D-MATH

# APPLIED STOCHASTIC PROCESSES

## Spring Term 2017


## Alain-Sol Sznitman


**August 2017**

# Contents

**0 Introduction** **1**
0.1 Poisson process . . . 1
0.2 Renewal processes . . . 2
0.3 Markov chains . . . 4
0.4 Brownian motion . . . 5

**1 Poisson process** **7**
1.1 Stationary Poisson process on $\mathbb{R}$ . . . 16
1.2 Superposition and thinning of Poisson processes . . . 19
1.2.1 Thinning . . . 20
1.2.2 Superposition . . . 22
1.3 Inhomogeneous Poisson process . . . 24
1.3.1 Construction via time change . . . 25
1.3.2 Construction by variable thinning . . . 25

**2 Renewal processes** **33**
2.1 The set-up . . . 33
2.2 Some properties of $M(t)$ . . . 34
2.3 Renewal with delays . . . 38
2.4 Blackwell's Renewal Theorem . . . 43
2.5 The renewal equation . . . 45
2.6 Examples . . . 46
2.6.1 The renewal function . . . 46
2.6.2 The age and excess distribution functions . . . 47
2.6.3 Cycles of operation and repair of a machine . . . 49
2.7 Well-posedness of the renewal equation . . . 51
2.8 Asymptotic behavior of solutions of the renewal equation . . . 52
2.9 Applications . . . 56
2.9.1 The age and excess distribution functions . . . 56
2.9.2 Cycles of operation and repair of a machine . . . 57
2.10 Renewal with defect . . . 57

**3 Discrete time Markov chains** **61**
3.1 Examples . . . 64
3.1.1 Simple random walk on $\mathbb{Z}^d$ . . . 64
3.1.2 Reflected random walk . . . 65
3.1.3 Galton-Watson chain . . . 65
3.1.4 Ehrenfest model of diffusion . . . 65
3.1.5 Residual waiting time . . . 66
3.2 Markov and strong Markov property . . . 67
3.3 Recurrence and transience . . . 68
3.4 Stationary distribution . . . 74

3.5 Asymptotic behavior . . . . . . . . . . 80

**4 Continuous time Markov chains** **83**
4.1 Construction of pure jump processes (with no explosion) . . . . . . . . . . 88
4.2 Backward and forward Kolmogorov equations, generator . . . . . . . . . . 95
4.3 Transience and recurrence . . . . . . . . . . 101
4.4 Stationary distributions . . . . . . . . . . 103
4.5 Stationary distributions and asymptotic behavior . . . . . . . . . . 106

**5 Brownian motion** **111**
5.1 Brownian motion as a Gaussian process . . . . . . . . . . 114
5.2 Brownian motion as a Markov process . . . . . . . . . . 115
5.3 Brownian motion as a continuous martingale . . . . . . . . . . 117

# 0 Introduction

The object of this course is to present some of the stochastic processes, which often occur in applications. Typical examples are the Poisson process, the renewal processes, Markov chains and Brownian motion.

The present introduction will discuss these objects in an informal way and give some flavor of these topics.

## 0.1 Poisson process

They are named after Siméon Denis Poisson (1781-1840), who introduced in his treatise "Recherches sur la probabilité des jugements" in 1837, the Poisson distribution, which is closely linked to Poisson processes.

Loosely speaking, Poisson processes come as follows:
One has random points on $[0, \infty)$, such that for any Borel set $A$ of $[0, \infty)$,

(0.1) $$N(A) \overset{\text{def}}{=} \text{number of points falling in } A$$

is distributed as a Poisson variable with parameter $\lambda|A|$ ($|A|$ = the Lebesgue measure of $A$), $\lambda > 0$ some constant, that is

(0.2) $$P[N(A) = k] = e^{-\lambda|A|} \, \frac{(\lambda|A|)^k}{k!}, \text{ for } k \in \mathbb{N} = \{0, 1, 2, \dots, \}, \text{ if } |A| < \infty;$$
$$\text{and } N(A) \overset{\text{a.s.}}{=} \infty, \text{ if } |A| = \infty.$$

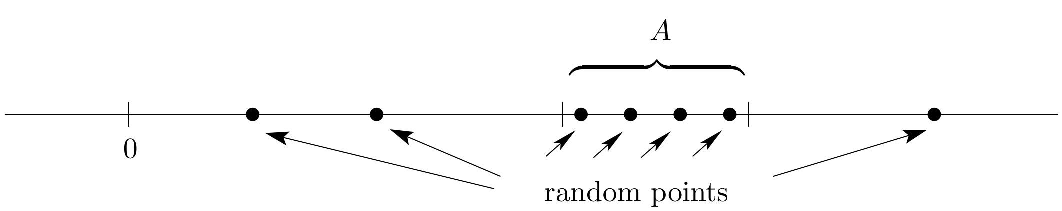



Fig. 0.1

and for $m \geq 1$ and $A_1, A_2, \dots, A_m$ pairwise disjoint Borel sets of $[0, \infty)$:

(0.3) $$N(A_1), \dots, N(A_m) \text{ are independent variables.}$$

The Poisson process records the number of points falling in the interval $[0, t]$, $t \geq 0$:

(0.4) $$N_t \overset{\text{def}}{=} N([0, t]), \; t \geq 0.$$

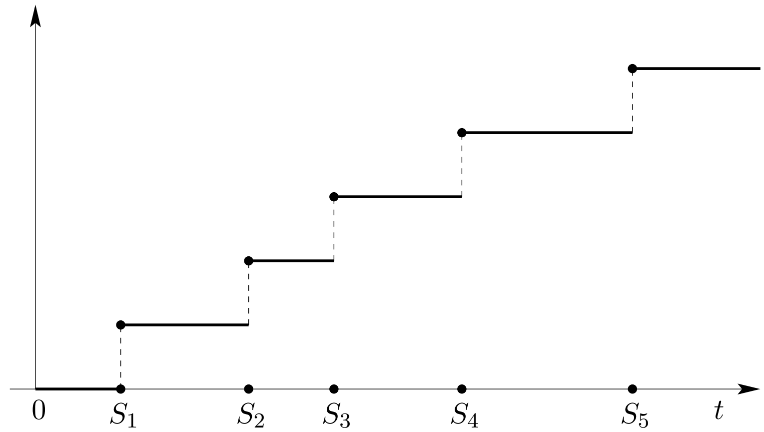



Fig. 0.2

Thus, it is a non-decreasing right-continuous function of $t$, which has jumps of size 1 at the location of the random points, and remains otherwise constant between consecutive jumps.

The Poisson process comes in many applications, for instance as a description of arrival times of customers in a queue, of times at which telephone calls arrive at a call center, of times at which claims arrive in an insurance company, of times of emission for $\alpha$-particles by a radioactive source, etc. ...

We will see different characterizations of the Poisson process. In particular if $S_i, i \geq 1$, denote the successive jump times of $N_t, t \geq 0$, (with $S_0 = 0$, by convention), and $T_i, i \geq 1$, are the inter-arrival times:

(0.5)
$$T_i = S_i - S_{i-1},\ i \geq 1,$$

we will see that

(0.6) $T_i, i \geq 1$, are independent exponential variables with parameter $\lambda$:

(0.7)
$$P[T_i \in B] = \int_B \lambda e^{-\lambda x} dx, \text{ for any Borel set } B \text{ in } [0, \infty).$$

This description makes the Poisson process a special (and important) example of another topic of interest in this course:

## 0.2 Renewal processes

One now considers i.i.d. variables $T_i, i \geq 1$, on $[0, \infty)$ with $P[T_i = 0] < 1$, and finite expectation:

(0.8)
$$\mu = E[T_i],\ (\text{for any } i \geq 1),$$

and one introduces the renewal times

(0.9)
$$\begin{aligned} &S_i = T_1 + \cdots + T_i,\ i \geq 1, \\ &S_0 = 0\,. \end{aligned}$$

The renewal process is now defined as

$$N_t = \sum_{i \geq 1} 1\{S_i \leq t\}, \text{ for } t \geq 0. \tag{0.10}$$

It counts the number of renewal times in the interval $[0, t]$.

Fig. 0.3

In view of (0.6), the renewal process is a generalization of the Poisson process. An important idea is that after a time $S_i$, "things start again and develop in the same fashion" (whence the terminology of "renewal times"):

$$(N_{S_i+t} - i)_{t \geq 0} \text{ is distributed like } (N_t)_{t \geq 0}, \text{ and independent of } T_1, \dots, T_i. \tag{0.11}$$

In some cases we will discuss delayed renewal processes, for which $S_0 \geq 0$ is not necessarily 0, and is independent of the i.i.d. sequence $(T_i)_{i \geq 1}$, so that

$$S_i = S_0 + T_1 + \dots + T_i, \ i \geq 1. \tag{0.12}$$

Of typical interest will be the large $t$ asymptotics of the renewal process. We will see that under suitable assumptions on the distribution of the variables $T_i, i \geq 1$ (namely "non-arithmeticity"), one has Blackwell's renewal theorem:

$$\lim_{t \to \infty} E[N_{t+h} - N_t] = \frac{h}{\mu}, \text{ for } h > 0 \text{ (and } \mu \text{ as in (0.8)).} \tag{0.13}$$

Another topic of interest will be the behavior of the "age" $A_t$ and "excess" $E_t$, when $t$ is a large deterministic time:

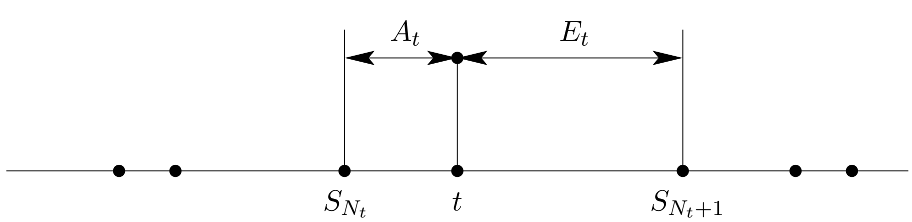


Fig. 0.4

The renewal processes show up in a number of applications. They can for instance describe the times of successive repairs of a machine, or of replacements of an electric component. They are linked to the notion of "regeneration", where things "start afresh" after certain renewal times. Renewal processes also show up in our next topic of discussion:

## 0.3 Markov chains

They are named after Andrei Markov (1856 - 1922), who introduced them. Typically, one considers a state space $E$ (in this course we assume that $E$ is an at most countable set). Then, a sequence $(X_n)_{n\geq 0}$ of random variables with values in $E$, on some probability space $(\Omega, \mathcal{A}, P)$, is a discrete time Markov chain when:

$$E[f(X_{n+1})|X_0,\dots,X_n] \overset{P-\text{a.s.}}{=} E[f(X_{n+1})|X_n], \tag{0.14}$$

for any bounded function $f\colon E \to \mathbb{R}$.

*Intuitively:* "The best prediction of the future of the sequence $(X_n)$, knowing its past, just relies on the knowledge of the present". This informal statement reflects the fact that when (0.14) holds, one can see that a similar statement holds as well with $f(X_{n+1})$ replaced by $g(X_{n+1}, X_{n+2}, \dots, X_{n+k})$, with $k \geq 1$ arbitrary, and $g$ bounded on $E^k$.

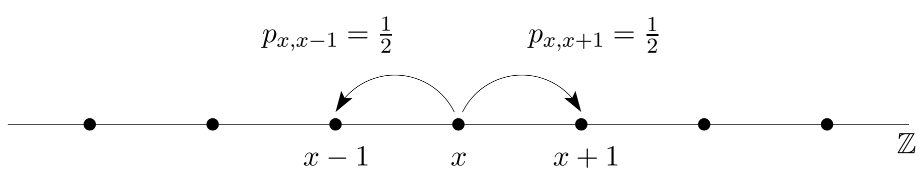
Of special interest for us will be the temporally homogeneous situation when:

$$P[X_{n+1} = y|X_n = x] = p_{x,y} \quad \begin{array}{l}\text{for any } n \geq 0,\ y \in E,\\ \text{and } x \in E \text{ with } P[X_n = x] > 0,\end{array} \tag{0.15}$$

where $(p_{x,y})_{x,y\in E}$ is a (fixed) transition probability on $E$:

$$p_{x,y} \geq 0, \text{ for } x, y \in E, \text{ and } \sum_{y\in E} p_{x,y} = 1, \text{ for all } x \in E. \tag{0.16}$$

Markov chains show up in many situations. For instance, the state space can be $E = \mathbb{Z}$ in the case of the simple random walk on $\mathbb{Z}$:




$p_{x,y} = 0$, if $y \notin \{x-1, x+1\}$

Fig. 0.5

But $E$ can also be of the form $E = \{1,\dots,K\}^\Lambda$, with $\Lambda = \{1,\dots,L\}^2 \subseteq \mathbb{Z}^2$, the space of "configurations of pixels" in a box $\Lambda$, so that $x \in E$ can be thought of as an image and $K \geq 2$, represents the number of grey levels of the pixels. There are indeed many examples of state spaces and Markov chains that one encounters in applications.

Typical questions concerning Markov chains have to do with their asymptotic behavior. For instance one is interested in:

- How does the distribution of $X_n$ look like for large $n$?

- Is a state $x \in E$ visited finitely many times or infinitely often by the chain (i.e. is the variable $\sum_{n=0}^{\infty} 1\{X_n = x\}$ finite or infinite)?

If the state $x \in E$ is visited infinitely often by the chain, then the successive times of visit:

$$
\begin{aligned}
S_0 &= \inf\{n \geq 0; X_n = x\},\ S_1 = \inf\{n > S_0; X_n = x\}, \dots \\
S_{k+1} &= \inf\{n > S_k : X_n = x\}, \dots
\end{aligned}
\tag{0.17}
$$

turn out to lead to a renewal process, possibly with delay (see (0.12)).

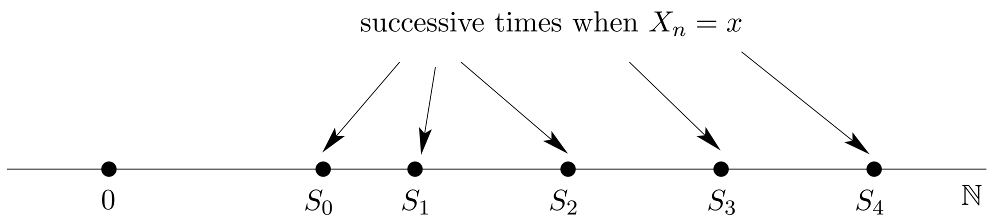



Fig. 0.6

An important role in the asymptotic analysis of the chain is played by the stationary distributions, i.e. the probabilities $\pi$ on $E$ such that

$$
\sum_{x \in E} \pi(x)\, p_{x,y} = \pi(y), \text{ for all } y \in E,
\tag{0.18}
$$

and one is naturally led to study the existence and uniqueness of such distributions.

One does also study Markov chains in continuous time, $(X_t)_{t \geq 0}$, where in place of (0.14) one now requires that

$$
E[f(X_{t_{n+1}})|X_{t_0}, X_{t_1}, \dots, X_{t_n}] = E[f(X_{t_{n+1}})|X_{t_n}],
\tag{0.19}
$$

for any $0 \leq t_0 < t_1 < \dots < t_n < t_{n+1}$, and bounded $f\colon E \to \mathbb{R}$.

One is interested in very similar questions as in the case of Markov chains in discrete time, and there are important links between the continuous time and the discrete time situation.

## 0.4 Brownian motion

It is named after Robert Brown, who in 1828, at the time director of the British botanical museum, discovered the disordered motion of pollen grains suspended in water.

One can construct Brownian motion as a limit of rescaled polygonal interpolations of a simple random walk, somewhat in the same way as the normal distribution in the weak limit of the distributions of rescaled sums of i.i.d. random signs (as a consequence of the central limit theorem).

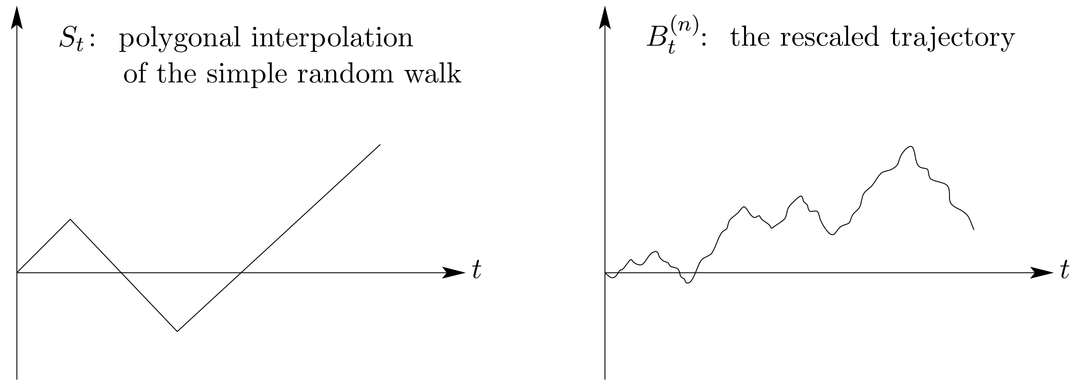



Fig. 0.7

$$
\begin{aligned}
&(0.20) \qquad X_1, \dots, X_n, \dots \text{ i.i.d. with } P[X_i = \pm 1] = \tfrac{1}{2}, \text{ for all } i \geq 1,\\
&\qquad\qquad S_n = X_1 + \dots + X_m,\ m \geq 1,\ S_0 = 0,\\
&\qquad\qquad S_t,\ t \geq 0, \text{ the polygonal interpolation of } S_m, m \geq 1,
\end{aligned}
$$

and

$$
(0.21) \qquad B_t^{(n)} = \tfrac{1}{n}\, S_{n^2 t},\ t \geq 0, \text{ the rescaled (time and space) trajectory, with } n \geq 1.
$$

The central limit theorem implies that

$$
(0.22) \qquad P[B_1^{(n)} \leq a] \underset{n \to \infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} \exp\Big\{ -\frac{x^2}{2} \Big\}\, dx, \text{ for all } a \in \mathbb{R}.
$$

Much more is true, and in a sense which can be made precise, the whole trajectory $B_{\cdot}^{(n)}$ converges to a limit object $B_{\cdot}$, the Brownian motion (this is a special case of Donsker's invariance principle).

Brownian motion turns out to be a fundamental stochastic process. The trajectories $t \to B_t(\omega)$ are continuous, but very rough (nowhere differentiable, of infinite variation on any proper interval). One can nevertheless build an "infinitesimal calculus" for Brownian motion and show that when $f$ is a $C^2$ function on $\mathbb{R}$:

$$
(0.23) \qquad f(B_t) = f(B_0) + \int_0^t f'(B_s)\, dB_s + \frac{1}{2} \int_0^t f''(B_s)\, ds,\ t \geq 0,\ \text{(Ito's formula)},
$$

where $\int_0^t f'(B_s) dB_s$ is a so-called "stochastic integral", ($t \to B_t$ has infinite variation on proper intervals and $\int_0^t f'(B_s) dB_s$ has no meaning as a Stieltjes integral). Note the surprising apparition of a term "$\frac{1}{2}\int_0^t f''(B_s) ds$" in (0.23), which reflects the Brownian trajectories.

# 1 Poisson process

We begin with some general definitions before getting to the heart of the matter.

**Definition 1.1.** *Given a probability space* $(\Omega, \mathcal{A}, P)$, *a set* $I \neq \emptyset$, *and a measurable space* $(E, \mathcal{E})$, *a* ***stochastic process*** *with* ***time parameter*** $\boldsymbol{I}$, *and* ***state space*** $\boldsymbol{E}$, *is a collection* $(X_t)_{t \in I}$ *of random variables* $X_t$ *on* $(\Omega, \mathcal{A}, P)$ *with values in* $E$. *For* $\omega \in \Omega$, *the application* $t \in I \to X_t(\omega)$ *is called* ***trajectory****, or* ***realization****, or* ***sample path*** *of the stochastic process.*

Typically, we will be interested in $I = \mathbb{N}(= \{0, 1, 2, \dots\})$ or $I = \mathbb{R}_+$, and in $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, the Borel $\sigma$-algebra on $\mathbb{R}$. If we do not explicitly specify the state space of the stochastic process, it is tacitly assumed to be $E = \mathbb{R}$ (and $\mathcal{E} = \mathcal{B}(\mathbb{R})$).

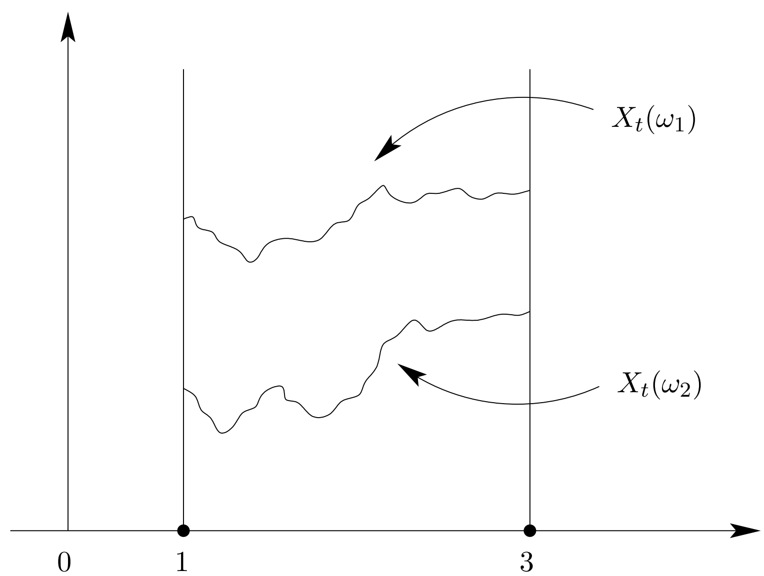



Fig. 1.1: Two sample paths of a stochastic process with $I = [1, 3]$

**Definition 1.2.** *A stochastic process* $(Y_t)_{t \geq 0}$ *defined on* $(\Omega, \mathcal{A}, P)$ *is said to have* ***independent increments****, if for any* $k \geq 1$, $0 = t_0 < t_1 < \dots < t_k$,

(1.1) $$Y_{t_1} - Y_{t_0}, \dots, Y_{t_k} - Y_{t_{k-1}} \text{ are independent.}$$

The stochastic process $(Y_t)_{t \geq 0}$ is said to have **stationary increments** if for any $k \geq 1$, $t_0 = 0 < t_1 < \dots < t_k$ and $h > 0$, the random vectors:

(1.2) $$(Y_{t_1} - Y_{t_0}, \dots, Y_{t_k} - Y_{t_{k-1}}) \text{ and } (Y_{t_1+h} - Y_{t_0+h}, \dots, Y_{t_k+h} - Y_{t_{k-1}+h})$$

have same distribution (on $\mathbb{R}^k$).

By a **counting process**, we mean a stochastic process $(N_t)_{t\ge 0}$, such that the trajectories

(1.3) $$t \ge 0 \to N_t(\omega) \text{ are non-decreasing, right-continuous with values in } \mathbb{N}.$$

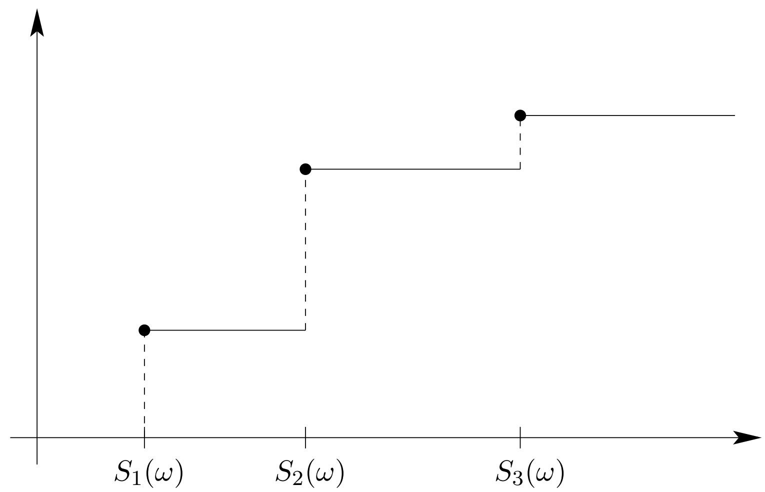



Fig. 1.2: A sample path of a counting process (for which $N_0 = 0$)

Given a counting process $(N_t)_{t\ge 0}$, one can define its successive jump times $(S_i)_{i\ge 1}$, via

(1.4)
$$\begin{aligned} S_1(\omega) &= \inf\{t \ge 0;\ N_t(\omega) > N_0(\omega)\} \le \infty, \\ &\quad \text{(if the set } \{\dots\} = \emptyset \text{, then } S_1(\omega) = \infty\text{), and inductively for } k \ge 1, \\ S_{k+1}(\omega) &= \inf\{t \ge S_k(\omega);\ N_t(\omega) > N_{S_k}(\omega)\} \\ &\quad \text{(this is understood as } \infty \text{ if } S_k(\omega) = \infty\text{).} \end{aligned}$$

In this way one has

(1.5)
$$\begin{cases} 0 < S_1(\omega) \le S_2(\omega) \le \dots \le S_k(\omega) \le \dots \quad \text{(each inequality is in fact strict if the left member is finite),} \\ \lim\limits_k S_k(\omega) = \infty \text{ (note that by (1.3), for any } T > 0 \text{ and } \omega \in \Omega, \; S_{N_T(\omega)+1}(\omega) > T\text{).} \end{cases}$$

We are now ready to prove some equivalent properties for a counting process, which will then characterize the Poisson processes.

**Theorem 1.3.** *Consider $\lambda > 0$, and a counting process $(N_t)_{t \geq 0}$, with $N_0 = 0$ and jumps of size 1 defined on some probability space $(\Omega, \mathcal{A}, P)$. The following properties are equivalent:*

(1.6) $(N_t)_{t\geq 0}$ *has independent and stationary increments, and as* $t \to 0$,

$$P[N_t = 1] = \lambda t + o(t),$$
$$P[N_t \geq 2] = o(t);$$

(1.7) $(N_t)_{t\geq 0}$ *has independent and stationary increments, and for all* $t \geq 0$, $N_t$ *has a Poisson*$(\lambda t)$ *distribution;*

(1.8) *the successive jump times* $S_i, i \geq 1$ *(see (1.4)), are* $P$*-a.s. finite and the* $T_i = S_i - S_{i-1}, i \geq 1$ *(where* $S_0 = 0$*, by convention) are i.i.d. exponentially*$(\lambda)$*-distributed variables;*

(1.9) *for each* $t > 0$, $N_t$ *has a Poisson*$(\lambda t)$ *distribution, and when* $k \geq 1$, *conditional on* $N_t = k$*, the variables* $(S_1, \dots, S_k)$ *are distributed as the non-decreasing reordering of* $k$ *independent variables* $U_1, \dots, U_k$ *uniformly distributed on* $[0, t]$*, i.e. they admit the density (conditional on* $\{N_t = k\}$*):* $f(s_1, \dots, s_k | N_t = k) = k!\,\frac{1}{t^k}\,1\{0 < s_1 < \cdots < s_k < t\}$.

The above theorem allows to introduce the following definition:

**Definition 1.4.** *A Poisson process with rate $\lambda > 0$ is a counting process $(N_t)_{t\geq 0}$, with $N_0 = 0$, and jumps of size 1 satisfying one of the equivalent properties (1.6), (1.7), (1.8), (1.9).*

*Proof of Theorem 1.3:*

- (1.7) $\Longrightarrow$ (1.6):

This is easy because:

$$P[N_t = 1] = e^{-\lambda t}\lambda t = \lambda t + o(t), \text{ as } t \to 0, \text{ and}$$
$$P[N_t \geq 2] = 1 - e^{-\lambda t} - e^{-\lambda t}\lambda t = o(t), \text{ as } t \to 0.$$

- (1.6) $\Longrightarrow$ (1.7):

Consider for $t > 0$ and $n \geq 1$:

$$M_{n,t} = \sum_{k=1}^{n} 1\{N_{\frac{kt}{n}} - N_{(k-1)\frac{t}{n}} \geq 1\}. \tag{1.10}$$

It follows from the independence and the stationarity of the increments of $N_.$ that

$$\text{(1.11)} \qquad M_{n,t} \text{ is distributed as a binomial } (n, \underset{\substack{\uparrow \\ \text{the success probability}}}{P[N_{\frac{t}{n}} \ge 1]}) \text{ variable.}$$

Moreover we have:

$$\text{(1.12)} \qquad n\,P[N_{\frac{t}{n}} \ge 1] \overset{(1.6)}{=} n\Big(\frac{\lambda t}{n} + o\big(\frac{t}{n}\big)\Big) \underset{n\to\infty}{\longrightarrow} \lambda t.$$

From the usual Poisson approximation result for binomial $(n, p_n)$ variables with $np_n \to$ const., see for instance (2.2.5), p. 47 of [14], we find that

$$\text{(1.13)} \qquad M_{n,t} \text{ converges in law to a Poisson}(\lambda t) \text{ distribution as } n \to \infty.$$

At the same time we know that

$$\text{(1.14)} \qquad \begin{aligned} &P[N_t \ne M_{n,t}] = P\Big[\bigcup_{k=1}^{n} \big\{N_{\frac{kt}{n}} - N_{(k-1)\frac{t}{n}} \ge 2\big\}\Big] \le \\ &\sum_{k=1}^{n} P\big[N_{\frac{kt}{n}} - N_{(k-1)\frac{t}{n}} \ge 2\big] \overset{(1.6)}{=} n\,o\big(\frac{t}{n}\big) \underset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

This fact together with (1.13) implies that $N_t$ is Poisson$(\lambda t)$-distributed (indeed for $f$ bounded continuous on $\mathbb{R}$:

$$\begin{aligned} &|E[f(N_t)] - E[f(M_{n,t})]| \le \|f\|_\infty\, P[N_t \ne M_{n,t}] \underset{n\to\infty}{\longrightarrow} 0, \text{ and} \\ &E[f(M_{n,t})] \underset{n\to\infty}{\overset{(1.13)}{\longrightarrow}} e^{-\lambda t} \sum_{k=0}^{\infty} f(k)\, \frac{(\lambda t)^k}{k!}, \text{ so that} \\ &E[f(N_t)] = e^{-\lambda t} \sum_{k=0}^{\infty} f(k)\, \frac{(\lambda t)^k}{k!}, \text{ for all } f \text{ as above ).} \end{aligned}$$

This proves (1.7).

- <u>(1.7) $\Longrightarrow$ (1.8)</u>:

We first give a **non-rigorous heuristic argument**:

$$P[T_1 > t] = P[N_t = 0] = e^{-\lambda t}, \text{ so } T_1 \text{ is exponential } (\lambda)\text{-distributed.}$$

$$\begin{aligned} P[T_2 > t | T_1 = s] &\text{“=”}\ P[N_{t+s} - N_s = 0 | N_u = 0,\ 0 \le u < s,\ N_s = 1] \\ &\underset{\substack{\uparrow \\ \text{independence} \\ \text{of increments}}}{\text{“=”}}\ P[N_{t+s} - N_s = 0] \underset{\substack{\uparrow \\ \text{stationarity} \\ \text{of increments}}}{=} P[N_t = 0] = e^{-\lambda t}, \end{aligned}$$

so that $T_2$ is independent of $T_1$ and exponential $(\lambda)$-distributed, and so on.

We now give a **rigorous argument**:

$$P[T_1 > t] = P[N_t = 0] = e^{-\lambda t}, \text{ for } t > 0, \tag{1.15}$$

so that $T_1 < \infty$, $P$-a.s., and $T_1$ is exponential($\lambda$)-distributed. Further, for $0 \le s_1 \le t_1 \le s_2 \le t_2$, one has:

$$\begin{aligned}
&P[s_1 < S_1 \le t_1,\, s_2 < S_2 \le t_2] = \\
&P[N_{s_1} = 0,\, N_{t_1} - N_{s_1} = 1,\, N_{s_2} - N_{t_1} = 0,\, N_{t_2} - N_{s_2} \ge 1] = \\
&\text{using independence and stationarity of increments} \\
&e^{-\lambda s_1} \times \lambda(t_1 - s_1)\, e^{-\lambda(t_1 - s_1)} \times e^{-\lambda(s_2 - t_1)} \times \big(1 - e^{-\lambda(t_2 - s_2)}\big) = \\
&\lambda(t_1 - s_1)(e^{-\lambda s_2} - e^{-\lambda t_2}) = \int_{\substack{\{s_1 < y_1 \le t_1, \\ s_2 < y_2 \le t_2\}}} \lambda^2\, e^{-\lambda y_2} dy_1 dy_2\,.
\end{aligned} \tag{1.16}$$

In the same way we have

$$\begin{aligned}
&P[s_1 < S_1 \le t_1,\, s_2 < S_2] = \lambda(t_1 - s_1)\, e^{-\lambda s_2} \underset{s_2 \to \infty}{\longrightarrow} 0, \text{ and hence} \\
&P[s_1 < S_1 \le t_1,\, S_2 = \infty] = 0\,.
\end{aligned}$$

Choosing $s_1 = 0$, and letting $t_1 \to \infty$, we find that $P[0 < S_1 < \infty, S_2 = \infty] = 0$, and from (1.15) and the first line of (1.5), we know that $P[0 < S_1 < \infty, S_1 < S_2] = 1$. Therefore:

$$P[0 < S_1 < S_2 < \infty] = 1\,. \tag{1.17}$$

If we now introduce the probability density on $\mathbb{R}_+^2$:

$$f(y_1, y_2) = \lambda^2\, e^{-\lambda y_2} 1\{0 < y_1 < y_2\}, \tag{1.18}$$

we deduce from (1.16) that

$$P[(S_1, S_2) \in A] = \int_A f(y_1, y_2)\, dy_1 dy_2 \text{ for } A \in \mathcal{B}(\mathbb{R}_+^2), \tag{1.19}$$

(the class of sets $A = (s_1, t_1] \times (s_2, t_2]$, with $0 \le s_1 \le t_1 \le s_2 \le t_2$ is a $\pi$-system of subsets of $U = \{(x, y);\ 0 < x < y < \infty\}$, which generates $\mathcal{B}(U)$. One can then apply Dynkin's lemma, cf. [12], p. 41, noting that both $f\, dy_1 dy_2$ and the law of $S_1, S_2$ are supported by $U$, cf. (1.17), (1.18)).

By an analogous argument, we obtain that for $k \ge 1$, $S_k < \infty$, $P$-a.s. and

$$\begin{aligned}
&f_{(S_1, \dots, S_k)}(y_1, \dots, y_k) = \lambda^k\, e^{-\lambda y_k} 1\{0 < y_1 < \dots < y_k\} \\
&\text{is the density of } (S_1, \dots, S_k).
\end{aligned} \tag{1.20}$$

Denote with $h$ the linear map on $\mathbb{R}^k$:

$$h(t_1, \dots, t_k) = (t_1, t_1 + t_2, \dots, t_1 + \dots + t_k), \tag{1.21}$$

and with $\mu_k$ the product measure on $\mathbb{R}^k$:

$$\mu_k(dt_1, \dots, dt_k) = \prod_{i=1}^{k} \lambda e^{-\lambda t_i} 1\{t_i > 0\}\, dt_1 \dots dt_k \,. \tag{1.22}$$

The image of $\mu_k$ under $h$ (which has determinant 1) admits the density $\lambda^k e^{-\lambda y_k} 1\{0 < y_1 < y_2 < \dots < y_k\}$ of (1.20). Looking at inter-arrival times, this proves that

$$\begin{aligned}(T_1, T_2, \dots, T_k) &= (S_1, S_2 - S_1, \dots, S_k - S_{k-1}) \\ &= h^{-1}(S_1, \dots, S_k) \text{ is } \mu_k\text{-distributed,}\end{aligned} \tag{1.23}$$

and (1.8) follows.

- (1.8) $\Longrightarrow$ (1.9):

If $g(t) = \lambda\, e^{-\lambda t} 1\{t > 0\}$, then $S_2 = T_1 + T_2$ admits the density:

$$\begin{aligned} g * g(t) &= \int_0^t \lambda\, e^{-\lambda(t-s)} \lambda e^{-\lambda s} ds = \lambda^2\, t\, e^{-\lambda t}, \text{ when } t > 0, \\ &= 0, \text{ when } t \le 0, \end{aligned} \tag{1.24}$$

and by induction $S_{k+1}$ admits the density

$$\begin{aligned} g^{*(k+1)}(t) = g * g^{*(k)}(t) &= \int_0^t \lambda\, e^{-\lambda(t-s)}\, \lambda^k\, \frac{s^{k-1}}{(k-1)!}\, e^{-\lambda s} ds \\ &= \lambda^{k+1}\, \frac{t^k}{k!}\, e^{-\lambda t}, \text{ if } t > 0, \\ &= 0, \text{ if } t \le 0. \end{aligned}$$

In other words, for each $k \ge 1$:

$$\lambda^k\, \frac{s^{k-1}}{(k-1)!}\, e^{-\lambda s} 1\{s > 0\} \text{ is the density of } S_k, \tag{1.25}$$

(i.e. $S_k$ is Gamma$(k, \lambda)$-distributed, where the general Gamma$(\nu, \lambda)$-density, $\nu, \lambda > 0$, cf. [6], p. 47, is defined by $f_{\lambda,\nu}(s) = \frac{1}{\Gamma(\nu)} \lambda^\nu s^{\nu-1} e^{-\lambda s} 1\{s > 0\}$).

