, d\mu \leq \alpha \mathscr{E}(f,f) + \int f^2\, d\mu \log \int f^2\, d\mu \quad \forall\, \varepsilon > 0$$

where $f^2 \log f^2 = 2\varepsilon g + \varepsilon^2 g^2 + \frac{1}{2}(2\varepsilon g)^2 + O(\varepsilon^3)$ and $\int f^2\, d\mu \log \int f^2\, d\mu = \varepsilon^2 \int g^2\, d\mu + O(\varepsilon^3)$.

$$x \log x = x - 1 + \frac{1}{2}(x-1)^2 + O\left(|x-1|^3\right)$$

which implies that

$$2\varepsilon^2 \int g^2\, d\mu + O(\varepsilon^3) \leq \alpha \varepsilon^2 \mathscr{E}(g,g) \quad \forall\, \varepsilon > 0$$
$$\Rightarrow 2 \int g^2\, d\mu \leq \alpha \mathscr{E}(g,g)$$

□

**Application to convergence to equilibrium:**

**Theorem 4.19** (Exponential decay of relative entropy)**.**

1. *$H(\nu p_t|\mu) \leq H(\nu|\mu)$ for all $t \geq 0$ and $\nu \in M_1(S)$.*
2. *If a logarithmic Sobolev inequality with constant $\alpha > 0$ holds then*

$$H(\nu p_t|\mu) \leq e^{-\frac{2}{\alpha}t} H(\nu|\mu)$$

*Proof for gradient diffusions.* $\mathscr{L} = \frac{1}{2}\Delta + b\nabla, b = \frac{1}{2}\nabla \log \varrho \in C(\mathbb{R}^n), \mu = \varrho\, dx$ probability measure, $\mathscr{A}_0 = \operatorname{span}\{C_0^\infty(\mathbb{R}^n), 1\}$
. The Logarithmic Sobolev Inequality implies that

$$\int f^2 \log \frac{f^2}{\|f\|^2_{L^2(\mu)}}\, d\mu \le \frac{\alpha}{2} \int |\nabla f|^2\, d\mu = \alpha \mathscr{E}(f,f)$$

(i) Suppose $\nu = g \cdot \mu,\ 0 < \varepsilon \le g \le \frac{1}{\varepsilon}$ for some $\varepsilon > 0$. Hence $\nu p_t \ll \mu$ with density $p_t g,\ \varepsilon \le p_t g \le \frac{1}{\varepsilon}$ (since $\int f\, d(\nu p_t) = \int p_t f\, d\nu = \int p_t f g d\mu = \int f p_t g\, d\mu$ by symmetry). This implies that

$$\frac{d}{dt} H(\nu p_t|\mu) = \frac{d}{dt} \int p_t g \log p_t g\, d\mu = \int \mathscr{L} p_t g (1 + \log p_t g)\, d\mu$$

by Kolmogorov. Using the fact that $(x \log x)' = 1 + \log x$ we get

$$\frac{d}{dt} H(\nu p_t|\mu) = -\mathscr{E}(p_t g, \log p_t g) = -\frac{1}{2} \int \nabla p_t g \cdot \nabla \log p_t g\, d\mu$$

where $\nabla \log p_t g = \frac{\nabla p_t g}{p_t g}$. Hence

$$\frac{d}{dt} H(\nu p_t|\mu) = -2 \int |\nabla \sqrt{p_t g}|^2\, d\mu \tag{4.7}$$

1. $-2 \int \left|\nabla \sqrt{p_t g}\right|^2\, d\mu \le 0$
2. The Logarithmic Sobolev Inequality yields that

$$-2 \int |\nabla \sqrt{p_t g}|^2\, d\mu \le -\frac{4}{\alpha} \int p_t g \log \frac{p_t g}{\int p_t g\, d\mu}\, d\mu$$

where $\int p_t g\, d\mu = \int g\, d\mu = 1$ and hence

$$-2 \int |\nabla \sqrt{p_t g}|^2\, d\mu \le -\frac{4}{\alpha} H(\nu p_t|\mu)$$

(ii) Now for a general $\nu$. If $\nu \not\ll \mu,\ H(\nu|\mu) = \infty$ and we have the assertion. Let $\nu = g \cdot \mu,\ g \in L^1(\mu)$ and

$$g_{a,b} := (g \vee a) \wedge b, \quad 0 < a < b,$$
$$\nu_{a,b} := g_{a,b} \cdot \mu.$$

Then by (i),

$$H(\nu_{a,b} p_t|\mu) \le e^{-\frac{2t}{\alpha}} H(\nu_{a,b}|\mu)$$

The claim now follows for $a \downarrow 0$ and $b \uparrow \infty$ by dominated and monotone convergence.

□

**Remark .** *1. The proof in the general case is analogous, just replace (4.7) by inequality*

$$4\mathscr{E}(\sqrt{f},\sqrt{f}) \leq \mathscr{E}(f,\log f)$$

*2. An advantage of the entropy over the $\chi^2$ distance is the good behavior in high dimensions. E.g. for product measures,*

$$H(\nu^d|\mu^d) = d \cdot H(\nu|\mu)$$

*grows only linearly in dimension.*

**Corollary 4.20** (Total variation bound)**.** *For all $t \geq 0$ and $\nu \in M_1(S)$,*

$$\begin{aligned} \|\nu p_t - \mu\|_{TV} &\leq \frac{1}{\sqrt{2}} e^{-\frac{t}{\alpha}} H(\nu|\mu)^{\frac{1}{2}} \\ \Big( &\leq \frac{1}{\sqrt{2}} \log \frac{1}{\min \mu(x)} e^{-\frac{t}{\alpha}} \quad \textit{if } S \textit{ is finite} \Big) \end{aligned}$$

*Proof.*

$$\|\nu p_t - \mu\|_{\text{TV}} \leq \frac{1}{\sqrt{2}} H(\nu p_t|\mu)^{\frac{1}{2}} \leq \frac{1}{\sqrt{2}} e^{-\frac{t}{\alpha}} H(\nu|\mu)^{\frac{1}{2}}$$

where we use Pinsker's Theorem for the first inequality and Theorem ??? for the second inequality. Since $S$ is finite,

$$H(\delta_x|\mu) = \log \frac{1}{\mu(x)} \leq \log \frac{1}{\min \mu} \quad \forall x \in S$$

which leads to

$$H(\nu|\mu) \leq \sum \nu(x) H(\delta_x|\mu) \leq \log \frac{1}{\min \mu} \quad \forall \nu$$

since $\nu = \sum \nu(x)\delta_x$ is a convex combination. □

**Consequence for mixing time:** (S finite)

$$\begin{aligned} T_{\text{mix}}(\varepsilon) &= \inf \{ t \geq 0 \ : \ \|\nu p_t - \mu\|_{\text{TV}} \leq \varepsilon \text{ for all } \nu \in M_1(S) \} \\ &\leq \alpha \cdot \log \frac{1}{\sqrt{2}\varepsilon} + \log \log \frac{1}{\min_{x \in S} \mu(x)} \end{aligned}$$

Hence we have $\log \log$ instead of $\log$ !