In particular, we see that

$$\begin{aligned} &\text{i)} \quad P[N_t = 0] = P[S_1 > t] \overset{(1.25)}{=} e^{-\lambda t} \\ &\text{ii)} \quad P[N_t = k] = P[S_k \le t,\, S_{k+1} > t] = P[S_k \le t] - P[S_{k+1} \le t] \\ &\qquad \overset{(1.25)}{=} \int_0^t \lambda e^{-\lambda s}\, \frac{(\lambda s)^{k-1}}{(k-1)!}\, ds - \underbrace{\int_0^t \lambda e^{-\lambda s}\, \frac{(\lambda s)^k}{k!}\, ds}_{\substack{\text{integrating by parts, this equals} \\ e^{-\lambda t} \frac{(\lambda t)^k}{k!} - \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{k-1}}{(k-1)!} ds}} \\ &\qquad = e^{-\lambda t}\, \frac{(\lambda t)^k}{k!}, \text{ for } k \ge 1. \end{aligned} \tag{1.26}$$

This shows that

$$N_t \text{ is Poisson } (\lambda t)\text{-distributed.} \tag{1.27}$$

If $h$ is as in (1.21), $(S_1, \dots, S_k) = h(T_1, \dots, T_k)$, when $k \ge 1$, as explained below (1.22) has density $\lambda^k e^{-\lambda s_k} 1\{0 < s_1 < \dots < s_k\}$. Applying this observation with $k+1$ in place of $k$ we find that $\lambda^{k+1} e^{-\lambda s_{k+1}} 1\{0 < s_1 < s_2 < \dots < s_{k+1}\}$ is the density of $(S_1, \dots, S_{k+1})$. Thus as in (1.26) ii), we see that the function of $s_1, \dots, s_{k+1}$ defined by

$$\begin{aligned} f(s_1, \dots, s_{k+1} \,|N_t = k) &= \frac{\lambda^{k+1}}{P[N_t = k]}\, e^{-\lambda s_{k+1}}\; 1\{0 < s_1 < \dots < s_{k+1}\} \\ &\qquad\qquad\qquad\qquad\quad 1\{s_k \le t < s_{k+1}\} \\ &\overset{(1.27)}{=} \lambda e^{-\lambda s_{k+1}}\, \frac{e^{\lambda t} k!}{t^k}\; 1\{0 < s_1 < \dots < s_k \le t < s_{k+1}\} \end{aligned} \tag{1.28}$$

is the density of $(S_1, \dots, S_{k+1})$ conditional on $\{N_t = k\}$. As a result, the density of $(S_1, \dots, S_k)$ conditional on $\{N_t = k\}$ is

$$\begin{aligned} & f(s_1, \dots, s_k | N_t = k) = \int_t^\infty f(s_1, \dots, s_k, s_{k+1} | N_t = k)\, ds_{k+1} = \\ & \frac{k!}{t^k}\, 1\{0 < s_1 < \dots < s_k \le t\} \int_t^\infty \lambda e^{-\lambda(s-t)} ds = \frac{k!}{t^k}\, 1\{0 < s_1 < \dots < s_k \le t\}, \end{aligned} \tag{1.29}$$

which is the density of the non-decreasing reordering of $U_1, U_2, \dots, U_k$ i.i.d. uniformly distributed variables on $[0, t]$ (see also [12], p. 323). This completes the proof of (1.9).

- (1.9) $\Longrightarrow$ (1.7):

We will use the following lemma:

**Lemma 1.5.** *Consider $\lambda_1, \dots, \lambda_n > 0$, $\lambda = \sum_{i=1}^n \lambda_i$, $p_i = \frac{\lambda_i}{\lambda}$, $1 \le i \le n$. Given $\mathbb{N}$-valued random variables $Z_1, \dots, Z_n$ and $Z = Z_1 + \dots + Z_n$, the following properties are equivalent:*

(1.30) *$Z_1, \dots, Z_n$ are independent, respectively Poisson$(\lambda_i)$-distributed, $1 \le i \le n$,*

(1.31) *$Z$ is Poisson$(\lambda)$-distributed and for any $k \ge 1$, conditional on $Z = k$, $(Z_1, \dots, Z_n)$ is multinomial $(k; p_1, \dots, p_n)$-distributed, that is:*

$$P[Z_1 = j_1, \dots, Z_n = j_n | Z = k] = \frac{k!}{j_1! \dots j_n!}\, p_1^{j_1} \dots p_n^{j_n},$$

*for $0 \le j_1, \dots, j_n$ with $j_1 + \dots + j_n = k$.*

*Proof of Lemma 1.5*:

- (1.30) $\Longrightarrow$ (1.31):

$$\begin{aligned} P[Z_1 = j_1, \dots, Z_n = j_n] &= e^{-\lambda_1}\, \frac{\lambda_1^{j_1}}{j_1!} \dots e^{-\lambda_n}\, \frac{\lambda_n^{j_n}}{j_n!}, \ \text{ for } j_1, \dots, j_n \ge 0 \\ &= e^{-\lambda}\, \frac{\lambda_1^{j_1} \dots \lambda_n^{j_n}}{j_1! \dots j_n!}\ . \end{aligned} \tag{1.32}$$

Then we have for $k \geq 0$:

$$(1.33) \qquad P[Z=k] = e^{-\lambda} \sum_{j_1+\cdots+j_n=k} \frac{\lambda_1^{j_1}\dots\lambda_n^{j_n}}{j_1!\dots j_n!} \underset{\uparrow}{=} \frac{e^{-\lambda}}{k!}\,(\lambda_1+\cdots+\lambda_n)^k = e^{-\lambda}\,\frac{\lambda^k}{k!}\,.$$

multiplying and dividing by $k!$ and
using the multinomial formula

So $Z$ is Poisson($\lambda$)-distributed.

Moreover, for $k \geq 1$, $j_1,\dots,j_n \geq 0$ with $j_1+\cdots+j_n = k$ we find

$$P[Z_1=j_1,\dots,Z_n=j_n|Z=k] \overset{(1.32)-(1.33)}{=} \frac{k!}{j_1!\dots j_n!}\Big(\frac{\lambda_1}{\lambda}\Big)^{j_1}\dots\Big(\frac{\lambda_n}{\lambda}\Big)^{j_n}$$

and this proves (1.31), since $\frac{\lambda_i}{\lambda} = p_i, 1 \leq i \leq n$.

- <u>(1.31) $\Longrightarrow$ (1.30)</u>:

For $j_1,\dots,j_n \geq 0$, with $k = j_1+\cdots+j_n \geq 1$, we have:

$$\begin{aligned} &P[Z_1=j_1,\dots,Z_n=j_n] = P[Z_1=j_1,\dots,Z_n=j_n|Z=k]\;P[Z=k] \\ &= \Big(\frac{k!}{j_1!\dots j_n!}\,p_1^{j_1}\dots p_n^{j_n}\Big)\Big(e^{-\lambda}\,\frac{\lambda^k}{k!}\Big) = e^{-\lambda_1+\cdots+\lambda_n}\,\frac{\lambda^{j_1+\cdots+j_n}}{j_1!\dots j_n!}\Big(\frac{\lambda_1}{\lambda}\Big)^{j_1}\dots\Big(\frac{\lambda_n}{\lambda}\Big)^{j_n} \\ &= \prod_{i=1}^n e^{-\lambda_i}\,\frac{\lambda_i^{j_i}}{j_i!}\,. \end{aligned}$$

Now, when $j_1 = \cdots = j_n = 0$, $P[Z_1 = \cdots = Z_n = 0] = P[Z=0] = e^{-\lambda} = e^{-\lambda_1}\dots e^{-\lambda_n}$, and (1.30) follows. □

We will now prove (1.7) assuming (1.9). Observe that when $n$, $k \geq 1$, $0 < t_1 < t_2 < \cdots < t_n = t$, $j_1,\dots,j_n \geq 0$, with $k = j_1+\cdots+j_n$,

$$P[N_{t_1}=j_1,\,N_{t_2}-N_{t_1}=j_2,\dots,\underbrace{N_{t_n}}_{=t}-N_{t_{n-1}}=j_n|N_t=k] \overset{(1.9)}{=}$$

with $U_1,\dots,U_k$ uniformly distributed on $[0,t]$ and independent

$$P\Big[\sum_{i=1}^k 1\{U_i\in[0,t_1]\} = j_1, \sum_{i=1}^k 1\{U_i\in(t_1,t_2]\} = j_2,\dots,\sum_{i=1}^k 1\{U_i\in(t_{n-1},t]\} = j_n\Big] =$$

$$\frac{k!}{j_1!\dots j_n!}\Big(\frac{t_1}{t}\Big)^{j_1}\Big(\frac{t_2-t_1}{t}\Big)^{j_2}\dots\Big(\frac{t-t_{n-1}}{t}\Big)^{j_n},$$

that is $N_{t_1}$, $N_{t_2}-N_{t_1},\dots,N_t-N_{t_{n-1}}$ conditional on $N_t = k$ are multinomial $(k; \frac{t_1}{t}, \frac{t_2-t_1}{t},\dots,\frac{t-t_{n-1}}{t})$-distributed. With Lemma 1.5 we see that $N_{t_1}$, $N_{t_2}-N_{t_1},\dots,N_t-N_{t_{n-1}}$ are independent and respectively Poisson($\lambda t_1$), Poisson($\lambda(t_2-t_1)$)$,\dots,$ Poisson($\lambda(t-t_{n-1})$)-distributed. This proves (1.7) and concludes the proof of Theorem 1.3. □

**Remark 1.6.** One can prove a variant of Theorem 1.3, for which one does not make any assumption on the jumps of a counting process $(N_t)_{t\geq 0}$, such that $N_0 = 0$. In this set-up one has

(1.6) $\Longleftrightarrow$ (1.7) $\Longleftrightarrow$ (1.8)' $\Longleftrightarrow$ (1.9)', where
(1.8)' is the condition (1.8) and $P$-a.s. the jumps of $(N_t)$ have size 1, and
(1.9)' is the condition (1.9) and $P$-a.s. the jumps of $(N_t)$ have size 1.

The only change in the proof given above concerns the step (1.7) $\Longrightarrow$ (1.8)', where one argues as for (1.7) $\Longrightarrow$ (1.8) and one also observes that for $t > 0$

$$\begin{aligned}&P\big[N_{\cdot} \text{ has a jump of size} \geq 2 \text{ on } [0,t]\big] \leq \\ &P\Big[\bigcup_{k=1}^{n} \big\{N_{k\frac{t}{n}} - N_{(k-1)\frac{t}{n}} \geq 2\big\}\Big] \leq n\, P[N_{\frac{t}{n}} \geq 2] \overset{\text{see } (1.7) \Longrightarrow (1.6)}{=} n\, o\Big(\frac{1}{n}\Big) \underset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

and hence $P[N_{\cdot}$ has a jump of size $\geq 2] = 0$, which allows to prove (1.8)'. $\square$

We will now discuss some properties of the Poisson process. We begin with the fact that it satisfies the Markov property (more about this topic in Chapter 4).

**Proposition 1.7.** *(time-homogeneous Markov property)*

*Consider $(N_t)_{t\geq 0}$ a Poisson process with rate $\lambda > 0$, $f\colon \mathbb{N} \to \mathbb{R}$, bounded, $0 = t_0 < t_1 < \dots < t_\ell = t$, $s \geq 0$. Then, one has*

$$E[f(N_{t+s}) \,|\, N_{t_0}, \dots, N_{t_\ell}] = (R_s\, f)(N_t), \ \mathbb{P}\text{-a.s.}, \tag{1.34}$$

*where for $n \geq 0$ and $u \geq 0$,*

$$(R_u f)(n) = \sum_{m\geq n} f(m)\, e^{-\lambda u}\, \frac{(\lambda u)^{m-n}}{(m-n)!} \qquad (= E[f(N_u + n)]). \tag{1.35}$$

*Moreover,*

$$\begin{aligned}&(R_u)_{u\geq 0} \text{ is a semigroup of bounded operators on } L^\infty(\mathbb{N}),\\ &(\text{i.e. } R_{u+v} = R_u\, R_v \text{ for } u,v \geq 0).\end{aligned} \tag{1.36}$$

*Proof.*

$$E[f(N_{t+s}) | N_{t_0}, \dots, N_{t_\ell}] = E\big[f(\underbrace{N_{t+s} - N_t}_{\text{independent from}} + \underbrace{N_t) \,|\, N_{t_0}, \dots, N_{t_\ell}\big]} \overset{P\text{-a.s.}}{=}$$

$$\sum_{k\geq 0} f(k + N_t)\, e^{-\lambda s}\, \frac{(\lambda s)^k}{k!} = (R_s f)(N_t),$$

and (1.34) follows. Note that for $s_1, s_2 \geq 0$, $n \geq 0$, and $f$ as above:

$$
\begin{aligned}
& R_{s_1}(R_{s_2} f)(n) = R_{s_1}\Big( \sum_{m_2 \geq \cdot} f(m_2)\, e^{-\lambda s_2} \, \frac{(\lambda s_2)^{m_2 - \cdot}}{(m_2 - \cdot)!} \Big)(n) = \\
& e^{-\lambda s_1} \sum_{m_1 \geq n} \frac{(\lambda s_1)^{m_1 - n}}{(m_1 - n)!} \sum_{m_2 \geq m_1} f(m_2)\, e^{-\lambda s_2} \, \frac{(\lambda s_2)^{m_2 - m_1}}{(m_2 - m_1)!} = \\
& e^{-\lambda(s_1+s_2)} \sum_{m_2 \geq m_1 \geq n} f(m_2) \, \frac{(\lambda s_1)^{(m_1 - n)}}{(m_1 - n)!} \, \frac{(\lambda s_2)^{(m_2 - m_1)}}{(m_2 - m_1)!} \underset{\text{binomial formula}}{\overset{\text{summing over } m_1}{=}} \\
& e^{-\lambda(s_1+s_2)} \sum_{m_2 \geq n} f(m_2) \, \frac{\big(\lambda(s_1 + s_2)\big)^{m_2 - n}}{(m_2 - n)!} = R_{s_1+s_2} f(n),
\end{aligned}
\tag{1.37}
$$

and this proves (1.36). $\square$

**Remark 1.8.** Note that for $t > 0$, $n \geq 0$, and $f$ bounded on $\mathbb{N}$ we have

$$
\frac{1}{t}\, (R_t f - f)(n) = \frac{1}{t}\, (e^{-\lambda t} - 1)\, f(n) + \lambda e^{-\lambda t} f(n+1) + \sum_{k \geq 2} \frac{e^{-\lambda t}}{t} \, \frac{(\lambda t)^k}{k!}\, f(n+k). \tag{1.38}
$$

As a result, letting $t \to 0$, only the first two terms in the right-hand side of (1.38) contribute and we find:

$$
\lim_{t \to 0} \frac{1}{t}\, (R_t f - f)(n) = Lf(n) \text{ (the convergence is uniform in } n) \tag{1.39}
$$

where

$$
Lf(n) = \lambda\big(f(n+1) - f(n)\big), \text{ for } n \geq 0, \tag{1.40}
$$

is a bounded operator on $L^\infty(\mathbb{N})$ called the **generator of the semigroup** $(R_t)_{t \geq 0}$. $\square$

## 1.1 Stationary Poisson process on $\mathbb{R}$

We consider $(N_t^+)_{t \geq 0}$ and $(N_t^-)_{t \geq 0}$, two independent Poisson processes with rate $\lambda$, with respective jump times $(S_i^+)_{i \geq 1}$, $(S_i^-)_{i \geq 1}$. We organize these two sequences into a doubly infinite sequence on $\mathbb{R}$, $(S_k)_{k \in \mathbb{Z}}$:

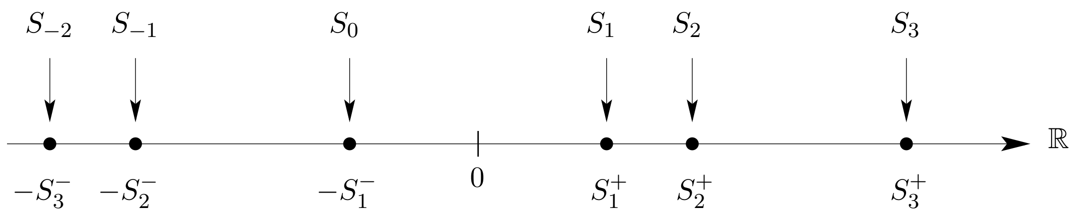



Fig. 1.3

$$S_k = S_k^+, \quad \text{if } k \ge 1, \qquad = -S^-_{1+|k|}, \quad \text{if } k \le 0. \tag{1.41}$$

As a result we see that the distribution of the $(S_k)_{k\in\mathbb{Z}}$ is determined by the property

$$\begin{array}{l} S_1, S_2 - S_1, \dots, S_{k+1} - S_k, \dots \\ -S_0, S_0 - S_{-1}, S_{-1} - S_{-2}, \dots \text{ are i.i.d. exponentially}(\lambda)\text{-distributed.} \end{array} \tag{1.42}$$

Of course the origin 0 plays a special role for the sequence $(S_k)_{k\in\mathbb{Z}}$ since one has:

$$S_0 < 0 < S_1\,. \tag{1.43}$$

We then choose $t > 0$ (the case $t < 0$ is similar) and let $t$ play the role of the new origin of time for the doubly infinite sequence. So we define

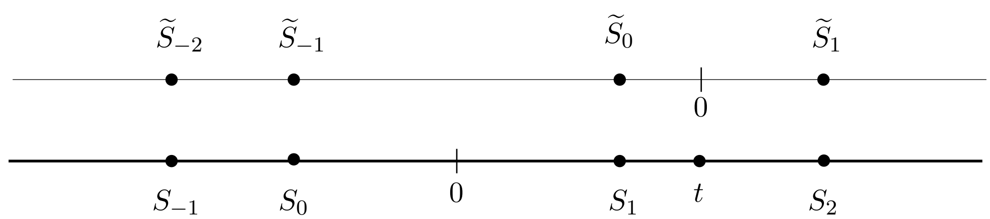



Fig. 1.4

$$\widetilde{S}_k = S_{N_t^+ + k} - t,\ k \in \mathbb{Z}\,. \tag{1.44}$$

**Theorem 1.9.** *(stationarity)*

$$(\widetilde{S}_k)_{k\in\mathbb{Z}} \textit{ has the same distribution as } (S_k)_{k\in\mathbb{Z}}. \tag{1.45}$$

*Proof.* We use the notation $N^+_{u-} \overset{\text{def}}{=} \lim_{v \uparrow u} N^+_v$, for $u > 0$, to denote the left-limit of $N^+_{\cdot}$ at $u$ (it only differs from $N^+_u$ when $u$ is a jump time of $N^+_{\cdot}$). We then define the counting processes:

$$\widetilde{N}^+_s \overset{\text{def}}{=} N^+_{t+s} - N^+_t,\ s \geq 0, \tag{1.46}$$

$$\widetilde{N}^-_s \overset{\text{def}}{=} N^+_t - N^+_{(t-s)-},\ \textbf{if } \boldsymbol{0 \leq s < t},\ \overset{\text{def}}{=}\ N^-_{s-t} + N^+_t,\ \textbf{if } \boldsymbol{s \geq t}. \tag{1.47}$$

Note that $\widetilde{N}^+_0 = 0$, but $\widetilde{N}^-_0 = 1$ possibly, when $t$ is a jump time of $N^+_{\cdot}$ (this event has probability 0, cf. (1.25)). When $t$ is not a jump time of $N^+_{\cdot}$, we have:

$$\begin{aligned} \widetilde{S}_k,\ k \geq 1, &\ \text{ are the jump times of } (\widetilde{N}^+_{\cdot}), \quad \text{and} \\ -\widetilde{S}_{1-k},\ k \geq 1, &\ \text{ are the jump times of } (\widetilde{N}^-_{\cdot}). \end{aligned} \tag{1.48}$$

Moreover, $(\widetilde{N}^+_{\cdot})$ and $(\widetilde{N}^-_{\cdot})$ have independent stationary increments and both $\widetilde{N}^+_s$, $\widetilde{N}^-_s$ are Poisson($\lambda s$)-distributed (note that for each $s \in [0,t)$, $P$-a.s., $\widetilde{N}^-_s = N^+_t - N^+_{t-s}$, since for each $u \geq 0$, $P[N^+_u \neq N^+_{u-}] = 0$). Also $N^+_{\cdot}$ and $N^-_{\cdot}$ are independent. With (1.7) we see that $(\widetilde{N}^+_s)_{s \geq 0}$ and $(\widetilde{N}^-_s\, 1\{\widetilde{N}^-_0 = 0\})_{s \geq 0}$ are two independent Poisson processes with rate $\lambda > 0$ (of course $P[\widetilde{N}^-_0 \neq 0] = P[t \text{ is a jump time of } N^+_{\cdot}] = 0$). From (1.41) and (1.48) we have expressed

$$(S_k)_{k \in \mathbb{Z}} \overset{(1.41)}{=} g\big((S^+_i)_{i \geq 1}, (S^-_i)_{i \geq 1}\big), \text{ and } P\text{-a.s. } (\widetilde{S}_k)_{k \in \mathbb{Z}} \overset{(1.48)}{=} g\big((\widetilde{S}^+_i)_{i \geq 1}, (\widetilde{S}^-_i)_{i \geq 1}\big),$$

with $\widetilde{S}^+_i$, $i \geq 1$, the jump times of $\widetilde{N}^+_{\cdot}$, $\widetilde{S}^-_i$, $i \geq 1$, the jump times of $\widetilde{N}^-_{\cdot}\, 1\{\widetilde{N}^-_0 = 0\}$. The claim (1.45) follows. □

We will now apply our discussion of the stationary Poisson process on $\mathbb{R}$ to the so-called **age** and **excess** processes.

We thus consider $(N_s)_{s \geq 0}$ a Poisson process with rate $\lambda > 0$, and for $t \geq 0$ we introduce (with the convention $S_0 = 0$ used for unilateral Poisson processes, cf. (1.8))

$$\begin{aligned} A_t &= t - S_{N_t} && \text{(age process), note that } A_t = t, \text{ when } N_t = 0, \\ E_t &= S_{N_t+1} - t && \text{(excess process).} \end{aligned} \tag{1.49}$$

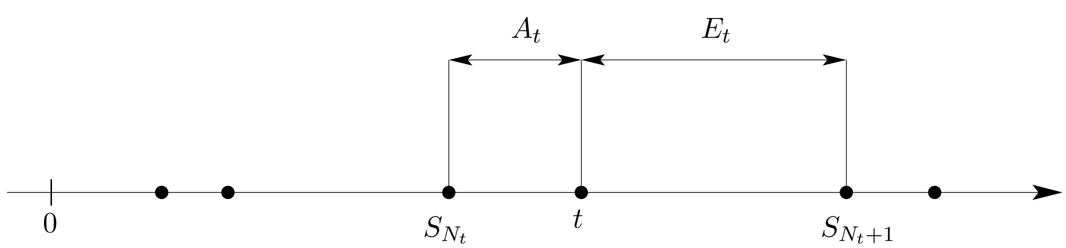



Fig. 1.5: An illustration of $A_t$ and $E_t$

**Proposition 1.10.** *($t > 0, (N_s)_{s\geq 0}$ Poisson process with rate $\lambda$)*

(1.50) *$(A_t, E_t)$ has same distribution as $(U \wedge t, V)$, where $U, V$ are independent exponential($\lambda$)-distributed random variables.*

*Proof.* If we choose an independent Poisson process with rate $\lambda$, $(N_s^-)_{s\geq 0}$, and set $N_s^+ = N_s$, for $s \geq 0$, we see that in the notation of (1.44) we have:

$$E_t = \widetilde{S}_1 \text{ and } A_t = (-\widetilde{S}_0) \wedge t, \tag{1.51}$$

and the claim follows from (1.45) of Theorem 1.9 and (1.42). □

**Remark 1.11.** The above result has an important consequence. It shows that the **length of the interval** between successive jump times, **which straddles $t$** (i.e. $L_t = E_t + A_t$) has not the same distribution as the $T_i, i \geq 1$ (the lengths between the successive jump times). **This length tends to be longer**, cf. [6], p. 12: "The waiting time paradox". Note that for $t \to \infty$:

$$\lim_{t\to\infty} E[L_t] = E[U] + E[V] = 2E[T_i] = \frac{2}{\lambda}\,.$$

□

## 1.2 Superposition and thinning of Poisson processes

We will now discuss some natural "operations" on Poisson processes. It will be useful to introduce the notion of "marking" of a Poisson process.

**Definition 1.12.** *Given a Poisson process $(N_t)_{t\geq 0}$, with rate $\lambda > 0$, an i.i.d. sequence $(X_n)_{n\geq 1}$ (possibly vector-valued), independent of $(N_t)_{t\geq 0}$, will be called a marking of the Poisson process.*

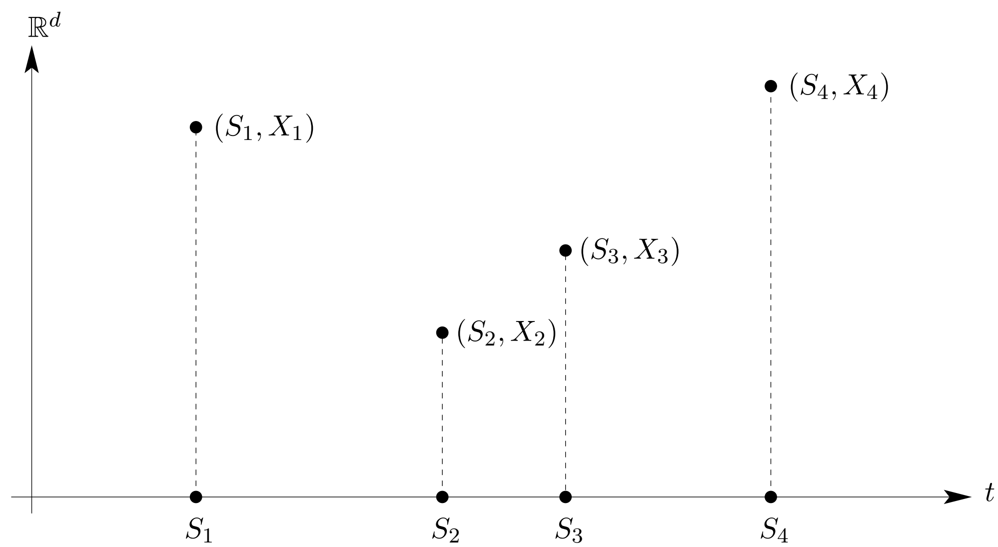



Fig. 1.6: A marked Poisson process

### 1.2.1 Thinning

We consider a Poisson process with rate $\lambda > 0$, $(N_t)_{t\ge 0}$, marked with Bernoulli variables $(X_n)_{n\ge 1}$, with success parameter $p \in (0,1)$. We define the **thinned processes**:

$$\begin{aligned} N_t^1 &= \sum_{k\ge 1} 1\{S_k \le t, X_k = 1\},\ t \ge 0, \\ N_t^0 &= \sum_{k\ge 1} 1\{S_k \le t, X_k = 0\},\ t \ge 0. \end{aligned} \tag{1.52}$$

**Remark 1.13.**

1) The Poisson process $(N_t)_{t\ge 0}$ can for instance describe the arrival times in a queue of clients, which can have two types "1" or "0". So the thinned processes $(N_t^1)_{t\ge 0}$ and $(N_t^0)_{t\ge 0}$ respectively describe the arrival times of clients of type 1 and of clients of type 0 in the queue.

2) Another interpretation of $(N_t^1)_{t\ge 0}$ is for instance when certain arrivals or events are not registered due to some defects (with the mark being 0 for such arrivals).

3) We of course have the identity

$$N_t = N_t^0 + N_t^1,\ t \ge 0. \tag{1.53}$$

□

The next proposition will be very useful.

**Proposition 1.14.** *Assume $(N_t)_{t\ge 0}$ is a Poisson process with rate $\lambda > 0$, with $\mathbb{R}^d$-valued marks $(X_n)_{n\ge 1}$, with common law $\mu$ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. For $A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d)$, with $A \subseteq [0,T] \times \mathbb{R}^d$, for some $T > 0$, define*

$$N(A) = \sum_{k\ge 1} 1\{(S_k, X_k) \in A\}. \tag{1.54}$$

*Then, setting $\nu(ds,dx) = \lambda\, 1\{s > 0\} ds \otimes \mu(dx)$,*

$$N(A) \text{ is Poisson}\big(\nu(A)\big)\text{-distributed}, \tag{1.55}$$

*and for $A_1, \dots, A_m \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d)$, pairwise disjoint (i.e. $A_i \cap A_j = \phi$, for $i \ne j$) and included in $[0,T] \times \mathbb{R}^d$, for some $T > 0$,*

$$N(A_1), \dots, N(A_m) \text{ are independent variables.} \tag{1.56}$$

*Proof.* We will compute the characteristic function of $(N(A_1), \dots, N(A_m))$. We thus consider $t_1, \dots, t_m \in \mathbb{R}$, $A_1, \dots, A_m$ as above and set

$$f(s,x) = \sum_{j=1}^m t_j\, 1_{A_j}(s,x), \text{ for } s \ge 0,\ x \in \mathbb{R}^d. \tag{1.57}$$

The characteristic function of the random vector $(N(A_1), \dots, N(A_m))$ is:

$$
\begin{aligned}
\varphi(t_1,\dots,t_m) \;&=\; E\Big[\exp\Big\{i \sum_{j=1}^m t_j\, N(A_j)\Big\}\Big] \\
&\overset{(1.54),(1.57)}{=} E\Big[\exp\Big\{i \sum_{n\ge 1} f(S_n, X_n)\Big\}\Big].
\end{aligned}
\tag{1.58}
$$

By (1.9) we know that $N_T$ is Poisson($\lambda T$) and conditional on $N_t = k$, $k \ge 1$, $S_1, \dots, S_k$ are obtained as a reordering of $U_1, \dots, U_k$, independent uniformly distributed variables on $[0,T]$.

As a result we find that:

$$
\varphi(t_1,\dots,t_m) = e^{-\lambda T} \sum_{k\ge 0} \frac{(\lambda T)^k}{k!}\, E\Big[\exp\Big\{i \sum_{n\ge 1} f(S_n,X_n)\Big\}\Big|\, N_T = k\Big]
\tag{1.59}
$$

with (1.9) and the independence of the i.i.d. $(X_n)_{n\ge 1}$:

$$
\begin{aligned}
&= e^{-\lambda T} \sum_{k\ge 0} \frac{(\lambda T)^k}{k!}\, \frac{1}{T^k} \int_{([0,T]\times\mathbb{R}^d)^k} e^{\,i\sum_{n=1}^k f(u_n,x_n)}\, du_1\, d\mu(x_1)\dots du_k\, d\mu(x_k) \\
&= e^{-\lambda T} \sum_{k\ge 0} \frac{\lambda^k}{k!} \Big(\int_{[0,T]\times\mathbb{R}^d} e^{if(u,x)} du\, d\mu(x)\Big)^k \\
&= \exp\Big\{-\lambda \int_{[0,T]\times\mathbb{R}^d} (1 - e^{if(u,x)})\, du\, d\mu(x)\Big\}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
1 - e^{if} &= 0 \;\text{ outside } \bigcup_{j=1}^m A_j \nwarrow \text{pairwise disjoint} \\
&= 1 - e^{it_j} \text{ on } A_j.
\end{aligned}
$$

In other words:

$$
1 - e^{if} = \sum_{j=1}^m (1 - e^{it_j})\, 1_{A_j},
$$

and as a result coming back to (1.59), we find that

$$
\begin{aligned}
\varphi(t_1,\dots,t_m) &= \exp\Big\{-\sum_{j=1}^m (1 - e^{it_j})\, \nu(A_j)\Big\} \\
&= \prod_{j=1}^m \underbrace{\exp\{-\nu(A_j)(1 - e^{it_j})\}}_{\substack{\text{characteristic function of a} \\ \text{Poisson}(\nu(A_j))\text{-variable at the point } t_j}}.
\end{aligned}
\tag{1.60}
$$

With the inversion formula for characteristic functions (cf. [4], p. 150), (1.55), (1.56) follow. □

We will now apply the above proposition to the study of the thinning of Poisson processes. We recall (1.52) for notation.

**Theorem 1.15.** *(thinning of Poisson processes)*

*$(N^0_t)_{t \ge 0}$ and $(N^1_t)_{t \ge 0}$ are independent Poisson processes with respective rates*

$$\lambda_0 = \lambda(1-p), \ \lambda_1 = \lambda p. \tag{1.61}$$

*Proof.* In view of (1.52), $(N^1_t)_{t \ge 0}$ and $(N^0_t)_{t \ge 0}$ are counting processes with jumps of size 1, moreover one has $N^1_{t=0} = 0 = N^0_{t=0}$. Further given $t_0 = 0 < t_1 < \dots < t_n$, in the notation of (1.54) we have

$$\begin{aligned} & N^1_{t_1} = N([0,t_1] \times \{1\}), \ N^1_{t_2} - N^1_{t_1} = N\big((t_1,t_2] \times \{1\}\big), \dots, N^1_{t_n} - N^1_{t_{n-1}} = \\ & \qquad N\big((t_{n-1},t_n] \times \{1\}\big) \ \text{ and} \\ & N^0_{t_1} = N([0,t_1] \times \{0\}), \ N^0_{t_2} - N^0_{t_1} = N\big((t_1,t_2] \times \{0\}\big), \dots, N^0_{t_n} - N^0_{t_{n-1}} = \\ & \qquad N\big((t_{n-1},t_n] \times \{0\}\big). \end{aligned} \tag{1.62}$$

As a result of (1.55), (1.56) these are independent Poisson variables, and the respective parameters are given as in (1.55) with $\nu(ds,dx) = \lambda 1\{s > 0\} ds \otimes (p\delta_1(dx) + (1-p)\,\delta_0(dx))$, so that $N^1_{t_1}$ is Poisson$(\lambda p t_1)$, $N^1_{t_2} - N^1_{t_1}$ is Poisson$(\lambda p(t_2 - t_1)), \dots, N^1_{t_n} - N^1_{t_{n-1}}$ is Poisson$(\lambda p(t_n - t_{n-1}))$.

Likewise $N^0_{t_1}$ is Poisson$(\lambda(1-p)t_1)$, $N^0_{t_2} - N^0_{t_1}$ is Poisson$(\lambda(1-p)(t_2-t_1)), \dots, N^0_{t_n} - N^0_{t_{n-1}}$ is Poisson$(\lambda(1-p)(t_n - t_{n-1}))$.

Therefore by Theorem 1.3 $(N^1_s)_{s \ge 0}$, $(N^0_s)_{s \ge 0}$ are independent Poisson processes with respective rates $\lambda_1 = \lambda p$, $\lambda_0 = \lambda(1-p)$. □

Our next item is the discussion of the superposition of Poisson processes.

### 1.2.2 Superposition

We now consider $(N^0_t)_{t \ge 0}$ and $(N^1_t)_{t \ge 0}$ two independent Poisson processes with respective rates $\lambda_0 > 0$ and $\lambda_1 > 0$, and the superposition process:

$$N_t = N^0_t + N^1_t, \ t \ge 0. \tag{1.63}$$

We denote by $S^0_k$, $k \ge 1$, and $S^1_k$, $k \ge 1$, the respective jump times of $(N^0_t)_{t \ge 0}$ and $(N^1_t)_{t \ge 0}$. In view of (1.25) we see that

$$P\text{-a.s., for } k_0, k_1 \ge 1, \ S^0_{k_0} \ne S^1_{k_1} \text{ and are both finite.} \tag{1.64}$$

We can then introduce the successive jump times $S_k$, $k \ge 1$, for $N_t$, $t \ge 0$, and note that:

$$\{S_k(\omega), k \ge 1\} = \{S^0_{k_0}(\omega); k_0 \ge 1\} \cup \{S^1_{k_1}(\omega); k_1 \ge 1\}. \tag{1.65}$$

$P$-a.s. disjoint subsets in $(0,\infty)$

We then introduce the random variables $X_k, k \ge 1$, via:

$$X_k(\omega) = 1\{\text{the } k\text{-th jump time of } (N_t)_{t\ge 0} \text{ is finite and a jump time of } (N_t^1)_{t\ge 0}\} \stackrel{\text{def}}{=} g_k\big((N_t^1)_{t\ge 0}, (N_t^0)_{t\ge 0}\big). \tag{1.66}$$

In view of (1.64),

$$P\text{-a.s., } 1 - X_k(\omega) = 1\{\text{the } k\text{-th jump time of } (N_t)_{t\ge 0} \text{ is finite and a jump time of } (N_t^0)_{t\ge 0}\}.$$

**Theorem 1.16.** *(superposition of Poisson processes)*

*$(N_t)_{t\ge 0}$ is (up to a change on a negligible set) a Poisson process with rate $\lambda = \lambda_0 + \lambda_1$, and $(X_k)_{k\ge 1}$ constitute a marking of this Poisson process with Bernoulli variables having success probability*

$$p = \frac{\lambda_1}{\lambda_0 + \lambda_1}. \tag{1.67}$$

*Proof.* We will use the result on thinning. We thus introduce on some auxiliary probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P})$ a Poisson process $(\widetilde{N}_t)_{t\ge 0}$ with rate $\lambda = \lambda_0 + \lambda_1$, and independent marks $(\widetilde{X}_k)_{k\ge 1}$, which are Bernoulli($p$)-distributed, where $p = \frac{\lambda_1}{\lambda_0+\lambda_1}$, as in (1.67). If we denote with $(\widetilde{N}_t^1)_{t\ge 0}$ and $(\widetilde{N}_t^0)_{t\ge 0}$, the thinned processes, cf. (1.52), we know from Theorem 1.15, that $(\widetilde{N}_t^0)_{t\ge 0}$ and $(\widetilde{N}_t^1)_{t\ge 0}$ are independent Poisson processes with respective rates $(\lambda_0 + \lambda_1)(1-p) = \lambda_0$ and $(\lambda_0 + \lambda_1)\,p = \lambda_1$. In particular:

$$\big((\widetilde{N}_t^0)_{t\ge 0}, (\widetilde{N}_t^1)_{t\ge 0}\big) \text{ has same distribution as } \big((N_t^0)_{t\ge 0}, (N_t^1)_{t\ge 0}\big). \tag{1.68}$$

As a result:

$$(\widetilde{N}_t)_{t\ge 0} = (\widetilde{N}_t^0 + \widetilde{N}_t^1)_{t\ge 0} \text{ has same distribution as } (N_t)_{t\ge 0} = (N_t^0 + N^1)_{t\ge 0},$$

and since $\widetilde{N}_t$ is a Poisson process with parameter $\lambda_0 + \lambda_1$, we already find that

$$(N_t)_{t\ge 0} \text{ is a Poisson process with rate } \lambda = \lambda_0 + \lambda_1. \tag{1.69}$$

With the same notation as in (1.66), it also follows from (1.68) that

$$\Big((\widetilde{N}_t^0)_{t\ge 0},\ (\widetilde{N}_t^1)_{t\ge 0},\ \big(g_k\big((\widetilde{N}_t^1)_{t\ge 0},\ (\widetilde{N}_t^0)_{t\ge 0}\big)\big)_{k\ge 1}\Big) \text{ has same distribution as } \big((N_t^0)_{t\ge 0},\ (N_t^1)_{t\ge 0},\ (X_k)_{k\ge 1}\big). \tag{1.70}$$

However, for $k \ge 1$,

$$g_k\big((\widetilde{N}_t^1)_{t\ge 0},\ (\widetilde{N}_t^0)_{t\ge 0}\big) = 1\big\{\text{the } k\text{-th jump of } (\widetilde{N}_t^0 + \widetilde{N}_t^1)_{t\ge 0} = (\widetilde{N}_t)_{t\ge 0} \text{ is finite and a jump of } (\widetilde{N}_t^1)_{t\ge 0}\big\} \stackrel{(1.52)}{=} \widetilde{X}_k.$$

As a result:
$$\big((N_t^0)_{t\ge 0},\,(N_t^1)_{t\ge 0},\,(X_k)_{k\ge 1}\big) \text{ has same distribution as } \big((\widetilde{N}_t^0)_{t\ge 0},\,(\widetilde{N}_t^1)_{t\ge 0},\,(\widetilde{X}_k)_{k\ge 1}\big),$$

and therefore since $(N_t)_{t\ge 0} = (N_t^0 + N_t^1)_{t\ge 0}$, $(\widetilde{N}_t)_{t\ge 0} = (\widetilde{N}_t^0 + \widetilde{N}_t^1)_{t\ge 0}$, we see that

$$\big((N_t)_{t\ge 0},\,(X_k)_{k\ge 1}\big) \text{ has same distribution as } \big((\widetilde{N}_t)_{t\ge 0},\,(\widetilde{X}_k)_{k\ge 1}\big). \tag{1.71}$$

This proves that $(X_k)_{k\ge 1}$ is a marking of $(N_t)_{t\ge 0}$, which is Bernoulli($p$)-distributed. □

**Remark 1.17.** We have in several instances invoked the notion of distribution of a counting process $(N_t)_{t\ge 0}$. What is meant is that we consider the canonical space

$$\Omega_c = \{\text{functions } w(\cdot) \text{ from } \mathbb{R}_+ \text{ into } \mathbb{N}, \text{ which are non-decreasing and right-continuous}\}, \tag{1.72}$$

endowed with the canonical $\sigma$-algebra $\mathcal{F}_c$ generated by the canonical coordinates $w \in \Omega_c \to w(t) \in \mathbb{N}$, where $t$ varies over $\mathbb{R}_+$:

$$\mathcal{F}_c = \sigma(w(t);\, t \ge 0). \tag{1.73}$$

A counting process $(N_t)_{t\ge 0}$ defined on $(\Omega, \mathcal{A}, P)$ can then be viewed as a random variable with values in $\Omega_c$, endowed with the $\sigma$-algebra $\mathcal{F}_c$, and the distribution of $(N_t)_{t\ge 0}$ is simply the image measure of $P$ under the measurable map $\omega \to (N_t(\omega))_{t\ge 0}$ from $(\Omega, \mathcal{A})$ to $(\Omega, \mathcal{F}_c)$.