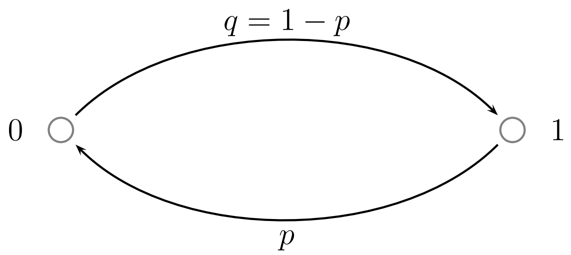
## 4.5 Logarithmic Sobolev inequalities: Examples and techniques

**Example .** ***Two-point space.*** $S = \{0,1\}$. *Consider a Markov chain with generator*

$$\mathscr{L} = \begin{pmatrix} -q & q \\ p & -p \end{pmatrix}, \qquad\qquad p,q \in (0,1),\ p+q=1$$

*which is symmetric with respect to the Bernoulli measure,*

$$\mu(0) = p, \quad \mu(1) = q$$




*Dirichlet form:*

$$\begin{aligned}\mathscr{E}(f,f) &= \frac{1}{2}\sum_{x,y}\left(f(y)-f(x)\right)^2 \mu(x)\mathscr{L}(x,y) \\ &= pq\cdot|f(1)-f(0)|^2 = \mathrm{Var}_\mu(f)\end{aligned}$$

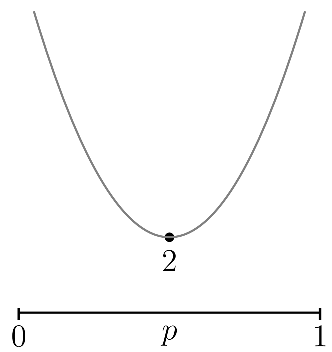
*Spectral gap:*

$$\lambda(p) = \inf_{f \textit{ not const.}} \frac{\mathscr{E}(f,f)}{\mathrm{Var}_\mu(f)} = 1 \quad \textit{independent of } p!$$

*Optimal Log Sobolev constant:*

$$\alpha(p) = \sup_{\substack{f\perp 1 \\ \int f^2\,d\mu = 1}} \frac{\int f^2 \log f^2\,d\mu}{\mathscr{E}(f,f)} = \begin{cases} 2 & \textit{if } p = \frac{1}{2} \\ \frac{\log q - \log p}{q-p} & \textit{else} \end{cases}$$

*goes to infinity as $p \downarrow 0$ or $p \uparrow \infty$!*



**Spectral gap and Logarithmic Sobolev Inequality for product measures:**

$$\mathrm{Ent}_\mu(f) := \int f \log f \, d\mu, \ f > 0$$

**Theorem 4.21** (Factorization property)**.** *$(S_i, \mathscr{S}_i, \mu_i)$ probability spaces, $\mu = \otimes_{i=1}^n \mu_i$. Then*

*1.*

$$\mathrm{Var}_\mu(f) \leq \sum_{i=1}^n \mathbb{E}_\mu \left[\mathrm{Var}_{\mu_i}^{(i)}(f)\right]$$

*where on the right hand side the variance is taken with respect to the i-th variable.*

*2.*

$$\mathrm{Ent}_\mu(f) \leq \sum_{i=1}^n \mathbb{E}_\mu \left[\mathrm{Ent}_{\mu_i}^{(i)}(f)\right]$$

*Proof.* 1. Exercise.

2.

$$\mathrm{Ent}_\mu(f) = \sup_{g \,:\, \mathbb{E}_\mu[e^g]=1} \mathbb{E}_\mu[fg], \quad \text{cf. above}$$

Fix $g\colon S^n \to \mathbb{R}$ such that $\mathbb{E}_\mu\left[e^g\right] = 1$. Decompose:

$$\begin{aligned} g(x_1,\dots,x_n) &= \log e^{g(x_1,\dots,x_n)} \\ &= \log \frac{e^{g(x_1,\dots,x_n)}}{\int e^{g(y_1,x_2,\dots,x_n)}\,\mu_1(dy_1)} + \log \frac{\int e^{g(y_1,x_2,\dots,x_n)}\,\mu_1(dy_1)}{\iint e^{g(y_1,y_2,x_3,\dots,x_n)}\,\mu_1(dy_1)\mu_2(dy_2)} + \cdots \\ &=: \sum_{i=1}^n g_i(x_1,\dots,x_n) \end{aligned}$$

and hence

$$\begin{aligned} &\mathbb{E}_{\mu_i}^i\left[e^{g_i}\right] = 1 \quad \forall, 1 \leq i \leq n \\ \Rightarrow \quad &\mathbb{E}_\mu[fg] = \sum_{i=1}^n \mathbb{E}_\mu\left[fg_i\right] = \sum_{i=1}^n \mathbb{E}_\mu\left[\mathbb{E}_{\mu_i}^{(i)}\left[fg_i\right]\right] \leq \mathrm{Ent}_{\mu_i}^{(i)}(f) \\ \Rightarrow \quad &\mathrm{Ent}_\mu[f] = \sup_{\mathbb{E}_\mu[e^g]=1} E_\mu[fg] \leq \sum_{i=1}^n \mathbb{E}_\mu\left[\mathrm{Ent}_{\mu_i}^{(i)}(f)\right] \end{aligned}$$

□

**Corollary 4.22.** *1. If the Poincaré inequalities*

$$\mathrm{Var}_{\mu_i}(f) \le \frac{1}{\lambda_i}\mathscr{E}_i(f,f) \quad \forall\, f \in \mathscr{A}_i$$

*hold for each $\mu_i$ then*

$$\mathrm{Var}_{\mu}(f) \le \frac{1}{\lambda}\mathscr{E}(f,f) \quad \forall\, f \in \bigotimes_{i=1}^{n} \mathscr{A}_i$$

*where*

$$\mathscr{E}(f,f) = \sum_{i=1}^{n} \mathbb{E}_\mu \left[\mathscr{E}_i^{(i)}(f,f)\right]$$

*and*

$$\lambda = \min_{1\le i\le n} \lambda_i$$

*2. The corresponding assertion holds for Logarithmic Sobolev Inequalities with $\alpha = \max \alpha_i$*

*Proof.*

$$\mathrm{Var}_{\mu}(f) \le \sum_{i=1}^{n} \mathbb{E}_\mu \left[\mathrm{Var}_{\mu_i}^{(i)}(f)\right] \le \frac{1}{\min \lambda_i}\mathscr{E}(f,f)$$

since

$$\mathrm{Var}_{\mu_i}^{(i)}(f) \le \frac{1}{\lambda_i}\mathscr{E}_i(f,f)$$

□

**Example .** $S = \{0,1\}^n$, $\mu^n$ *product of* $\mathrm{Bernoulli}(p)$,

$$\begin{aligned}&\mathrm{Ent}_{\mu^n}(f)\\ &\le \alpha(p)\cdot p\cdot q\cdot \sum_{i=1}^{n} \int |f(x_1,\dots,x_{i-1},1,x_{i+1},\dots,x_n) - f(x_1,\dots,x_{i-1},0,x_{i+1},\dots,x_n)|^2 \; \mu^n(dx_1,\dots,dx_n)\end{aligned}$$

*independent of* $n$.

**Example .** *Standard normal distribution* $\gamma = N(0,1)$,

$$\varphi_n\colon \{0,1\}^n \to \mathbb{R}, \qquad \varphi_n(x) = \frac{\sum_{i=1}^{n}\left(x_i - \frac{1}{2}\right)}{\sqrt{\frac{n}{4}}}$$

*The Central Limit Theorem yields that* $\mu = \mathrm{Bernoulli}(\frac{1}{2})$ *and hence*

$$\mu^n \circ \varphi_n^{-1} \xrightarrow{w} \gamma$$

*Hence for all $f \in C_0^\infty(\mathbb{R})$,*

$$\begin{aligned}\mathrm{Var}_\gamma(f) &= \lim_{n\to\infty} \mathrm{Var}_{\mu^n}(f\circ\varphi_n)\\ &\le \liminf \frac{1}{2}\sum_{i=1}^n \int |\Delta_i f\circ\varphi_n|^2 \, d\mu^n\\ &\le \cdots \le 2\cdot \int |f'|^2\, d\gamma\end{aligned}$$

*Central Limit Theorem with constant $\alpha = 2$.*
*Similarly: Poincaré inequality with $\lambda = 1$ (Exercise).*

## Central Limit Theorem with respect to log concave probability measures:

Stochastic gradient flow in $\mathbb{R}^n$:

$$dX_t = dB_t - (\nabla H)(X_t)\,dt, \quad H \in C^2(\mathbb{R}^n)$$

Generator:

$$\mathscr{L} = \frac{1}{2}\Delta - \nabla H\cdot\nabla$$
$$\mu(dx) = e^{-H(x)}\,dx \text{ satisfies } \mathscr{L}^*\mu = 0$$

***Assumption:*** There exists a $\kappa > 0$ such that

$$\partial^2 H(x) \ge \kappa\cdot I \quad \forall\, x\in\mathbb{R}^n$$
$$\text{i.e.} \quad \partial^2_{\xi\xi} H \ge \kappa\cdot|\xi|^2 \quad \forall\,\xi\in\mathbb{R}^n$$

**Remark .** *The assumption implies the inequalities*

$$x\cdot\nabla H(x) \ge \kappa\cdot|x|^2 - c, \tag{4.8}$$
$$H(x) \ge \frac{\kappa}{2}|x|^2 - \tilde{c} \tag{4.9}$$

*with constants $c, \tilde{c} \in \mathbb{R}$. By (4.8) and a Lyapunov argument it can be shown that $X_t$ does not explode in finite time and that $p_t(\mathscr{A}_0) \subseteq \mathscr{A}$ where $\mathscr{A}_0 = \mathrm{span}\,(\mathbb{C}_0^\infty(\mathbb{R}^n), 1)$, $\mathscr{A} = \mathrm{span}\,(\mathscr{S}(\mathbb{R}^n), 1)$. By (4.9), the measure $\mu$ is finite, hence by our results above, the normalized measure is a stationary distribution for $p_t$.*

**Lemma 4.23.** *If* $\mathrm{Hess}\, H \ge \kappa I$ *then*

$$|\nabla p_t f| \le e^{-\kappa t} p_t\,|\nabla f| \quad f\in C_b^1(\mathbb{R}^n)$$

**Remark .** *1. Actually, both statements are equivalent.*

*2. If we replace $\mathbb{R}^n$ by an arbitrary Riemannian manifold the same assertion holds under the assumption*

$$\mathrm{Ric} + \mathrm{Hess}\, H \geq \kappa \cdot I$$

*(Bochner-Lichnerowicz-Weitzenböck).*

*Informal analytic proof:*

$$\begin{aligned}\nabla \mathscr{L} f &= \nabla \left(\frac{1}{2}\Delta - \nabla H \cdot \nabla\right) f \\ &= \left(\frac{1}{2}\Delta - \nabla H \cdot \nabla - \partial^2 H\right) \nabla f \\ &=: \vec{\mathscr{L}} \text{ operator on one-forms (vector fields)}\end{aligned}$$

This yields to the Evolution equation for $\nabla p_t f$:

$$\frac{\partial}{\partial t}\nabla p_t f = \nabla \frac{\partial}{\partial t} p_t f = \nabla \mathscr{L} p_t f = \vec{\mathscr{L}}\, \nabla p_t f$$

and hence

$$\begin{aligned}\frac{\partial}{\partial t}\left|\nabla p_t f\right| &= \frac{\partial}{\partial t}\left(\nabla p_t f \cdot \nabla p_t f\right)^{\frac{1}{2}} = \frac{\left(\frac{\partial}{\partial t}\nabla p_t f\right) \cdot \nabla p_t f}{\left|\nabla p_t f\right|} \\ &= \frac{\left(\vec{\mathscr{L}}\, \nabla p_t f\right) \cdot \nabla p_t f}{\left|\nabla p_t f\right|} \leq \frac{\mathscr{L} \nabla p_t f \cdot \nabla p_t f}{\left|\nabla p_t f\right|} - \kappa \cdot \frac{\left|\nabla p_t f\right|^2}{\left|\nabla p_t f\right|} \\ &\leq \cdots \leq \mathscr{L}\left|\nabla p_t f\right| - \kappa \left|\nabla p_t f\right|\end{aligned}$$

We get that $v(t) := e^{\kappa t} p_{s-t}\left|\nabla p_t f\right|$ with $0 \leq t \leq s$ satisfies

$$v'(t) \leq \kappa v(t) - p_{s-t}\mathscr{L}\left|\nabla p_t f\right| + p_{s-t}\mathscr{L}\left|\nabla p_t f\right| - \kappa p_{s-t}\left|\nabla p_t f\right| = 0$$

and hence

$$e^{\kappa s}\left|\nabla p_s f\right| = v(s) \leq v(0) = p_s\left|\nabla f\right|$$

□

- The proof can be made rigorous by approximating $|\cdot|$ by a smooth function, and using regularity results for $p_t$, cf. e.g. Deuschel, Stroock[8].
- The assertion extends to general diffusion operators.

*Probabilistic proof:* $p_t f(x) = \mathbb{E}[f(X_t^x)]$ where $X_t^x$ is the solution flow of the stochastic differential equation

$$dX_t = dB_t - (\nabla H)(X_t)\,dt, \quad \text{i.e.,}$$
$$X_t^x = x + B_t - \int_0^t (\nabla H)(X_s^x)\,ds$$

By the assumption on $H$ one can show that $x \to X_t^x$ is smooth and the derivative flow $Y_t^x = \nabla_x X_t$ satisfies the differentiated stochastic differential equation

$$dY_t^x = -(\partial^2 H)(X_t^x)Y_t^x\,dt,$$
$$Y_0^x = I$$

which is an ordinary differential equation. Hence if $\partial^2 H \geq \kappa I$ then for $v \in \mathbb{R}^n$,

$$\frac{d}{dt}|Y_t \cdot v|^2 = -2\left(Y_t \cdot v,\ (\partial^2 H)(X_t)Y_t \cdot v\right)_{\mathbb{R}^n} \leq \kappa \cdot |Y_t \cdot v|^2$$

where $Y_t \cdot v$ is the derivative of the flow in direction $v$. Hence

$$|Y_t \cdot v|^2 \leq e^{-2\kappa t}|v|$$
$$\Rightarrow \quad |Y_t \cdot v| \leq e^{-\kappa t}|v|$$

This implies that for $f \in C_b^1(\mathbb{R}^n)$, $p_t f$ is differentiable and

$$v \cdot \nabla p_t f(x) = \mathbb{E}\left[(\nabla f(X_t^x);\ Y_t^x \cdot v)\right]$$
$$\leq \mathbb{E}\left[|\nabla f(X_t^x)|\right] \cdot e^{-\kappa t} \cdot |v| \quad \forall\, v \in \mathbb{R}^n$$

i.e.