In the same vein, the distribution of $((\widetilde{N}_t^0)_{t\ge 0},\,(\widetilde{N}_t^1)_{t\ge 0})$ in (1.68), is the image measure of $\widetilde{P}$ on $\Omega_c \times \Omega_c$ endowed with the product $\sigma$-algebra $\mathcal{F}_c \otimes \mathcal{F}_c$ under the measurable map $\widetilde{\omega} \in \widetilde{\Omega} \to ((\widetilde{N}_t^0(\widetilde{\omega}))_{t\ge 0},\,(\widetilde{N}_t^1(\widetilde{\omega}))_{t\ge 0}) \in \Omega_c \times \Omega_c$, and so on and so forth. □

## 1.3 Inhomogeneous Poisson process

**Definition 1.18.** *Given a continuous function $\rho$: $\mathbb{R}_+ \to (0,\infty)$, a counting process $(N_t)_{t\ge 0}$, with $N_0 = 0$, and jumps of size 1, is called inhomogeneous Poisson process with (instantaneous) rate $\rho(t)$, if $(N_t)_{t\ge 0}$ has independent increments and uniformly for bounded $t$, as $h \to 0$:*

$$P[N_{t+h} - N_t = 1] = \rho(t)\, h + o(h), \tag{1.74}$$

$$P[N_{t+h} - N_t \ge 2] = o(h). \tag{1.75}$$

*(This generalizes (1.6), which pertains to Poisson processes with constant rates).*

We will first discuss the existence of such processes.

### 1.3.1 Construction via time change

We introduce the function

$$R(t) = \int_0^t \rho(s)\, ds,\ t \geq 0, \tag{1.76}$$

and consider $(\widetilde{N}_t)_{t\geq 0}$ a Poisson process with rate $\lambda = 1$.

**Proposition 1.19.**

$$N_t \overset{\text{def}}{=} \widetilde{N}_{R(t)},\ t \geq 0, \text{ is an inhomogeneous Poisson process with rate } \rho(t). \tag{1.77}$$

*Proof.* Observe that $(N_t)_{t\geq 0}$ has independent increments, since $(\widetilde{N}_s)_{s\geq 0}$ also has independent increments. Moreover, as $h \to 0$, uniformly for $t \leq T$, one has:

$$\begin{aligned} P[N_{t+h} - N_t = 1] &= P[\underbrace{\widetilde{N}_{R(t+h)} - \widetilde{N}_{R(t)}}_{\nwarrow \text{ Poisson } (R(t+h)-R(t))} = 1] \\ &= \big(R(t+h) - R(t)\big)\, e^{-(R(t+h)-R(t))} \\ &= \Big(\int_t^{t+h} \rho(u)du\Big)\ \exp\Big\{-\int_t^{t+h} \rho(u)\, du\Big\} = \rho(t)\, h + o(h)\,. \end{aligned} \tag{1.78}$$

In a similar fashion, as $h \to 0$, uniformly for $t \leq T$,

$$\begin{aligned} &P[N_{t+h} - N_t \geq 2] = P[\widetilde{N}_{R(t+h)} - \widetilde{N}_{R(t)} \geq 2] = \\ &P[\widetilde{N}_{\underbrace{R(t+h)-R(t)}_{\leq C_T h}} \geq 2] \leq P[\widetilde{N}_{C_T h} \geq 2] = o(h). \\ &\qquad \nwarrow \text{ with } C_T = \sup\{\rho(s);\, 0 \leq s \leq T+1\} < \infty, \text{ and } h \leq 1 \end{aligned} \tag{1.79}$$

This proves the proposition. $\square$

When the function $\rho(\cdot)$ is bounded, we now provide another way to construct an inhomogeneous Poisson process with rate $\rho(\cdot)$.

### 1.3.2 Construction by variable thinning

We now assume in addition that

$$\sup_{t\geq 0} \rho(t) \leq C < \infty\,. \tag{1.80}$$

We consider $(\widetilde{N}_t)_{t\geq 0}$ a Poisson process with rate $C$, which is marked by i.i.d. variables $(\widetilde{X}_n)_{n\geq 1}$, uniformly distributed on $[0,1]$. We can then introduce $(N_t)_{t\geq 0}$, which is a **variable thinning** of $(\widetilde{N}_t)_{t\geq 0}$:

$$\begin{aligned} N_t = &\sum_{k\geq 1} 1\Big\{\widetilde{S}_k \leq t,\ \widetilde{X}_k \leq \frac{\rho(\widetilde{S}_k)}{C}\Big\} \text{ (compare with (1.52))}, \\ &\text{with } \widetilde{S}_k, k \geq 1 \text{ the jump times of } \widetilde{N}_{\cdot}\,. \end{aligned} \tag{1.81}$$

**Proposition 1.20.**

$$(N_t)_{t\geq 0} \textit{ is an inhomogeneous Poisson process with rate } \rho(t). \tag{1.82}$$

*Proof.* Define as in (1.55) for $A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d)$ with $A \subseteq [0,T] \times \mathbb{R}^d$, for some $T > 0$:

$$\widetilde{N}(A) = \sum_{k\geq 1} 1\{(\widetilde{S}_k, \widetilde{X}_k) \in A\}. \tag{1.83}$$

Then we have

$$N_t = \widetilde{N}(A), \text{ with } A = \Big\{(s,x) \in [0,t] \times [0,1];\ 0 \leq x \leq \frac{\rho(s)}{C}\Big\}, \tag{1.84}$$

and as a result of (1.55), where $\nu(ds,dx) = C\,1\{s > 0,\ 0 \leq x \leq 1\}\,ds\,dx$, we find:

$$N_t \text{ is Poisson } \Big(C \int_0^t ds \int_0^1 1\Big\{0 \leq x \leq \frac{\rho(s)}{C}\Big\}\, dx = R(t)\Big). \tag{1.85}$$

In the same way $N_{t+h} - N_t$ is Poisson$(R(t+h) - R(t))$-distributed, and with (1.56):

$$(N_t)_{t\geq 0} \text{ has independent increments.} \tag{1.86}$$

The claim now follows just as in (1.78), (1.79). □

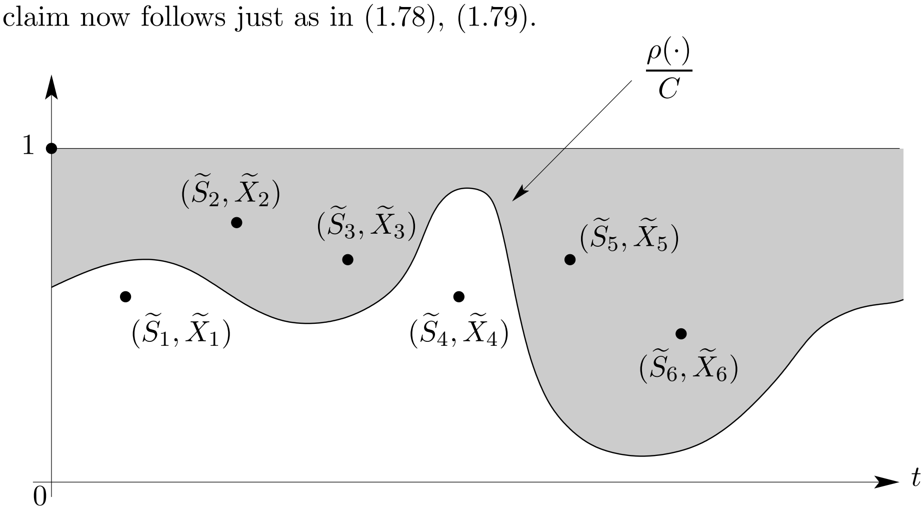



Fig. 1.7: The jump times $(S_k)_{k\geq 1}$ of $(N_t)_{t\geq 0}$ obtained by variable thinning of the jump times of $(\widetilde{N}_t)_{t\geq 0}$

We now return to the Definition 1.18 and investigate the structure of Poisson processes with rate $\rho(\cdot)$.

**Theorem 1.21.** *Let $(N_t)_{t\geq 0}$ be a Poisson process with rate $\rho(\cdot)$, then*

(1.87) $\quad N_t - N_s$ *is Poisson* $(R(t)-R(s))$*-distributed for* $0 \leq s < t$,

(1.88) $\quad$ *for $t > 0$ and $k \geq 1$, given $\{N_t = k\}$, the distribution of $(S_1, \dots S_k)$ is*
$$\mu(ds_1 \dots ds_k) = k!\, \rho(s_1) \dots \rho(s_k)\, R(t)^{-k}\, 1\{0 < s_1 < \dots < s_k \leq t\}\, ds_1 \dots ds_k.$$

*Proof.*

- (1.87): The argument is analogous to the proof of (1.6) $\Longrightarrow$ (1.7).

We define for $0 \leq s < t$, $n \geq 1$:

$$M_n = \sum_{k=1}^{n} 1\big\{N_{s+(t-s)\,\frac{k}{n}} - N_{s+(t-s)\,\frac{(k-1)}{n}} \geq 1\big\}. \tag{1.89}$$

It is a sum of independent Bernoulli variables with possibly different success parameters. Its characteristic function is

$$\begin{aligned}&\varphi_n(u) = E[\exp\{i\,u\,M_n\}] \overset{\text{indep.}}{=} \prod_{k=1}^{n} \big(1 + p_{n,k}(e^{iu}-1)\big), \text{ for } u \in \mathbb{R}, \text{ with}\\ &p_{n,k} = P[N_{s+(t-s)\,\frac{k}{n}} - N_{s+(t-s)\,\frac{(k-1)}{n}} \geq 1],\ 1 \leq k \leq n,\ n \geq 1.\end{aligned} \tag{1.90}$$

With the assumptions (1.74), (1.75), we find that

$$\sup_{1\leq k\leq n} p_{n,k} \underset{n\to\infty}{\longrightarrow} 0, \text{ and } \sum_{k=1}^{n} p_{n,k} \underset{n\to\infty}{\longrightarrow} \int_s^t \rho(v)\, dv = R(t) - R(s). \tag{1.91}$$

We use the analytic function $\log(1+z)$, for $|z| < 1$, and write for large $n$:

$$\varphi_n(u) = \exp\Big\{\sum_{k=1}^{n} \log\big(1 + p_{n,k}(e^{iu}-1)\big)\Big\}.$$

Note that when $n$ is large

$$\Big|\sum_{k=1}^{n} \log(1 + p_{n,k}(e^{iu}-1)\big) - \sum_{k=1}^{n} p_{n,k}(e^{iu}-1)\Big| \leq C \sum_{k=1}^{n} p_{n,k}^2 \underset{n\to\infty}{\overset{(1.91)}{\longrightarrow}} 0\,.$$

This fact and (1.91) now yield that for $u \in \mathbb{R}$

$$\varphi_n(u) \underset{n\to\infty}{\longrightarrow} \exp\big\{\big(R(t) - R(s)\big)(e^{iu}-1)\big\}\,. \tag{1.92}$$

characteristic function of the
Poisson$(R(t)-R(s))$-distribution
at the point $u \in \mathbb{R}$

From (1.92) we deduce with the "Continuity Theorem" for characteristic functions, cf. [4], p. 97, that

(1.93) $\qquad M_n$ converges in law to a Poisson $(R(t) - R(s))$-distribution.

The remainder of the proof of (1.87) is similar to (1.13), (1.14) and the explanation following (1.14). The claim (1.87) follows.

- (1.88): For $k \geq 1$, $0 \leq s_1 \leq t_1 < s_2 \leq t_2 < \cdots < s_k \leq t_k \leq t$:

(1.94)
$$\begin{aligned}
&P[s_1 < S_1 \leq t_1, s_2 < S_2 \leq t_2, \dots, s_k < S_k \leq t_k \,|\, N_t = k] = \\
&P[s_1 < S_1 \leq t_1, s_2 < S_2 \leq t_2, \dots, s_k < S_k \leq t_k, S_{k+1} > t]/P[N_t = k] = \\
&P[N_{s_1} = 0, N_{t_1} - N_{s_1} = 1, N_{s_2} - N_{t_1} = 0, \\
&\quad N_{t_2} - N_{s_2} = 1, \dots, N_{t_k} - N_{s_k} = 1, N_t - N_{t_k} = 0]/P[N_t = k] =
\end{aligned}$$

using (1.87) and the independence of increments

$$\begin{aligned}
&\exp\{-R(t)\} \prod_{i=1}^{k} \big(R(t_i) - R(s_i)\big) \times \frac{1}{\exp\{-R(t)\} \, \frac{R(t)^k}{k!}} = \\
&k! \prod_{i=1}^{k} \big(R(t_i) - R(s_i)\big) \, R(t)^{-k} = \mu\big((s_1, t_1] \times (s_2, t_2] \times \cdots \times (s_k, t_k]\big),
\end{aligned}$$

using the notation of (1.88).

The class of sets $A = (s_1, t_1] \times \cdots \times (s_k, t_k]$, with $0 \leq s_1 \leq t_1 < \cdots < s_k \leq t_k \leq t$, is a $\pi$-system of subsets of $U_t \overset{\text{def}}{=} \{(x_1, \dots, x_k);\, 0 < x_1 < \cdots < x_k \leq t\}$, which generate $\mathcal{B}(U_t)$. Note that the conditional law of $S_1, \dots, S_k$ given $\{N_t = k\}$ gives measure 1 to $U_t$ (indeed $P[S_\ell = S_{\ell+1}$ for some $\ell < k$ and $N_t = k] \leq \sum_{m=1}^{n} P[N_{t\frac{m}{n}} - N_{t\frac{(m-1)}{n}} \geq 2] \underset{n}{\longrightarrow} 0$ by (1.75)) and $\mu$ (from (1.88)) also gives measure 1 to $U_t$. So, with the help of Dynkin's lemma, see [12], p. 41, it follows from (1.94) that these two probabilities are equal. This concludes the proof of (1.88). □

**Remark 1.22.**

1) Note that the **assumption of uniformity over $t$ bounded in (1.74), (1.75)** is **important**. Indeed, consider $(N_t)_{t \geq 0}$ a Poisson process with rate 1, and

$$\widetilde{N}_t = N_t + [t] = N_t + \sum_{k \geq 1} 1\{k \leq t\}, \text{ for } t \geq 0.$$

Then $(\widetilde{N}_t)_{t \geq 0}$ has independent increments, and for any $t \geq 0$, $\widetilde{N}_{t+h} - \widetilde{N}_t$ is Poisson$(h)$ if $h > 0$ is small enough, and as a result **for each $t \geq 0$**:

$$P[\widetilde{N}_{t+h} - \widetilde{N}_t = 1] = h + o(h), P[\widetilde{N}_{t+h} - \widetilde{N}_t \geq 2] = o(h), \text{ as } h \to 0.$$

On the other hand, $\widetilde{N}_t$ is not Poisson($t$)-distributed, and one can modify $(\widetilde{N}_t)_{t\ge 0}$ on a set of measure 0 so that it only has jumps of size 1. This shows the importance of the uniformity assumption over bounded $t$ in Definition 1.18 if one wants the conclusions of Theorem 1.21 to hold.

2) On the other hand, in Definition 1.18, the **assumption** that $(N_t)_{t\ge 0}$ only has **jumps of size 1** is **not usual**. If this assumption is omitted, a counting process satisfying the other assumptions of the definition can be modified on a set of null-probability, in order to also satisfy the assumption of only having jumps of size 1, see Remark 1.6 for analogous considerations. □

We now continue our investigation of **inhomogeneous Poisson processes**. We are now ready to see that they all **have the form of a time change of a Poisson process with rate 1**, as in (1.77).

**Corollary 1.23.** *Let $\rho(\cdot)$: $\mathbb{R}_+ \longmapsto (0,\infty)$ be a continuous function, and $(N_t)_{t\ge 0}$ be a Poisson process with rate $\rho(\cdot)$. Then for a suitable Poisson process $(\widetilde{N}_s)_{s\ge 0}$ with rate 1, on a set of full probability one has:*

$$N_t = \widetilde{N}_{R(t)},\ t \ge 0. \tag{1.95}$$

*Proof.* As before, $R(t) = \int_0^t \rho(u)\,du$, $t \ge 0$.

a) When $\lim_{t\to\infty} R(t) = \infty$, we define

$$A_s = (\text{the inverse function of } R)(s),\ s \ge 0 \tag{1.96}$$

(incidentally note that $A'_s = \frac{1}{R'\circ A_s} = \frac{1}{\rho(A_s)}$), and

$$\widetilde{N}_s = N_{A_s}. \tag{1.97}$$

Then, by (1.87), we see that for $s > 0$, $\widetilde{N}_s$ is Poisson $(R(A_s) = s)$-distributed. Moreover, conditional on $\widetilde{N}_s = k \ge 1$, the variables $\widetilde{S}_1, \dots, \widetilde{S}_k$ (i.e. the first $k$ jumps of $\widetilde{N}_.$) coincide with $R(S_1), \dots, R(S_k)$ (where $S_1, \dots, S_k$ are the first $k$ jumps of $N_.$) and by (1.88) have distribution $k!\,1(0 < \widetilde{s}_1 < \dots < \widetilde{s}_k \le s)\, s^{-k}\, d\widetilde{s}_1 \dots d\widetilde{s}_k$. It now follows from (1.9) that $(\widetilde{N}_s)_{s\ge 0}$ is a Poisson process with rate 1. Then (1.97) implies that

$$\widetilde{N}_{R(t)} = N_{A_{R(t)}} = N_t, \text{ for } t \ge 0, \text{ and (1.95) follows.}$$

b) In case $\lim_{t\to\infty} R(t) = R_\infty < \infty$, we then pick an independent Poisson process with rate 1, $(N'_u)_{u\ge 0}$ (this means we are now working on an "enlarged space' of the form $(\Omega \times \Omega', \mathcal{A} \otimes \mathcal{A}', P \times P')$, if $(\Omega, \mathcal{A}, P)$ is the probability space where $(N_t)_{t\ge 0}$ is defined. Then we set

$$\begin{aligned} &A_s = (\text{the inverse function of } R)(s), \text{ so that} \\ &A_. \text{ is an increasing bijection } [0, R_\infty) \to [0,\infty). \end{aligned} \tag{1.98}$$

Note that, using for instance the characteristic function $\exp\{R(t)(e^{iu}-1)\}$ of $N_t$, $N_t$ converges in law to a Poisson($R_\infty$)-distribution as $t \to \infty$, so that $N_\infty = \lim_{t\to\infty} N_t$ (recall $N_.$ is non-decreasing) is a.s. finite. We then define

$$
\begin{aligned}
\widetilde{N}_s &= 0, \;\text{ if } N_\infty = \infty \text{ (this event has 0-probability)},\\
&= N_{A_s}, \;\text{ if } N_\infty < \infty \text{ and } s < R_\infty,\\
&= N_\infty + N'_{s-R_\infty}, \;\text{ if } N_\infty < \infty \text{ and } s \ge R_\infty\,.
\end{aligned} \tag{1.99}
$$

With the independence of the increments of $N_.$ and (1.87), we see that for $0 < s_1 < \cdots < s_n < R_\infty$,

$$
\widetilde{N}_{s_1}, \widetilde{N}_{s_2} - \widetilde{N}_{s_1}, \dots, \widetilde{N}_{s_n} - \widetilde{N}_{s_{n-1}} \;\text{ are independent, respectively Poisson}(s_k - s_{k-1})\text{-distributed, } k = 1, \dots, n. \tag{1.100}
$$

Moreover, $\widetilde{N}_{R_\infty} - \widetilde{N}_{s_n} = \lim_{s\uparrow R_\infty} \widetilde{N}_s - \widetilde{N}_{s_n} \overset{\text{a.s.}}{=} N_\infty - \widetilde{N}_{s_n}$ is Poisson($R_\infty - s_n$)-distributed and independent of the increments in (1.100).

It then follows that $(\widetilde{N}_s)_{s\ge 0}$ is a counting process with $\widetilde{N}_0 = 0$, jumps of size 1, which also fulfills (1.7), with $\lambda = 1$. It is therefore a Poisson process with rate 1. Then we see from (1.98), (1.99) that a.s.

$$
N_t = \widetilde{N}_{R(t)}, \text{ for all } t \ge 0.
$$

This concludes the proof of (1.95). □

**Example 1.24.** (Record values of an i.i.d. sequence)
We consider $X_i, i \ge 1$, i.i.d. positive variables on some $(\Omega, \mathcal{A}, P)$ with density:

$$
f\colon \mathbb{R}_+ \to (0,\infty), \text{ which is continuous.} \tag{1.101}
$$

We introduce the process counting the **number of record values of the sequence $\boldsymbol{X_i, i \ge 1}$, up to level $\boldsymbol{t}$**:

$$
N_t = \sum_{i\ge 1} 1\{X_i \le t,\; \underbrace{X_i > \max(X_1, X_2, \dots, X_{i-1})}_{\text{“}X_i \text{ is a record”}}\}. \tag{1.102}
$$

**Proposition 1.25.**

$$
\begin{aligned}
&(N_t)_{t\ge 0} \textit{ is an inhomogeneous Poisson process with rate}\\
&\rho(t) = f(t)/(1-F(t)), \textit{ where } F(t) = \int_0^t f(u)du \textit{ for } t \ge 0.
\end{aligned} \tag{1.103}
$$

*Proof.* We first prove (1.74), (1.75). We pick $T > 0$, and write for $t \le T$, as $h \to 0$, considering the first index $i \ge 1$, such that $X_i > t$:

$$(1.104)\qquad P[N_{t+h} - N_t \ge 1] = \sum_{i \ge 1} P\big[\underbrace{X_i \in (t, t+h] \text{ and } \max(X_1, \dots, X_{i-1}) \le t}_{\text{disjoint events as } i \text{ varies}}\big] \overset{\text{indep.}}{=}$$

$$\sum_{i \ge 1} \int_t^{t+h} f(u) du \Big(\int_0^t f(u)\, du\Big)^{i-1} = \int_t^{t+h} f(u) du \big(1 - F(t)\big)^{-1} =$$

$$\frac{f(t)}{1 - F(t)}\, h + o(h) \text{ (uniformly in } t \le T, \text{ as } h \to 0).$$

In the same fashion we have

$$(1.105)\qquad \begin{aligned} &P[N_{t+h} - N_t \ge 2] \le \sum_{1 \le i < j} P\big[X_1 \le t, \dots, X_{i-1} \le t, X_i \in (t, t+h], \\ &X_{i+1} \le t+h, \dots, X_{j-1} \le t+h, X_j \in (t, t+h]\big] \overset{\text{indep.}}{=} \\ &\Big(\int_t^{t+h} f(u)\, du\Big)^2 \sum_{1 \le i < j} \Big(\int_0^t f(u)\, du\Big)^{i-1} \Big(\int_0^{t+h} f(u)\, du\Big)^{j-i-1} = \\ &\Big(\int_t^{t+h} f(u)\, du\Big)^2 \times \big(1 - F(t)\big)^{-1} \big(1 - F(t+h)\big)^{-1} = o(h), \\ &\text{(uniformly in } t \le T, \text{ as } h \to 0). \end{aligned}$$

With (1.104) and (1.105), the properties (1.74), (1.75) readily follow.

We will now see that

$$(1.106)\qquad (N_t)_{t \ge 0} \text{ has independent increments.}$$

We consider $0 < u < v$, and we introduce the successive times where $X_i$, $i \ge 1$, is above level $u$:

$$T_1 = \inf\{i \ge 1; X_i > u\},\ T_2 = \inf\{i > T_1, X_i > u\}, \dots, T_{j+1} = \inf\{i > T_j; X_i > u\}, \dots$$

as well as the $\sigma$-algebra:

$$(1.107)\qquad \begin{aligned} \mathcal{F} = \big\{A \in \mathcal{A};\ &\text{for each } k \ge 1, \text{ there is } B_k \in \sigma(X_1, \dots, X_{k-1}) \text{ with} \\ &A \cap \{T = k\} = B_k \cap \{T = k\}\big\}. \end{aligned}$$

Note for instance that for $r \le u$, no $X_i$ with $i \ge T_1$ can be a record below level $r$, and:

$$(1.108)\qquad N_r = \sum_{1 \le i < T_1} 1\{X_i \le r, X_i > \max(X_1, \dots, X_{i-1})\} \text{ is } \mathcal{F}\text{-measurable}$$

(to see this, one expresses $N_r$ on the event $\{T_1 = k\}$ as a measurable function of $X_1, \dots, X_{k-1}$).

Moreover, we also see that

$$\text{the sequence } X_{T_1}, \dots, X_{T_j}, \dots \text{ is i.i.d., independent of } \mathcal{F}. \tag{1.109}$$

Indeed, if $A \in \mathcal{F}$, and $f_1, \dots, f_j$ are bounded measurable on $\mathbb{R}$:

$$E[f_1(X_{T_1}) \dots f_j(X_{T_j}), A] = \sum_{k \ge 1} E\big[f_1(X_{T_1}) \dots f_j(X_{T_j}),\, B_k \cap \underbrace{\{T_1 = k\}}_{\substack{\| \\ \{X_k > u, X_1 \le u, \dots, X_{k-1} \le u\}}}\big]. \tag{1.110}$$

and using the i.i.d. character of the $X_i$, and the fact that on $\{T_1 = k\}$ $T_2, \dots, T_j, \dots$ are the successive times where $X_{k+1}$, $X_{k+2}, \dots$ is above level $u$, the last expression equals

$$\begin{aligned}
&\sum_{k \ge 1} P[B_k, X_1 \le u, \dots, X_{k-1} \le u]\; E[f_1(X_k), X_k > u]\; E[f_2(X_{T_1}) \dots f_j(X_{T_{j-1}})] = \\
&\underbrace{\Big(\sum_{k \ge 1} P[B_k \cap \{T_1 = k\}]\Big)}_{\substack{\| \\ P[A]}} E[f_1(X_1) \,|\, X_1 > u]\, E[f_2(X_{T_1}) \dots f_j(X_{T_{j-1}})] = \\
&P[A]\; E[f_1(X_1) \,|\, X_1 > u]\; E[f_2(X_{T_1}) \dots f_j(X_{T_{j-1}})].
\end{aligned}$$

Applying this identity with $A = \Omega$ and $j = 1, 2, \dots$, we see that

$$E[f_1(X_{T_1}) \dots f_j(X_{T_j}), A] = P[A] \prod_{1 \le \ell \le j} E[f_\ell(X_1) \,|\, X_1 > u], \tag{1.111}$$

and this proves (1.109).

We can now observe that in view of (1.109), (1.108)

$$N_v - N_u = \sum_{j \ge 1} 1\{X_{T_j} \le v;\; X_{T_j} > \max(X_{T_1}, \dots, X_{T_{j-1}})\}, \tag{1.112}$$

is independent of $\mathcal{F}$ and hence of $(N_r)_{0 \le r \le u}$ (see (1.108)). The claim (1.106) now follows. This concludes the proof of (1.103). $\square$

Note that when the $X_i, i \ge 1$ are i.i.d. exponential($\lambda$)-distributed, then $f(t) = \lambda\, e^{-\lambda t}$, $t \ge 0$, and $F(t) = \int_0^t f(u)\, du = 1 - e^{-\lambda t}$, so that $\rho(t) \equiv \lambda$. In other words:

$$\boxed{\begin{array}{l}\text{for i.i.d. exponential}(\lambda)\text{-variables the process of} \\ \text{record values is a Poisson process with rate } \lambda.\end{array}} \tag{1.113}$$

# 2 Renewal processes

In this chapter we will discuss a class of counting processes, which generalizes Poisson processes and allows more general inter-arrival distributions.

## 2.1 The set-up

**Definition 2.1.** *Given $(T_i)_{i\geq 1}$, i.i.d. variables with values in $[0,\infty)$, for which*

$$P[T_i = 0] < 1, \tag{2.1}$$

$$\mu = E[T_i] \leq \infty. \tag{2.2}$$

*The process with values in $\mathbb{N} \cup \{\infty\}$ defined by:*

$$N_t = \sum_{k\geq 1} 1\{S_k \leq t\} = \sup\{n \geq 0;\, S_n \leq t\},\ t \geq 0, \tag{2.3}$$

*where*

$$S_n = \sum_{i=1}^{n} T_i,\ n \geq 1,\ \textit{and}\ S_0 = 0, \tag{2.4}$$

*is called renewal process with inter-arrival distribution function $F(\cdot) = P[T_i \leq \cdot]$ (this last expression does not depend on $i \geq 1$).*

Observe that

$$(N_t)_{t\geq 0} \text{ is non-decreasing, right-continuous,} \tag{2.5}$$

$$P[N_t = \infty] = P[\text{for all } n;\, S_n \leq t] = 0, \text{ for any } t \geq 0, \tag{2.6}$$

thanks to the second lemma of Borel-Cantelli and the fact that $P[T_i > \alpha] > 0$, for some $\alpha > 0$, (c.f. (2.1)).

Note also that $N_t \geq n$ for $t \geq S_n$, and hence

$$\lim_{t\to\infty} N_t = \infty. \tag{2.7}$$

In view of (2.6), we see that we can always modify $(N_t)_{t\geq 0}$ on a set of measure 0, to obtain a counting process in the sense of (1.3). However $(N_t)_{t\geq 0}$ **may have jumps of size > 1**, since the variables $T_i, i \geq 1$, may take the value 0.

**Definition 2.2.**

$$M(t) = E[N_t],\ t \geq 0,\ \textit{is called the } \textbf{\textit{renewal function}}. \tag{2.8}$$

**Notation:**
We write $F^{*k}(\cdot)$ to denote the distribution function of $S_k$:

$$F^{*k}(t) = P[S_k \le t],\ t \in \mathbb{R},\ k \ge 0, \tag{2.9}$$

(note that $F^{*0}(t) = 1\{t \ge 0\}$ is the "Heavyside function").

As a result for $k \ge 1$:

$$\begin{aligned} F^{*k}(t) &= \int_{\mathbb{R}} F^{*(k-1)}(t-s)\,dF(s) = \int_0^t F^{*(k-1)}(t-s)\,dF(s),\ \text{if } t \ge 0,\\ &= 0,\ \text{if } t < 0. \end{aligned} \tag{2.10}$$

## 2.2 Some properties of $M(t)$

**Lemma 2.3.**

(2.11) $M(t) = \sum_{k=1}^{\infty} F^{*k}(t)$ *is non-decreasing right-continuous, and* $dM(t) = \sum_{k=1}^{\infty} dF^{*k}(t)$ *(where* $dF^{*k}$ *is the law of* $S_k$*),*

(2.12) *for* $t \ge 0$, $r \ge 1$, $E[N_t^r] < \infty$, *(in particular* $M(t) = E[N_t] < \infty$*),*

*for* $s \ge 0$, $\widehat{M}(s) \overset{\text{def}}{=} \int_0^\infty e^{-sx} dM(x)$ *is the Laplace transform of* $dM$ *and one has*

(2.13) $\widehat{M}(s) = \dfrac{\widehat{F}(s)}{1-\widehat{F}(s)}$, *where* $\widehat{F}(s) \overset{\text{def}}{=} \int_0^\infty e^{-sx} dF(x)$ $(= E[e^{-sT_1}])$ *is the Laplace transform of* $dF$.

*Proof.*

- (2.12):

As below (2.6) we choose $\alpha > 0$ with $P[T_i > \alpha] > 0$, and define for $i \ge 1$ the i.i.d. variables

$$\overline{T}_i \overset{\text{def}}{=} \alpha\, 1\{T_i \ge \alpha\} \le T_i,\ \text{for } i \ge 1, \tag{2.14}$$

as well as the corresponding renewal process:

$$\overline{N}_t = \sup\{n \ge 0; \overline{S}_n \le t\}, \tag{2.15}$$

where of course we have set $\overline{S}_0 = 0$, and $\overline{S}_n = \overline{T}_1 + \cdots + \overline{T}_n$, for $n \ge 1$.

Since $\overline{T}_i \le T_i$ for $i \ge 1$, it follows that $\overline{S}_n \le S_n$ for all $n \ge 0$, and hence:

$$\overline{N}_t \ge N_t,\ \text{for } t \ge 0. \tag{2.16}$$

Note that $\tau_1 = \inf\{i \ge 1; \overline{T}_i > 0\}$, $\tau_2 = \inf\{i > \tau_1; \overline{T}_i > 0\}, \dots \tau_{j+1} = \inf\{i > \tau_j; \overline{T}_i > 0\}$, are such that

(2.17) $\tau_1, \tau_2 - \tau_1, \dots \tau_{j+1} - \tau_j, \dots$ are i.i.d. geometric($P[T_1 \ge \alpha]$)-variables (i.e. $P[\tau = k] = p(1-p)^{k-1}$, for $k \ge 1$, where $p = P[T_1 \ge \alpha]$), and

(2.18) $\overline{S}_{\tau_j} = j\,\alpha$, when $\tau_j < \infty$ (which is $P$-a.s. the case).

As a result we see that for $j \ge 0$:

$$N_{j\alpha} \le \overline{N}_{j\alpha} \le \tau_{j+1}, \text{ (because } \overline{S}_{\tau_{j+1}} > j\,\alpha). \tag{2.19}$$

Since $E[\tau_{j+1}^r] < \infty$ for all $r \ge 1$, $j \ge 1$, in view of (2.17), the claim (2.12) follows.

- (2.11):

$M(t)$ is non-decreasing and right-continuous, thanks to the monotone convergence theorem and the properties of $(N_t)_{t \ge 0}$ (non-decreasing, right-continuous and integrable by (2.12)). Moreover:

$$\begin{aligned} M(t) &= E\Big[\sum_{k=1}^{\infty} 1\{S_k \le t\}\Big] \overset{\substack{\text{monotone}\\ \text{convergence}}}{=} \sum_{k \ge 1} P[S_k \le t] \\ &\overset{(2.9)}{=} \sum_{k=1}^{\infty} F^{*k}(t), \end{aligned} \tag{2.20}$$

and $dM(t) = \sum_{k=1}^{\infty} dF^{*k}(t)$ follows by Lebesgue-Stieltjes.

- (2.13):

$$\begin{aligned} \widehat{M}(s) &= \int_0^\infty e^{-sx} dM(x) \overset{(2.11)}{=} \int_0^\infty e^{-sx} \sum_{k=1}^{\infty} dF^{*k}(x) \\ &= \sum_{k \ge 1} \int_0^\infty e^{-sx} dF^{*k}(x) = \sum_{k \ge 1} E[e^{-s\,S_k}] \\ &\overset{\text{indep.}}{=} \sum_{k \ge 1} E[e^{-sT_1}]^k = \sum_{k \ge 1} \widehat{F}(s)^k = \frac{\widehat{F}(s)}{1 - \widehat{F}(s)}, \;\text{ if } s > 0. \end{aligned} \tag{2.21}$$

The equality extends to $s = 0$, since both members are infinite. □

**Remark 2.4.** The renewal function $M(t) = \sum_{k=1}^{\infty} F^{*k}(t)$ is typically difficult to compute, and the Laplace transform $\widehat{M}(s)$ is easier to calculate. However, the inversion of the Laplace transform is not a straightforward operation, see [6], p. 442. □

**Theorem 2.5.** *(elementary renewal theorem)*

$$\lim_{t \to \infty} \frac{M(t)}{t} = \frac{1}{\mu} \Big( = \frac{1}{E[T_1]} \Big) \in [0, \infty). \tag{2.22}$$

*Proof.* Before giving the proof we recall the **Wald identity**, cf. [4], p. 158. When $X_i, i \ge 1$ are i.i.d. variables with $E[|X_i|] < \infty$, and $\tau$ is a stopping time of the filtration $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$, $n \ge 1$, $\mathcal{F}_0 = \{\phi, \Omega\}$, with $E[\tau] < \infty$, then:

$$\begin{aligned} E[S_\tau] = E[X_1]\, E[\tau], \text{ with } S_n &= X_1 + \dots + X_n, \ n \ge 1 \\ &= 0, \ n = 0. \end{aligned} \tag{2.23}$$

We now continue the proof of (2.22). We begin with the case

- $\underline{\mu < \infty}$: Then $P$-a.s., for $t \ge 0$,

$$S_{N_t} \le t < S_{N_t+1}, \tag{2.24}$$

and moreover

$$\begin{aligned} &\tau \overset{\text{def}}{=} N_t + 1 = \inf\{k \ge 0, \, S_k > t\} \text{ is an } (\mathcal{F}_n)\text{-stopping time,} \\ &\qquad (\text{where } \mathcal{F}_n = \sigma(T_1, \dots, T_n), \text{ for } n \ge 1, = \{\phi, \Omega\}, \text{ for } n = 0). \end{aligned} \tag{2.25}$$

As a result of (2.23), (2.24), (2.25)

$$E[S_{N_t}] \le t < E[S_{N_t+1}] \overset{(2.23)}{=} \underbrace{E[N_t + 1]}_{\substack{\uparrow \\ \text{finite by (2.12)}}} \times \mu \overset{(2.8)}{=} \big(M(t) + 1\big)\, \mu.$$

Hence

$$\varliminf_{t \to \infty} \frac{M(t)}{t} \ge \frac{1}{\mu}\,. \tag{2.26}$$

Moreover, if we define for $c > 0$, the truncated variables

$$\overline{T}_i = T_i \wedge c \; (= \min(T_i, c)), \text{ for } i \ge 1, \tag{2.27}$$

the variables $\overline{T}_i$, $i \ge 1$, are i.i.d., $E[\overline{T}_i] < \infty$, and $\overline{T}_i \le T_i$, so that

$$\overline{N}_t \ge N_t \text{ and } \overline{M}(t) \ge M(t), \tag{2.28}$$

with $\overline{N}_t$, resp. $\overline{M}(t)$, the renewal process, resp. the renewal function, attached to the variables $\overline{T}_i$, $i \ge 1$. Moreover, one has $P$-a.s.

$$\overline{S}_{\overline{N}_t+1} \le t + c, \tag{2.29}$$

so that

$$t + c \ge E[\overline{S}_{\overline{N}_t+1}] \overset{(2.23)}{=} E[\overline{T}_1](\overline{M}(t) + 1) \overset{(2.28)}{\ge} E[\overline{T}_1](M(t) + 1), \tag{2.30}$$

and hence

$$\overline{\lim_{t\to\infty}} \; \frac{M(t)}{t} \le \frac{1}{E[\overline{T}_1]} = \frac{1}{E[T_1 \wedge c]}\,. \tag{2.31}$$

Letting $c \to \infty$, we find that $E[T_1 \wedge c] \uparrow E[T_1]$ so that

$$\overline{\lim_{t\to\infty}} \; \frac{M(t)}{t} \le \frac{1}{\mu}\,. \tag{2.32}$$

Together with (2.26) this proves (2.22).