$$|\nabla p_t f(x)| \leq e^{-\kappa t} p_t |\nabla f|(x)$$

□

**Theorem 4.24** (BAHRY-EMERY). *Suppose that*

$$\partial^2 H \geq \kappa \cdot I \quad \textit{with } \kappa > 0$$

*Then*

$$\int f^2 \log \frac{f^2}{\|f\|_{L^2(\mu)}^2}\,d\mu \leq \frac{1}{\kappa}\int |\nabla f|^2\,d\mu \quad \forall\, f \in C_0^\infty(\mathbb{R}^n)$$

**Remark .** *The inequality extends to $f \in H^{1,2}(\mu)$ where $H^{1,2}(\mu)$ is the closure of $C_0^\infty$ with respect to the norm*

$$\|f\|_{1,2} := \left(\int |f|^2 + |\nabla f|^2\,d\mu\right)^{\frac{1}{2}}$$

*Proof.* $g \in \operatorname{span}(C_0^\infty, 1),\ g \geq \delta \geq 0.$
***Aim:***

$$\int g \log g \, d\mu \leq \frac{1}{\kappa} \int |\nabla \sqrt{g}|^2 \, d\mu + \int g \, d\mu \log \int g \, d\mu$$

Then $g = f^2$ and we get the assertion.
***Idea:*** Consider

$$u(t) = \int p_t g \log p_t g \, d\mu$$

***Claim:***

(i) $u(0) = \int g \log g \, d\mu$

(ii) $\lim_{t \uparrow \infty} u(t) = \int g \, d\mu \log \int g \, d\mu$

(iii) $-u'(t) \leq 2e^{-2\kappa t} \int \left|\nabla \sqrt{g}\right|^2 \, d\mu$

By (i), (ii) and (iii) we then obtain:

$$\begin{aligned} \int g \log g \, d\mu - \int g \, d\mu \log \int g \, d\mu &= \lim_{t \to \infty} (u(0) - u(t)) \\ &= \lim_{t \to \infty} \int_0^t -u'(t) \, ds \\ &\leq \int |\nabla \sqrt{g}|^2 \, d\mu \end{aligned}$$

where $2 \int_0^\infty e^{-2\kappa s} \, ds = \frac{1}{\kappa}$.

*Proof of claim:* (i) Obvious.

(ii) Ergodicity yields to

$$p_t g(x) \to \int g \, d\mu \quad \forall\, x$$

for $t \uparrow \infty$.
In fact:

$$|\nabla p_t g| \leq e^{-\kappa t} p_t |\nabla g| \leq e^{-\kappa t} |\nabla g|$$

and hence

$$|p_t g(x) - p_t g(y)| \leq e^{-\kappa t} \sup |\nabla g| \cdot |x - y|$$

which leads to

$$\begin{aligned} \left| p_t g(x) - \int g \, d\mu \right| &= \left| \int (p_t g(x) - p_t g(y)) \ \mu(dy) \right| \\ &\leq e^{-\kappa t} \sup |\nabla g| \cdot \int |x - y| \, \mu(dy) \to 0 \end{aligned}$$

Since $p_t g \geq \delta \geq 0$, dominated convergence implies that

$$\int p_t g \log p_t \delta \, d\mu \to \int g \, d\mu \log \int g \, d\mu$$

(iii) *Key Step!* By the computation above (decay of entropy) and the lemma,

$$\begin{aligned}
-u'(t) &= +\frac{1}{2}\int \nabla p_t g \cdot \nabla \log p_t g \, d\mu = \frac{1}{2}\int \frac{|\nabla p_t g|^2}{p_t g} \, d\mu \\
&\leq \frac{1}{2} e^{-\kappa t} \int \frac{|p_t \nabla g|^2}{p_t g} \, d\mu \leq \frac{1}{2} e^{-2\kappa t} \int p_t \frac{|\nabla g|^2}{g} \, d\mu \\
&= \frac{1}{2} e^{-2\kappa t} \int \frac{|\nabla g|^2}{g} \, d\mu = 2e^{-2\kappa t} \int |\nabla \sqrt{g}|^2 \, d\mu
\end{aligned}$$

□

□

**Example .** ***An Ising model with real spin:*** *(Reference: Royer [19])*
$S = \mathbb{R}^\Lambda = \{(x_i)_{i\in\Lambda} \mid x_i \in \mathbb{R}\}$, $\Lambda \subset \mathbb{Z}^d$ *finite.*

$$\begin{aligned}
\mu(dx) &= \frac{1}{Z} \exp(-H(x)) \, dx \\
H(x) &= \sum_{i\in\Lambda} \underbrace{V(x_i)}_{\text{potential}} - \frac{1}{2} \sum_{i,j\in\Lambda} \underbrace{\vartheta(i-j)}_{\text{interactions}} x_i x_j - \sum_{i\in\Lambda, j\in\mathbb{Z}^d\setminus\Lambda} \vartheta(i-j) x_i z_j,
\end{aligned}$$

*where* $V\colon \mathbb{R} \to \mathbb{R}$ *is a non-constant polynomial, bounded from below, and* $\vartheta\colon \mathbb{Z} \to \mathbb{R}$ *is a function such that* $\vartheta(0) = 0$, $\vartheta(i) = \vartheta(-i)\ \forall\, i$, *(symmetric interactions),* $\vartheta(i) = 0\ \forall\, |i| \geq R$ *(finite range),* $z \in \mathbb{R}^{\mathbb{Z}^d\setminus\Lambda}$ *fixed boundary condition.*
***Glauber-Langevin dynamics:***

$$dX_t^i = -\frac{\partial H}{\partial x_i}(X_t) \, dt + dB_t^i, \quad i \in \Lambda \tag{4.10}$$

***Dirichletform:***

$$\mathscr{E}(f,g) = \frac{1}{2} \sum_{i\in\Lambda} \int \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_i} \, d\mu$$

**Corollary 4.25.** *If*

$$\inf_{x\in\mathbb{R}} V''(x) > \sum_{i\in\mathbb{Z}} |\vartheta(i)|$$

*then* $\mathscr{E}$ *satisfies a log Sobolev inequality with constant independent of* $\Lambda$.

*Proof.*

$$\begin{aligned}
&\frac{\partial^2 H}{\partial x_i \partial x_j}(x) = V''(x_i) \cdot \delta_{ij} - \vartheta(i-j) \\
\Rightarrow \quad &\partial^2 H \geq \left( \inf V'' - \sum_i |\vartheta(i)| \right) \cdot I
\end{aligned}$$

in the sense of ???. □

**Consequence:** There is a unique Gibbs measure on $\mathbb{Z}^d$ corresponding to $H$, cf. Royer [19]. What can be said if $V$ is not convex?