- $\underline{\mu = \infty}$:

Using the same truncation technique, cf. (2.27), we obtain (2.31) and letting $c \to \infty$, we find

$$\lim_{t\to\infty} \frac{M(t)}{t} = 0, \tag{2.33}$$

and this proves (2.22). □

We right away state a refinement of the above result using the strong law of large numbers.

**Theorem 2.6.**

1) *If* $\mu \le \infty$,

$$P\text{-}a.s.,\ \lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{\mu} \;\big(\in [0,\infty)\big). \tag{2.34}$$

2) *If* $E[T_1^2] < \infty$ *and* $\sigma > 0$ *(writing* $\sigma^2 = \mathrm{var}(T_1)$*), then as* $t \to \infty$,

$$\frac{N_t - t/\mu}{\sigma\sqrt{t/\mu^3}} \text{ converges in law to an } N(0,1)\text{-distribution.} \tag{2.35}$$

*Proof.*

- For (2.34), see [4], p. 66, or [14], p. 40.

- For (2.35), see [12], p. 189. □

**Example 2.7.** (renewal reward process)

A non-profit organization is receiving at the times $S_i$, $i \ge 1$, of a renewal process the donations $D_i$, $i \ge 1$, which are independent and i.i.d. distributed ($D_i \ge 0$, with $E[D_i] < \infty$). The cumulative wealth received by the organization at time $t$ is:

$$R(t) = \sum_{i\ge 1} D_i \, 1\{S_i \le t\} \quad \text{(reward process)}. \tag{2.36}$$

By the strong law of large numbers and (2.34), $P$-a.s.,

$$\lim_{t\to\infty} \frac{R(t)}{t} = \lim_{t\to\infty} \underbrace{\frac{1}{N_t}\sum_{i=1}^{N_t} D_i}_{\substack{\downarrow \\ E[D_1]}} \times \underset{\searrow \frac{1}{\mu}}{\frac{N_t}{t}} \quad\longrightarrow\quad \frac{E[D_1]}{\mu},$$

that is

$$P\text{-a.s.} \quad \lim_{t\to\infty} \frac{R(t)}{t} = \frac{E[D_1]}{\mu}\,. \tag{2.37}$$

This result can of course be extended to the case where the $D_i$ are not necessarily non-negative (gains and costs), with $E[|D_i|] < \infty$. $\square$

## 2.3 Renewal with delays

This corresponds to the situation where the distribution of $T_1$ does not necessarily coincide with the common distribution of the $T_i, i \geq 2$:

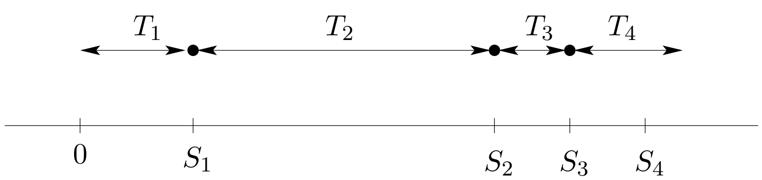



Fig. 2.1

(2.38) $T_i$, $i \geq 1$, are independent $\mathbb{R}$-valued, with $T_i, i \geq 2$, identically distributed and such that (2.1) holds (i.e. $P[T_i = 0] < 1$, for $i \geq 2$).

**Remark 2.8.**

1) Interpretation in terms of delay:

   If $\widetilde{T}_i$, $i \geq 1$, are i.i.d. $\mathbb{R}_+$-valued variables satisfying (2.1), and $\widetilde{S}_0$ is an independent $\mathbb{R}_+$-valued variable, "the delay", we can set as in (0.12), for $i \geq 1$,

$$\widetilde{S}_i = \widetilde{S}_0 + \widetilde{T}_1 \cdots + \widetilde{T}_i\,.$$

   To fall back on (2.38) we simply define

$$S_i = \widetilde{S}_{i-1}, i \geq 1 \text{ and } T_1 = \widetilde{S}_0,\ T_i = \widetilde{T}_{i-1}, \text{ for } i \geq 2.$$

   This explains the terminology "renewal with delay".

2) Given $(N_s)_{s\geq 0}$ a renewal process, one is very naturally led to a renewal process with delay, if one for instance considers the shifted renewal process after time $t$:

$$\overline{N}_s = N_{s+t} - N_t,\ s \geq 0 \qquad (P\text{-a.s. well-defined, cf. (2.6)}). \tag{2.39}$$

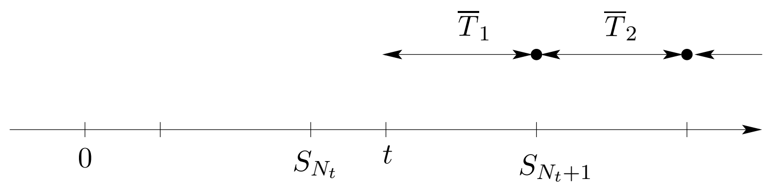



Fig. 2.2

So we see that for $s \geq 0$

$$\begin{aligned} \overline{N}_s = \sum_{k\geq 1} 1\{\overline{S}_k \leq s\}, \quad & \text{with } \overline{S}_k = \overline{T}_1 + \cdots + \overline{T}_k, \text{ for } k \geq 1, \text{ and} \\ & \overline{T}_1 = S_{N_t+1} - t, \text{ the “excess at time } t\text{”}, \\ & \overline{T}_i = T_{N_t+i}, \text{ for } i \geq 2. \end{aligned} \tag{2.40}$$

With (2.25), it is straightforward to infer that $\overline{T}_i$, $i \geq 1$, defined above satisfy (2.38). □

For a **renewal with delay** we use the following **notation**:

$$\begin{aligned} G(t) &: \text{ the distribution function of } T_1 \\ F(t) &: \text{ the distribution function of } T_i,\ i \geq 2, \\ \underset{\substack{\uparrow \\ \text{renewal function}}}{M(t)} &= E[N_t], \text{ with } \underset{\substack{\uparrow \\ \text{renewal process}}}{N_t} = \sum_{k\geq 1} 1\{S_k \leq t\} \end{aligned}$$

(note that $E[N_t] < \infty$ because of (2.12) and the interpretation of “delay”).

In analogy to (2.11), (2.13) of Lemma 2.3 we now have:

**Lemma 2.9.** *(renewal with delay)*

$$M(t) = \sum_{k\geq 0} G * F^{*k}(t),\ t \geq 0,\ \textit{and}\ dM(t) = \sum_{k=0}^{\infty} dG * F^{*k}(t) \tag{2.41}$$

$$\begin{aligned} \textit{(with the notation } G * F^{*k}(t) = & \int_0^t G(t-s)\, dF^{*k}(s),\ t \geq 0, \\ & \textit{for } k \geq 0 \textit{ and } t \geq 0, \\ = &\ 0,\ \textit{for } t < 0\textit{)}. \end{aligned}$$

$$\widehat{M}(s) = \frac{\widehat{G}(s)}{1 - \widehat{F}(s)}\,,\ \textit{for } s \geq 0. \tag{2.42}$$

□

In the case where

$$\mu(= E[T_2]) < \infty, \tag{2.43}$$

an **important special role** is played by the following **specific delay distribution function**:

$$\begin{aligned} G_*(x) &= \frac{1}{\mu} \int_0^x \big(1 - F(t)\big)\, dt,\ x \geq 0 \\ &= 0,\ x < 0. \end{aligned} \tag{2.44}$$

**Remark 2.10.**

1) Note that

$$\int_0^\infty \big(1 - F(t)\big)\, dt = \int_0^\infty P[T_2 > t]\, dt \overset{\text{Fubini}}{=} E\Big[\int_0^\infty 1\{T_2 > t\}\, dt\Big] = E[T_2] = \mu.$$

As a consequence:

$$\lim_{x \to \infty} G_*(x) = 1, \text{ and } G_* \text{ is indeed a distribution function.}$$

2) Note that in the case of the exponential($\lambda$)-distribution (i.e. $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$, $= 0$, when $x < 0$),

$$G_*(x) = \lambda \int_0^x e^{-\lambda t} dt = 1 - e^{-\lambda x}, \text{ for } x \geq 0, = 0, \text{ for } x < 0,$$

and $G_*$ coincides with $F$ in this special case. □

As we now see the delay distribution function $G_*$ induces a stationary behavior of the renewal process with delay (see also (1.45), in the case of the Poisson process).

**Theorem 2.11.** *($\mu < \infty$, stationary behavior)*

*For the renewal with delay corresponding to $G_*$ in (2.44),*

$$M(t) = \frac{t}{\mu}\,,\ \textit{for } t \geq 0. \tag{2.45}$$

Moreover, given $t > 0$, if one considers the inter-arrival times after time $t$:

$$\overline{T}_1 = S_{N_t+1} - t,\ \overline{T}_i = T_{N_t+i},\ i \geq 2, \quad \text{(see also (2.40))}, \tag{2.46}$$

then

$$(\overline{T}_i)_{i\geq 1} \text{ has same distribution as } (T_i)_{i\geq 1}. \tag{2.47}$$

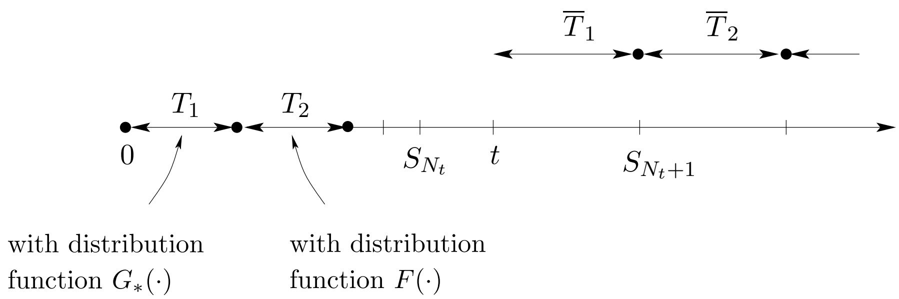



Fig. 2.3

*Proof.*

- (2.45): Note that for $s > 0$,

$$\widehat{G}_*(s) = \int_0^\infty e^{-sx}\, dG_*(x) \overset{(2.44)}{=}$$

$$\int_0^\infty \big(1 - F(x)\big)\, e^{-sx}\, \frac{dx}{\mu} = \frac{1}{s\mu} \int_0^\infty \underbrace{\big(1 - F(x)\big)}_{\substack{\| \\ P[T_2 > x]}} s\, e^{-sx} dx$$

(2.48)

$$= \frac{1}{s\mu} \int_0^\infty P[T_2 > x]\, se^{-sx} dx \overset{\text{Fubini}}{=} \frac{1}{s\mu}\, E\Big[\int_0^{T_2} se^{-sx} dx\Big] = \frac{1}{s\mu}\, E[1 - e^{-sT_2}]$$

$$= \frac{1 - \widehat{F}(s)}{s\mu}\,.$$

Therefore with (2.42) we find that

$$\widehat{M}(s) = \frac{\widehat{G}_*(s)}{1 - \widehat{F}(s)} = \frac{1 - \widehat{F}(s)}{s\mu}\, \frac{1}{1 - \widehat{F}(s)} = \frac{1}{s\mu} = \frac{1}{\mu} \int_0^\infty e^{-sx} dx. \tag{2.49}$$

This equality of the Laplace transforms shows that $M(t) = \frac{t}{\mu}$, $t \geq 0$, and proves (2.45), (see also [6], p. 432).

- Proof of (2.47):

Using the same argument as below (2.40) (which uses the independence of the $T_i$, $i \geq 1$, and the fact that $N_t + 1$ is an $(\mathcal{F}_n)$-stopping time, for $\mathcal{F}_n = \sigma(T_1, \dots, T_n)$, $n \geq 1$, $\mathcal{F}_0 = \{\phi, \Omega\}$), we deduce that

(2.50) $\quad \overline{T}_i$, $i \geq 1$, are independent and $\overline{T}_i$, $i \geq 2$, are distributed like $T_2$.

So we only need to check that $\overline{T}_1$ has same law as $T_1$ (i.e. has the distribution function $G_*(\cdot)$, from (2.44)). For $x > 0$, one has:

$$
\tag{2.51}
\begin{array}{l}
P[\overline{T}_1 > x] = P[\overline{T}_1 > x, N_t = 0] + P[\overline{T}_1 > x, N_t > 0] \overset{(2.46)}{=} \\
P[T_1 > x + t] + \sum\limits_{k \ge 1} P[S_k \le t, S_k + T_{k+1} > t + x] = ,
\end{array}
$$

conditioning on $S_k$ in the last sum this equals

$$
1 - G_*(x+t) + \sum_{k \ge 1} E[S_k \le t, 1 - F(t + x - S_k)]
$$

and since $S_k$ has distribution function $G_* * F^{*(k-1)}(\cdot)$, see (2.41), this equals

$$
\begin{array}{l}
1 - G_*(x+t) + \sum\limits_{k \ge 1} \displaystyle\int_0^t \big(1 - F(t + x - y)\big)\, d(G_* * F^{*(k-1)})(y) \overset{(2.41)}{=} \\[2ex]
1 - G_*(x+t) + \displaystyle\int_0^t \big(1 - F(t + x - y)\big)\, dM(y) \overset{(2.45)}{=} 1 - G_*(x+t) \\[2ex]
\qquad\qquad + \displaystyle\int_0^t \big(1 - F(t + x - y)\big)\, \frac{dy}{\mu}\,.
\end{array}
$$

Setting $u = t + x - y$ in the last integral, and using (2.44) for the first term, the above equals

$$
\begin{array}{l}
\dfrac{1}{\mu} \displaystyle\int_{x+t}^\infty \big(1 - F(u)\big)\, du + \dfrac{1}{\mu} \int_x^{x+t} \big(1 - F(u)\big)\, du = \\[2ex]
\dfrac{1}{\mu} \displaystyle\int_x^{+\infty} \big(1 - F(u)\big)\, du = P[T_1 > x].
\end{array}
$$

We thus see that $\overline{T}_1$ and $T_1$ have the same distribution and using (2.50) our claim (2.47) follows. □

**Remark 2.12.** Note that $P$-a.s.,

$$
\tag{2.52}
\overline{N}_s \overset{\text{def}}{=} N_{t+s} - N_t = \sum_{k \ge 1} 1\{\underset{\overline{T}_1 + \cdots + \overline{T}_k}{\underset{\uparrow}{\overline{S}_k}} \le s\}, \text{ for all } s \ge 0,
$$

and we see that the renewal process with delay $(\overline{N}_s)_{s \ge 0}$ has the same distribution as $(N_s)_{s \ge 0}$, when the delay has distribution $G_*$, this regardless of $t > 0$. This amplifies the statement (2.47) about stationarity of the renewal with delay (2.44). Note also that for $t, s \ge 0$

$$
\tag{2.53}
M(t+s) - M(t) = \frac{s}{\mu} = \overline{M}(s),
$$

(with $\overline{M}(s)$ the renewal function attached to $\overline{N}_.$). □

## 2.4 Blackwell's Renewal Theorem

This is an important strengthening of the elementary renewal theorem, cf. Theorem 2.5. We begin with

**Definition 2.13.** *A probability $\nu$ on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ with $\nu(0) < 1$ is called* ***arithmetic*** *if for some $a > 0$,*

$$\nu(\{0, a, 2a, 3a, \dots\}) = 1. \tag{2.54}$$

*The largest $a$ such that (2.54) holds is called the span of $\nu$. If there is no $a > 0$ for which (2.54) holds, then $\nu$ is called* ***non-arithmetic****.*

We say that a distribution function $F$ is arithmetic or non-arithmetic if the corresponding statement holds for the probability $dF$.

**Theorem 2.14.** *(Blackwell's renewal theorem)*

*If the distribution function $F$ of inter-arrival times is non-arithmetic, then*

$$\lim_{t\to\infty} M(t+h) - M(t) = \frac{h}{\mu}, \text{ for } h > 0, \text{ (with } \mu = \int_0^\infty x\,dF(x) \le \infty). \tag{2.55}$$

**Remark 2.15.**

1) A version of Blackwell's theorem also holds for arithmetic distributions, (cf. [12], pp. 221 and 238 or [6], p. 360). In this case one simply needs to restrict $h$ to multiples of the span of $dF$.

2) The convergence in (2.55) can be rephrased in terms of vague convergence of Radon measures on $\mathbb{R}_+$. Namely, if $\rho_t$ is the measure on $\mathbb{R}_+$ such that $\rho_t([0,s]) = M(t+s) - M(t)$, for $s \ge 0$, then

$$\lim_{t\to\infty} \int f(s)\, d\rho_t(s) = \int_0^\infty f(s)\, \frac{ds}{\mu}\ , \text{ for any } f \text{ continuous on } \mathbb{R}_+ \text{ with compact support} \tag{2.56}$$

((2.55) corresponds to $f = 1_{[0,h]}$, which is not continuous, but the equivalence can be proved along similar arguments as for the various definitions/characterizations of weak convergence, cf. [14], p. 50 and 53, and (2.56) expresses the fact that $\rho_t(ds)$ converges vaguely to $1_{[0,\infty)}\frac{ds}{\mu}$, as $t \to \infty$).

3) Blackwell's theorem immediately implies the elementary renewal theorem. Indeed, for $t \ge 1$,

$$\frac{M(t)}{t} = \frac{M(t) - M([t])}{t} + \sum_{k=0}^{[t]-1} \frac{M(k+1) - M(k)}{[t]} \times \underbrace{\frac{[t]}{t}}_{\substack{\searrow \text{ as } t\to\infty \\ 1}} + \frac{M(0)}{t}$$

$$\text{But: } \frac{M(t) - M([t])}{t} \le \frac{M([t]+1) - M([t])}{t} \overset{(2.55)}{\longrightarrow} 0\,, \text{ and } \frac{M(0)}{t} \to 0$$

and, since convergence of a sequence implies the convergence in the sense of Cesàro of the same sequence, $\frac{1}{[t]}\sum_{k=0}^{[t]-1} M(k+1) - M(k) \underset{t\to\infty}{\overset{(2.55)}{\longrightarrow}} \frac{1}{\mu}$. We thus recover the elementary renewal theorem:

$$\lim_{t\to\infty} \frac{M(t)}{t} = \frac{1}{\mu}\,.$$

□

For the time being we will content ourselves with a heuristic description of one of the proofs of Blackwell's theorem using coupling, initially due to T. Lindvall, see [12], p. 243. An other proof more analytical in spirit can be found in [6], p. 364.

*Sketch of proof* by "coupling" ($\mu < \infty$).

The idea is to consider the basic renewal process as well as an independent renewal process with the delay distribution $G_*$ of (2.44), $(\overline{T}_i)_{i\geq 1}$.

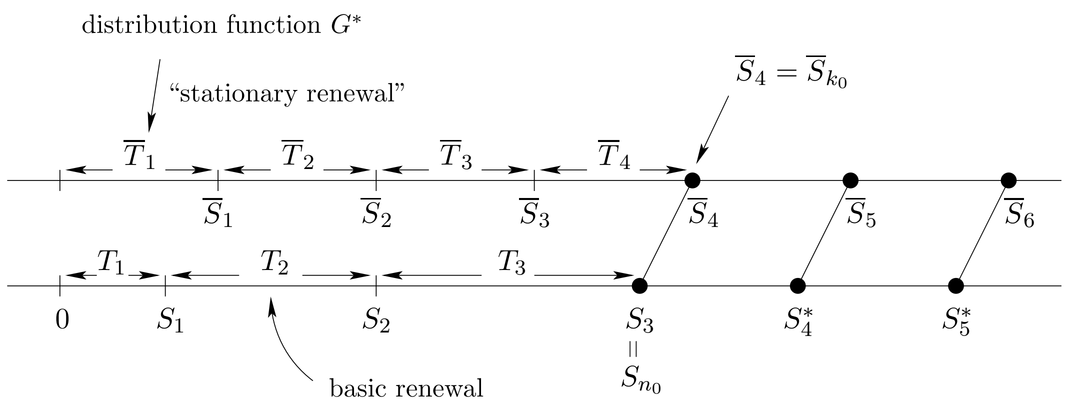



Fig. 2.4

At some point $S_{n_0}$ of the basic renewal process, there is "very closely" afterwards a point $\overline{S}_{k_0}$ in the stationary renewal. One builds a new sequence $S_n^*$ with

$$S_n^* = S_n, \text{ for } n \leq n_0,$$
$$S_{n_0+i}^* = S_{n_0} + \overline{S}_{k_0+i} - \overline{S}_{k_0}, \text{ for } i \geq 1.$$

One checks that $(S_n^*)_{n\geq 1}$ has same distribution as $S_n$. On the other hand, almost surely for large $t$,

$$\begin{aligned} N_{t+h}^* - N_t^* &= \sum_{k\geq 1} 1\{t < S_k^* \leq t+h\} \\ &\cong \sum_{j\geq 1} 1\{t < \overline{S}_j \leq t+h\} = \overline{N}_{t+h} - \overline{N}_t. \end{aligned}$$

But we also know that, cf. (2.45) or (2.53),

$$\overline{M}(t+h) - \overline{M}(t) \overset{\text{def}}{=} E[\overline{N}_{t+h} - \overline{N}_t] = \frac{h}{\mu}\,,$$

and in this way one infers that

$$\begin{aligned} M(t+h) - M(t) &= E[N_{t+h}] - E[N_t] = E[N^*_{t+h}] - E[N^*_t] \\ &\simeq \overline{M}(t+h) - \overline{M}(t) = \frac{h}{\mu}\,. \end{aligned}$$

Of course the above lines are not a proof, but merely carry out the intuition of the real proof. □

## 2.5 The renewal equation

An important feature of renewal processes is that "things start afresh" after the times $S_k$, $k \ge 1$. This element plays a crucial role in the derivation of so-called renewal equations for a variety of quantities of interest.

We consider

(2.57) $h\colon \mathbb{R}_+ \to \mathbb{R}$, measurable and bounded on compact intervals,

(2.58) $F$ the distribution function of a non-negative variable, not a.s. equal to zero.

**The $(h, F)$-renewal equation:**

One looks for measurable functions $g$ on $\mathbb{R}$, vanishing on $(-\infty, 0)$ with $g(t - \cdot) \in L^1(dF)$ for each $t \ge 0$, such that

$$\begin{aligned} g(t) &= h(t) + \int_0^t g(t-s)\, dF(s), \text{ for } t \ge 0, \\ &= 0, \text{ for } t < 0 \end{aligned} \tag{2.59}$$

(the notation "$\int_0^t$" means "$\int_{[0,t]}$"). In compact notation (2.59) is

$$g = h + g * F \text{ (where } g * F \text{ stands for the convolution of } g \text{ with } dF).$$

There is a rich collection of examples where such renewal equations occur. They are related with the widespread occurrence of mechanisms of "regeneration" or "renewal". We will later introduce a more restrictive notion of solution that will offer a more convenient functional framework to solve renewal equations, see Section 2.7 in this chapter.

## 2.6 Examples

### 2.6.1 The renewal function

We consider $M(t) = E[N_t] \overset{(2.11)}{=} \sum_{k=1}^{\infty} F^{*k}(t)$.

**Proposition 2.16.**

$$\begin{aligned} M(t) &= F(t) + \int_0^t M(t-s)\, dF(s),\ \textit{for } t \geq 0, \\ &= 0,\ \textit{for } t < 0 \end{aligned} \tag{2.60}$$

*(i.e. $M(\cdot)$ is solution of the $(F,F)$-equation).*

*Proof.* We give two arguments:

1) <u>analytic</u>: for $t \geq 0$, one has

(2.61)

$$\begin{aligned} M(t) &= \sum_{k=1}^{\infty} F^{*k}(t) = F(t) + \sum_{k\geq 2} F^{*(k-1)} * F(t) \\ &= F(t) + \sum_{k\geq 2} \int_0^t F^{*(k-1)}(t-s)\, dF(s) \overset{\substack{\text{monotone} \\ \text{convergence}}}{=} F(t) + \int_0^t M(t-s)\, dF(s), \end{aligned}$$

and (2.60) readily follows.

2) <u>probabilistic</u>: for $t \geq 0$, one has

$$\begin{aligned} M(t) = E[N_t] &= E\Big[\sum_{k\geq 1} 1\{T_1 + \cdots + T_k \leq t\}\Big] \\ &= P[T_1 \leq t] + E\Big[\sum_{k\geq 2} 1\{T_1 + S_k - S_1 \leq t\}\Big]\,. \end{aligned}$$

independent

Note that $T_1$ and $(S_{n+1}-S_1)_{n\geq 1}$ are independent and $(S_{n+1}-S_1)_{n\geq 1}$ has same distribution as $(S_n)_{n\geq 1}$. As a result conditioning on $T_1$ we find that

$$E\Big[\sum_{k\geq 2} 1\{T_1 + S_k - S_1 \leq t\} \,\Big|\, T_1\Big] \overset{P-\text{a.s.}}{=} \underset{\substack{\uparrow \\ \text{this is } = 0 \text{ for } T_1 > t}}{M(t - T_1)}\,. \tag{2.62}$$

We thus find that

$$\begin{aligned} M(t) &= P[T_1 \leq t] + E\Big[E\big[\sum_{k\geq 2} 1\{T_1 + S_k - S_1 \leq t\} \,\big|\, T_1\big]\Big] \\ &\overset{(2.62)}{=} P[T_1 \leq t] + E[M(t-T_1)] \\ &= P[T_1 \leq t] + \int_0^t M(t-s)\, dF(s), \end{aligned}$$

and (2.60) follows. The probabilistic argument we just described highlights the regeneration, which occurs after time $T_1 (= S_1)$. $\square$

**Remark 2.17.** (renewal with delay)

In the case of a renewal with delay distribution function $G(\cdot)$, one finds instead that $M(t) = E[N_t]$ satisfies the $(G, F)$-renewal equation:

$$\begin{aligned} M(t) &= G(t) + \int_0^t M(t-s)\, dF(s),\ t \ge 0, \\ &= 0,\ t < 0 \end{aligned} \tag{2.63}$$

(with similar arguments as above).

### 2.6.2 The age and excess distribution functions

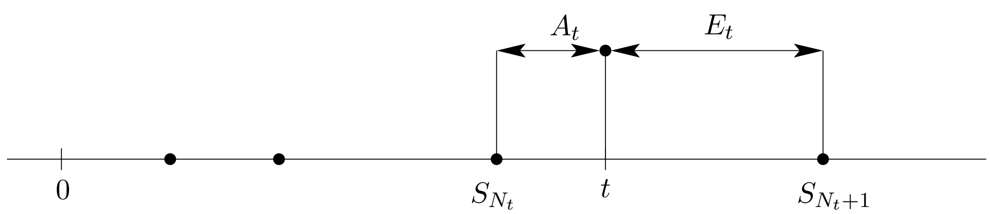



Fig. 2.5: $(N_s)_{s \ge 0}$ renewal process

For $x \ge 0$, we define

$$\begin{aligned} a_x(t) &= P[A_t \le x] \quad (= 0 \text{ for } t < 0, \text{ by convention}) \\ e_x(t) &= P[E_t \le x] \quad (= 0 \text{ for } t < 0, \text{ by convention}). \end{aligned} \tag{2.64}$$

**Proposition 2.18.**

$$\begin{aligned} a_x(t) &= 1_{\{t \le x\}} \big(1 - F(t)\big) + \int_0^t a_x(t-s)\, dF(s),\ t \ge 0, \\ &= 0,\ t < 0 \end{aligned} \tag{2.65}$$

*(i.e. $a_x(\cdot)$ is the solution of the $(1_{\{\cdot \le x\}}(1 - F(\cdot)), F)$-renewal equation),*

$$\begin{aligned} e_x(t) &= F(t+x) - F(t) + \int_0^t e_x(t-s)\, dF(s),\ t \ge 0, \\ &= 0,\ t < 0 \end{aligned} \tag{2.66}$$

*(i.e. $e_x(\cdot)$ is the solution of the $(F(\cdot + x) - F(\cdot), F)$-renewal equation).*

*Proof.*

- $\underline{(2.65)}$: Pick $t \geq 0$, $x \geq 0$, and write:

$$
\begin{aligned}
a_x(t) = P[A_t \leq x] &= P[S_1 > t, \underset{\substack{\| \\ t \text{ on } \{S_1 > t\}}}{A_t} \leq x] + P[S_1 \leq t, A_t \leq x] \\
&= 1_{\{t \leq x\}}\big(1 - F(t)\big) + P[S_1 \leq t, t - S_{N_t} \leq x]\,.
\end{aligned}
\tag{2.67}
$$

We also have since $\{S_1 \leq t\} = \{N_t \geq 1\}$,

$$
P[S_1 \leq t,\, t - S_{N_t} \leq x] = \sum_{n=1}^{\infty} P[N_t = n,\, t - S_n \leq x] =
$$

$$
\sum_{n=1}^{\infty} P[t - x \leq S_n \leq t < S_{n+1}] = \sum_{n=1}^{\infty} P[t - x \leq T_1 + S_n - \underbrace{S_1 \leq t < T_1 + S_{n+1}}_{\text{independent}} - S_1]\,.
$$

So, conditioning on $T_1$, we obtain

$$
= \sum_{n=1}^{\infty} \int_0^t P\big[t - x \leq s + \underbrace{S_n - S_1}_{} \leq t < s + \underbrace{S_{n+1} - S_1}_{}\,\big]\, dF(s)
$$

distributed as $(S_{n-1}, S_n)$

$$
\begin{aligned}
= \sum_{n=1}^{\infty} \int_0^t P\big[\underbrace{t - s - x \leq S_{n-1} \leq t - s < S_n}_{\substack{\| \\ \{N_{t-s} = n-1,\, A_{t-s} \leq x\}}}\,\big]\, dF(s) &= \int_0^t P[A_{t-s} \leq x]\, dF(s) \\
&= \int_0^t a_x(t-s)\, dF(s)\,.
\end{aligned}
$$

Coming back to (2.67) we find (2.65).

- $\underline{(2.66)}$: We follow an analogous strategy. We pick $t \geq 0$, $x \geq 0$, and write:

$$
\begin{aligned}
e_x(t) = P[E_t \leq x] &= \underset{\substack{\| \\ 1\{t < T_1 \leq t + x\}}}{P[S_1 > t, E_t \leq x]} + P[S_1 \leq t, E_t \leq x] \\
&= F(t+x) - F(t) + P[S_1 \leq t, S_{N_t+1} \leq t + x],
\end{aligned}
$$

and since $\{S_1 \leq t\} = \{N_t \geq 1\}$:

$$
\begin{aligned}
&= F(t+x) - F(t) + \sum_{n=1}^{\infty} P\big[\underbrace{S_n \leq t < S_{n+1}}_{\substack{\| \\ N_t = n}},\, S_{n+1} \leq t + x\big] \\
&= F(t+x) - F(t) + \sum_{n=1}^{\infty} P[S_n - S_1 \leq t - T_1 < S_{n+1} - S_1 \leq t - T_1 + x]
\end{aligned}
$$

and then conditioning on $T_1$:

$$= F(t+x) - F(t) + \sum_{n=1}^{\infty} \int_0^t P\big[\,\underbrace{S_n - S_1} \le t - s < \underbrace{S_{n+1} - S_1} \le t - s + x\big]\, dF(s)$$

same law as $(S_{n-1}, S_n)$

$$= F(t+x) - F(t) + \sum_{n=1}^{\infty} \int_0^t \, P\big[\,\underbrace{S_{n-1} \le t - s < S_n \le t - s + x}_{\substack{\| \\ \{N_{t-s} = n-1,\, E_{t-s} \le x\}}}\,\big]\, dF(s)$$

$$= F(t+x) - F(t) + \int_0^t P[E_{t-s} \le x]\, dF(s)$$

$$= F(t+x) - F(t) + \int_0^t e_x(t-s)\, dF(s),$$

and (2.66) follows. □

We continue our discussion of examples of quantities satisfying a **renewal equation**.

### 2.6.3 Cycles of operation and repair of a machine

We assume that $(U_i, V_i)_{i \ge 1}$ are i.i.d. $\mathbb{R}_+^2$-valued (but $U_1, V_1$ are possibly dependent, for instance $V_i = U_i^2 + 1$, with the $U_i \ge 0$, $i \ge 1$, i.i.d. variables), and set for $i \ge 1$

(2.68)
$$\begin{aligned} &T_i = U_i + V_i, \text{ by assumption not a.s. equal to } 0, \\ &F(t) = P[T_1 \le t]. \end{aligned}$$

We consider the times

(2.69)
$$S_k = T_1 + T_2 + \cdots + T_k,\ k \ge 1,\ S_0 = 0.$$

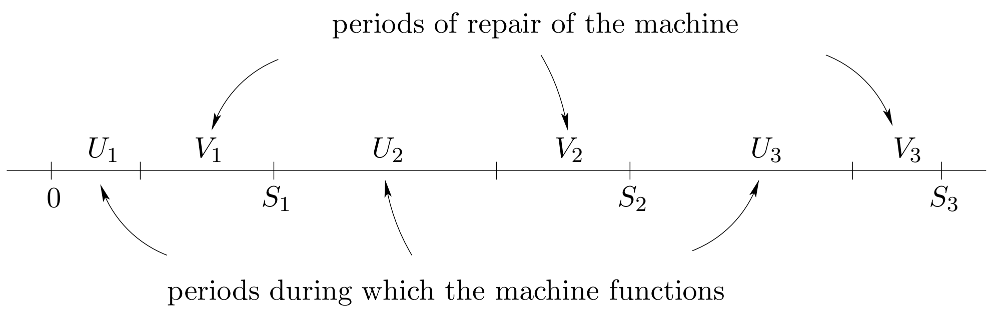



Fig. 2.6

We interpret the $U_i, i \geq 1$, as successive periods during which a given machine is operational and $V_i$ as the repair time of the machine consecutive to the period of length $U_i$ during which the machine was operational.

We introduce $g(t)$ the probability that the machine is operational at time $t \geq 0$:

$$
\begin{aligned}
&g(t) = P[Y_t = 1] = \sum_{k \geq 0} P[S_k \leq t < S_k + U_{k+1}], \text{ where} \\
&Y_s = 1\{s \in \bigcup_{i \geq 0} [S_i, S_i + U_{i+1})\}, \text{ for } s \geq 0.
\end{aligned}
\tag{2.70}
$$

By convention we set $g(t) = 0$, for $t < 0$.

**Proposition 2.19.**

$$
\begin{aligned}
g(t) &= P[U_1 > t] + \int_0^t g(t-s)\, dF(s), \text{ for } t \geq 0, \\
&= 0, \text{ for } t < 0
\end{aligned}
\tag{2.71}
$$

*(i.e. $g(\cdot)$ is the solution of the $(P[U_1 > \cdot], F)$-renewal equation).*

*Proof.* For $t \geq 0$, we have

$$
\begin{aligned}
g(t) &= P\big[\underbrace{Y_t = 1, T_1 > t}_{\{U_1 > t\}}\big] + P[Y_t = 1, T_1 \leq t] \\
&\qquad \text{and since } \{T_1 \leq t\} = \{N_t \geq 1\}, \\
&= P[U_1 > t] + E\Big[\sum_{k \geq 1} 1\{\underbrace{S_k}_{T_1 + S_k - S_1} \leq t < \underbrace{S_k + U_{k+1}}_{T_1 + S_k - S_1 + U_{k+1}}\}\Big].
\end{aligned}
\tag{2.72}
$$

Conditioning on $T_1$ we find that due to the independence properties

$$
E\Big[\sum_{k \geq 1} 1\{T_1 + S_k - S_1 \leq t < T_1 + S_k - S_1 + U_{k+1}\} \,|\, T_1\Big] = \Phi(T_1), \text{ where}
$$

$$
\begin{aligned}
\Phi(s) &= E\Big[\sum_{k \geq 1} 1\{s + S_k - S_1 \leq t < s + S_k - S_1 + U_{k+1}\Big] \\
&= E\Big[\sum_{k \geq 1} 1\{S_{k-1} \leq t - s < S_{k-1} + U_k\}\Big] = g(t-s).
\end{aligned}
\tag{2.73}
$$

Coming back to (2.72) we find that

$$
\begin{aligned}
g(t) &= P[U_1 > t] + E[g(t - T_1)] \\
&= P[U_1 > t] + \int_0^t g(t-s)\, dF(s),\ t \geq 0 \\
&\qquad \text{(and } g(t) = 0, \text{ for } t < 0, \text{ by convention).}
\end{aligned}
$$

This proves (2.71). □

## 2.7 Well-posedness of the renewal equation

Our next step is to provide an existence and uniqueness result for solutions of the renewal equation.

**Theorem 2.20.** *(existence and uniqueness)*

*Let $h(\cdot)$ vanish on $(-\infty, 0)$ be measurable, locally bounded, and $F(\cdot)$ with $F(0) < 1$, be the distribution function of a non-negative variable not a.s. equal to 0. There is a unique $g$ measurable, locally bounded, vanishing on $(-\infty, 0)$, solution of the $(h, F)$-equation:*

$$g = h + g * F, \tag{2.74}$$

*namely*

$$g = h + h * M \;\; (= h * U, \; \textit{if}\, U(t) \stackrel{\text{def}}{=} 1\{t \geq 0\} + M(t)\,). \tag{2.75}$$

*Proof.*

• <u>Existence</u>: Define

$$\begin{aligned} g_0(t) = (h + h * M)(t) &= h(t) + \int_0^t h(t-s)\, dM(s), \text{ if } t \geq 0, \\ &= 0, \text{ if } \; t < 0. \end{aligned}$$

Then $g_0$ is indeed measurable, locally bounded (because $h$ is locally bounded), vanishes on $(-\infty, 0)$. Moreover, one has:

$$\begin{aligned} &h + g_0 * F = h + (h + h * M) * F = h + h * F + (h * M) * F = \\ &h + h * F + h * (M * F) = h + h * \big(\underbrace{F + M * F}_{\substack{||\ (2.60) \\ M}}\big) = h + h * M = g_0, \end{aligned} \tag{2.76}$$

and $g_0$ is a solution of (2.74). Incidentally, the equality $(h*M)*F = h*(M*F)$ follows by noting that both members coincide with the "distribution function" of the image measure of $h(u)\, du \otimes dM(v) \otimes dF(w)$ under the map $(u, v, w) \to u + v + w$.

• <u>Uniqueness</u>:

If $g_1, g_2$ are two solutions of (2.74) which are measurable and locally bounded, then:

$$\begin{aligned} g_1 - g_2 &= (g_1 - g_2) * F, \;\text{ and iterating} \\ &= (g_1 - g_2) * F^{*n} \;\text{ for all } n \geq 1\,. \end{aligned} \tag{2.77}$$

Therefore, for $t \geq 0$,

$$\begin{aligned} &|g_1(t) - g_2(t)| = \Big| \int_0^t (g_1 - g_2)(t-s)\, dF^{*n}(s) \Big| \leq \\ &\big(\sup_{[0,t]} |g_1(\cdot)| + \sup_{[0,t]} |g_2(\cdot)|\big)\; P[N_t \geq n] \underset{n \to \infty}{\longrightarrow} 0, \\ &\text{since } E[N_t] = M(t) \stackrel{(2.12)}{<} \infty\,. \end{aligned} \tag{2.78}$$

Hence, $g_1 = g_2$ and the theorem is proved. □

## 2.8 Asymptotic behavior of solutions of the renewal equation

We are now going to discuss the large $t$ behavior of the solution $g(\cdot)$ of the renewal equation (cf. Theorem 2.20 for the precise formulation)

$$g = h + g * F\,.$$

We begin by the discussion of an example.

**Example 2.21.** When $h = 1_{[a,b)}$, where $0 \le a < b < \infty$, we know from (2.75) that the unique locally bounded measurable solution $g$ of

$$\begin{aligned} g(t) &= h(t) + \int_0^t g(t-s)\,dF(s), \text{ for } t \ge 0,\\ &= 0, \text{ for } t < 0, \end{aligned}$$

is given by

$$\begin{aligned} g(t) &= h(t) + \int_0^t h(t-s)\,dM(s) \overset{\substack{\text{since } h(t-s) = 1\{a \le t-s < b\}\\ \downarrow}}{=} h(t) + \int_{[0,t]} 1_{(t-b,t-a]}(s)\,dM(s) \\ &= 1_{[a,b)}(t) + M(t-a) - M(t-b) \text{ (by convention } M(u) = 0 \text{ for } u < 0). \end{aligned} \tag{2.79}$$

As a result of Blackwell's renewal theorem, cf. (2.55), we see that when $F$ is a non-arithmetic distribution function, then

$$\lim_{t\to\infty} g(t) \overset{(2.55)}{=} \frac{b-a}{\mu} = \frac{1}{\mu} \int_0^\infty h(u)\,du. \tag{2.80}$$

Clearly, a similar statement "$\lim_{t\to\infty} g(t) = \frac{1}{\mu}\int_0^\infty h(u)du$", will hold when $h$ is a finite linear combination of indicator functions of intervals of the form $1_{[a_k,b_k)}$, $1 \le k \le \ell$, with $0 \le a_k < b_k < \infty$, for $1 \le k \le \ell$. We are now going to vastly generalize the class of functions $h$ for which this statement holds. □

**Definition 2.22.** *$h$: $\mathbb{R}_+ \to \mathbb{R}_+$, measurable, is called* ***directly Riemann-integrable*** *(abbreviation: d.R.i.) if*

$$\textit{for all } \Delta > 0,\ \sum_{k=0}^\infty \sup_{t\in[k\Delta,(k+1)\Delta)} h(t) < \infty, \tag{2.81}$$

*and*

$$\lim_{\Delta\to 0} \Delta \sum_{k=0}^\infty \sup_{t\in[k\Delta,(k+1)\Delta)} h(t) = \lim_{\Delta\to 0} \Delta \sum_{k=0}^\infty \inf_{t\in[k\Delta,(k+1)\Delta)} h(t). \tag{2.82}$$

*$h$: $\mathbb{R}_+ \to \mathbb{R}$ is said to be d.R.i. when $h_+ = \max(h,0)$ and $h_- = \max(-h,0)$ are both d.R.i.*

**Remark 2.23.** $h(t) = \sum_{k=1}^{\infty} 1_{[k,k+2^{-k}]}(t)$ is not d.R.i. (indeed (2.81) does not hold, pick $\Delta = 1$). On the other hand

$$\int_0^\infty h(t)dt = \lim_{u\to\infty} \int_0^u \underset{\substack{\uparrow \\ \text{Riemann-integral}}}{h(u)du}$$

exists, and $h$ is Riemann-integrable. □

**Lemma 2.24.**

(2.83) *$h \geq 0$, d.R.i., then $h$ is bounded, continuous at a.e. point of $\mathbb{R}_+$, and $\lim_{t\to\infty} h(t) = 0$.*

(2.84) *$h$ measurable bounded on $\mathbb{R}_+$, vanishing outside a compact and continuous at a.e. point of $\mathbb{R}_+$ is d.R.i..*

(2.85) *$h \geq 0$, non-increasing with $\int_0^\infty h(s)\,ds < \infty$, is d.R.i..*

(2.86) *$0 \leq h \leq H$, $H$ d.R.i., $h$ continuous at a.e. points of $\mathbb{R}_+$, measurable, then $h$ is d.R.i..*

*Proof.*

- (2.83):

  Because of (2.81) with $\Delta = 1$, $h$ is bounded and $\lim_{t\to\infty} h(t) = 0$. Moreover, (2.82) implies that $h$ is Riemann-integrable on each $[0,T]$.

  This fact implies that $h$ is continuous at a.e. point of $[0,T]$, see [7], p. 184. This proves (2.83).

- (2.84):

  This implies that $h$ is Riemann-integrable on $[0,T]$ and equal to 0 on $[0,T]^c$, for some $T > 0$ (same reference), and (2.81), (2.82) follow.

- (2.85):

$$\infty > \int_0^\infty h(s)\,ds = \sum_1^\infty \int_{(n-1)\Delta}^{n\Delta} h(s)\,ds \geq \sum_1^\infty \Delta h(n\Delta) = \Delta \sum_{n\geq 1} \sup_{[n\Delta,(n+1)\Delta)} h(\cdot),$$

and (2.81) follows. Moreover, we have:

$$\underbrace{\Delta \sum_{n=0}^\infty \sup_{[n\Delta,(n+1)\Delta)} h(\cdot)}_{= \, \Delta \sum_{n=0}^\infty h(n\Delta)} \geq \int_0^\infty h(s)\,ds \geq \underbrace{\Delta \sum_{n=0}^\infty \inf_{[n\Delta,(n+1)\Delta)} h(\cdot)}_{\geq \, \Delta \sum_{n=0}^\infty h((n+1)\Delta)} .$$

However,

$$\Delta \sum_{n=0}^{\infty} h(n\Delta) - h\big((n+1)\Delta\big) = \Delta h(0) \to 0, \text{ as } \Delta \to 0,$$

and hence (2.82) follows.

- (2.86):

  $h$ is then Riemann-integrable on any finite interval and (2.81) holds. Then (2.82) easily follows. □

We now come to our main result concerning the asymptotic behaviour of solutions of the renewal equation.

**Theorem 2.25.** *(Smith's key renewal theorem)*

*Let **h be d.R.i.** and $F$ be a **non-arithmetic** distribution function, then the unique measurable, locally bounded, vanishing on $(-\infty, 0)$, solution $g$ of the equation $g = h + g * F$ satisfies*

$$\lim_{t\to\infty} g(t) = \frac{1}{\mu} \int_0^\infty h(u)\, du. \tag{2.87}$$

*Proof.* We know from (2.75) that

$$g(t) = h(t) + \int_0^t h(t-s)\, dM(s), \quad \text{for } t \geq 0,$$

and (2.83) implies that $\lim_{t\to\infty} h(t) = 0$. We thus only have to show that:

$$\lim_{t\to\infty} \int_0^t h(t-s)\, dM(s) = \frac{1}{\mu} \int_0^\infty h(u)\, du. \tag{2.88}$$

- Assume first that:

$$h(t) = \sum_{n\geq 1} c_n\, 1_{[(n-1)\Delta, n\Delta)}(t), \text{ with } c_n \geq 0,\ \sum_n c_n < \infty \text{ and } F(\Delta) < 1. \tag{2.89}$$

Then, by monotone convergence and the calculation of (2.79), we find that

$$\int_0^t h(t-s)\, dM(s) = \sum_{n\geq 1} c_n \big(M(t-(n-1)\Delta) - M(t-n\Delta)\big). \tag{2.90}$$

We will dominate the terms of the above series with the help of the following

**Lemma 2.26.**

$$\sup_{u\geq 0} \big(M(u) - M(u-\Delta)\big) \leq \big(1 - F(\Delta)\big)^{-1} \textit{ (recall that } M(v) = 0, \textit{ for } v < 0\textit{)}. \tag{2.91}$$

*Proof.* $M$ satisfies the $(F,F)$-renewal equation, cf. (2.60), and since $M * F = F * M$, we find:

$$M(t) = F(t) + \int_0^t F(t-s)\,dM(s), \text{ for } t \geq 0,$$

and hence

$$\begin{aligned} 1 \geq F(t) &= M(t) - \int_0^t F(t-s)\,dM(s) \\ &= \int_0^t \big(1 - F(t-s)\big)\,dM(s) \geq \int_0^t \big(1-F(\Delta)\big)\,1_{(t-\Delta,t]}(s)\,dM(s) \\ &= \big(1-F(\Delta)\big)\big(M(t) - M(t-\Delta)\big), \end{aligned} \tag{2.92}$$

and (2.91) follows. ☐

Coming back to (2.90), we see that each term in the series in the right-hand side of (2.90) is dominated by const$(\Delta)\,c_n$, which is summable. By Blackwell's renewal theorem, cf. (2.55), each term converges to $c_n\,\frac{\Delta}{\mu}$, as $t$ tends to infinity. Thus, with dominated convergence in the right-hand side of (2.90), we find that

$$\lim_{t\to\infty} \int_0^t h(t-s)\,dM(s) = \sum_{n\geq 1} c_n\,\frac{\Delta}{\mu} = \frac{1}{\mu} \int_0^\infty h(u)\,du. \tag{2.93}$$

This proves (2.88) and hence (2.87) when $h$ is of the form (2.89).

- Assume $h \geq 0$ is d.R.i.: We write

$$\underline{h}_\Delta(t) = \sum_{n\geq 0} \inf_{[n\Delta,(n+1)\Delta)} h\ \ 1_{[n\Delta,(n+1)\Delta)}(t)$$

$$\overline{h}_\Delta(t) = \sum_{n\geq 0} \sup_{[n\Delta,(n+1)\Delta)} h\ \ 1_{[n\Delta,(n+1)\Delta)}(t),$$

so that $\underline{h}_\Delta \leq h \leq \overline{h}_\Delta$.

Using the previous step we find that when $\Delta > 0$:

$$\begin{array}{ccccc} \displaystyle\int_0^t \underline{h}_\Delta(t-s)\,dM(s) & \leq & \displaystyle\int_0^t h(t-s)\,dM(s) & \leq & \displaystyle\int_0^t \overline{h}_\Delta(t-s)\,dM(s) \\ {\scriptstyle (2.93)}\big\downarrow{\scriptstyle t\to\infty} & & & & {\scriptstyle (2.93)}\big\downarrow{\scriptstyle t\to\infty} \\ \frac{1}{\mu}\int_0^\infty \underline{h}_\Delta(u)du & & & & \frac{1}{\mu}\int_0^\infty \overline{h}_\Delta(u)du\,. \end{array} \tag{2.94}$$

Moreover, in view of (2.82), we see that

$$\lim_{\Delta\to 0} \int_0^\infty \underline{h}_\Delta(u)\,du = \lim_{\Delta\to 0} \int_0^\infty \overline{h}_\Delta(u)\,du = \int_0^\infty h(u)\,du.$$

We thus find from (2.94) that

$$\lim_{t\to\infty} \int_0^t h(t-s)\, dM(s) = \frac{1}{\mu} \int_0^\infty h(u)\, du,$$

i.e. (2.88) holds and (2.87) follows.

- The general case of $h$ d.R.i.:

  We simply write $h = h_+ - h_-$, with $h_+ = \max(h, 0)$, $h_- = \max(-h, 0)$, so (2.88), and hence (2.87), follow from the previous step. □

## 2.9 Applications

### 2.9.1 The age and excess distribution functions

We keep the same notations as in Section 2.6.2. We know that for $x \ge 0$, $a_x(t) = P[A_t \le x]$, $e_x(t) = P[E_t \le x]$ are solutions of $(h, F)$-renewal equations, where

- In the case of $a_x(\cdot)$:

  $h(t) = 1_{[0,x]}(t)\big(1 - F(t)\big)$ (d.R.i. thanks to (2.84)).

- In the case of $e_x(\cdot)$:

  $h(t) = F(t+x) - F(t)(\ge 0)$, which is d.R.i. when $\mu < \infty$, because

$$h(t) = \overbrace{1 - F(t)}^{\substack{\text{non-increasing}\\ \downarrow}} - \Big(\overbrace{1 - F(t+x)}^{\substack{\text{non-increasing}\\ \downarrow}}\Big) \quad \text{and}$$

$$\int_0^\infty \big(1 - F(t)\big)\, dt \overset{\text{Remark 2.10 1)}}{=} \mu < \infty, \ \int_0^\infty \big(1 - F(t+x)\big)\, dt = \int_x^\infty \big(1 - F(s)\big)\, ds \le \mu < \infty,$$

so that with (2.85) and (2.86), we find that $h$ is d.R.i. Note also that

$$\int_0^\infty h(t)\, dt = \int_0^x \big(1 - F(t)\big)\, dt.$$

Thus, when $F$ is non-arithmetic, we find that

$$\lim_{t\to\infty} a_x(t) = \frac{1}{\mu} \int_0^x \big(1 - F(t)\big)\, dt = G_*(x) \text{ in the notation of (2.44)}\,; \tag{2.95}$$

$$\lim_{t\to\infty} e_x(t) = \frac{1}{\mu} \int_0^\infty \big(F(t+x) - F(t)\big)\, dt = \frac{1}{\mu} \int_0^x \big(1 - F(x)\big)\, dt = G_*(x). \tag{2.96}$$

In particular, when $\mu < \infty$, we see that when $t$ tends to infinity, both $A_t$ and $E_t$ converge in distribution to a random variable with distribution function $G_*$.
(Incidentally, the statements (2.95), (2.96) hold even when $\mu = \infty$, because in fact

$$\begin{aligned} e_x(t) &= P[E_t \le x] = P[N_{t+x} - N_t \ge 1] \ \le M(t+x) - M(t) \ \underset{t\to\infty}{\overset{(2.55)}{\longrightarrow}} 0, \text{ and} \\ a_x(t) &= P[A_t \le x] \le P[N_t - N_{t-2x} \ge 1] \le M(t) - M(t-2x) \ \underset{t\to\infty}{\overset{(2.55)}{\longrightarrow}} 0 \end{aligned}$$

when $\mu = \infty$.)

### 2.9.2 Cycles of operation and repair of a machine

We keep the same notations as in Section 2.6.3. So the probability that the machine is operational at time $t$ is $g(t)$ which is the solution of the $(h, F)$-renewal equation with

$$h(t) = P[U_1 > t], \text{ which is d.R.i. when } E[U_1] < \infty \text{ (cf. (2.85)).}$$

Therefore, when $F$ is non-arithmetic and $E[U_1] < \infty$, we find that

$$\lim_{t\to\infty} g(t) = \frac{1}{\mu} \int_0^\infty P[U_1 > t]\, dt = \frac{E[U_1]}{E[U_1] + E[V_1]} \,, \tag{2.97}$$

($\mu \overset{(2.68)}{=} E[U_1]+E[V_1]$ is the expectation of the duration of a cycle of operation and repair).

## 2.10 Renewal with defect

This corresponds to the case where the random variables $(T_i)_{i\ge1}$ are i.i.d., $[0,\infty]$-valued with $P[T_1 = \infty] = 1 - F(\infty) > 0$, and for $t \in \mathbb{R}$, $F(t) = P[T_1 \le t]$. Just as in (2.3), (2.4), (2.8) one defines the renewal process with defect:

$$N_t = \sum_{k\ge1} 1\{S_k \le t\} = \sup\{n \ge 0; S_n \le t\}, \text{ for } t \ge 0, \tag{2.98}$$

where

$$S_n = T_1 + T_2 + \cdots + T_n, \text{ when } n \ge 1, \text{ and } S_0 = 0,$$

as well as the corresponding renewal function:

$$M(t) = E[N_t], \text{ for } t \ge 0. \tag{2.99}$$

The renewal equation

$$g = h + g * F, \tag{2.100}$$

with $\boldsymbol{F(\infty) < 1}$, **is called renewal equation with defect**.

**Remark 2.27.** The renewal equation with defect often occurs in practice when considering equations of the type

$$g = h + \gamma\, g * F,$$

where $F(\cdot)$ is a distribution, with no defect (such that $F(u) = 0$, for $u < 0$, $F(0) < 1$, and $F(\infty) = 1$), and $\gamma$ some parameter in $(0, 1)$. $\square$

**Proposition 2.28.** *Let $h(\cdot)$ be a function vanishing on $(-\infty, 0)$, which is measurable and locally bounded. There is a unique $g$ measurable, locally bounded, vanishing on $(-\infty, 0)$, solution of (2.100), namely:*

$$g = h + h * M. \tag{2.101}$$

*Proof.* Just like the proof of (2.74). $\square$

**Proposition 2.29.** *Let $h(\cdot)$ vanishing on $(-\infty, 0)$, be measurable and bounded with*

$$h(t) \text{ tending to } h(\infty) \text{ as } t \text{ tends to infinity.} \tag{2.102}$$

*The unique measurable, locally bounded solution of (2.100) vanishing on $(-\infty, 0)$ satisfies*

$$\lim_{t\to\infty} g(t) = \frac{h(\infty)}{1 - F(\infty)} \overset{\text{def}}{=} g(\infty). \tag{2.103}$$

*Proof.*

$$M(t) = \sum_{k=1}^{\infty} F^{*k}(t) = \sum_{k=1}^{\infty} P[T_1 + \cdots + T_k \le t], \text{ for } t \ge 0. \tag{2.104}$$

Observe that

$$\begin{aligned} \lim_{t\to\infty} P[T_1 + \cdots + T_k \le t] &= P[T_1 + \cdots + T_k < \infty] \\ &= P[T_\ell < \infty, \text{ for } 1 \le \ell \le k] = F(\infty)^k. \end{aligned} \tag{2.105}$$

Coming back to (2.104), it follows from monotone convergence that

$$\lim_{t\to\infty} M(t) = M(\infty) \overset{\text{def}}{=} \sum_{k=1}^{\infty} F(\infty)^k = \frac{F(\infty)}{1 - F(\infty)}\ . \tag{2.106}$$

As a result we see that

$$g(t) \overset{(2.101)}{=} h(t) + \int_0^t h(t-s)\, dM(s) = h(t) + \int_0^\infty \underbrace{1_{[0,t]}(s) h(t-s)}_{\substack{t\to\infty \,\downarrow\, (2.102) \\ h(\infty)}}\, dM(s)$$

and using dominated convergence and (2.102):

$$\lim_{t\to\infty} g(t) = h(\infty) + h(\infty) \int_0^\infty dM(s) = h(\infty)\big(1 + M(\infty)\big) \overset{(2.106)}{=} \frac{h(\infty)}{1 - F(\infty)}\ , \tag{2.107}$$

and this proves (2.103). $\square$

Sometimes, if for some $\alpha > 0$,

$$\int_0^\infty e^{\alpha x} dF(x) = 1 \tag{2.108}$$

(we recall that on the interval of values of $\alpha$ where $\int_0^\infty e^{\alpha x} dF(x) < \infty$, the function $\alpha \to \log(\int_0^\infty e^{\alpha x} dF(x))$ is convex, as follows from Hölder's inequality), one can go further than (2.103). Indeed, if $h$ is as in (2.102) and $g$ as in (2.103), we find that for $t \geq 0$:

$$\begin{aligned} g(t) - g(\infty) &= h(t) + \int_0^t g(t-s)\, dF(s) - g(\infty) \\ &= h(t) + \int_0^t \big(g(t-s) - g(\infty)\big)\, dF(s) - g(\infty)\big(1 - F(t)\big) \\ &\overset{(2.107)}{=} h(t) - h(\infty)\, \frac{(1 - F(t))}{1 - F(\infty)} + \int_0^t \big(g(t-s) - g(\infty)\big)\, dF(s). \end{aligned} \tag{2.109}$$

If we multiply both members of (2.109) by $e^{\alpha t}$, with $\alpha$ as in (2.108) and set

$$\begin{aligned} \widetilde{g}_\alpha(t) &= \big(g(t) - g(\infty)\big)\, e^{\alpha t} 1\{t \geq 0\}, \\ \widetilde{h}_\alpha(t) &= \Big(h(t) - h(\infty)\, \frac{1 - F(t)}{1 - F(\infty)}\Big)\, e^{\alpha t} 1\{t \geq 0\}, \\ dF_\alpha(x) &= e^{\alpha x} dF(x), \end{aligned}$$

we obtain:

$$\begin{aligned} \widetilde{g}_\alpha(t) &= \widetilde{h}_\alpha(t) + \int_0^t \widetilde{g}_\alpha(t-s)\, dF_\alpha(s), \text{ for } t \geq 0 \\ &= 0, \text{ for } t < 0. \end{aligned} \tag{2.110}$$

In other words: $\widetilde{g}_\alpha$ solves the $(\widetilde{h}_\alpha, F_\alpha)$-renewal equation. We can then apply Smith's key renewal theorem to find

**Proposition 2.30.** *When $h$ is as in (2.102), $g$ as in (2.101), and for some $\alpha > 0$, $\int_0^\infty e^{\alpha x} dF(x) = 1$, $F_\alpha$ is non-arithmetic and $\widetilde{h}_\alpha$ d.R.i., then*

$$\lim_{t \to \infty} \big(g(t) - g(\infty)\big)\, e^{\alpha t} = \frac{1}{\int_0^\infty x e^{\alpha x} dF(x)} \int_0^\infty \widetilde{h}_\alpha(u)\, du. \tag{2.111}$$

**Complement:** 1) A beautiful application of the above result concerns the so-called risk-process, cf. [12], pp. 205 and 259. Claims arrive at an insurance company according to a Poisson process of rate $\lambda > 0$, and the claims are marks of this Poisson process given by i.i.d. variables $X_1, X_2, \dots, X_n, \dots$, which are non-negative.

The capital at time 0 of the insurance company is $x$, and it receives $ct$ premiums by time $t$. So the fortune at time $t$ of the insurance company is

$$f(t) = x + ct - \sum_{k \geq 1} X_k\, 1\{S_k \leq t\}. \tag{2.112}$$

An important quantity is the so-called probability of no-ruin of the insurance company:

$$R(x) = P[f(t) > 0, \text{ for all } t > 0]. \tag{2.113}$$

In the above reference it is shown that when

$$\lambda E[X_1] < c, \text{ then: } \begin{aligned} &R(\infty) = 1, \ R(0) = 1 - \frac{\lambda}{c} E[X_1] \\ &R(t) = R(0)\big(1 + M(t)\big), \ t \geq 0, \end{aligned} \tag{2.114}$$

with $M(t)$ the defective renewal function attached to

$$F(t) = \int_0^t \frac{\lambda}{c} P[X_1 > u] \, du. \tag{2.115}$$

One can then use (2.111) to find a rate of convergence to zero of $1 - R(t)$ as $t \to \infty$. If $\alpha > 0$ can be chosen such that $\frac{\lambda}{c} \int_0^\infty e^{\alpha x} P[X_1 > x] \, dx = 1$, then

$$1 - R(t) \underset{t \to \infty}{\sim} \frac{e^{-\alpha t}}{\dfrac{\alpha \lambda}{c} \displaystyle\int_0^\infty x e^{\alpha x} P[X_1 > x] \, dx}.$$

2) In the case of an equation,

$$g = h + g * F, \tag{2.116}$$

where $h, g$ vanish on $(-\infty, 0)$, and $F(u) = 0$, $u < 0$, with $\boldsymbol{F(0) < 1 < F(\infty) < \infty}$**, one can choose** $\boldsymbol{\alpha < 0}$ so that

$$\int_0^\infty e^{\alpha x} dF(x) = 1.$$

Then, multiplying both members of (2.116) with $e^{\alpha x}$, one obtains the equation

$$\begin{aligned} g_\alpha &= h_\alpha + g_\alpha * F_\alpha, \ \text{ where} \\ h_\alpha(t) &= e^{\alpha t} h(t), \\ g_\alpha(t) &= e^{\alpha t} g(t), \\ dF_\alpha(t) &= e^{\alpha t} dF(t), \end{aligned} \tag{2.117}$$

which is a usual renewal equation. □

# 3 Discrete time Markov chains

We consider the state space

$$(3.1) \qquad E\text{: an at most denumerable non-empty set.}$$

We begin with the discussion of **discrete time** Markov chains on $E$.

**Definition 3.1.** *A sequence $(X_n)_{n\geq 0}$ of random variables with values in $E$ defined on some $(\Omega, \mathcal{A}, P)$ is a* ***discrete time Markov chain*** *with state space $E$ when:*

$$(3.2) \qquad E[f(X_{n+1}) \,|\, X_0, X_1, \dots, X_n] \overset{P\text{-}a.s.}{=} E[f(X_{n+1}) \,|\, X_n], \text{ for } n \geq 0,$$

*for any bounded function $f\colon E \to \mathbb{R}$.*

Intuitively:

| |
|---|
| The best prediction of the future of the sequence $(X_n)$ knowing its past only relies on the information contained in the present. |

Of special interest to us will be the situation when the chain is time-homogeneous and we have a fixed **transition probability** on $E$:

$$(3.3) \qquad (r_{x,y})_{x,y\in E}, \text{ with } \begin{array}{l} r_{x,y} \geq 0, \text{ for } x, y \in E, \\ \sum\limits_{y\in E} r_{x,y} = 1, \text{ for } x \in E. \\ \nwarrow \text{ "time-homogeneous"} \end{array}$$

In this case one has

**Definition 3.2.** *A sequence $(X_n)_{n\geq 0}$ of random variables with values in $E$ on some $(\Omega, \mathcal{A}, P)$ is a Markov chain with state space $E$ and transition probability $(r_{x,y})_{x,y\in E}$ when:*

$$(3.4) \qquad E[f(X_{n+1})|X_0, \dots, X_n] \overset{P\text{-}a.s.}{=} \sum_{y\in E} r_{X_n,y} f(y),$$

*for all $n \geq 0$, and bounded functions $f\colon E \to \mathbb{R}$.*

Given the transition probability $(r_{x,y})$ on $E$ one defines

$$r_{x,y}(n) \text{ for } n \geq 0,\ x, y \text{ in } E, \text{ via:}$$

$$(3.5) \qquad \begin{array}{rcl} r_{x,y}(0) & = & \underset{\swarrow \text{ Kronecker symbol}}{\delta_{x,y}}, \ r_{x,y}(1) = r_{x,y}, \text{ and by induction:} \\ r_{x,y}(n+1) & = & \sum\limits_{z\in E} r_{x,z}(n)\, r_{z,y}\,. \end{array}$$

**Proposition 3.3.** *Given a Markov chain $(X_n)_{n\ge 0}$, with state space $E$, transition probability $(r_{x,y})$, and initial distribution*

$$\mu(x) = P[X_0 = x],\ x \in E, \tag{3.6}$$

*one has*

$$r_{x,y}(n+m) = \sum_{z\in E} r_{x,z}(n)\, r_{z,y}(m),\ \text{for } n,m \ge 0,\ x,y \in E, \tag{3.7}$$

*(the so-called Chapman-Kolmogorov equation)*

$$r_{x,y}(n) = (R^n 1_{\{y\}})(x),\ n \ge 0,\ x,y \in E,\ \text{with} \tag{3.8}$$

*$R$ the linear operator on the set of bounded functions on $E$ defined by $Rf(x) = \sum_{z\in E} r_{x,z}\, f(z)$ for $f$ bounded function on $E$,*

$$\text{for } x_0, x_1, \dots, x_n \in E,\ P[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] = \mu(x_0)\, r_{x_0,x_1} r_{x_1,x_2} \dots r_{x_{n-1},x_n}, \tag{3.9}$$

$$\text{for } y \in E, n \ge 0,\ P[X_n = y] = (\mu R^n)(y),\ \text{with } (\mu R^n)(y) \overset{\text{def}}{=} \sum_z \mu(z)(R^n 1_{\{y\}})(z). \tag{3.10}$$

*Proof.*

- (3.7): One fixes $n$ and proves (3.7) by induction on $m$.
- (3.8): The claim holds for $n = 0$ and $n = 1$, and then by induction

$$r_{x,y}(n+1) \overset{(3.5)}{=} \sum_{z\in E} r_{x,z}(n)\, r_{z,y} = [R^n(R\, 1_{\{y\}})](x) = (R^{n+1} 1_{\{y\}})(x),$$

and (3.8) follows.

- (3.9): The claim holds for $n = 0$, and then for $n \ge 1$,

$$\begin{aligned} P[x_0 = x_0, \dots, X_n = x_n] &= E\big[P[X_n = x_n \,|\, X_0, \dots, X_{n-1}],\ X_0 = x_0, \dots, X_{n-1} = x_{n-1}\big] \\ &\overset{(3.4)}{=} E[r_{X_{n-1},x_n}, X_0 = x_0, \dots, X_{n-1} = x_{n-1}] = r_{x_{n-1},x_n} P[X_0 = x_0, \dots, X_{n-1} = x_{n-1}] \\ &\overset{\text{induction}}{=} \big(\mu(x_0) r_{x_0,x_1} \dots r_{x_{n-2},x_{n-1}}\big)\, r_{x_{n-1},x_n}, \end{aligned}$$

and this proves (3.9).

• (3.10): The claim holds for $n = 0$, and then for $n \geq 1$:

$$P[X_n = y] = E\big[P[X_n = y \,|\, X_0, \dots, X_{n-1}]\big] \overset{(3.4)}{=} E[r_{X_{n-1},y}] =$$

$$\sum_{z \in E} P[X_{n-1} = z]\, r_{z,y} \underset{(3.8)}{\overset{\text{induction}}{=}} \sum_{x,z} \mu(x)\, r_{x,z}(n-1)\, r_{z,y} \overset{(3.5)}{=} \sum_x \mu(x)\, r_{x,y}(n) \overset{(3.8)}{=} (\mu R^n)(y),$$

and this proves (3.10). □

One can in fact construct a canonical homogeneous Markov chain as follows. One introduces

$$\begin{aligned} &\Omega_0 = E^{\mathbb{N}} = \{\text{sequences } (x_i)_{i \geq 0}, \text{ with } x_i \in E, \text{ for all } i \geq 0\}, \\ &X_n(\omega) \overset{\text{def}}{=} \omega(n),\ n \geq 0, \text{ the canonical coordinates}, \\ &\mathcal{A}_0 = \sigma(X_0, X_1, \dots), \text{ the canonical } \sigma\text{-algebra}, \\ &\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n),\ n \geq 0, \text{ the canonical filtration.} \end{aligned} \tag{3.11}$$

**Proposition 3.4.** *Given $(r_{x,y})$ transition probability on $E$, for each $z \in E$, there is a unique probability $P_z$ on $(\Omega_0, \mathcal{A}_0)$, under which the $(X_n)_{n \geq 0}$ are a Markov chain on $E$ with transition probability $(r_{x,y})$ and initial distribution $\delta_z$ (= point mass at $z$). Moreover, for $\mu$ probability on $E$, under*

$$P_\mu \overset{\text{def}}{=} \sum_{x \in E} \mu(x)\, P_x, \tag{3.12}$$

*$(X_n)_{n \geq 0}$ is a Markov chain on $E$ with transition probability $(r_{x,y})$ and initial distribution $\mu$ (i.e. $X_0$ has law $\mu$ under $P_\mu$).*

*Proof.* First given $z \in E$, with the help of (3.9), we only need to show that there is a unique probability $P_z$ on $(\Omega_0, \mathcal{A}_0)$ such that for any $n \geq 0$, $x_0, x_1, \dots, x_n$,

$$P_z[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] = \delta_z(x_0)\, r_{x_0,x_1}, \dots, r_{x_{n-1},x_n} \tag{3.13}$$

(such a probability then automatically makes $(X_n)_{n \geq 0}$, a Markov chain with transition $(r_{x,y})$ and initial law $\delta_z$).

• The uniqueness of $P_z$ follows from Dynkin's lemma, cf. [12], p. 41.

• The existence of $P_z$:
Denote by $Q$ the Lebesgue measure on $((0,1), \mathcal{B}(0,1))$, and for each $x \in E$, let

(3.14) $\Phi(x, \cdot)$ be a measurable function $(0,1) \to E$ under which $Q$ has image measure $r_{x,\cdot}$ on $E$,
(this is easily done by partitioning $(0,1)$ into intervals with respective lengths $r_{x,y}$, with $y \in E$, on which the function $\Phi(x, \cdot)$ takes the value $y$).

Consider now on some probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P})$ an i.i.d. sequence $U_i$, $i \geq 1$, of $(0,1)$-valued uniformly distributed variables and set

$$\widetilde{X}_0 = z,\ \widetilde{X}_1 = \Phi(z, U_1),\ \widetilde{X}_2 = \Phi(\widetilde{X}_1, U_2), \dots, \widetilde{X}_{n+1} = \Phi(\widetilde{X}_n, U_{n+1}), \dots . \tag{3.15}$$

Then for any $n \geq 0$, $x_0, \dots, x_n \in E$, one has

$$\widetilde{P}[\widetilde{X}_0 = x_0,\ \widetilde{X}_1 = x_1, \dots, \widetilde{X}_n = x_n] = \delta_z(x_0)\, r_{x_0,x_1}, \dots, r_{x_{n-1},x_n},$$

as follows from induction from (3.14) and the fact that the variables $U_i$, $i \geq 1$ are i.i.d. uniformly distributed on $(0, 1)$.

Then $P_z \overset{\text{def}}{=}$ the law of $(\widetilde{X}_0, \widetilde{X}_1, \dots, \widetilde{X}_n \dots)$ on $(\Omega_0, \mathcal{A}_0)$, satisfies (3.13).

- To check the last statement, let $\mu$ be a probability on $E$, then for $f$ bounded $E \to \mathbb{R}$ and $n \geq 0$, $x_0$, $x_1, \dots, x_n$ in $E$,

$$E_\mu[f(X_{n+1}),\ X_0 = x_0, \dots, X_n = x_n] \overset{(3.12)}{=} \sum_{x \in E} \mu(x)\, E_x[f(X_{n+1}), X_0 = x_0, \dots, X_n = x_n],$$

$P_\mu$-expectation and since $(X_n)_{n \geq 0}$ is a Markov chain under $P_x$,

$$= \sum_{x \in E} \mu(x)\, P_x[X_0 = x_0, \dots, X_n = x_n]\, (Rf)(x_n)$$

$$= \sum_{x \in E} \mu(x)\, E_x[X_0 = x_0, \dots, X_n = x_n,\ Rf(X_n)]$$

$$= E_\mu[(Rf)(X_n),\ X_0 = x_0, \dots, X_n = x_n].$$

The Markov property of $(X_n)_{n \geq 0}$ under $P_\mu$ follows. Moreover, $P_\mu[X_0 = x_0] = \mu(x_0)$, and $\mu$ is the initial distribution (under $P_\mu$). $\square$

## 3.1 Examples

### 3.1.1 Simple random walk on $\mathbb{Z}^d$

$E = \mathbb{Z}^d$, $d \geq 1$, the corresponding transition probability is

$$\begin{aligned} r_{x,y} &= \frac{1}{2d}, \text{ if } |y - z| = 1,\ x, y \text{ in } \mathbb{Z}^d \\ &= 0, \quad \text{otherwise.} \end{aligned} \tag{3.16}$$

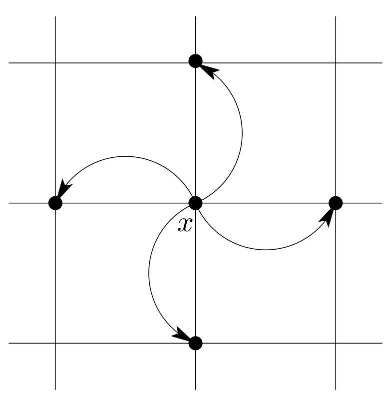



Fig. 3.1

### 3.1.2 Reflected random walk

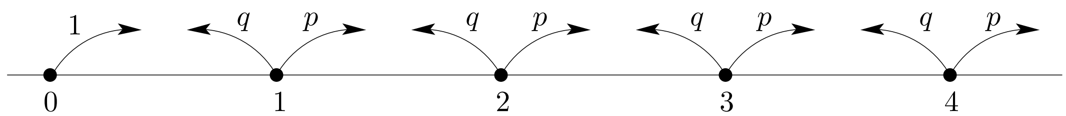



Fig. 3.2

The state space is $E = \mathbb{N}$, $p, q > 0$ are such that $p + q = 1$, and the transition probability is:

$$
\begin{aligned}
r_{x,y} &= 1, \ \text{if } x = 0, \ y = 1, \\
&= p, \ \text{if } x \geq 1, \ y = x + 1, \\
&= q, \ \text{if } x \geq 1, \ y = x - 1, \\
&= 0, \ \text{otherwise.}
\end{aligned} \tag{3.17}
$$

### 3.1.3 Galton-Watson chain

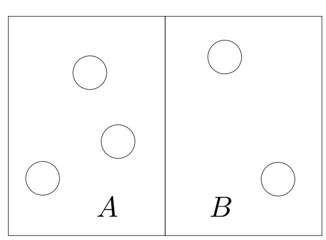
The state space is $E = \mathbb{N}$, we have a probability $\pi$ on $\mathbb{N}$ with $\pi(0) \neq 1$, $\pi(1) \neq 1$, so that each individual has $k$ descendants with probability $\pi(k)$.