**Theorem 4.26** (Bounded perturbations). $\mu, \nu \in M_1(\mathbb{R}^n)$ *??? absolut continuous,*

$$\frac{d\nu}{d\mu}(x) = \frac{1}{Z} e^{-U(x)}.$$

*If*

$$\int f^2 \log \frac{f^2}{\|f\|^2_{L^2(\mu)}} \, d\mu \leq \alpha \cdot \int |\nabla f|^2 \, d\mu \quad \forall f \in \mathbb{C}_0^\infty$$

*then*

$$\int f^2 \log \frac{f^2}{\|f\|^2_{L^2(\nu)}} \, d\nu \leq \alpha \cdot e^{\operatorname{osc}(U)} \cdot \int |\nabla f|^2 \, d\nu \quad \forall f \in C_0^\infty$$

*where*

$$\operatorname{osc}(U) := \sup U - \inf U$$

*Proof.*

$$\int f^2 \log \frac{|f|^2}{\|f\|^2_{L^2(\nu)}} \, d\nu \leq \int \left( f^2 \log f^2 - f^2 \log \|f\|^2_{L^2(\mu)} - f^2 + \|f\|^2_{L^2(\mu)} \right) \, d\nu \tag{4.11}$$

since

$$\int f^2 \log \frac{|f|^2}{\|f\|^2_{L^2(\nu)}} \, d\nu \leq \int f^2 \log f^2 - f^2 \log t^2 - f^2 + t^2 \, d\nu \quad \forall t > 0$$

Note that in (4.11) the integrand on the right hand side is non-negative. Hence

$$\begin{aligned} \int f^2 \log \frac{|f|^2}{\|f\|^2_{L^2(\nu)}} \, d\nu &\leq \frac{1}{Z} \cdot e^{-\inf U} \int \left( f^2 \log f^2 - f^2 \log \|f\|^2_{L^2(\mu)} - f^2 + \|f\|^2_{L^2(\mu)} \right) \, d\mu \\ &= \frac{1}{Z} e^{-\inf U} \cdot \int f^2 \log \frac{f^2}{\|f\|^2_{L^2(\mu)}} \, d\mu \\ &\leq \frac{1}{Z} \cdot e^{-\inf U} \alpha \int |\nabla f|^2 \, d\mu \\ &\leq e^{\sup U - \inf U} \alpha \int |\nabla f|^2 \, d\nu \end{aligned}$$

□

**Example .** *We consider the Gibbs measures $\mu$ from the example above*

1. ***No interactions:***

$$H(x) = \sum_{i\in\Lambda}\left(\frac{x_i^2}{2} + V(x_i)\right), \qquad V\colon \mathbb{R}\to\mathbb{R} \text{ bounded}$$

*Hence*

$$\mu = \bigotimes_{i\in\Lambda}\mu_V$$

*where*

$$\mu_V(dx) \propto e^{-V(x)}\gamma(dx)$$

*and $\gamma(dx)$ is the standard normal distribution. Hence $\mu$ satisfies the logarithmic Sobolev inequality with constant*

$$\alpha(\mu) = \alpha(\mu_V) \leq e^{-\operatorname{osc}(V)}\alpha(\gamma) = 2\cdot e^{-\operatorname{osc}(V)}$$

*by the factorization property. Hence we have independence of dimension!*

2. ***Weak interactions:***

$$H(x) = \sum_{i\in\Lambda}\left(\frac{x_i^2}{2} + V(x_i)\right) - \vartheta\sum_{\substack{i,j\in\Lambda\\|i-j|=1}} x_i x_j - \vartheta\sum_{\substack{i\in\Lambda\\ j\notin\Lambda\\|i-j|=1}} x_i z_j,$$

$\vartheta\in\mathbb{R}$*. One can show:*

**Theorem 4.27.** *If $V$ is bounded then there exists $\beta > 0$ such that for $\vartheta\in[-beta,\beta]$ a logarithmic Sobolev inequality with constant independent of $\lambda$ holds.*

*The proof is based on the exponential decay of correlations $\operatorname{Cov}_\mu(x_i, x_j)$ for Gibbs measure, cf. ???, Course ???.*

3. ***Discrete Ising model:*** *One can show that for $\beta < \beta_c$ (???) a logarithmic Sobolev inequality holds on$\{-N,\dots,N\}^d$ with constant of Order $O(N^2)$ independent of the boundary conditions, whereas for $\beta > \beta_c$ and periodic boundary conditions the spectral gap, and hence the log Sobolev constant, grows exponentially in $N$, cf. [???].*

## 4.6 Concentration of measure

$(\Omega,\mathscr{A},P)$ probability space, $X_i\colon\Omega\to\mathbb{R}^d$ independent identically distributed, $\sim\mu$.
Law of large numbers:

$$\frac{1}{N}\sum_{i=1}^{N} U(X_i) \to \int U\,d\mu \qquad U\in\mathscr{L}^1(\mu)$$

Cramér:

$$P\left[\left|\frac{1}{N}\sum_{i=1}^{N}U(X_i)-\int U\,d\mu\right|\geq r\right]\leq 2\cdot e^{-NI(r)},$$
$$I(r)=\sup_{t\in\mathbb{R}}\left(tr-\log\int e^{tU}\,d\mu\right)\quad \text{LD rate function.}$$

Hence we have

- Exponential concentration around mean value provided $I(r)>0\;\forall\, r\neq 0$
-
$$P\left[\left|\frac{1}{N}\sum_{i=1}^{N}U(X_i)-\int U\,d\mu\right|\geq r\right]\leq e^{-\frac{Nr^2}{c}}\text{ provided } I(r)\geq\frac{r^2}{c}$$

  *Gaussian concentration.*

When does this hold? Extension to non independent identically distributed case? This leads to: Bounds for $\log\int e^{tU}\,d\mu$ !

**Theorem 4.28** (HERBST)**.** *If $\mu$ satisfies a logarithmic Sobolev inequality with constant $\alpha$ then for any Lipschitz function $U\in C_b^1(\mathbb{R}^d)$:*

*(i)*
$$\frac{1}{t}\log\int e^{tU}\,d\mu\leq\frac{\alpha}{4}t+\int U\,d\mu\quad\forall\, t>0\tag{4.12}$$

*where $\frac{1}{t}\log\int e^{tU}\,d\mu$ can be seen as the* free energy at inverse temperature *$t$, $\frac{\alpha}{4}$ as a* bound for entropy *and $\int U\,d\mu$ as the* average energy.

*(ii)*
$$\mu\left(U\geq\int U\,d\mu+r\right)\leq e^{-\frac{r^2}{\alpha}}$$

Gaussian concentration inequality

*In particular,*

*(iii)*
$$\int e^{\gamma|x|^2}\,d\mu<\infty\quad\forall\,\gamma<\frac{1}{\alpha}$$

**Remark .** *Statistical mechanics:*
$$F_t=t\cdot S-\langle U\rangle$$
*where $F_t$ is the* free energy, *$t$ the* inverse temperature, *$S$ the* entropy *and $\langle U\rangle$ the* potential.

*Proof.* WLOG, $0 \leq \varepsilon \leq U \leq \frac{1}{\varepsilon}$. Logarithmic Sobolev inequality applied to $f = e^{\frac{tU}{2}}$:

$$\int tUe^{tU}\, d\mu \leq \alpha \int \left(\frac{t}{2}\right)^2 |\nabla U|^2 e^{tU}\, d\mu + \int e^{tU}\, d\mu \log \int e^{tU}\, d\mu$$

For $\Lambda(t) := \log \int e^{tU}\, d\mu$ this implies

$$t\Lambda'(t) = \frac{\int tUe^{tU}\, d\mu}{\int e^{tU}\, d\mu} \leq \frac{\alpha t^2}{4} \frac{\int |\nabla U|^2 e^{tU}\, d\mu}{\int e^{tU}\, d\mu} + \Lambda(t) \leq \frac{\alpha t^2}{4} + \Lambda(t)$$

since $|\nabla U| \leq 1$. Hence

$$\frac{d}{dt}\frac{\Lambda(t)}{t} = \frac{t\Lambda'(t) - \Lambda(t)}{t^2} \leq \frac{\alpha}{4} \quad \forall\, t > 0$$

Since

$$\Lambda(t) = \Lambda(0) + t \cdot \Lambda'(0) + O(t^2) = t \int U\, d\mu + O(t^2),$$

we obtain

$$\frac{\Lambda(t)}{t} \leq \int U\, d\mu + \frac{\alpha}{4} t,$$

i.e. (i).