The chain describes the evolution of a population generation after generation, with the assumption that each individual in the population has a certain number of descendants, which has distribution $\pi$, independently from the other individuals.

The transition probability of the chain is:

$$
\begin{aligned}
r_{x,y} &= 1, \ \text{if } x = 0, y = 0, \\
&= \pi^{*x}(y), \ \text{if } x \geq 1, \qquad \Big(\pi^{*x} \overset{\text{def}}{=} \underbrace{\pi \overset{\text{convolution}}{*} \pi * \cdots * \pi}_{x \text{ times}}\Big) \\
&= 0, \ \text{if } x = 0, y \neq 0.
\end{aligned} \tag{3.18}
$$

### 3.1.4 Ehrenfest model of diffusion




$N = 5$

Fig. 3.3

$N$ molecules are distributed in two containers $A$ and $B$, and at each step a molecule is chosen at random and placed in the other container. The chain describes the number of molecules in container $A$.

The state space is $E = \{0, 1, 2, \dots, N\}$, and the transition probability is

$$
\begin{aligned}
r_{x,y} = &\ 1 - \frac{x}{N}, && \text{if } x < N \text{ and } y = x + 1, \\
= &\ \frac{x}{N}, && \text{if } x > 0 \text{ and } y = x - 1, \qquad (3.19) \\
= &\ 0, && \text{otherwise.}
\end{aligned}
$$

### 3.1.5 Residual waiting time

$\pi$ is a probability on $\mathbb{N}\backslash\{0\}$, and the state space is $E = \mathbb{N}\backslash\{0\}$.

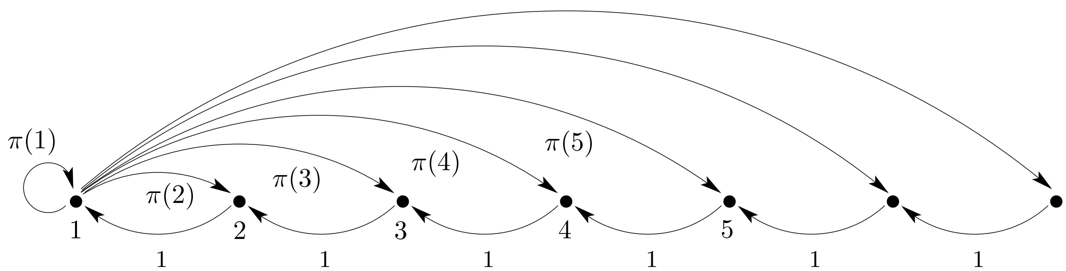



Fig. 3.4

The transition probability is

$$
\begin{aligned}
r_{x,y} = &\ \pi(y), && \text{if } x = 1 \text{ and } y \geq 1, \\
= &\ 1, && \text{if } x > 1 \text{ and } y = x - 1, \qquad (3.20) \\
= &\ 0, && \text{otherwise.}
\end{aligned}
$$

The chain describes the excess process $E_t$, at integer times $t \in \mathbb{N}$, if the inter-arrival times $(T_i)_{i \geq 1}$ have distribution $\pi$ (the initial distribution of the chain is $\pi$)

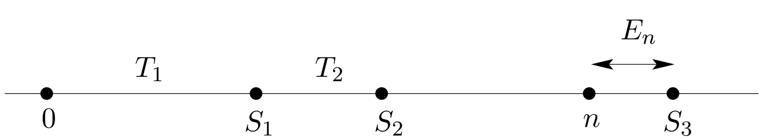



Fig. 3.5

## 3.2 Markov and strong Markov property

We consider the canonical shift on $\Omega_0$, see (3.11),

$$\begin{aligned}&\theta_n\colon \Omega_0 \to \Omega_0, \text{ for } n \geq 0, \text{ defined via } \theta_n(\omega)(\cdot) = \omega(n+\cdot),\\ &\quad \text{for } \omega \in \Omega_0, \text{ (the "trajectory } \omega(\cdot) \text{ shifted by } n \text{ units of time").}\end{aligned} \tag{3.21}$$

**Proposition 3.5.** *(simple Markov property)*

*If $\mu$ is a probability on $E$, $Y$ is a bounded $\mathcal{A}_0$-measurable variable on $\Omega_0$, then for $n \geq 0$,*

$$E_\mu[\underbrace{Y \circ \theta_n}_{\substack{\uparrow \\ \text{function of the "future after time } n\text{"}}} \,|\mathcal{F}_n] \overset{P_\mu\text{-a.s.}}{=} E_{X_n}[Y], \text{ (see (3.11), (3.12) for notation).} \tag{3.22}$$

*Proof.* It suffices to prove (3.22) for $Y = 1_A$, with $A \in \mathcal{A}_0$. Indeed, (3.22) then follows for linear combinations of indicator functions and then by monotone convergence for $Y \geq 0$, bounded and $\mathcal{A}_0$-measurable, and then for general $Y$ as claimed in (3.22) by taking differences.

Since $\mathcal{F}_n$ is generated by an at most countable partition of $\Omega_0$ into sets $\{X_0 = x_0, X_1 = x_1, \dots, X_n = x_n\}$, with $x_0, \dots, x_n$ in $E$, we only need to prove that for $x_0, \dots, x_n$ in $E$:

$$E_\mu[1_A \circ \theta_n; X_0 = x_0, \dots, X_n = x_n] = P_\mu[X_0 = x_0, \dots, X_n = x_n]\, P_{x_n}[A]. \tag{3.23}$$

When $A = \{X_0 = y_0, \dots, X_m = y_m\}$, one has

$$1_A \circ \theta_n = 1\{X_n = y_0, X_{n+1} = y_1, \dots, X_{n+m} = y_m\},$$

and (3.23) follows from (3.9).

The class of sets $A$, which are either empty or of the above form is a $\pi$-system generating $\mathcal{A}_0$. By Dynkin's lemma, see [12], p. 41, the claim (3.23) follows for general $A \in \mathcal{A}_0$, and this concludes the proof of (3.22). □

We will now replace $n$ in (3.22) with an $(\mathcal{F}_n)$-stopping time, to obtain the so-called strong Markov property. We recall that for $N$ an $(\mathcal{F}_n)$-stopping time (i.e. $N$ is $\mathbb{N} \cup \{\infty\}$-valued and $\{N = n\} \in \mathcal{F}_n$, for each $n \geq 0$), one defines the $\sigma$-algebra $\mathcal{F}_N$ of the "past of $N$", via:

$$\mathcal{F}_N \overset{\text{def}}{=} \{A \in \mathcal{A}_0;\, A \cap \{N = n\} \in \mathcal{F}_n, \text{ for each } n \geq 0\}. \tag{3.24}$$

**Theorem 3.6.** *(strong Markov property)*

*Let $N$ be an $(\mathcal{F}_n)$-stopping time, $Y$ a bounded $\mathcal{A}_0$-measurable variable, $\mu$ a probability on $E$, then one has:*

$$E_\mu[Y \circ \theta_N \,|\, \mathcal{F}_N] \overset{P_\mu\text{-a.s.}}{=} E_{X_N}[Y] \text{ on } \{\underset{\substack{\uparrow \\ \mathcal{F}_N\text{-measurable}}}{N < \infty}\}, \tag{3.25}$$

*(here $Y \circ \theta_N$ is understood as $Y \circ \theta_{N(\omega)}(\omega)$, if $N(\omega) < \infty$, and 0 otherwise, and $E_{X_N}[Y]$ is understood as $E_{X_{N(\omega)}(\omega)}[Y]$, when $N(\omega) < \infty$, and 0 otherwise).*

*Proof.* Observe that on $\{N = n\}$, $E_{X_N}[Y] = E_{X_n}[Y] \leftarrow \mathcal{F}_n$-measurable, and hence

(3.26) $$X_N \text{ is } \mathcal{F}_N\text{-measurable.}$$

Moreover, for $A \in \mathcal{F}_N$, we have

(3.27) $$\begin{array}{l} E_\mu[Y \circ \theta_N;\, A \cap \{N < \infty\}] = \sum\limits_{n \ge 0} E_\mu[Y \circ \theta_N;\, A \cap \{N = n\}] = \\ \qquad\qquad\qquad\qquad\qquad \text{simple Markov property} \\ \qquad\qquad\qquad\qquad\qquad \swarrow \\ \sum\limits_{n \ge 0} E_\mu\big[Y \circ \theta_n;\, \underbrace{A \cap \{N = n\}}_{\in \mathcal{F}_n}\big] \overset{(3.22)}{=} \sum\limits_{n \ge 0} E_\mu[E_{X_n}[Y];\, A \cap \{N = n\}] = \\ E_\mu[E_{X_N}[Y];\, A \cap \{N < \infty\}], \text{ and this proves (3.25).} \end{array}$$

□

## 3.3 Recurrence and transience

We consider the canonical Markov chain with state space $E$ and transition probability $(r_{x,y})_{x,y \in E}$, cf. (3.11), (3.12). Given $x \in E$, we define the **hitting time** of $x$:

(3.28) $$\widetilde{H}_x = \inf\{n \ge 1;\, X_n = x\} \le \infty\,.$$

It is an $(\mathcal{F}_n)$-stopping time (indeed one has:

$\{\widetilde{H}_x = k\} = \{X_1 \ne x, \dots, X_{k-1} \ne x,\ X_k = x\} \in \mathcal{F}_k$, for $k \ge 1$, and $\{\widetilde{H}_x = 0\} = \phi \in \mathcal{F}_0$.)

It is also useful to consider the **entrance time** in $x \in E$:

(3.29) $$H_x = \inf\{n \ge 0;\ X_n = x\} \le \infty\,.$$

It is also an $(\mathcal{F}_n)$-stopping (via similar arguments as below (3.28)). The difference between hitting and entrance times only has to do with whether the state of the chain at time 0 is taken into account or not.

**Definition 3.7.**

(3.30) $$\rho_{x,y} \overset{\text{def}}{=} P_x[\widetilde{H}_y < \infty], \text{ for } x, y \in E.$$

*A state $x \in E$ is said* ***recurrent*** *if*

(3.31) $$\rho_{x,x} = 1\,.$$

*It is said* ***transient*** *if*

(3.32) $$\rho_{x,x} < 1.$$

In general it is not easy to decide whether a state is recurrent or transient. Given $x \in E$, one defines the **successive times of visit** of the chain to $x$, $\widetilde{H}_x^n$, $n \geq 0$, via:

$$\begin{aligned} &\widetilde{H}_x^0 = 0, \widetilde{H}_x^1 = \widetilde{H}_x \leq \infty, \text{ and for } n \geq 1, \\ &\widetilde{H}_x^{n+1} = \widetilde{H}_x^n + \widetilde{H}_x \circ \theta_{\widetilde{H}_x^n} \leq \infty \text{ (understood as } +\infty \text{ on } \{\widetilde{H}_x^n = \infty\}). \end{aligned} \tag{3.33}$$

**Proposition 3.8.** $(x, y \in E)$

$$P_x[\widetilde{H}_y^n < \infty] = \begin{cases} \rho_{x,y}, & \textit{if } n = 1, \\ \rho_{x,y}\, \rho_{y,y}^{n-1}, & \textit{if } n > 1. \end{cases} \tag{3.34}$$

*Proof.* For $n = 1$, this is the definition (3.30). For $n > 1$, we write:

$$P_x[\widetilde{H}_y^n < \infty] \overset{(3.33)}{=} P_x[\widetilde{H}_y^{n-1} < \infty \text{ and } \widetilde{H}_y \circ \theta_{\widetilde{H}_y^{n-1}} < \infty],$$

applying the strong Markov property (3.25) we find

$$= E_x\big[\widetilde{H}_y^{n-1} < \infty,\ P_{\underbrace{X_{\widetilde{H}_y^{n-1}}}_{\nearrow \atop \text{on } \{\widetilde{H}_y^{n-1} < \infty\}, = y}}[\widetilde{H}_y < \infty]\big] \overset{(3.30)}{=} \rho_{y,y}\, P_x[\widetilde{H}_y^{n-1} < \infty] \overset{\text{induction}}{=} \rho_{x,y}\, \rho_{y,y}^{n-1},$$

and this proves (3.34). □

**Remark 3.9.** (link with renewal processes)

Note that as a consequence of the strong Markov property, under $P_x$,

$$S_n \overset{\text{def}}{=} \widetilde{H}_x^n,\ n \geq 0, \tag{3.35}$$

has same distribution as the sum

$$T_1 + \dots + T_n, \text{ when } n \geq 1, \text{ and equals } 0, \text{ when } n = 0,$$

for i.i.d. variables $(T_i)_{i \geq 1}$, with values in $\mathbb{N} \cup \{\infty\}$ having same distribution as $\widetilde{H}_x$ under $P_x$. So the counting function

$$N_t = \sum_{k \geq 1} 1\{S_k \leq t\} = \sup\{n \geq 0;\ S_n \leq t\},\ t \geq 0, \tag{3.36}$$

is a renewal process, which is defective when $x$ is transient, and has no defect when $x$ is recurrent. □

Given a state $x \in E$, we define

$$N_x = \sum_{k \geq 1} 1\{X_k = x\} = \sum_{n \geq 1} 1\{\widetilde{H}_x^n < \infty\}. \tag{3.37}$$

**Proposition 3.10.** $(y \in E)$

*If $y$ is recurrent, then*

$$P_y\text{-a.s., } N_y = \infty. \tag{3.38}$$

*If $y$ is transient, then for $x \in E$,*

$$E_x[N_y] = \frac{\rho_{x,y}}{1-\rho_{y,y}} < \infty. \tag{3.39}$$

*Proof.*

- <u>(3.38)</u>: With (3.34) we see that $P_y[\widetilde{H}_y^n < \infty] = 1$, for $n \geq 1$, and (3.38) follows.
- <u>(3.39)</u>:

$$E_x[N_y] = \sum_{n\geq 1} P_x[\underbrace{N_y \geq n}_{=\{\widetilde{H}_y^n<\infty\}}] = \sum_{n\geq 1} P_x[\widetilde{H}_y^n < \infty]$$

$$\overset{(3.34)}{=} \sum_{n\geq 1} \rho_{x,y}\, \rho_{y,y}^{n-1} = \frac{\rho_{x,y}}{1-\rho_{y,y}} < \infty\ .$$

□

One immediate consequence is the following

**Corollary 3.11.**

$$\textit{When } E \textit{ is finite, at least one } y \in E \textit{ is recurrent.} \tag{3.40}$$

*Proof.* Indeed, otherwise all $y \in E$ are transient and for any $x \in E$

$$E_x\Big[\sum_{y\in E} N_y\Big] \overset{(3.39)}{=} \sum_{y\in E} \frac{\rho_{x,y}}{1-\rho_{y,y}} \overset{(E \text{ finite})}{<\infty}\ .$$

On the other hand:

$$\sum_{y\in E} N_y \overset{(3.37)}{=} \sum_{y\in E}\sum_{k\geq 1} 1\{X_k = y\} = \sum_{k\geq 1} 1 = \infty,$$

a contradiction. □

We will now devise a **decomposition of the state space**. To this end we will first see that "anything that can be reached from a recurrent state is recurrent as well".

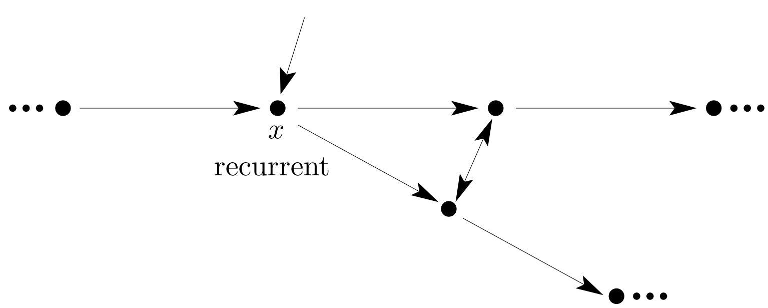



Fig. 3.6

**Proposition 3.12.**

(3.41) *If $x$ is recurrent and $\rho_{x,y} > 0$, then $y$ is recurrent and $\rho_{y,x} = 1$ (and therefore, switching the role of $x$ and $y$, $\rho_{x,y} = 1$ as well).*

*Proof.*

- We show that $\rho_{y,x} = 1$.

We assume $y \neq x$, otherwise this is clear (since $x$ is recurrent, cf. (3.31)). Let $k \geq 1$ be the smallest integer such that $P_x[X_k = y] > 0$. Then

$$0 < P_x[X_k = y] = \sum_{x_1,\dots,x_{k-1} \in E} \underbrace{P_x[X_1 = x_1, \dots, X_{k-1} = x_{k-1}, X_k = y]}_{\substack{\uparrow\,(3.9) \\ r_{x,x_1} \dots r_{x_{k-1},y}}} \,. \tag{3.42}$$

One of the above terms is $> 0$, and since $k$ is minimal, the corresponding $x_1, \dots, x_{k-1}$ are different from $x$ and $y$:

$$r_{x,x_1}\, r_{x_1,x_2} \dots r_{x_{k-1},y} > 0, \text{ for some } x_1, \dots, x_{k-1} \text{ different from } x, y. \tag{3.43}$$

Since $x$ is recurrent, we find that

different from $x$

$$\begin{aligned} 0 = P_x[\widetilde{H}_x = \infty] &\geq P_x[X_1 = x_1, \dots, X_k = y,\ \widetilde{H}_x \circ \theta_k = \infty] \\ &\overset{\text{Markov},(3.22)}{=} r_{x,x_1}, \dots, r_{x_{k-1},y}\, \underbrace{P_y[\widetilde{H}_x = \infty]}_{\uparrow\, 1-\rho_{y,x}} \end{aligned}$$

and therefore we find that

$$\rho_{y,x} = 1. \tag{3.44}$$

- We show that $y$ is recurrent (recall $y \neq x$, otherwise this is trivial). Since $\rho_{y,x} = 1$, there is an $\ell \geq 1$, such that $P_y[X_\ell = x] > 0$. Then for $n \geq 1$, and $k$ as above (3.42) we find

$$P_y[X_{\ell+n+k} = y] \geq P_y[X_\ell = x,\ X_{\ell+n} = x,\ X_{\ell+n+k} = y]$$

with the Markov property at time $\ell + n$, cf. (3.22), this equals

$$\begin{aligned} &= P_y[X_\ell = x,\ X_{\ell+n} = x]\ P_x[X_k = y] \overset{(3.22)\ \text{at time}\ \ell}{=} \\ &P_y[X_\ell = x]\ P_x[X_n = x]\ P_x[X_k = y]. \end{aligned}$$

Therefore with the notation of (3.37) we find

$$
(3.45) \qquad \begin{aligned} E_y[N_y] &\geq \sum_{n\geq 1} P_y[X_{\ell+n+k} = y] \geq P_y[X_\ell = x]\; P_x[X_k = y] \;\sum_{n\geq 1} P_x[X_n = x] \\ &= \underbrace{P_y[X_\ell = x]}_{>0}\; \underbrace{P_x[X_k = y]}_{>0} \quad \underbrace{E_x[N_x]}_{=\infty,\,\text{(recurrent)}} \; = \infty. \end{aligned}
$$

In view of (3.39) this show that $y$ is recurrent. □

**Example 3.13.** Consider the Galton-Watson chain (see Section 3.1.3), and assume that the probability $\pi$ on $\mathbb{N}$ describing the number of descendants of an individual satisfies:

$$
(3.46) \qquad \pi(0) > 0.
$$

Under (3.46) all states $x \geq 1$ are transient and 0 is the only recurrent state. Indeed, 0 is clearly recurrent, cf. (3.18), and for $x \geq 1$, we have

$$
\rho_{x,0} \geq \pi(0)^x > 0 \text{ (“none of the } x \text{ individuals has a descendant”)}.
$$

Thus if $x$ is recurrent, then $\rho_{0,x} \overset{(3.41)}{=} 1$, which is a contradiction since $\rho_{0,x} = P_0[\widetilde{H}_x < \infty] = 0$, in view of (3.18). □

We continue the discussion of recurrence and transience and will introduce a certain decomposition of the state space.

**Definition 3.14.** *Two states $x, y \in E$ are communicating (one writes $x \leftrightarrow y$), if*

$$
(3.47) \qquad P_x[H_y < \infty] > 0 \;\; \textit{and} \;\; P_y[H_x < \infty] > 0,
$$

*(in other words: $x \leftrightarrow y$ means $x = y$ or $\rho_{x,y}\, \rho_{y,x} > 0$).*

Note that **$x \leftrightarrow y$, $x, y \in E$, is an equivalence relation** on $E$. The only point to check is the transitivity, which follows from the fact that for $x, y, z \in E$,

$$
P_x[H_z < \infty] \geq P_x[H_y < \infty \;\text{ and }\; H_z \circ \theta_{H_y} < \infty] \overset{(3.25)}{=} P_x[H_y < \infty]\, P_y[H_z < \infty].
$$

When there is only one equivalence class in $E$, the chain is called **irreducible**, (i.e. for all $x, y \in E$, $x \leftrightarrow y$).

**Proposition 3.15.** *(state space decomposition)*

*One can partition $E$ into*

$$
(3.48) \qquad E = T \cup R_1 \cup R_2 \cup \dots\,,
$$

*where $T$ is the collection of transient states and each $R_i$ is an equivalence class for “$\leftrightarrow$” of recurrent states.*

*Proof.* The only point to observe is that an equivalence class for "$\leftrightarrow$" is either included in $T$ or in $T^c \overset{\text{def}}{=} R$ (the set of recurrent states), as follows from (3.41). $\square$

**Remark 3.16.**

1) It follows from (3.41) that for $x \in R_i$ (one of the recurrent classes) one has

$$P_x[X_n \in R_i, \text{ for all } n \geq 0] = 1 \tag{3.49}$$

(because for $x$ recurrent $\rho_{x,y} > 0 \overset{(3.41)}{\Longrightarrow} x \leftrightarrow y$). Hence, when the process starts in $R_i$, it remains in $R_i$ for ever.

2) When the process starts in $T$ it may remain in $T$ or at some point enter one of the $R_i$ (which it then never leaves).

$\square$

**Example 3.17.** Consider again the Galton-Watson chain, cf. (3.18), and assume that

$$0 < \pi(0) < 1, \text{ and } m = \sum_{k \geq 0} k\pi(k) \in (1, \infty), \text{ i.e. the "supercritical case"}. \tag{3.50}$$

Then, it is known (see for instance [14], p. 101) that

$$P_x[X_n \neq 0, \text{ for all } n \geq 0] > 0, \text{ for any } x \geq 1. \tag{3.51}$$

At the same time, we have seen that $\rho_{x,0} > 0$, cf. Example 3.13. This is a situation where

$$E = \mathbb{N} \text{ is partitioned into } E = T \cup \underset{\substack{\uparrow \\ R_1}}{R}, \quad \text{with } T = \mathbb{N} \backslash \{0\},\ R = \underset{\substack{\nearrow \\ \text{called "absorbing state"}}}{\{0\}}.$$

When starting in $T$, the chain may either remain for ever in $T$, or at some point reach $R = \{0\}$. $\square$

**Definition 3.18.** *A recurrent state $x \in E$, is called **positive recurrent** if*

$$E_x[\widetilde{H}_x] < \infty, \tag{3.52}$$

*it is called **null recurrent** if*

$$E_x[\widetilde{H}_x] = \infty. \tag{3.53}$$

*(To explain the terminology, recall that in view of the link with renewal processes in (3.36) and the law of large numbers in (2.34), when $x$ is recurrent, then one has*

$$P_x\text{-a.s., } \frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\} \underset{n \to \infty}{\longrightarrow} \frac{1}{E_x[\widetilde{H}_x]} \begin{array}{l} > 0, \quad \textit{if } x \textit{ is positive recurrent} \\ = 0, \quad \textit{if } x \textit{ is null recurrent .)} \end{array} \tag{3.54}$$

## 3.4 Stationary distribution

A probability $\pi$ on $E$ is called a stationary distribution of the Markov chain with transition probability $(r_{x,y})$ if

$$\text{for all } y \in E,\ \pi(y) = \sum_{x \in E} \pi(x)\, r_{x,y}. \tag{3.55}$$

**Remark 3.19.** Note that $\pi$ being a stationary distribution is equivalent to:

$$\underbrace{\theta_n \circ P_\pi}_{\nwarrow} = P_\pi, \text{ for any } n \geq 0. \tag{3.56}$$

image of $P_\pi$ under $\theta_n$ (on $\Omega_0$), in the notation of (3.11)

Indeed, (3.56) implies (3.55) because applying (3.56) with $n = 1$

$$\begin{aligned}\pi(y) = P_\pi(X_0 = y) &\overset{(3.56)}{=} \theta_1 \circ P_\pi[X_0 = y] = P_\pi[X_0 \circ \theta_1 = y] \\ = P_\pi[X_1 = y] &\overset{(3.10)}{=} \sum_{x \in E} \pi(x)\, r_{x,y}.\end{aligned}$$

Conversely, (3.55) implies (3.56) because for any $A \in \mathcal{A}_0$

$$\begin{aligned}\theta_1 \circ P_\pi[A] &= E_\pi[1_A \circ \theta_1] = E_\pi\big[E_\pi[1_A \circ \theta_1 \,|\, \mathcal{F}_1]\big] \overset{(3.22)}{=} E_\pi\big[P_{X_1}[A]\big] \\ &\underset{\substack{\text{(swapping} \\ \text{summations)}}}{\overset{(3.10)}{=}} \sum_{y \in E} \sum_{x \in E} \pi(x)\, r_{x,y}\, P_y[A] \overset{(3.55)}{=} \sum_{y \in E} \pi(y)\, P_y[A] = P_\pi[A].\end{aligned}$$

Since $\theta_n = (\theta_1)^n$, and we have $\theta_1 \circ P_\pi = P_\pi$, the identity (3.56) follows. Thus (3.55) and (3.56) are equivalent.

Thus (3.56) provides an interpretation for the terminology "stationary distribution": such an initial distribution $\pi$ will make $P_\pi$ invariant under the shifts $\theta_n$, $n \geq 0$. $\square$

We continue our discussion of **stationary distributions** of Markov chains. They play an **important role** in the study of the **asymptotic behavior** of the chain.

A special class of stationary distributions comes in the next definition.

**Definition 3.20.** *A probability $\pi$ on $E$ is called a* **reversible distribution** *of the Markov chain with transition probability $(r_{x,y})$ when*

$$\pi(x)\, r_{x,y} = \pi(y)\, r_{y,x}, \text{ for all } x, y \in E; \tag{3.57}$$

*(this is called the* **detailed balance** *condition).*

**Proposition 3.21.**

(3.58) *A reversible distribution is a stationary distribution.*

*Proof.* Let $\pi$ be a reversible distribution, then for $y \in E$,

$$\sum_{x \in E} \pi(x)\, r_{x,y} \overset{(3.57)}{=} \sum_{x \in E} \pi(y)\, r_{y,x} = \pi(y) \underbrace{\sum_{x \in E} r_{y,x}}_{1} = \pi(y),$$

i.e. $\pi$ satisfies (3.55). $\square$

**Remark 3.22.** The detailed balance condition (3.57) is easier to check than (3.55) (because the condition is local in nature, in the sense that it involves only two states $x, y$ at a time, as opposed to (3.55), which may involve many states at a time). The terminology "detailed balance" comes from the fact that when $\pi$ is reversible

$$\pi(x)\, r_{x,y} = \underbrace{P_\pi[X_0 = x, X_1 = y]}_{\substack{\text{"flux from } x \text{ to } y \\ \text{under the } \theta\text{-invariant} \\ \text{measure } P_\pi\text{"}}} = \underbrace{P_\pi[X_0 = y, X_1 = x]}_{\substack{\text{"flux from } y \text{ to } x \\ \text{under the } \theta\text{-invariant} \\ \text{measure } P_\pi\text{"}}} = \pi(y)\, r_{y,x}\,.$$

$\square$

**Example 3.23.** Consider the Markov chain on $E = \{1, \dots N\}$, with transition probability

$$\begin{aligned} r_{x,y} &= 1, \text{ if } x = N,\ y = 1 \\ &\quad 1, \text{ if } x < N,\ y = x + 1 \\ &\quad 0, \text{ otherwise.} \end{aligned}$$

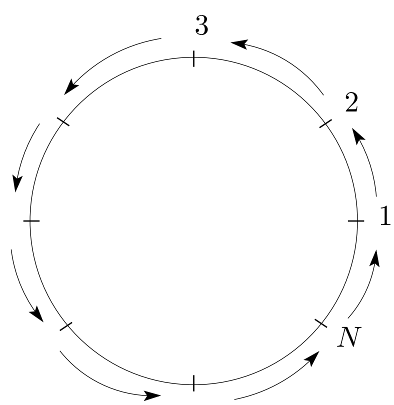



Fig. 3.7

Note that $r_{x,y}\, r_{y,x} \equiv 0$, and this Markov chain cannot have a reversible distribution. On the other hand, the uniform distribution on $E$:

$$\pi(x) = \frac{1}{N},\ x \in \{1, \dots, N\},$$

is clearly stationary.

**Example 3.24.** *Reflected random walk* (cf. (3.17)):

We look for a reversible distribution $\pi$ so that:

$$\pi(0) = q\pi(1), \text{ and for } x \geq 1,\ p\pi(x) = q\pi(x+1). \tag{3.59}$$

Necessarily, we have for $x \geq 1$:

$$\begin{aligned}
\pi(x+1) =& \Big(\frac{p}{q}\Big)\,\pi(x) \overset{\text{induction}}{=} \Big(\frac{p}{q}\Big)^x \pi(1) = \Big(\frac{p}{q}\Big)^x\, \frac{1}{q}\,\pi(0), \text{ and} \\
\pi(1) =& \frac{1}{q}\,\pi(0).
\end{aligned}$$

As a result:

$$\text{when } p < q, \text{ we can define } \pi(0) = c,\ \pi(x) = \frac{c}{q}\,\Big(\frac{p}{q}\Big)^{x-1}, \text{ for } x \geq 1, \tag{3.60}$$

with $c$ the constant defined by

$$1 = c\Big(1 + \frac{1}{q}\sum_{n\geq 0}\Big(\frac{p}{q}\Big)^n\Big) = c\Big(1 + \frac{1}{q}\,\frac{1}{1-\frac{p}{q}}\Big) = c\Big(1 + \frac{1}{q-p}\Big), \text{ i.e.}$$

$$c = \Big(1 + \frac{1}{q-p}\Big)^{-1} = \frac{q-p}{2q} \text{ (recall that } p + q = 1). \tag{3.61}$$

On the other hand, when

$$p \geq q, \tag{3.62}$$

one can prove with the help of results discussed below (see Proposition 3.26) that the Markov chain has no stationary distribution.

**Example 3.25.** *Ehrenfest model of diffusion* (cf. Section 3.1.4):

The state space is $E = \{0, 1, \dots, N\}$.

We consider the binomial distribution with parameter $N$ and $p = \frac{1}{2}$, i.e.

$$\pi(x) = \binom{N}{x}\, 2^{-N},\ 0 \leq x \leq N. \tag{3.63}$$

(Recall the total number of molecules is $N$, and the chain records the number of molecules in the container $A$. The choice of $\pi$ in (3.63) corresponds to deciding at time 0 in an i.i.d. fashion for each molecule to place it with equal probability in container $A$ or in container $B$.)

(3.64) $\pi$ is a reversible distribution for the Markov chain with transition probability (3.19).

Indeed, we only need to check (3.57) for $0 \le x < y = x+1 \le N$. Observe that

$$\pi(x)\, r_{x,x+1} = 2^{-N} \binom{N}{x}\Big(1 - \frac{x}{N}\Big) = 2^{-N}\ \frac{N!}{x!(N-x)!}\ \frac{N-x}{N} = 2^{-N}\binom{N-1}{x},$$

and that

$$\pi(x+1)\, r_{x+1,x} = 2^{-N} \binom{N}{x+1}\frac{x+1}{N} = 2^{-N}\ \frac{N!}{(x+1)!(N-1-x)!}\ \frac{x+1}{N} = 2^{-N}\binom{N-1}{x},$$

and this proves (3.64).

We will now see that a stationary distribution can only give positive mass to positive recurrent states, cf. (3.52).

**Proposition 3.26.** *Let $\pi$ be a stationary distribution, then for $x \in E$,*

$$\pi(x) > 0 \textit{ implies that } x \textit{ is positive recurrent.} \tag{3.65}$$

*Proof.*

$$\pi(x) \overset{(3.56)}{=} \frac{1}{n} \sum_{k=1}^{n} P_\pi[X_k = x] = E_\pi\Big[\frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\}\Big]. \tag{3.66}$$

We will now introduce a lemma that will also be useful later.

**Lemma 3.27.** *(this does not assume the existence of $\pi$).*

$$E_y\Big[\frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\}\Big] \underset{n\to\infty}{\longrightarrow} \frac{\rho_{y,x}}{E_x[\widetilde{H}_x]}, \textit{ for } x, y \in E. \tag{3.67}$$

$$\begin{aligned}
&E_y\Big[\frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\}\Big] = E_y[\widetilde{H}_x < \infty,\ \frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\}\Big] = \\
&\sum_{m\ge 1} E_y\Big[\widetilde{H}_x = m, \frac{1}{n}\Big( \underbrace{\sum_{k=1}^{m\wedge n} 1\{X_k = x\}}_{1\{m\le n\}} + \Big(\sum_{\ell=1}^{(n-m)_+} 1\{X_\ell = x\}\Big) \circ \theta_m\Big)\Big] \\
&\overset{\substack{\text{Markov property (3.22)}\\ \text{at time } m}}{=} \\
&\underbrace{\frac{1}{n} P_y[\widetilde{H}_x \le n]}_{\xrightarrow{n\to\infty} 0} + \sum_{m\ge 1} P_y[\widetilde{H}_x = m] E_x\Big[\underbrace{\frac{1}{n} \sum_{\ell=1}^{(n-m)_+} 1\{X_\ell = x\}}_{\substack{n\to\infty\ \downarrow\ \text{(3.54) when } x \text{ recurrent}\\ \text{(3.39) when } x \text{ transient}\\ E_x[\widetilde{H}_x]^{-1}}}\Big].
\end{aligned} \tag{3.68}$$

As a result, using dominated convergence, we find that

$$E_y\Big[\frac{1}{n}\sum_{k=1}^{n} 1\{X_k = x\}\Big] \underset{n\to\infty}{\longrightarrow} P_y[\widetilde{H}_x < \infty]\,\frac{1}{E_x[\widetilde{H}_x]},$$

and this proves (3.67).

Coming back to (3.66), we see using dominated convergence that

$$\begin{aligned}\pi(x) = \sum_{y\in E}\pi(y)\ \underbrace{E_y\Big[\frac{1}{n}\sum_{k=1}^{n} 1\{X_k = x\}\Big]}_{\le 1} &\underset{n\to\infty}{\longrightarrow} \sum_{y\in E}\pi(y)\,\rho_{y,x}\,E_x[\widetilde{H}_x]^{-1} \\ &= P_\pi[\widetilde{H}_x < \infty]\,E_x[\widetilde{H}_x]^{-1}.\end{aligned} \tag{3.69}$$

However, when $x$ is transient or null recurrent $E_x[\widetilde{H}_x] = \infty$, and hence $\pi(x) = 0$. □

**Theorem 3.28.** *Consider an irreducible Markov chain on $E$. One has the equivalences*

(3.70) *some $x \in E$ is positive recurrent,*

(3.71) *all $x \in E$ are positive recurrent,*

(3.72) *there is a stationary distribution.*

*Furthermore, if one of the above equivalent conditions holds, then*

$$\pi(x) = \frac{1}{E_x[\widetilde{H}_x]}\,(>0),\ x\in E,\ \textit{is the unique stationary distribution.} \tag{3.73}$$

• <u>(3.70) $\Longrightarrow$ (3.71)</u>:

Let $x_0 \in E$ be a positive recurrent state. From (3.41) or (3.48) we know that all states in $E$ are recurrent and communicating. Set for $x \in E$,

$$\nu(x) \overset{\substack{\text{well-defined by}\\ (3.67)}}{=} \lim_{n\to\infty} \underbrace{E_z}_{\text{arbitrary } z\in E}\Big[\frac{1}{n}\sum_{k=1}^{n} 1\{X_k = x\}\Big] \overset{(3.67)}{=} E_x[\widetilde{H}_x]^{-1}\ (\text{recall } \rho_{z,x} = 1). \tag{3.74}$$

From Fatou's lemma, we know that

$$\sum_{x\in E}\nu(x) \le 1 = \lim_n \underbrace{\sum_{x\in E} E_z\Big[\frac{1}{n}\sum_{k=1}^{n} 1\{X_k = x\}\Big]}_{=1}. \tag{3.75}$$

Since we have for any $z \in E$

(3.76)
$$\underbrace{E_z\Big[\frac{1}{n+1}\sum_{k=1}^{n+1} 1\{X_k = x\}\Big]}_{\substack{n\to\infty \downarrow \\ \nu(x)}} = \frac{1}{n+1}\sum_{k=1}^{n+1} P_z[X_k = x]$$
$$= \frac{1}{n+1}\sum_{k=1}^{n+1}\sum_{y\in E} P_z[X_{k-1} = y, X_k = x]$$
$$\overset{\substack{\text{Markov}\\\text{property}}}{=} \frac{1}{n+1}\sum_{k=1}^{n+1}\sum_{y\in E} P_z[X_{k-1} = y]\, r_{y,x} = \sum_{y\in E} \underbrace{E_z\Big[\frac{1}{n+1}\sum_{k=0}^{n} 1\{X_k = y\}\Big]}_{\substack{n\to\infty \downarrow (3.74) \\ \nu(y)}} r_{y,x},$$

we can again apply Fatou's lemma and find:

(3.77)
$$\nu(x) \geq \sum_{y\in E} \nu(y)\, r_{y,x}, \text{ for all } x \in E.$$

Summing over $x$ in $E$ we obtain:

$$\sum_{x\in E} \nu(x) \geq \sum_{y\in E} \nu(y) \sum_{x\in E} r_{y,x} = \sum_{y\in E} \nu(y),$$

and in view of (3.75), we see that (3.77) has to be an equality:

(3.78)
$$\nu(x) = \sum_{y\in E} \nu(y)\, r_{y,x}, \text{ for all } x \in E.$$

Assume that $\overline{x}$ is a null recurrent state. Since $\rho_{x_0,\overline{x}} = 1$, as in (3.42), we can find $x_0, x_1, \dots, x_n = \overline{x}$ with $r_{x_i,x_{i+1}} > 0$, $0 \leq i < n$. Since $\nu(\overline{x}) \overset{(3.74)}{=} E_{\overline{x}}[\widetilde{H}_{\overline{x}}]^{-1} = 0$, it follows from (3.77) with $\overline{x}$ in place of $x$ that $\nu(x_{n-1}) = 0$ and by induction that

$$\nu(\overline{x}) = \nu(x_{n-1}) = \cdots = \nu(x_0) = 0, \text{ a contradiction since } x_0 \text{ is positive recurrent.}$$

- <u>(3.71) $\Longrightarrow$ (3.72)</u>:

We see that $\nu$ in (3.74) has a positive mass $\leq 1$, and it satisfies (3.78). As a result

$$\pi = \frac{1}{\sum_{x\in E} \nu(x)}\, \nu \text{ is a stationary distribution.}$$

- <u>(3.72) $\Longrightarrow$ (3.70)</u>: This follows directly from (3.65).