(ii) follows from (i) by the Markov inequality, and (iii) follows from (ii) with $U(x) = |x|$. □

**Corollary 4.29** (Concentration of empirical measures)**.** *$X_i$ independent identically distributed, $\sim \mu$. If $\mu$ satisfies a logarithmic Sobolev inequality with constant $\alpha$ then*

$$P\left[\left|\frac{1}{N}\sum_{i=1}^N U(X_i) - \mathbb{E}_\mu[U]\right| \geq r\right] \leq 2 \cdot e^{-\frac{Nr^2}{4\alpha}}$$

*for any Lipschitz function $U \in C_b^1(\mathbb{R}^d)$, $N \in \mathbb{N}$ and $r > 0$.*

*Proof.* By the factorization property, $\mu^N$ satisfies a logarithmic Sobolev inequality with constant $\alpha$ as well. Now apply the theorem to

$$\tilde{U}(x) := \frac{1}{\sqrt{N}} \sum_{i=1}^N U(x_i)$$

noting that

$$\nabla \tilde{U}(x_1, \ldots, x_n) = \frac{1}{\sqrt{N}} \begin{pmatrix} \nabla U(x_1) \\ \vdots \\ \nabla U(x_N) \end{pmatrix}$$

hence since $U$ is Lipschitz,

$$\left|\nabla\tilde{U}(x)\right| = \frac{1}{\sqrt{N}}\left(\sum_{i=1}^{N}|\nabla U(x_i)|^2\right)^{\frac{1}{2}} \leq 1$$

$\square$

# Chapter 5

## 5.1 Ergodic averages

$(X_t, P_x)$ canonical Markov process on $(\Omega, \mathcal{F})$, i.e.

$$X_t(\omega) = \omega(t)$$

$p_t$ transition semigroup,

$$\begin{aligned} \Theta_t(\omega) &= \omega(\cdot + t) \quad \text{shift operator} \\ X_t(\Theta_s(\omega)) &= X_{t+s}(\omega) \end{aligned}$$

$\mu$ stationary distribution for $p_t$. Hence $(X_t, P_\mu)$ is a stationary process, i.e.

$$P_\mu \circ \Theta_t^{-1} = P_\mu \quad \forall\, t \geq 0,$$

$(\Omega, \mathcal{F}, P_\mu, (\Theta_t)_{t\geq 0})$ is a dynamical system where $\Theta_t$ are measure preserving maps, $\Theta_{t+s} = \Theta_t \circ \Theta_s$.

**Definition 5.1.**

$$\vartheta := \left\{ A \in \mathcal{F} \;:\; \Theta_t^{-1}(A) = A\, \forall\, t \geq 0 \right\}$$

*$\sigma$-algebra of shift-invariant events. The dynamical system $(\Omega, \mathcal{F}, P_\mu, (\Theta_t)_{t\geq 0})$ is called* ***ergodic*** *if and only if*

$$P_\mu[A] \in \{0, 1\} \quad \forall\, A \in \vartheta$$

*or, equivalently,*

$$F \in \mathscr{L}^2(P_\mu), \quad F \circ \Theta_t = F \; P_\mu\text{-a.s.} \quad \Rightarrow \quad F = \text{const. } P_\mu\text{-a.s.}$$

**Theorem 5.2** (Ergodic theorem)**.**

$$\frac{1}{t} \int_0^t F(\Theta_s(\omega))\, ds \to \mathbb{E}_\mu[F \mid \vartheta](\omega)$$

*a)* $P_\mu$*-a.s. and* $L^1(P_\mu)$ *for all* $F \in L^1(P_\mu)$

*b)* *in* $L^2(P_\mu)$ *for* $F \in L^2(P_\mu)$

*In particular*

$$\frac{1}{t}\int_0^t F \circ \Theta_s \, ds \to \mathbb{E}_\mu[F] \quad P_\mu\text{-a.s. if ergodic.}$$

*Proof.* cf. e.g. Stroock[21]. □

**Remark .** *1. The Ergodic theorem implies* $P_x$*-a.s. convergence for* $\mu$*-almost every* $x$ *(since* $P_\mu = \int P_x \, \mu(dx)$*).*

*2. In general* $P_x$*-a.s. convergence for fixed* $x$ *does not hold!*

**Example .** *Ising model with Glauber dynamics on* $\mathbb{Z}^2$*,* $\beta > \beta_{crit}$ *(low temperature regime). It follows that there exist two extremal stationary distributions* $\mu_\beta^+$ *and* $\mu_\beta^-$*.* $P_{\mu_\beta^+}$ *and* $P_{\mu_\beta^-}$ *are both ergodic. Hence*

$$\frac{1}{t}\int_0^t F \circ \Theta_s \, ds \to \begin{cases} \mathbb{E}_{\mu_\beta^+}[F] & P_{\mu_\beta^+}\text{-a.s.} \\ \mathbb{E}_{\mu_\beta^-}[F] & P_{\mu_\beta^-}\text{-a.s.} \end{cases}$$

*No assertion for the initial distribution* $\nu \perp \mu_\beta^+, \mu_\beta^-$*.*

When are stationary Markov processes ergodic?
Let $(L, \mathrm{Dom}(L))$ denote the generator of $(p_t)_{t\geq 0}$ on $L^2(\mu)$.

**Theorem 5.3.** *The following assertions are equivalent:*

*(i)* $P_\mu$ *is ergodic*

*(ii)* $\ker L = \mathrm{span}\{1\}$*, i.e.*

$$h \in \mathscr{L}^2(\mu) \text{harmonic} \quad \Rightarrow \quad h = \text{const. } \mu\text{-a.s.}$$

*(iii)* $p_t$ *is* $\mu$*-irreducible, i.e.*

$$B \in \mathscr{S} \text{ such that } p_t I_B = I_B \quad \mu\text{-a.s. } \forall\, t \geq 0 \quad \Rightarrow \quad \mu(B) \in \{0, 1\}$$

*If reversibility holds then (i)-(iii) are also equivalent to:*

*(iv)* $p_t$ *is* $L^2(\mu)$*-ergodic, i.e.*

$$\left\| p_t f - \int f \, d\mu \right\|_{L^2(\mu)} \to 0 \quad \forall\, f \in L^2(\mu)$$

*Proof.* (i)⇒(ii) If $h$ is harmonic then $h(X_t)$ is a martingale. Hence when we apply the $L^2$ martingale convergence theorem,

$$h(X_t) \to M_\infty \text{ in } L^2(P_\mu), \quad M_\infty \circ \Theta_t = M_\infty$$

and since ergodicity holds

$$M_\infty = \text{const.} \quad P_\mu\text{-a.s.}$$

hence

$$h(X_0) = \mathbb{E}_\mu[M_\infty \mid \mathcal{F}_0] = \text{const.} \quad P_\mu\text{-a.s.}$$

and we get that $h$ =const. $\mu$-a.s.

(ii)⇒(iii) $h = I_B$

(iii)⇒(i) If $A \in \vartheta$ then $I_A$ is shift-invariant. Hence $h(x) = \mathbb{E}_x[I_A]$ is harmonic since, applying the Markov property,

$$p_t h(x) = \mathbb{E}_x\left[\mathbb{E}_{X_t}[I_A]\right] = \mathbb{E}_x\left[I_A \circ \Theta_t\right] = \mathbb{E}_x[I_A] = h(x).$$

Applying the Markov property once again gives

$$h(X_t) = \mathbb{E}_{X_t}[I_A] = \mathbb{E}_\mu\left[I_A \circ \Theta_t \mid \mathcal{F}_t\right] \to I_A \quad P_\mu\text{-a.s.}$$

if $t \uparrow \infty$. Hence, applying the stationarity,

$$\begin{aligned} & \mu \circ h^{-1} = P_\mu \circ (h(X_t))^{-1} \to P_\mu \circ I_A^{-1} \\ \Rightarrow \quad & h \in \{0,1\} \quad \mu\text{-a.s.} \\ \Rightarrow \quad & \exists\, B \in \mathscr{S} \ : \ h = I_B \quad \mu\text{-a.s.}, \quad p_t I_B = I_B \quad \mu\text{-a.s.} \end{aligned}$$

and irreducibility gives

$$h = I_B = \text{const.} \quad \mu\text{-a.s.}$$

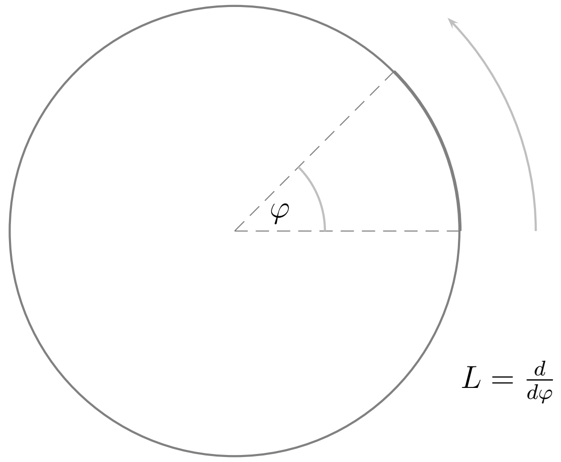
(iii)⇔(iv) If reversibility holds, the assertion follows from the spectral theorem:

$p_t$ symmetric $C_0$ semigroup on $L^2(\mu)$, generator $L$ self-adjoint and negative definite. Hence

$$p_t f = e^{tL} f = \int_{-\infty}^{0} e^{t\lambda} \, dP_{((-\infty,\lambda])}(f) \to P_{\{0\}} f = \text{Projection of } f \text{ onto } \ker L$$

□

**Example .** *1. Rotation on $S^1$:*




*Uniform distribution is stationary and ergodic but*

$$p_t f(x) = f\left(e^{it}x\right)$$

*does not converge.*

2. *$S = \mathbb{R}^n$, $p_t$ irreducible (i.e. there is a $t \geq 0$ such that for all $x \in \mathbb{R}^n, U \subset \mathbb{R}^n$ open: $p_t(x, U) > 0$) and strong Feller (i.e. if $f$ is bounded and measurable then $p_t f$ is continuous) then $p_t$ is $\mu$-irreducible and $P_\mu$ is ergodic.*

   *E.g. for Itô diffusions: $a_{ij}(x), b(x)$ locally Hölder continuous and $(a_{ij})$ non-degenerate, then ergodicity holds.*

## 5.2 Central Limit theorem for Markov processes

Let $(M_t)_{t\geq 0}$ be a continuous square-integrable $(\mathcal{F}_t)$ martingale and $\mathcal{F}_t$ a filtration satisfying the usual conditions. Then $M_t^2$ is a submartingale and there exists a unique natural (e.g. continuous) increasing process $\langle M\rangle_t$ such that

$$M_t^2 = \text{martingale} + \langle M\rangle_t$$

(Doob-Meyer decomposition, cf. e.g. Karatzas, Shreve [12]).

**Example .** *If $N_t$ is a Poisson process then*

$$M_t = N_t - \lambda t$$

*is a martingale and*

$$\langle M\rangle_t = \lambda t$$

*almost sure.*

**Note:** For discontinuous martingales, $\langle M\rangle_t$ is *not* the quadratic variation of the paths!

### Identification of bracket process for martingales corresponding to Markov processes:

$(X_t, P_\mu)$ stationary Markov process, $L_L^{(2)}$, $L^{(1)}$ generator on $L^2(\mu),\ L^1(\mu)$, $f \in \mathrm{Dom}(L^{(1)}) \supseteq \mathrm{Dom}(L^{(2)})$. Hence

$$f(X_t) = M_t^f + \int_0^t (L^{(1)}f)(X_s)\,ds \quad P_\mu\text{-a.s.}$$

and $M^f$ is a martingale.

**Theorem 5.4.** *Suppose $f \in \mathrm{Dom}(L^{(2)})$ with $f^2 \in \mathrm{Dom}(L^{(1)})$. Then*

$$\langle M^f\rangle_t = \int_0^t \Gamma(f,f)(X_s)\,ds \quad P_\mu\text{-}a.s.$$

*where*

$$\Gamma(f,g) = L^{(1)}(f\cdot g) - fL^2 g - gL^{(2)}f \in L^1(\mu)$$

*is called* ***Carré du champ (square field) operator***.

*Proof.* We write $A \sim B$ if and only if $A - B$ is a martingale. Hence

$$\begin{aligned}\left(M_t^f\right)^2 &= \left(f(X_t) - \int_0^t Lf(X_s)\,ds\right)^2 \\ &= f(X_t)^2 - 2f(X_t)\int_0^t Lf(X_s)\,ds + \left(\int_0^t Lf(X_s)\,ds\right)^2\end{aligned}$$

where

$$f(X_t)^2 \sim \int_0^t Lf^2(X_s)\,ds$$

and, applying Itô,

$$2f(X_t)\int_0^t Lf(X_s)\,ds = 2\int_0^t f(X_s)Lf(X_r)\,dr + 2\int_0^t\int_0^r Lf(X_s)\,ds\,df(X_r)$$

where

$$f(X_r) \sim \int_0^r Lf(X_s)\,ds$$

Hence

$$\left(M_t^f\right)^2 \sim 2\int_0^t f(X_r)Lf(X_r)\,dr + \left(\int_0^t Lf(X_s)\,ds\right)^2$$

□

**Example .** *Diffusion in* $\mathbb{R}^n$*,*

$$L = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + b(x)\cdot\nabla$$

*Hence*

$$\Gamma(f,g)(x) = \sum_{i,j} a_{ij}(x)\frac{\partial f}{\partial x_i}(x)\frac{\partial g}{\partial x_j}(x) = \left|\sigma^T(x)\nabla f(x)\right|^2_{\mathbb{R}^n}$$

*for all* $f, g \in C_0^\infty(\mathbb{R}^n)$*. Results for gradient diffusions on* $\mathbb{R}^n$ *(e.g. criteria for log Sobolev) extend to general state spaces if* $|\nabla f|^2$ *is replaced by* $\Gamma(f,g)$*!*

**Connection to Dirichlet form:**

$$\mathscr{E}(f,f) = -\int fL^{(2)}f\,d\mu + \underbrace{\left(\frac{1}{2}\int L^{(1)}f^2\,d\mu\right)}_{=0} = \frac{1}{2}\int \Gamma(f,f)\,d\mu$$

*Reference:* Bouleau, Hirsch [6].

**Application 1:** Maximal inequalities.

$$\mathbb{E}\left[\sup_{s\le t}|M_s^f|^p\right] \le C_p\cdot\mathbb{E}\left[\langle M^f\rangle_t^{\frac{p}{2}}\right] \le C_p\cdot t^{\frac{p}{2}-1}\int \Gamma(f,f)^{\frac{p}{2}}\,d\mu$$

This is an important estimate for studying convergence of Markov processes!