- <u>(3.71) $\Longrightarrow$ (3.73)</u>:

We consider $\pi$ a stationary distribution, and we know that $\rho_{y,x} = 1$, for all $y, x \in E$. Then with (3.69) $\pi(x) = E_x[\widetilde{H}_x]^{-1}$, for $x \in E$, and this proves the claim. $\square$

**Remark 3.29.** One can give an alternative formula to (3.73), namely:

$$\pi(x) = \frac{E_y\Big[\sum\limits_{k=0}^{\widetilde{H}_y-1} 1\{X_k = x\}\Big]}{E_y[\widetilde{H}_y]}, \quad x, y \in E, \text{ (when } y \text{ is chosen as } x, \text{ this coincides with (3.73)),} \tag{3.79}$$

where $\pi$ is the unique stationary distribution under the assumptions of Theorem 3.28. Indeed, note that $P_y$-a.s., when for $m \geq 1$, $\widetilde{H}_y^m \leq n < \widetilde{H}_y^{m+1}$ (cf. (3.33)), one has

$$\begin{aligned} &\frac{m}{\widetilde{H}_y^{m+1}} \cdot \frac{1}{m} \sum_{\ell=0}^{m-1} \Big(\sum_0^{\widetilde{H}_y-1} 1\{X_k = x\}\Big) \circ \theta_{\widetilde{H}_y^\ell} \leq \\ &\frac{1}{n} \sum_{k=0}^{n} 1\{X_k = x\} \leq \frac{1}{m+1} \sum_{\ell=0}^{m} \Big(\sum_0^{\widetilde{H}_y-1} 1\{X_k = x\}\Big) \circ \theta_{\widetilde{H}_y^\ell} \cdot \frac{m+1}{\widetilde{H}_y^m}\,. \end{aligned} \tag{3.80}$$

By the strong Markov property $(\sum_0^{\widetilde{H}_y-1} 1\{X_k = x\}) \circ \theta_{\widetilde{H}_y^\ell}$, $\ell = 0, 1, \dots$ are i.i.d. variables under $P_y$ and $\widetilde{H}_y \circ \theta_{\widetilde{H}_y^\ell}$, $\ell = 0, 1, \dots$ are also i.i.d. variables. As a result of the strong law of large numbers, the expressions on the left and the right of (3.80) converge $P_y$-a.s. to the expression in the right-hand side of (3.79), when $n$ tends to infinity. On the other hand, we know from (3.67), (3.73) that

$$E_y\Big[\frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\}\Big] \underset{n\to\infty}{\longrightarrow} \pi(x).$$

The formula (3.79) now follows. $\square$

## 3.5 Asymptotic behavior

We know from (3.67) that for $y, x$ in $E$,

$$E_y\Big[\frac{1}{n} \sum_{k=1}^{n} 1\{X_k = x\}\Big] = \frac{1}{n} \sum_{k=1}^{n} P_y[X_k = x] \underset{n\to\infty}{\longrightarrow} \rho_{y,x}\, E_x[\widetilde{H}_x]^{-1}.$$

We are going to derive further results concerning the asymptotic behavior of $P_y[X_n = x]$, as $n \to \infty$, and replace, when possible, Cesàro convergence (as above), with convergence. We will see that there are some obstructions to this program. We begin with the

**Proposition 3.30.** *Assume $x$ is transient, then for $y \in E$,*

$$P_y[X_n = x] \underset{n\to\infty}{\longrightarrow} 0. \tag{3.81}$$

*Proof.* We know that, cf. (3.39), $P_y$-a.s., $\sum_{k\geq 1} 1\{X_k = x\} < \infty$, and therefore

$$\begin{aligned} &L_x \overset{\text{def}}{=} \sup\{k \geq 1;\, X_k = x\} < \infty,\ P_y\text{-a.s., “time of last visit to } x\text{”}, \\ &\qquad (L_x = 0, \text{ by convention when } \{\dots\} = \emptyset). \end{aligned} \tag{3.82}$$

But for $n \geq 1$, $P_y[X_n = x] \leq P_y[L_x \geq n] \underset{n\to\infty}{\longrightarrow} 0$, and (3.81) follows. $\square$

The above proposition handles the case of transient states. We will need the

**Definition 3.31.** *For $x \in E$, the period of $x$ is*

$$d_x = \text{greatest common divisor of } I_x \overset{\text{def}}{=} \Big\{n \geq 1;\ \underbrace{P_x[X_n = x]}_{\substack{\uparrow \\ r_{x,x}(n),\ cf.\ (3.5)}} > 0\Big\} \quad \text{(with the convention } d_x = 1, \text{ if } I_x = \emptyset) \tag{3.83}$$

**Proposition 3.32.** *If $x$ and $y$ are communicating (see Definition 3.14), then*

$$d_x = d_y. \tag{3.84}$$

*Proof.* We can assume $x \neq y$. Let $k, \ell \geq 1$, be such that $r_{x,y}(k) > 0$, $r_{y,x}(\ell) > 0$, cf. (3.5). Then $r_{y,y}(\ell + k) \overset{(3.7)}{\geq} r_{y,x}(\ell)\, r_{x,y}(k) > 0$, so that $\ell + k \in I_y$ and thus

$$d_y \text{ divides } \ell + k. \tag{3.85}$$

If $n \in I_x$, then $r_{y,y}(\ell + n + k) \overset{(3.7)}{\geq} r_{y,x}(\ell)\, r_{x,x}(n)\, r_{x,y}(k) > 0$, and $d_y$ divides $\ell + n + k$, and in view of (3.85), $d_y$ divides $n$. So $d_y$ divides $d_x$. Interchanging $x$ and $y$, $d_x$ divides $d_y$, and the claim (3.84) follows. □

**Definition 3.33.** *An irreducible Markov chain is called* ***aperiodic*** *if $d_x = 1$, for all $x \in E$.*

**Theorem 3.34.** *Consider an* ***irreducible*** *recurrent Markov chain. Then,* ***if*** *it is* ***null recurrent*** *(cf. Theorem 3.28),*

$$P_y[X_n = x] \underset{n \to \infty}{\longrightarrow} 0, \text{ for all } x, y \in E. \tag{3.86}$$

*If instead it is* ***positive recurrent*** *and* ***aperiodic***

$$P_y[X_n = x] \underset{n \to \infty}{\longrightarrow} E_x[\widetilde{H}_x]^{-1} (= \underset{\substack{\nwarrow \\ \text{unique stationary distribution}}}{\pi(x)},\ cf.\ (3.73)), \text{ for all } y, x \in E. \tag{3.87}$$

*Proof.* For $n \geq 1$, $y, x \in E$, one has

$$\begin{aligned} P_y[X_n = x] &= P_y[\widetilde{H}_x \leq n, X_n = x] = \sum_{m=1}^{n} P_y[\widetilde{H}_x = m,\ X_{n-m} \circ \theta_m = x] \\ &\overset{\substack{\text{Markov} \\ \text{property}}}{=} \sum_{m=1}^{n} P_y[\widetilde{H}_x = m]\, P_x[X_{n-m} = x]. \end{aligned} \tag{3.88}$$

Applying Blackwell's theorem in the case of an arithmetic distribution, cf. Remark 2.15 1) and Remark 3.9, we know that

$$P_x[X_{n-m} = x] \underset{n \to \infty}{\longrightarrow} 0, \text{ if } x \text{ is null recurrent.} \tag{3.89}$$

On the other hand, if the chain is aperiodic and $x$ positive recurrent,

$$P_x[X_{n-m} = x] \underset{n\to\infty}{\longrightarrow} E_x[\widetilde{H}_x]^{-1}. \tag{3.90}$$

Using dominated convergence in (3.88), the claim (3.86), (3.87) follow. □

**Remark 3.35.**

1) We based the proof on Blackwell's theorem for arithmetic distributions. We also refer to Resnick [12], p. 134, 128, or to Durrett [4], p. 263 for further details.

2) In case the chain is not aperiodic, then (3.87) can clearly break down. For instance this is the case for Example 3.23 (deterministic motion on the discrete circle).

   This example captures what happens when the chain is not aperiodic because:
   when the **chain** is **irreducible and recurrent**, and **all states** have **period *d***, one can partition $E$ into

$$\begin{aligned} &E = E_0 \cup E_1 \cup \dots \cup E_{d-1}, \\ &\text{so that when } x \in E_i,\ r_{x,y} > 0 \quad \begin{aligned}&\text{implies } y \in E_{i+1}, \text{ if } i < d-1, \\ &\text{implies } y \in E_0, \text{ if } i = d-1.\end{aligned}\end{aligned} \tag{3.91}$$

   If one considers the chain at times, which are integer multiples of $d$, it is a Markov chain on $E$ with transition kernel $(r_{x,y}(d))_{x,y\in E}$, which has recurrence classes $E_0, E_1, \dots, E_{d-1}$, and each of these recurrence classes is aperiodic (so one can apply the previous theorem to this chain). The decomposition (3.91) is called "cyclic decomposition", cf. [4], p. 286.

□

# 4 Continuous time Markov chains

We consider as in the previous chapter the state space

$$E\text{: an at most denumerable non-empty set.} \tag{4.1}$$

**Definition 4.1.** *A collection $(X_t)_{t\ge 0}$ of random variables with values in $E$ defined on some $(\Omega, \mathcal{A}, P)$ is a* **continuous time Markov chain** *with state space $E$, when:*

$$\begin{aligned} &E[f(X_{t_{n+1}}) \,|\, X_{t_0}, \dots, X_{t_n}] \overset{P\text{-a.s.}}{=} E[f(X_{t_{n+1}}) \,|\, X_{t_n}], \\ &\text{for any } f \text{ bounded } E \to \mathbb{R},\ n \ge 0, \text{ and } 0 \le t_0 < t_1 < \dots < t_n < t_{n+1}. \end{aligned} \tag{4.2}$$

*Of special interest to us in this chapter will be the study of the so-called* **pure jump processes** *(with no explosion).*

Loosely speaking this corresponds to temporally homogeneous Markov chains that remain at a location for a positive duration and perform only finitely many jumps during a finite time. More precisely, we endow $E$ with the discrete topology and introduce the canonical space

$$\begin{aligned} \Omega = \big\{&\text{functions } \omega(\cdot) \text{ from } \mathbb{R}_+ \text{ into } E, \text{ right-continuous with} \\ &\text{finitely many jumps on each compact interval}\big\} \end{aligned} \tag{4.3}$$

(in other words; an $\omega \in \Omega$ is a function of the form:

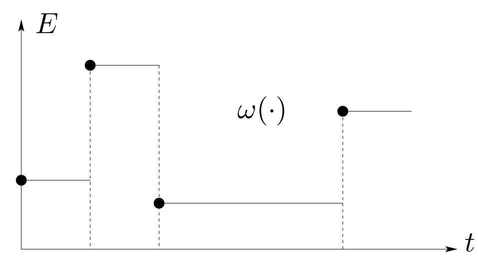



Fig. 4.1

).

One has a canonical $\sigma$-algebra on $\Omega$:

$$\begin{aligned} \mathcal{F} = \sigma\big(\omega(s), s \ge 0\big)&, \text{ i.e. the smallest } \sigma\text{-algebra on } \Omega, \text{ for which all canonical} \\ &\text{coordinates: } \omega(s)\colon \Omega \to E,\ s \ge 0, \text{ are measurable,} \end{aligned} \tag{4.4}$$

a canonical filtration on $\Omega$:

$$\mathcal{F}_t = \sigma\big(\omega(s), 0 \le s \le t\big),\ t \ge 0, \tag{4.5}$$

a canonical process:

$$X_t(\omega) = \omega(t),\ t \ge 0,\ \text{for } \omega \in \Omega, \tag{4.6}$$

as well as a canonical shift

$$\theta_t(\omega)(\cdot) = \omega(t+\cdot) \in \Omega,\ \text{for } t \ge 0,\ \omega \in \Omega. \tag{4.7}$$

**Definition 4.2.** *A pure jump process (with no explosion) on the state space $E$, is a collection $P_x$, $x \in E$, of probability measures on $(\Omega, \mathcal{F})$, such that:*

$$\begin{cases} \text{i)} & E_x[f(X_{t_{n+1}}) \,|\, X_{t_0}, \dots, X_{t_n}] \overset{P_x\text{-a.s.}}{=} E_{X_{t_n}(\omega)}[f(X_{t_{n+1}-t_n})], \\ & \textit{for all } x \in E,\ n \ge 0,\ 0 \le t_0 < t_1 < \dots < t_n < t_{n+1},\ \textit{and } f\colon \\ & E \to \mathbb{R},\ \textit{bounded, and such that} \\ \text{ii)} & P_x[X_0 = x] = 1,\ \textit{for all } x \in E. \end{cases} \tag{4.8}$$

One can replace (4.8) i) with the equivalent formulation:

$$E_x[f(X_{t+h}) | \mathcal{F}_t] \overset{P_x\text{-a.s.}}{=} E_{X_t(\omega)}[f(X_h)], \tag{4.8$'$ i)}$$

for any $x \in E$, $t, h \ge 0$, and $f$: $E \to \mathbb{R}$, bounded, (4.8)' i) $\Longrightarrow$ (4.8) i) is immediate and (4.8) i) $\Longrightarrow$ (4.8)' i) uses Dynkin's lemma). These conditions (as in (3.22)) yield the

**(simple) Markov property:**

$$E_x[Y \circ \theta_t \,|\, \mathcal{F}_t] \overset{P_x\text{-a.s.}}{=} E_{X_t}[Y],\ \text{for } t \ge 0,\ Y \text{ bounded } \mathcal{F}\text{-measurable.} \tag{4.9}$$

One can then define the **transition probability** of the chain:

$$r_{x,y}(t) = P_x[X_t = y],\ \text{for } t \ge 0,\ x, y \in E. \tag{4.10}$$

As a direct consequence of the above definition, we obtain

**Proposition 4.3.** *Given a pure jump process (with no explosion) in $E$, one has*

$$r_{x,y}(t) \ge 0,\ \sum_{z \in E} r_{x,z}(t) = 1,\ \textit{for } t \ge 0,\ x, y \in E, \tag{4.11}$$

$$r_{x,y}(t+s) = \sum_{z \in E} r_{x,z}(t)\, r_{z,y}(s),\ \textit{for } t, s \ge 0,\ x, y \in E \quad \textit{(Chapman-Kolmogorov equations)}, \tag{4.12}$$

$$\lim_{t \to 0} r_{x,y}(t) = 1_{\{x=y\}} = r_{x,y}(0),\ \textit{for } x, y \in E, \tag{4.13}$$

$$R_t\, f(x) \overset{\text{def}}{=} \sum_{z \in E} r_{x,z}(t)\, f(z) = E_x[f(X_t)],\ t \ge 0, f:\ E \to \mathbb{R},\ \textit{bounded, defines} \tag{4.14}$$

*a semi-group of bounded operators on $L^\infty(E)$ (i.e. $R_{t+s} = R_t\, R_s, t, s \ge 0$).*

*Proof.*

- (4.11): obvious.

- (4.12): $r_{x,y}(t+s) = P_x[X_{t+s} = y] = E_x\big[P_x[X_{t+s} = y \mid X_t]\big]$
$$\overset{(4.8)i)}{=} E_x\big[P_{X_t(\omega)}[X_s = y]\big] \overset{(4.10)}{=} \sum_{z \in E} r_{x,z}(t)\, r_{z,y}(s).$$

- (4.13): $r_{x,y}(t) = P_x[X_t = y] \xrightarrow[t \to 0]{\text{dominated convergence}} P_x[X_0 = y] = 1_{\{x=y\}}.$

- (4.14): This is a direct application of (4.12). $\square$

We now continue the **investigation of pure jump processes** (with no explosion) introduced in Definition 4.1. To this end we further introduce on $\Omega$, cf. (4.3),

$$T = \inf\{s \geq 0; X_s \neq X_0\} (\leq \infty). \tag{4.15}$$

$T$ is an $(\mathcal{F}_t)$-stopping time, indeed for any $t \geq 0$,

$$\{T \leq t\} = \bigcup_{r \in (0,t] \cap \mathbb{Q}} \{X_r \neq X_0\} \cup \{X_t \neq X_0\} \in \mathcal{F}_t.$$

rational numbers

**Proposition 4.4.** *Given a pure jump process (with no explosion) on $E$, then for $x \in E$,*

(4.16) *there is $\lambda(x) \in [0,\infty)$ such that $P_x[T > t] = e^{-\lambda(x)\, t}$, for $t \geq 0$ (i.e. when $\lambda(x) > 0$, $T$ is exponential($\lambda(x)$)-distributed, and when $\lambda(x) = 0$, $T$ is $P$-a.s. infinite, and $x$ is then called "absorbing").*

*If $\lambda(x) > 0$, then*

(4.17) *$T$ and $X_T\big(= X_{T_{(\omega)}}(\omega)\big)$ are independent under $P_x$,*

(4.18) *for $A \in \mathcal{F}$, $E_x[1_A \circ \theta_T \mid T, X_T] \overset{P_x\text{-a.s.}}{=} P_{X_T}[A]$.*

*Proof.*

- (4.16): $P_x[T > t+s] = E_x[\underbrace{T > t}_{\in \mathcal{F}_t},\ 1_{\{T \geq s\}} \circ \theta_t] \overset{(4.9)}{=}$
$$E_x\big[T > t, P_{X_t}[T > s]\big] = P_x[T > t]\, P_x[T > s].$$
$x$ on $\{T > t\}$

As a result, $\varphi(t) = P_x[T > t]$ is non-increasing and tends to 1 as $t \to 0$, and satisfies $\varphi(t+s) = \varphi(t)\,\varphi(s)$. The claim easily follows.

- <u>(4.17)</u>: Let $f\colon E \to \mathbb{R}$ be bounded, then for $t \ge 0$:

$$
\begin{aligned}
&E_x[T > t, f(X_T)] = E_x[T > t,\, f(X_T)\circ\theta_t] \overset{(4.9)}{=} \\
&E_x\big[T > t, \underset{\substack{\nwarrow \\ x \text{ on } \{T>t\}}}{E_{X_t}}[f(X_T)]\big] = P_x[T>t]\,E_x[f(X_T)].
\end{aligned}
\tag{4.19}
$$

The claim now follows with the help of Dynkin's lemma.

- <u>(4.18)</u>: Define for $n \ge 1$:

$$
T^{(n)} = \sum_{k=0}^{\infty} \frac{k+1}{2^n}\,1\Big\{\frac{k}{2^n} \le T < \frac{k+1}{2^n}\Big\} + \infty\, \underset{\swarrow\, (P_x\text{-negligible event})}{1\{T=\infty\}}, \text{ so that}
\tag{4.20}
$$

$$
T^{(n)} \text{ are } (\mathcal{F}_t)\text{-stopping times, and } T^{(n)} \downarrow T, \text{ as } n \to \infty.
\tag{4.21}
$$

With $h$ bounded continuous on $\mathbb{R}_+$, $f$, $f_i$ bounded on $E$, $0 \le i \le m$, and $Y = \prod_{i=0}^m f_i(X_{t_i})$, where $t_0 = 0 < t_1 < \dots < t_m$, we have (note that $P_x$-a.s., all $T_n$ are finite)

$$
\begin{aligned}
&E_x[h(T^{(n)})\, f(X_{T^{(n)}})\, Y\circ\theta_{T^{(n)}}] = \\
&\sum_{k\ge 0} E_x\Big[h\Big(\frac{k+1}{2^n}\Big)\, 1\Big\{\frac{k}{2^n} \le T < \frac{k+1}{2^n}\Big\}\, f\Big(X_{\frac{k+1}{2^n}}\Big) Y\circ\theta_{\frac{k+1}{2^n}}\Big] \overset{(4.9)}{=} \\
&\sum_{k\ge 0} E_x\Big[h\Big(\frac{k+1}{2^n}\Big)\, 1\Big\{\frac{k}{2^n} \le T < \frac{k+1}{2^n}\Big\}\, f\Big(X_{\frac{k+1}{2^n}}\Big)\, E_{X_{\frac{k+1}{2^n}}}[Y]\Big] = \\
&E_x\big[h(T^{(n)})\, f(X_{T^{(n)}})\, E_{X_{T^{(n)}}}[Y]\big].
\end{aligned}
\tag{4.22}
$$

Note that $P_x$-a.s.,

$$
\begin{aligned}
&h(T^{(n)})\, f(X_{T^{(n)}})\, Y\circ\theta_{T^{(n)}} = h(T^{(n)})\, f(X_{T^{(n)}}) \prod_{i=0}^m f_i(X_{T^{(n)}+t_i}) \xrightarrow[n\to\infty]{\substack{h \text{ is continuous, } s \to X_s(\omega) \text{ right-continuous,} \\ \downarrow \text{ and (4.21)}}} \\
&h(T)\, f(X_T)\, Y\circ\theta_T, \text{ as well as} \\
&h(T^{(n)})\, f(X_{T^{(n)}})\, E_{X_{T^{(n)}}}[Y] \underset{n\to\infty}{\longrightarrow} h(T)\, f(X_T)\, E_{X_T}[Y],
\end{aligned}
$$

and all these quantities are bounded. It thus follows by dominated convergence in (4.22) that

$$E_x[h(T)\,f(X_T)\,Y\circ\theta_T] = E_x\big[h(T)\,f(X_T)\,E_{X_T}[Y]\big]. \tag{4.23}$$

Using Dynkin's lemma, (4.18) follows. □

In view of the above proposition, given a pure jump process (with no explosion) on $E$, we can introduce

$$\lambda(x)\in[0,\infty), x\in E, \text{ such that } P_x[T>t]=e^{-\lambda(x)t}, \text{ for } t\ge 0, \tag{4.24}$$

the **jump rate**,

$$\begin{aligned} q_{x,y} &= P_x[X_T=y], \text{ for } x\in E, \text{ with } \lambda(x)>0, \text{ and } y\in E,\\ &= 1_{\{x=y\}}, \text{ for } x\in E, \text{ with } \lambda(x)=0, \text{ and } y\in E, \end{aligned} \tag{4.25}$$

the **jump transition probability**.

**Remark 4.5.** With the help of (4.18) it is not hard to see that **given a jump rate function $\boldsymbol{\lambda(\cdot)}$**: $E\to[0,\infty)$ and **a jump transition probability $\boldsymbol{(q_{x,y})}$ on $\boldsymbol{E}$, compatible with $\boldsymbol{\lambda(\cdot)}$**, i.e., such that:

$$q_{x,y}\ge 0,\ \sum_{y\in E} q_{x,y}=1, \text{ and } q_{x,x}=0, \text{ if } \lambda(x)>0,\ q_{x,x}=1, \text{ if } \lambda(x)=0, \tag{4.26}$$

then **there is at most one pure jump process with no explosion** which has jump rate $\lambda(\cdot)$ and jump transition probability $(q_{x,y})_{x,y\in E}$.

This simply comes from the iteration of (4.18), which shows that the law of the $(S_n, X_{S_n})_{n\ge 0}$, under each $P_x$, where we set

$$S_0=0, S_1=T,\dots,S_{n+1}=T\circ\theta_{S_n}+S_n\le\infty, \tag{4.27}$$

are the successive times of jump, and by convention we set

$$X_{S_n}=X_{S_{n-1}}, \text{ if } S_n=\infty, \text{ (i.e. } X_{S_n}=X_{S_k} \text{ with } k\le n, \text{ the largest integer such that } S_k<\infty),$$

is uniquely determined (to this end, note that $S_{n+1}=T+S_n\circ\theta_T$ and $X_{S_{n+1}}=X_{S_n}\circ\theta_T$, for each $n\ge 0$, and use induction together with (4.18)).

Since the **canonical space $\boldsymbol{\Omega}$** in (4.3) **rules out explosions**, i.e. $\lim_n S_n(\omega)=\infty$, for $\omega\in\Omega$, $\omega$ is fully determined by $(S_n, X_{S_n})_{n\ge 0}$. Indeed, one has

$$\omega(t)=X_{S_n} \text{ on } \{S_n\le t<S_{n+1}\},\ t\ge 0,\ n\ge 0. \tag{4.28}$$

This settles the question of the uniqueness of pure jump processes with no explosion. □

**Example 4.6.** (Poisson process with rate $\lambda$)

We consider $E = \mathbb{N} = \{0, 1, 2, \dots\}$ and for $x \in \mathbb{N}$,

$$P_x = \text{the law on } \Omega \text{ of } (x + N_t)_{t \geq 0}, \tag{4.29}$$

where $(N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$. Then we deduce from (1.34) that $P_x$ defines a pure jump process on $\mathbb{N}$ with

$$\begin{aligned} r_{x,y}(t) =& e^{-\lambda t}\, \frac{(\lambda t)^{y-x}}{(y-x)!}, \text{ if } y \geq x \text{ in } \mathbb{N},\, t > 0 \\ & 0, \text{ if } y < x. \end{aligned} \tag{4.30}$$

Since for a Poisson process the jump inter-arrival times are exponential($\lambda$)-distributed, and the jumps have size 1, we see that

$$\lambda(\cdot) \equiv \lambda \text{ is the jump rate,} \tag{4.31}$$

$$q_{x,y} = 1_{\{y=x+1\}}, \text{ is the jump transition probability.} \tag{4.32}$$

□

## 4.1 Construction of pure jump processes (with no explosion)

As explained in the Remark 4.5, given a rate function $\lambda(\cdot)$ on $E$ and a jump transition probability $(q_{x,y})$ satisfying the requirements (4.26), there is at most one pure jump process with no explosion satisfying (4.24), (4.25).

We will now investigate the existence of such an object. This will lead to additional conditions to rule out "explosions" (i.e. accumulations of jumps in finite time). We thus consider

$$\lambda(\cdot) : E \to [0, \infty) \tag{4.33}$$

$$\begin{aligned} &(q_{x,y})_{x,y\in E}, \text{ with } q_{x,y} \geq 0, \sum_{y\in E} q_{x,y} = 1, \text{ for } x \in E, \text{ and} \\ &\qquad q_{x,x} = 0, \text{ if } \lambda(x) > 0, \; q_{x,x} = 1, \text{ if } \lambda(x) = 0. \end{aligned} \tag{4.34}$$

The transition probability $(q_{x,y})$ on $E$ enables us to construct the laws $\overline{P}_x$, $x \in E$, on $\overline{\Omega} = E^{\mathbb{N}}$ of the canonical discrete time Markov chain with transition probability $(q_{x,y})$, see Proposition 3.4. We denote by $\overline{X}_n$, $n \geq 0$, the canonical process on $\overline{\Omega}$.

We then consider on some auxiliary probability space $(\Omega_{\text{aux}}, \mathcal{A}_{\text{aux}}, P)$, independent variables $T_n(y)$, $n \geq 0$, $y \in E$, with

$$\begin{aligned} &P[T_n(y) > t] = e^{-\lambda(y)t}, \text{ when } \lambda(y) > 0, \text{ and} \\ &T_n(y) \equiv \infty, \text{ when } \lambda(y) = 0. \end{aligned} \tag{4.35}$$

We then define on $\overline{\Omega} \times \Omega_{\text{aux}}$ the variables

$$\begin{aligned} &S_0 = 0,\ S_1 = T_0(\overline{X}_0),\ S_2 = T_0(\overline{X}_0) + T_1(\overline{X}_1), \dots, \\ &S_n = T_0(\overline{X}_0) + \dots + T_{n-1}(\overline{X}_{n-1}), \dots . \end{aligned} \tag{4.36}$$

The idea is to exploit (4.28) to try to construct the pure jump process attached to (4.33), (4.34), with $S_n$ as in (4.36) and $\overline{X}_n$ playing the role of $X_{S_n}$ in (4.28). We now make a **crucial assumption**:

**Non-explosion assumption**:

$$\text{For all } x \in E,\ \underbrace{\overline{P}_x \times P}_{\substack{\| \text{ def} \\ Q_x}}\text{-a.s.},\quad \lim_n \underset{\substack{\uparrow \\ \text{in (4.36)}}}{S_n} = \infty\,. \tag{4.37}$$

**Lemma 4.7.** *The assumption (4.37) is equivalent to*

$$\text{for } x \in E,\ \overline{P}_x\text{-a.s.},\ \sum_{n \ge 0} \lambda(\overline{X}_n)^{-1} = \infty. \tag{4.38}$$

*Proof.*

- <u>(4.37) $\Longrightarrow$ (4.38)</u>: We prove the claim by contradiction.

Assume that for some $x \in E$, $M > 0$, $\overline{P}_x\big[\overbrace{\textstyle\sum_{n\ge 0} \lambda(\overline{X}_n)^{-1} \le M}^{\substack{A_M \\ \| \text{ def}}}\big] > 0$. Observe that

$$E^{Q_x}[S_n \,|\, (\overline{X}_.)] \overset{(4.36)}{=} \sum_{\ell=0}^{n-1} E^{Q_x}[T_\ell(\overline{X}_\ell) \,|\, (\overline{X}_.)] \overset{(4.35)}{=} \sum_{\ell=0}^{n-1} \lambda(\overline{X}_\ell)^{-1}. \tag{4.39}$$

Hence, using monotone consequence, we find that

$$\begin{aligned} &E^{Q_x}[\lim_n\, S_n \,|\, (\overline{X}_.)] = \sum_{\ell=0}^{\infty} \lambda(\overline{X}_\ell)^{-1} \le M, \text{ on } A_M, \text{ and therefore} \\ &E^{Q_x}[\lim_n\, S_n, A_M] < \infty, \text{ so that } \lim_n\, S_n < \infty,\ Q_x\text{-a.s. on } A_M, \\ &\text{with } Q_x[A_M] = \overline{P}_x[A_M] > 0, \text{ a contradiction.} \end{aligned}$$

As as result (4.38) holds.

- <u>(4.38) $\Longrightarrow$ (4.37)</u>: Note that for $s > 0$, $n \ge 1$, one has:

$$\begin{aligned} E^{Q_x}[\exp\{-s\, S_n\}] &= E^{Q_x}\Big[E^{Q_x}\Big[\exp\Big\{-s \sum_{\ell=0}^{n-1} T_\ell(\overline{X}_\ell)\Big\} \,\Big|\, (\overline{X}_.)\Big]\Big] \\ &\overset{(4.35)}{=} E^{\overline{P}_x}\Big[\prod_{\ell=0}^{n-1} \int_0^\infty \lambda(\overline{X}_\ell)\, \exp\{-\big(s + \lambda(\overline{X}_\ell)\big)\, u\}\, du\Big] \\ &= E^{\overline{P}_x}\Big[\prod_{\ell=0}^{n-1} \Big(\frac{\lambda(\overline{X}_\ell)}{\lambda(\overline{X}_\ell) + s}\Big)\Big] = E^{\overline{P}_x}\Big[\prod_{\ell=0}^{n-1}\Big(1 - \frac{s}{\lambda(\overline{X}_\ell) + s}\Big)\Big] \end{aligned}$$

and since $1 - a \le e^{-a}$

$$\le \; E^{\overline{P}_x}\Big[\exp\Big\{-\sum_{\ell=0}^{n-1} \frac{s}{\lambda(\overline{X}_\ell)+s}\Big\}\Big].$$

Letting $n \to \infty$, we can apply monotone convergence and find:

$$\text{(4.40)} \qquad \begin{array}{l} E^{Q_x}[\exp\{-s \lim\limits_n \; S_n\}] \le E^{\overline{P}_x}\Big[\exp\Big\{-\sum\limits_{\ell \ge 0} \dfrac{s}{\lambda(\overline{X}_\ell)+s}\Big\}\Big] \overset{(4.38)}{=} 0 \\ \text{(as we explain below).} \end{array}$$

Indeed, on $\{\sum_{\ell \ge 0}(\lambda(\overline{X}_\ell)+s)^{-1} < \infty\}$, $\overline{P}_x$-a.s. $\lambda(\overline{X}_\ell) \neq 0$ for all $\ell \ge 0$ (since $\lambda(y) = 0$ implies $y$ is absorbing for the discrete time chain, so that $\overline{P}_x$-a.s., on $\{\lambda(\overline{X}_\ell) = 0\}$, $\lambda(\overline{X}_m) = 0$, for all $m \ge \ell$), and also $\lambda(\overline{X}_\ell) \to \infty$, so that $(\lambda(\overline{X}_\ell)+s)^{-1} \underset{\ell \to \infty}{\sim} \lambda(\overline{X}_\ell)^{-1}$. This shows that

$$\begin{array}{l} \overline{P}_x\text{-a.s. on } \Big\{\sum\limits_{\ell \ge 0}(\lambda(\overline{X}_\ell)+s)^{-1} < \infty\Big\}, \text{ one has } \sum\limits_{\ell \ge 0} \lambda(\overline{X}_\ell)^{-1} < \infty, \\ \text{so that by (4.38), } \overline{P}_x\Big[\sum\limits_{\ell \ge 0}(\lambda(\overline{X}_\ell)+s)^{-1} < \infty\Big] = 0. \end{array}$$

So (4.40) shows that $Q_x$-a.s., $\lim S_n = \infty$, i.e. (4.37). $\square$

We will now see that under the equivalent conditions (4.37) or (4.38) we can construct a pure jump process with no explosion attached to $\lambda(\cdot)$ and $(q_{x,y})$ in the sense of (4.8), (4.9), (4.24), (4.25).

**Theorem 4.8.** *Given $\lambda(\cdot)$, $(q_{x,y})$ as in (4.33), (4.34), for which the non-explosion assumption (4.37) (or (4.38)) holds. Then, there exists a unique pure jump process with no explosion, cf. Definition 4.2, attached to $\lambda(\cdot)$ and $(q_{x,y})$ (i.e. so that (4.24), (4.25) hold).*

*Proof.* We already know the uniqueness part of the statement, cf. Remark 4.5. We turn to the existence part. We define on $\overline{\Omega} \times \Omega_{\text{aux}}$, cf. above (4.36),

$$\text{(4.41)} \qquad Z_t = \overline{X}_n \text{ on } \{S_n \le t < S_{n+1}\}, \text{ for } t \ge 0,\ n \ge 0.$$

Since the non-explosion condition (4.37) holds, for each $x \in E$, $Q_x$-a.s., $Z_.$ is an element of $\Omega$ (cf. (4.3)), and in fact $Z_.$ is a measurable map from $\{\lim_n S_n = \infty\}$ (of full $Q_x$-measure) into $\Omega$, in view of (4.41). We thus define

$$\text{(4.42)} \qquad P_x = \text{the image measure of } Q_x \text{ under } Z_..$$

Note that

$$\text{(4.43)} \qquad P_x[X_0 = x] \overset{(4.42)}{=} Q_x[Z_0 = x] \overset{(4.41)}{=} Q_x[\overline{X}_0 = x] = 1.$$

We will now check that

$$\text{(4.44)} \qquad E^{Q_x}\big[f_0(Z_{s_0}) \dots f_k(Z_{s_k})\; f(Z_{s_k+h})\big] = E^{Q_x}\big[f_0(Z_{s_0}) \dots f_k(Z_{s_k})\, E^{Q_{Z_{s_k}}}[f(Z_h)]\big],$$

for any $x \in E$, $0 \le s_0 < s_1 < \cdots < s_k$, $h > 0$, $f_0, f_1, \dots, f_k$, $f$: $E \to \mathbb{R}$, bounded. By (4.42) it will then follow that $P_x$, $x \in E$, fulfill (4.8), and this will complete the proof of our claim. We will only prove (4.44) when

$$\lambda(x) > 0, \text{ for all } x \in E. \tag{4.45}$$

The general case when there can be absorption points is handled in a similar fashion, simply the notation is a bit more cumbersome. We introduce, cf. (4.36),

$$N_t = \sup\{n \ge 0; S_n \le t\},$$

and denoting by $A$ the left-hand side of (4.44), we find with (4.41) that:

$$\begin{aligned} A &= E^{Q_x}[f_0(\overline{X}_{N_{s_0}}) \dots f_k(\overline{X}_{N_{s_k}})\, f(\overline{X}_{N_{s_k+h}})] \\ &= \sum_{\substack{0 \le n_0 \le n_1 \le \dots \le n_k \\ 0 \le n}} E^{Q_k}[f_0(\overline{X}_{n_0}) \dots f_k(\overline{X}_{n_k})\, f(\overline{X}_{n_k+n}), C] \\ &\quad \text{with } C = \{N_{s_0} = n_0, \dots, N_{s_k} = n_k, N_{s_k+h} = n_k + n\} \\ &\quad \text{(summing over possible values of } N_{s_0}, \dots, N_{s_k+h}). \end{aligned} \tag{4.46}$$

For $x_0, x_1, \dots, x_m$ in $E$ we write

$$\underset{\substack{\uparrow \\ \text{probability on } (0,\infty)^{m+1}}}{\nu_{x_0, \dots, x_m}(dt_0, \dots, dt_m)} = \prod_{0 \le i \le m} \underset{\substack{\uparrow \\ \text{exponential } (\lambda(x_i))\text{-distribution}}}{\underbrace{\mu_{\lambda(x_i)}}} (dt_i) \tag{4.47}$$

and we define

$$\begin{aligned} &h(x_0, \dots, x_{n_k+n}) = \\ &\nu_{x_0, \dots, x_{n_k+n}}\big(t_0 + \dots + t_{n_i - 1} \le s_i < t_0 + \dots + t_{n_i}, \text{ for } 0 \le i \le k, \text{ and} \\ &\qquad t_0 + \dots + t_{n_k+n-1} \le s_k + h < t_0 + \dots + t_{n_k+n}\big) \\ &\qquad \text{(we use the convention } t_0 + \dots + t_{m-1} = 0, \text{ when } m = 0). \end{aligned} \tag{4.48}$$

From (4.35), (4.36), we see that under $Q_x$, conditionally on $\overline{X}_0, \overline{X}_1, \dots, \overline{X}_{n_k+n}$, the variables $T_0(\overline{X}_0), \dots, T_{n_k+n}(\overline{X}_{n_k+n})$ are independent, respectively exponential $(\lambda(\overline{X}_i))$-distributed, $0 \le i \le n_k + n$. Hence we find that

$$Q_x[\underset{\substack{\uparrow \\ \text{last line of (4.46)}}}{C} \,|\, \overline{X}_0, \dots, \overline{X}_{n_k+n}] = h(\overline{X}_0, \dots, \overline{X}_{n_k+n}). \tag{4.49}$$

In the right-hand side of (4.48), we can express the condition on $t_{n_k}$ as

$$t_{n_k} > \underbrace{s_k - (t_0 + \dots + t_{n_k - 1})}_{\substack{\| \text{ def} \\ u \ge 0}} \text{ and } \begin{aligned}[t] &t_{n_k} - u + t_{n_k+1} + \dots + t_{n_k+n-1} \le \\ &h < t_{n_k} - u + \dots + t_{n_k+n}. \end{aligned} \tag{4.50}$$

Note also that for $\rho > 0$, $\varphi(\cdot)$ bounded measurable, $u \ge 0$,

$$\int_{t>u} \varphi(t-u)\,\rho\, e^{-\rho t} dt = e^{-\rho u} \int_0^\infty \varphi(s)\, e^{-\rho s} ds \tag{4.51}$$

(this reflects the "lack of memory of the exponential distribution").