**Application 2:** Central limit theorem for ergodic averages.

**Theorem 5.5** (Central limit theorem for martingales)**.** $(M_t)$ *square-integrable martingale on* $(\Omega, \mathcal{F}, P)$ *with stationary increments (i.e.* $M_{t+s} - M_s \sim M_t - M_0$*),* $\sigma > 0$*. If*

$$\frac{1}{t}\langle M\rangle_t \to \sigma^2 \quad \textit{in } L^1(P)$$

*then*

$$\frac{M_t}{\sqrt{t}} \xrightarrow{\mathscr{D}} N(0,\sigma^2)$$

For the proof cf. e.g. Landim [14], Varadhan [22].

**Corollary 5.6** (Central limit theorem for Markov processes (elementary version)). *Let* $(X_t, P_\mu)$ *be a stationary ergodic Markov process. Then for* $f \in \operatorname{Range}(L),\ f = Lg$:

$$\frac{1}{\sqrt{t}} \int_0^t f(X_s)\, ds \xrightarrow{\mathscr{D}} N(0, \sigma_f^2)$$

*where*

$$\sigma_f^2 = 2 \int g(-L)g\, d\mu = 2\mathscr{E}(g, g)$$

**Remark .** 1. *If* $\mu$ *is stationary then*

$$\int f\, d\mu = \int Lg\, d\mu = 0$$

*i.e. the random variables* $f(X_s)$ *are centered.*

2. $\ker(L) = \operatorname{span}\{1\}$ *by ergodicity*

$$(\ker L)^\perp = \left\{ f \in L^2(\mu)\ :\ \int f\, d\mu = 0 \right\} =: L_0^2(\mu)$$

*If* $L\colon L_0^2(\mu) \to L^2(\mu)$ *is bijective with* $G = (-L)^{-1}$ *then the Central limit theorem holds for all* $f \in L^2(\mu)$ *with*

$$\sigma_f^2 = 2(Gf, (-L)Gf)_{L^2(\mu)} = 2(f, Gf)_{L^2(\mu)}$$

*(*$H^{-1}$ *norm if symmetric).*

**Example .** $(X_t, P_\mu)$ *reversible, spectral gap* $\lambda$*, i.e.,*

$$\operatorname{spec}(-L) \subset \{0\} \cup [\lambda, \infty)$$

*hence there is a* $G = (-L\big|_{L_0^2(\mu)})^{-1}$, $\operatorname{spec}(G) \subseteq [0, \frac{1}{\lambda}]$ *and hence*

$$\sigma_f^2 \leq \frac{2}{\lambda} \|f\|_{L^2(\mu)}^2$$

*is a bound for asymptotic variance.*

*Proof of corollary.*

$$\frac{1}{\sqrt{t}}\int_0^t f(X_s)\,ds = \frac{g(X_t)-g(X_0)}{\sqrt{t}} + \frac{M_t^g}{\sqrt{t}}$$

$$\langle M^g\rangle_t = \int_0^t \Gamma(g,g)(X_s)\,ds \quad P_\mu\text{-a.s.}$$

and hence by the ergodic theorem

$$\frac{1}{t}\langle M^g\rangle_t \overset{t\uparrow\infty}{\to} \int \Gamma(g,g)\,d\mu = \sigma_f^2$$

The central limit theorem for martingales gives

$$M_t^g \overset{\mathscr{D}}{\to} N(0,\sigma_f^2)$$

Moreover

$$\frac{1}{\sqrt{t}}\,(g(X_t)-g(X_0)) \to 0$$

in $L^2(P_\mu)$, hence in distribution. This gives the claim since

$$X_t \overset{\mathscr{D}}{\to} \mu, \quad Y_t \overset{\mathscr{D}}{\to} 0 \quad \Rightarrow \quad X_t + Y_t \overset{\mathscr{D}}{\to} \mu$$

□

**Extension:** $\mathrm{Range}(L) \neq L^2$, replace $-L$ by $\alpha - L$ (bijective), then $\alpha \downarrow 0$. Cf. Landim [14].

# Bibliography


[1] ALDOUS, D. & FILL, J. *Reversible Markov Chains and Random Walks on Graph.* Available online: `http://www.stat.berkeley.edu/~aldous/RWG/book.html`

[2] ANÉ, C. & BLACHÈRE, S. & CHAFAÏ, D. & FOUGÈRES, P. & GENTIL, I. & MALRIEU, F. & ROBERTO, C. & SCHEFFER, G. (2000) *Sur les inégalités de Sobolev logarithmiques* Panoramas et Synthèses 10, Société Mathématique de France.

[3] APPLEBAUM, D. (2004) *Lévy processes and stochastic calculus* Cambridge University Press.

[4] ASMUSSEN, S. (2003) *Applied probability and queues.* Springer.

[5] BERTOIN, J. (1998) *Lévy processes.* Cambridge University Press.

[6] BOULEAU, N. & HIRSCH, F. (1991) *Dirichlet Forms and Analysis on Wiener Space.* Gruyter.

[7] BRÉMAUD, P. (2001) *Markov chains : Gibbs fields, Monte Carlo simulation, and queues.* Springer.

[8] DEUSCHEL, J.-D. & STROOCK, D.W. (1989) *Large deviations.* Series in Pure and Applied Mathematics Vol. 137, Academic Press Inc.

[9] DURETT, R. (1993) *Ten Lectures on Particle Systems.* St. Flour Lecture notes. Springer.

[10] FELLER, W. (1991) *An Introduction to Probability Theory and Its Applications, Volume 1 and 2.* John Wiley & Sons.

[11] GROSS, L. & FABES, E. & FUKUSHIMA, M. & KENIG, C. & RÖCKNER, M. & STROOCK, D.W. & GIANFAUSTO DELL'ANTONIO & UMBERTO MOSCO (1992) *Dirichlet Forms: Lectures given at the 1st Session of the Centro Internazionale Matematico Estivo (C.I.M.E.) held in Varenna, Italy, June 8-19, 1992 (Lecture Notes in Mathematics).* Springer.

[12] KARATZAS, I. & SHREVE, S. E. (2005) *Brownian Motion and Stochastic Calculus.* Springer.

[13] KIPNIS, C. & LANDIM, C. (1998) *Scaling limits of interacting particle systems.* Springer.

[14] LANDIM, C. (2003) *Central Limit Theorem for Markov processes,* from Classical to Modern Probability CIMPA Summer School 2001, Picco, Pierre; San Martin, Jaime (Eds.), Progress in Probability 54, 147207, Birkhäuser.

[15] LIGGETT, T.M. (2004) *Interacting particle systems.* Springer.

[16] NORRIS, J.R. (1998) *Markov chains.* Cambridge University Press.

[17] REVUZ, D. & YOR, M. (2004) *Continuous martingales and Brownian motion.* Springer.

[18] ROGERS, L.C.G. & WILLIAMS, D. (2000) *Diffusions, Markov Processes and Martingales.* Cambridge University Press.

[19] ROYER, G. (2007) *An initiation to logarithmic Sobolev inequalities.* American Mathematical Society.

[20] STROOCK, D.W. (2005) *An introduction to Markov processes.* Springer.

[21] STROOCK, D.W. (2000) *Probability Theory, an Analytic View.* Cambridge University Press.

[22] VARADHAN, S.R.S. (2001) *Probability Theory.* Oxford University Press.

# Index