Applying this identity with $\rho = \lambda(x_{n_k})$, $t = t_{n_k}$ in (4.48) and using (4.50) we find:

$$\begin{aligned} & h(x_0, \dots, x_{n_k+n}) = h_1(x_0, \dots, x_{n_k})\, h_2(x_{n_k}, \dots, x_{n_k+n}), \;\; \text{with} \\ & h_1(x_0, \dots, x_{n_k})) = \nu_{x_0,\dots,x_{n_k}}(t_0 + \dots + t_{n_i-1} \le \\ & s_i < t_0 + \dots + t_{n_i};\, 0 \le i \le n_k), \text{ and} \\ & h_2(y_0, \dots, y_n) = \nu_{y_0,\dots,y_n}(t_0 + \dots + t_{n-1} \le h < t_0 + \dots + t_n). \end{aligned} \tag{4.52}$$

Coming back to (4.49) we see that we have the **crucial factorization**:

$$Q_x[C \,|\, \overline{X}_0, \dots, \overline{X}_{n_k+n}] = h_1(\overline{X}_0, \dots, \overline{X}_{n_k})\, h_2(\overline{X}_{n_k}, \dots, \overline{X}_{n_k+n}). \tag{4.53}$$

Therefore the term under summation in the last line of (4.46) equals

$$\begin{aligned} & E^{Q_x}\big[\underbrace{f_0(\overline{X}_{n_0}) \dots f_k(\overline{X}_{n_k})\, f(\overline{X}_{n_k+n})}_{\substack{\| \text{ def} \\ F}}, C\big] = \\ & E^{\overline{P}_x}[F\, h_1(\overline{X}_0, \dots, \overline{X}_{n_k})\, h_2(\overline{X}_{n_k}, \dots, \overline{X}_{n_k+n})] \\ & \text{and using the Markov property of the discrete time Markov chain } \overline{X}_. \\ & \text{at time } n_k\text{:} \\ & = E^{\overline{P}_x}\big[f_0(\overline{X}_{n_0}) \dots f_k(\overline{X}_{n_k})\, h_1(\overline{X}_0, \dots, \overline{X}_{n_k})\, \underbrace{E^{\overline{P}_{x_{n_k}}}[f(\overline{X}_n)\, h_2(\overline{X}_0, \dots, \overline{X}_n)]}_{\substack{\| \text{ def} \\ G_n(\overline{X}_{n_k})}}\big] \\ & = E^{\overline{P}_x}\big[f_0(\overline{X}_{n_0}) \dots f_k(\overline{X}_{n_k})\, G_n(\overline{X}_{n_k})\, h_1(\overline{X}_0, \dots, \overline{X}_{n_k})] \\ & \text{and using a similar identity for } h_1 \text{ in place of } h \text{ in (4.49), we obtain} \\ & = E^{Q_x}\big[f_0(\overline{X}_{n_0}) \dots f_k(\overline{X}_{n_k})\, G_n(\overline{X}_{n_k}),\; N_{s_0} = n_0, \dots, N_{s_k} = n_k\big]. \end{aligned} \tag{4.54}$$

As a result, inserting this identity in the last line of (4.46), we find that

$$A = \sum_{n \ge 0} E^{Q_x}[f_0(\overline{X}_{N_{s_0}}) \dots f_k(\overline{X}_{N_{s_k}})\, G_n(\overline{X}_{N_{s_k}})]. \tag{4.55}$$

Similarly, we see that for $y \in E$, $n \ge 0$,

$$G_n(y) = E^{Q_y}[f(\overline{X}_{N_h}),\; N_h = n], \tag{4.56}$$

so that

$$
\begin{aligned}
A &= E^{Q_x}\big[f_0(\overline{X}_{N_{s_0}})\dots f_k(\overline{X}_{N_{s_k}})\; E^{Q_{\overline{X}_{N_{s_k}}}}[f(\overline{X}_{N_h})]\big] \\
&\overset{(4.41)}{=} E^{Q_x}\big[f_0(Z_{s_0})\dots f_k(Z_{s_k})\, E^{Q_{Z_{s_k}}}\big[f(Z_h)\big]\big],
\end{aligned}
$$

and this proves (4.44)! $\square$

**Terminology**:

The canonical discrete time Markov chain $\overline{X}_{n,n\geq 0}$, with transition probability $(q_{x,y})_{x,y\in E}$ is sometimes called the **discrete skeleton** of the pure jump process $(X_t)_{t\geq 0}$.

As an application of the main theorem concerning the construction of pure jump processes with no explosion, i.e. Theorem 4.8, we discuss the so-called birth and death processes.

**Example 4.9.** (Birth and death processes)

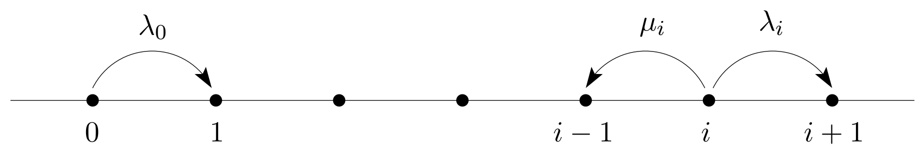



Fig. 4.2

We introduce the jump rate function and the jump transition probability

$$
\begin{aligned}
&\lambda(i) = \lambda_i + \mu_i > 0, \text{ for } i \in \mathbb{N}, \text{ with } \lambda_i, \mu_i \geq 0, \mu_0 = 0, \text{ and} \\
&q_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i},\; q_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}, \text{ when } i \geq 1, \text{ and } q_{0,1} = 1.
\end{aligned} \tag{4.57}
$$

An easy sufficient condition for (4.38) is for instance:

$$
\sum_{j\geq 0} \frac{1}{\lambda_j + \mu_j} = \infty\,. \tag{4.58}
$$

Indeed, if we set $V_j = \sum_{n\geq 0} 1\{\overline{X}_n = j\}$, the total number of visits to $j$ of the discrete time Markov chain, we have for any $i \geq 0$,

$$
\overline{P}_i\text{-a.s.},\; \sum_{n\geq 0} \lambda(\overline{X}_n)^{-1} = \sum_{j\geq 0} \frac{V_j}{\lambda_j + \mu_j}\,.
$$

For any given $i$, we partition the canonical space $\overline{\Omega}$ (where $(\overline{X}_n)_{n\geq 0}$ is defined) into the two events $\{V_j > 0$, for all $j > i\}$ and $\{V_j = 0$, for some $j > i\}$. Then, on $\{V_j > 0$, for all $j > i\}$, $P_i$-a.s.,

$$
\sum_{j\geq 0} \frac{V_j}{\lambda_j + \mu_i} \geq \sum_{j\geq i} \frac{1}{\lambda_j + \mu_j} \overset{(4.58)}{=} \infty.
$$

On the other hand, for $j > i$, on $\{V_j = 0\}$, $P_i$-a.s. $V_\ell = \infty$, for some $\ell < j$, and hence $P_i$-a.s. on $\{V_j = 0$, for some $j > i\}$, $V_\ell = \infty$, for some $\ell \geq 0$, so that $\sum_{n\geq 0} \lambda(\overline{X}_n)^{-1} = \infty$. So we see that when (4.58) holds, the condition (4.38) holds as well.

We will see below that for any $i \geq 1$, as $t \to 0$, in the notation of (4.10),

$$\begin{aligned} r_{i,i+1}(t) &= \lambda_i\, t + o(t), \quad r_{i,i-1}(t) = \mu_i\, t + o(t), \\ r_{0,1}(t) &= \lambda_0\, t + o(t). \end{aligned} \tag{4.59}$$

This gives the interpretation of $\lambda_i$ as a "birth rate in state $i$", and of $\mu_i$ as a "death rate in state $i$". $\square$

We now discuss the **small $\boldsymbol{t}$ behaviour of the transition probability** $r_{x,y}(t)$.

**Proposition 4.10.** *Given a pure jump process with no explosion on $E$, we have for $x \in E$*

$$r_{x,x}(t) = 1 - \lambda(x)\, t + o(t), \ \text{as } t \to 0 \ \text{(see (4.10) for notation)}, \tag{4.60}$$

$$r_{x,y}(t) = \lambda(x)\, q_{x,y}\, t + o(t), \ \text{as } t \to 0, \text{ for } y \neq x \text{ in } E. \tag{4.61}$$

*Proof.* We assume $\lambda(x) > 0$, otherwise the claim is obvious. We first show that

$$P_x\big[\underbrace{T + T \circ \theta_T \leq t}_{\text{"two jumps before time } t\text{"}}\big] = o(t), \ \text{as } t \to 0. \tag{4.62}$$

Indeed, one has for $t \to 0$:

$$\begin{aligned} &P_x[T + T \circ \theta_T \leq t] \leq P_x[T \leq t, T \circ \theta_T \leq t] \overset{(4.18)}{=} E_x\big[T \leq t, P_{X_T}[T \leq t]\big] \overset{(4.17)}{=} \\ &P_x[T \leq t]\, E_x\big[P_{X_T}[T \leq t]\big] \underset{(4.25)}{\overset{(4.16)}{=}} \underbrace{(1 - e^{-\lambda(x)t})}_{\leq \lambda(x)t} \underbrace{\sum_{y\in E} q_{x,y}(1 - e^{-\lambda(y)t})}_{\substack{\text{dominated} \\ \text{convergence}} \ \downarrow\, t\to 0} = o(t), \end{aligned} \tag{4.63}$$

and (4.62) follows.

- <u>(4.60)</u>:

$$\begin{aligned} r_{x,x}(t) &= P_x[T > t] + P_x[X_t = x, \text{ and at least two jumps occur before time } t] \\ &= e^{-\lambda(x)t} + o(t) = 1 - \lambda(x)\, t + o(t), \ \text{as } t \to 0. \end{aligned}$$

- <u>(4.61)</u>:

$$\begin{aligned} r_{x,y}(t) &\overset{(4.62)}{=} P_x[X_t = y, \text{ at most one jump occurs up to time } t] + o(t) \\ &= P_x[X_T = y, T \leq t] + o(t) \underset{(4.25)}{\overset{(4.16)}{=}} (1 - e^{\lambda(x)t})\, q_{x,y} + o(t) = \lambda(x)\, q_{x,y}\, t + o(t), \end{aligned}$$

as $t \to 0$. This concludes the proof of the proposition. $\square$

**Remark 4.11.**

$$\lambda_{x,y} \overset{\text{def}}{=} \lambda(x)\, q_{x,y}, \text{ for } x \neq y \text{ in } E, \tag{4.64}$$

can thus be viewed as the "**rate of jump of the chain from $\boldsymbol{x}$ to $\boldsymbol{y}$**".

We have already seen a similar infinitesimal description as in (4.60), (4.61) for the Poisson process, cf. (1.6). □

## 4.2 Backward and forward Kolmogorov equations, generator

We consider a pure jump process (with no explosion) on $E$, and we are going to derive integral and differential equations for the transition probability $r_{x,y}(t)$ of the chain, cf. (4.10) for the notation.

**Proposition 4.12.** *($\lambda(\cdot), q_{x,y}$) as in (4.24), (4.25))*

*Given a pure jump process (with no explosion) on $E$, then for $x, y \in E$, $t \geq 0$, we have (with $\delta_{x,y}$ the Kronecker symbol)*

$$r_{x,y}(t) = \delta_{x,y}\, e^{-\lambda(x)t} + \int_0^t \lambda(x) e^{-\lambda(x)s} \sum_{z \neq x} q_{x,z}\, r_{z,y}(t-s)\, ds, \tag{4.65}$$

(**Backward integral equation**),

$$r_{x,y}(t) = \delta_{x,y}\, e^{-\lambda(x)t} + \int_0^t \sum_{z \neq y} r_{x,z}(s)\, \lambda(z)\, q_{z,y}\, e^{-\lambda(y)(t-s)}\, ds, \tag{4.66}$$

(**Forward integral equation**).

*Proof.* Loosely speaking, the **backward equation** will **correspond** to **conditioning** in the **first jump** of the chain before time $t$, and the **forward equation** to conditioning on the **last jump** of the chain before time $t$.

- (4.65):

$$\begin{aligned} r_{x,y}(t) &= P_x[X_t = y] \\ &= \underbrace{P_x[X_t = y, T > t]}_{\overset{(4.16)}{=}\; e^{-\lambda(x)t}\delta_{x,y}} + P_x[X_t = y, T \leq t]. \end{aligned} \tag{4.67}$$

So we only need to concentrate on the last term. We note that $P_x$-a.s.:

$$1_{\{X_t = y, T \leq t\}} \underset{\substack{\nearrow \\ \text{the trajectory is} \\ \text{right-continuous}}}{=} \lim_{n \to \infty} \sum_{k=1}^n 1 \underbrace{\left\{\frac{(k-1)}{n}\, t < T \leq \frac{k}{n}\, t\right\}}_{\text{pairwise disjoint as } k \text{ varies}} 1\left\{X_{\frac{(n-k+1)}{n}t} = y\right\} \circ \theta_T \tag{4.68}$$

Hence it follows from dominated convergence that

$$
\begin{aligned}
&P_x[X_t = y, T \le t] = \\
&\lim_{n\to\infty} \sum_{k=1}^n E_x\Big[\frac{(k-1)}{n}\, t < T \le \tfrac{k}{n}\, t,\ 1\{X_{\frac{(n-k+1)}{n}t} = y\} \circ \theta_T\Big] \\
&\overset{(4.18)}{=} \lim_{n\to\infty} \sum_{k=1}^n\ E_x\Big[\frac{(k-1)}{n}\, t < T \le \frac{k}{n}\, t,\ P_{X_T}\big[X_{\frac{(n-k+1)}{n}\,t} = y\big]\Big] \\
&\overset{(4.17)}{=} \lim_{n\to\infty} \sum_{k=1}^n P_x\Big[\frac{(k-1)}{n}\, t < T \le \frac{k}{n}\, t\Big]\ E_x\Big[r_{X_t,y}\ \Big(\frac{(n-k+1)}{n}\ t\Big)\Big] \qquad (4.69)\\
&\underset{(4.25)}{\overset{(4.16)}{=}} \lim_{n\to\infty} \sum_{k=1}^n \Big(\int_{\frac{(k-1)}{n}\,t}^{\frac{k}{n}\,t} \lambda(x)\, e^{-\lambda(x)u}\, du\Big) \sum_{z\neq x} q_{x,z}\, r_{z,y}\Big(t - \frac{(k-1)}{n}\, t\Big) \\
&= \lim_{n\to\infty} \int_0^t \lambda(x)\, e^{-\lambda(x)u} \underbrace{\sum_{z\neq x} q_{x,z}\, r_{z,y}(t-u_n)}_{\le 1, \text{ with } u_n = \sum_1^n \frac{(k-1)}{n}\, t\, 1\{\frac{(k-1)}{n} t < u \le \frac{k}{n} t\}}\ du
\end{aligned}
$$

but since $X_.$ is right-continuous, piecewise-constant, $0 \le s \to r_{z,y}(s) = P_z[X_s = y]$ is right-continuous, so that

$$
\underset{\text{convergence}}{\overset{\text{dominated}}{=}} \int_0^t \lambda(x)\, e^{-\lambda(x)u} \sum_{z\neq x} q_{x,z}\, r_{z,y}(t-u)\, du\,.
$$

Inserting the above identity in the last line of (4.67), we obtain (4.65). One can actually have a quicker proof if one proves a slightly more general statement that (4.18) and uses it to handle (4.69)!

- (4.66): We start as in (4.67) and write

$$
r_{x,y}(t) = e^{-\lambda(x)t}\delta_{x,y} + P_x[X_t = y, T \le t]. \tag{4.70}
$$

Then, by (4.41), (4.42) we write

$$
P_x[X_t = y, T \le t] = \sum_{n\ge 1} Q_x[\overline{X}_n = y,\ S_n \le t < S_{n+1}].
$$

Note that for $n \ge 1$, by the tower property of conditional expectations

$$
\begin{aligned}
&Q_x[\overline{X}_n = y,\ S_n \le t < S_{n+1} \,|\, \overline{X}_0, \dots, \overline{X}_{n-1},\ S_1, \dots, S_{n-1}] = \\
&E^{Q_x}\big[\underbrace{Q_n[\overline{X}_n = y, S_n \le t < S_{n+1} \,|\, \overline{X}_0, \dots, \overline{X}_n,\ S_1, \dots, S_n]}_{\substack{(4.35)\ ||\ (4.36) \\ \exp\{-\lambda(y)(t-S_n)\}\, 1\{\overline{X}_n = y, S_n \le t\}}} \,\big|\, \overline{X}_0, \dots, \overline{X}_{n-1},\ S_1, \dots, S_{n-1}\big] = \qquad (4.71)\\
&E^{Q_x}\big[\exp\{-\lambda(y)(t-S_n)\}\, 1\{\overline{X}_n = y, S_n \le t\} \,|\, \overline{X}_0, \dots, \overline{X}_{n-1}, S_1, \dots, S_{n-1}\big] \\
&\underset{(4.36)}{\overset{(4.35)}{=}} q_{\overline{X}_{n-1},y}\ 1\{S_{n-1} \le t\} \int_{S_{n-1}}^t \lambda(\overline{X}_{n-1})\ \exp\{-\lambda(\overline{X}_{n-1})(s - S_{n-1}) \\
&-\lambda(y)(t-s)\}\, ds.
\end{aligned}
$$

As a result, we have obtained that

$$
\begin{aligned}
&P_x[X_t = y, T \le t] = \\
&\sum_{n\ge 1} E^{Q_x}\Big[S_{n-1} \le t, q_{\overline{X}_{n-1},y} \int_{S_{n-1}}^t \lambda(\overline{X}_{n-1}) e^{-\lambda(\overline{X}_{n-1})(s-S_{n-1})-\lambda(y)(t-s)} ds\Big] \overset{\substack{\text{(with}\\ m=n-1)}}{=} \\
&\sum_{z\neq y} \sum_{m\ge 0} E^{Q_x}\Big[S_m \le t, \overline{X}_m = z, \int_{S_m}^t e^{-\lambda(z)(s-S_m)-\lambda(y)(t-s)} ds\Big]\, \lambda(z)\, q_{z,y} \\
&\overset{\text{Fubini}}{=} \sum_{z\neq y} \lambda(z)\, q_{z,y} \int_0^t \sum_{m\ge 0} \underbrace{E^{Q_x}[1\{S_m \le s, \overline{X}_m = z\}\, e^{-\lambda(z)(s-S_m)}]}_{\substack{\|\\ Q_x[\overline{X}_m = z, S_m \le s < S_{m+1}] \text{ by (4.35),(4.36)}}} e^{-\lambda(y)(t-s)} ds \\
&\overset{(4.41)}{\underset{(4.42)}{=}} \sum_{z\neq y} \lambda(z)\, q_{z,y} \int_0^t r_{x,z}(s)\, e^{-\lambda(y)(t-s)} ds.
\end{aligned}
\tag{4.72}
$$

Inserting this identity in (4.70), we obtain(4.66). □

We are now ready to derive the so-called Kolmogorov backward and forward equations. We recall the notation $\lambda_{x,y} \overset{(4.64)}{=} \lambda(x)\, q_{x,y}$, for $x \neq y$.

**Theorem 4.13.** *Given a pure jump process (with no explosion) on $E$, then for $x, y \in E$,*

$$r_{x,y}(\cdot) \textit{ is continously differentiable on } \mathbb{R}_+, \tag{4.73}$$

*and for $t \ge 0$, one has:*

$$\frac{d}{dt}\, r_{x,y}(t) = \sum_{z\neq x} \lambda_{x,z}\, r_{z,y}(t) - \lambda(x)\, r_{x,y}(t), \tag{4.74}$$

(**Kolmogorov backward equation**),

$$\frac{d}{dt}\, r_{x,y}(t) = \sum_{z\neq y} r_{x,z}(t)\, \lambda_{z,y} - r_{x,y}(t)\, \lambda(y), \textit{ for a.e. } t \tag{4.75}$$

(**Kolmogorov forward equation**).

**Remark 4.14.** One can **informally** obtain (4.74) and (4.75) as follows. One first writes:

$$\frac{d}{dt}\, r_{x,y}(t) = \lim_{h\downarrow 0} \frac{1}{h}\, \big(r_{x,y}(t+h) - r_{x,y}(t)\big)$$

and then with the Chapman-Kolmogorov equation (4.12), one writes

- for (4.74) (using $R_{t+h} = R_h\, R_t$)

$$\frac{1}{h}\, \big(r_{x,y}(t+h) - r_{x,y}(t)\big) = \sum_{z\neq x} \underbrace{\frac{r_{x,z}(h)}{h}}_{\underset{(4.61)}{\overset{h\to 0}{\longrightarrow}} \lambda_{x,z}} r_{z,y}(t) + \underbrace{\frac{r_{x,x}(h)-1}{h}}_{\underset{(4.60)}{\overset{h\to 0}{\longrightarrow}} -\lambda(x)} r_{x,y}(t)$$

and assuming one can exchange limits and sums (we are only informally arguing at this stage) the above for $h \to 0$ tends to $\sum_{z \neq x} \lambda_{x,z}\, r_{z,y}(t) - \lambda(x)\, r_{x,y}(t)$, i.e. this "yields" (4.74).

- for (4.75) (using instead $R_{t+h} = R_t\, R_h$)

$$\frac{1}{h}\,\big(r_{x,y}(t+h) - r_{x,y}(t)\big) = \sum_{z \neq y} r_{x,z}(t)\, \frac{r_{z,y}(h)}{h} + r_{x,y}(t)\, \frac{(r_{y,y}(h) - 1)}{h},$$

which (assuming again that one can exchange limit $h \to 0$ and sum over $z$), with (4.61), (4.60), leads to a right-hand side, which should tend to $\sum_{z \neq y} r_{x,z}(t)\, \lambda_{z,y} - r_{x,y}(t)\, \lambda(y)$, for $h \to 0$. In other words, this "yields" (4.75).

The integral equations (4.65), (4.66) will enable us to bypass the difficulty of making the above informal arguments rigorous. □

*Proof.*

- <u>(4.73)</u>:

We can rewrite (4.65), changing $s$ into $t - s$ in the integral, as:

$$\text{(4.76)} \qquad \begin{aligned} r_{x,y}(t) &= \delta_{x,y}\, e^{-\lambda(x)t} + \int_0^t \lambda(x)\, e^{-\lambda(x)(t-s)} \sum_{z \neq x} q_{x,z}\, r_{z,y}(s)\, ds \\ &= e^{-\lambda(x)t}\Big(\delta_{x,y} + \int_0^t \underbrace{\lambda(x)\, e^{-\lambda(x)s} \sum_{z \neq x} q_{z,z}\, r_{z,y}(s)\, ds}_{\text{bounded function}}\Big). \end{aligned}$$

We first see from (4.76) that $r_{x,y}(\cdot)$ is a continuous function. Then we can inject again this information into the right-hand side of (4.76), and see that the "bounded function" is in addition continuous. It now follows that $r_{x,y}(\cdot)$ is continuously differentiable. (We just employed that is known as a "bootstrap argument".) This proves (4.73).

- <u>(4.74)</u>:

We differentiate the last line of (4.76) and obtain:

$$\begin{aligned} \frac{d}{dt}\, r_{x,y}(t) &= \underbrace{-\lambda(x)\, r_{x,y}(t)}_{\substack{\uparrow \\ \text{derivative of the } e^{-\lambda(x)t} \text{ factor}}} + e^{-\lambda(x)t} \lambda(x)\, e^{\lambda(x)t} \sum_{z \neq x} q_{x,y}\, r_{z,y}(t) \\ &= -\lambda(x)\, r_{x,y}(t) + \sum_{z \neq x} \lambda_{x,z}\, r_{z,y}(t). \end{aligned}$$

This proves (4.74).

• (4.75):

We rewrite (4.66), noting that $\delta_{x,y}\, e^{-\lambda(x)t} = \delta_{x,y}\, e^{-\lambda(y)t}$, as

$$r_{x,y}(t) = e^{-\lambda(y)t}\Big(\delta_{x,y} + \int_0^t \sum_{z\neq y} r_{x,z}(s)\,\lambda(z)\,q_{z,y}\,e^{\lambda(y)s}ds\Big). \tag{4.77}$$

From the Lebesgue differentiation theorem, we find that

$$\begin{aligned}\frac{d}{dt}\, r_{x,y}(t) &= -\lambda(y)\, r_{x,y}(t) + e^{-\lambda(y)t} \sum_{z\neq y} r_{x,z}(t)\,\lambda(z)\,q_{z,y}\,e^{\lambda(y)t}, \text{ a.e. } t,\\ &= -\lambda(y)\, r_{x,y}(t) + \sum_{z\neq y} r_{z,y}(t)\,\lambda_{z,y}.\end{aligned}$$

This proves (4.75). □

**Remark 4.15.**

1) One can in fact argue that $\sum_{z\neq y} r_{x,z}(t)\,\lambda_{z,y}$ is continuous in $t$, so that one has (4.75) for all $t$ (thanks to (4.73)), and one does not need to write for a.e. $t$, see for instance [3], p. 247. To establish the continuity of $\sum_{z\neq y} r_{x,z}(t)\,\lambda_{z,y}$, one can also express $r_{x,z}(t)$ for $z \neq x$ with the help of (4.76) as a function of the form $e^{-\lambda(x)t}\int_0^t \psi(s)ds$, with $\psi \geq 0$, integrable over bounded intervals.

2) One can also use the integral equations (4.65) and (4.66) to show that when $\widetilde{r}_{x,y}(t)$, $t \geq 0$, $x, y \in E$, are transition probabilities such that $\widetilde{r}_{x,y}(t) \to \delta_{x,y}$, as $t \to 0$, for $x, y$ in $E$, and $\widetilde{r}_{x,y}(t)$ satisfies the backward (resp. the forward) equation attached to $\lambda(\cdot)$, and $(q_{x,y})$ (which we tacitly assume to determine a pure jump process with no explosion), then necessarily:
$$\widetilde{r}_{x,y}(t) = r_{x,y}(t), \text{ for all } t \geq 0,\ x, y \in E,$$
see [3], p. 253. □

We now define the **generator matrix** of the pure jump process

$$\begin{aligned} A_{x,y} = &\quad \lambda_{x,y}\,, \quad \text{when } y \neq x,\\ = &\ -\lambda(x)\,, \quad \text{when } y = x. \end{aligned} \tag{4.78}$$

Note that $A_{x,y}$ satisfies

$$\sum_{y\in E} A_{x,y} = 0, \text{ for } x \in E. \tag{4.79}$$

(The **generator matrix is sometimes called the $Q$-matrix** of the Markov chain).

When **$E$ is finite**, we can write the backward Kolmogorov equation in matrix notation as

$$\frac{d}{dt}\, r(t) = A\, r(t),\ t \geq 0. \tag{4.80}$$

and the forward Kolmogorov equation as

$$\frac{d}{dt}\, r(t) = r(t)\, A, \quad t \ge 0. \tag{4.81}$$

The solution to (4.80) or (4.81) with $r(0) = Id$, is then

$$r(t) = \exp\{t\,A\} \stackrel{\text{def}}{=} \sum_{n \ge 0} \frac{t^n}{n!}\, A^n. \tag{4.82}$$

**Example 4.16.** A machine is operative for exponential times with parameter $\lambda > 0$, and then inoperative for exponential times with parameter $\mu > 0$. If $E = \{0,1\}$, with 0: operative, 1: inoperative, we have

$$\lambda_{0,1} = \lambda = \lambda(0), \;\; \lambda_{1,0} = \mu = \lambda(1), \tag{4.83}$$

and the generator matrix is

$$A = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}. \tag{4.84}$$

The backward Kolmogorov equations yield:

$$r'_{0,0}(t) = \lambda\big(r_{1,0}(t) - r_{0,0}(t)\big) \;\; (= -r'_{0,1}(t), \text{ because } r_{0,0}(t) + r_{0,1}(t) = 1), \tag{4.85}$$

$$r'_{1,0}(t) = \mu\big(r_{0,0}(t) - r_{1,0}(t)\big) \;\; (= -r'_{1,1}(t), \text{ for similar reasons as above}). \tag{4.86}$$

As a result we see that $\mu\, r'_{0,0}(t) + \lambda\, r'_{1,0}(t) = 0$ so that

$$\mu\, r_{0,0}(t) + \lambda\, r_{1,0}(t) = \text{const.} \stackrel{t=0}{=} \mu. \tag{4.87}$$

Therefore $\lambda\, r_{1,0}(t) = \mu\big(1 - r_{0,0}(t)\big)$ and inserting in (4.85):

$$r'_{0,0}(t) = \mu\big(1 - r_{0,0}(t)\big) - \lambda\, r_{0,0}(t) = \mu - (\lambda + \mu)\, r_{0,0}(t). \tag{4.88}$$

As a result $g(t) = r_{0,0}(t) - \frac{\mu}{\lambda+\mu}$ satisfies the differential equation:

$$g'(t) = -(\lambda + \mu)\, g(t), \tag{4.89}$$

and hence $g(t) = c\, e^{-(\lambda+\mu)t}$, for $t \ge 0$.

It follows that $r_{0,0}(t) = c\, e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu}$, and for $t = 0$, this forces the relation

$$1 = c + \frac{\mu}{\lambda + \mu}\,, \text{ i.e. } c = \frac{\lambda}{\lambda + \mu}\,.$$

We thus obtain:

$$\text{(4.90)} \qquad \begin{cases} r_{0,0}(t) = \dfrac{\lambda}{\lambda+\mu}\, e^{-(\lambda+\mu)t} + \dfrac{\mu}{\lambda+\mu}, \\[2ex] r_{0,1}(t) = 1 - r_{0,0}(t) = \dfrac{-\lambda}{\lambda+\mu}\, e^{-(\lambda+\mu)t} + \dfrac{\lambda}{\lambda+\mu}, \\[2ex] r_{1,0}(t) = \dfrac{\mu}{\lambda}\, r_{0,1}(t) = \dfrac{-\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} + \dfrac{\mu}{\lambda+\mu}, \\[2ex] r_{1,1}(t) = \dfrac{\mu}{\lambda+\mu}\, e^{-(\lambda+\mu)t} + \dfrac{-\lambda}{\lambda+\mu}\,. \end{cases}$$

- The forward Kolmogorov equations yield:

$$\text{(4.91)} \qquad r'_{0,0}(t) = -r_{0,0}(t)\, \lambda + r_{0,1}(t)\, \mu,$$

$$\text{(4.92)} \qquad r'_{1,0}(t) = -r_{1,0}(t)\, \lambda + r_{1,1}(t)\, \mu.$$

Since $r_{0,1}(t) = 1 - r_{0,0}(t)$, we find

$$r'_{0,0}(t) = -(\lambda+\mu)\, r_{0,0}(t) + \mu\,.$$

This equation coincides with (4.88), and we then proceed as below (4.88).

As it often happens, the forward equations are easier to solve in concrete examples. □

## 4.3 Transience and recurrence

In the case of pure jump process with no explosion, the notions of **recurrence or transience** of a state $x \in E$ **turn out to be governed** by the **corresponding notions** for the discrete time Markov chain $\overline{X}_n$, $n \ge 0$, with transition probability $(q_{x,y})_{x,y \in E}$, cf. Theorem 4.8 (this chain is also sometimes called the **discrete skeleton** of the pure jump process $(X_t)_{t \ge 0}$). More precisely, one introduces the successive return times of the pure jump process to the site $y \in E$, via

$$\text{(4.93)} \qquad \underset{\swarrow}{\overset{\text{hitting time of } y}{}}\widetilde{H}_y = \inf\{t > 0;\ X_t = y \text{ and there exists } s \in (0,t) \text{ with } X_s \neq y\}$$

(note that $P_y$-a.s., $\widetilde{H}_y = \infty$, when $y$ is absorbing), and

$$\text{(4.94)} \qquad \widetilde{H}^0_y = 0, \text{ and for } m \ge 0,\ \widetilde{H}^{m+1}_y = \widetilde{H}_y \circ \theta_{\widetilde{H}^m_y} + \widetilde{H}^m_y,$$

so that $\widetilde{H}^m_y$, $m \ge 0$, are the **successive return times of $(X_t)_{t\ge 0}$ to $y$**.

**Definition 4.17.** *A state $y \in E$ is said **recurrent** for $(X_t)_{t \ge 0}$, if*

$$\text{(4.95)} \qquad P_y[\widetilde{H}_y < \infty] = 1, \ \textit{or when } y \textit{ is absorbing},$$

*and it is said **transient** if*

$$\text{(4.96)} \qquad P_y[\widetilde{H}_y < \infty] < 1, \ \textit{and } y \textit{ is not absorbing},$$

From (4.41), (4.42) that link $(\overline{X}_n)_{n\geq 0}$ and $(X_t)_{t\geq 0}$, it is clear that

$$\text{(4.97)} \qquad \begin{aligned} &y \text{ is transient for } (X_t)_{t\geq 0} \Longleftrightarrow y \text{ is transient for } (\overline{X}_n)_{n\geq 0}, \\ &y \text{ is recurrent for } (X_t)_{t\geq 0} \Longleftrightarrow y \text{ is recurrent for } (\overline{X}_n)_{n\geq 0}. \end{aligned}$$

In a similar fashion $x, y \in E$ are communicating for the pure jump process with no explosion $(X_t)_{t\geq 0}$ if

$$\text{(4.98)} \qquad \begin{aligned} &P_x[H_y < \infty]\; P_y[H_x < \infty] > 0, \text{ where} \\ &H_z \stackrel{\text{def}}{=} \inf\{t \geq 0, X_t = z\} \text{ denotes the } \textbf{entrance time} \text{ of } X_{.} \text{ in } z. \end{aligned}$$

Then with (3.47) and (4.98)

$$\text{(4.99)} \qquad \begin{aligned} &x \text{ and } y \text{ are communicating for } (X_t)_{t\geq 0} \Longleftrightarrow x \text{ and } y \text{ are} \\ &\text{communicating for } (\overline{X}_n)_{n\geq 0}. \end{aligned}$$

As a result of (4.97) and (4.99) we can directly import the partition of $E$ in (3.48):

$$\text{(4.100)} \qquad E = T \cup R_1 \cup R_2 \cup \dots,$$

where $T$ is the set of transient states in $E$ and $R_1, R_2, \dots$ pairwise disjoint equivalence classes of recurrent states.

One also defines positive and null recurrent states via:

**Definition 4.18.** *A recurrent state $x \in E$ is said* ***positive recurrent*** *when*

$$\text{(4.101)} \qquad E_x[\widetilde{H}_x] < \infty, \text{ or when } x \text{ is absorbing (i.e. } \lambda(x) = 0\text{)}.$$

*It is said* ***null recurrent*** *when*

$$\text{(4.102)} \qquad E_x[\widetilde{H}_x] = \infty, \text{ and } x \text{ is not absorbing.}$$

**Remark 4.19.** Unlike what happens in the case of recurrence, or transience, the **notions of positive and null recurrence** for the **continuous chain and its discrete skeleton are in general different**. For instance, see the exercises for an example of

- a site $x$ which is positive recurrent for the discrete skeleton but null recurrent for the pure jump process.

On the other hand, if one looks at $E = \mathbb{Z}$, with

$$q_{x,y} = \frac{1}{2} \text{ if } |x-y| = 1, \; \lambda(x) = 1 + x^2,$$

then the discrete skeleton is simple random walk on $\mathbb{Z}$, which is null recurrent. In particular, (4.38) holds and $(X_t)_{t\geq 0}$ is well-defined. We will see in Example 4.29 that:

- any $x \in \mathbb{Z}$ is null recurrent for the discrete skeleton but positive recurrent for $(X_t)_{t\geq 0}$.

□

## 4.4 Stationary distributions

**Definition 4.20.** *A probability $\pi$ on $E$ is called a* **stationary distribution** *of $(X_t)_{t\ge 0}$, pure jump process (with no explosion) on $E$, when*

$$\sum_{x\in E} \pi(x) \underbrace{r_{x,y}(t)}_{\substack{\uparrow \\ \text{transition probability of} \\ \text{the pure jump process}}} = \pi(y), \text{ for all } t \ge 0, \text{ and } y \in E. \tag{4.103}$$

*It is called a* ***reversible distribution*** *of $(X_t)_{t\ge 0}$, when*

$$\pi(x)\, r_{x,y}(t) = \pi(y)\, r_{y,x}(t), \text{ for all } t \ge 0, \text{ and } x, y \in E. \tag{4.104}$$

**Remark 4.21.**

1) As in the discrete time setting,

$$\pi \text{ reversible} \Longrightarrow \pi \text{ stationary.} \tag{4.105}$$

Indeed $\sum_{x\in E} \pi(x)\, r_{x,y}(t) \overset{(4.104)}{=} \sum\limits_{x\in E} \pi(y)\, r_{y,x}(t) = \pi(y)$, for all $y \in E$, $t \ge 0$.

2) By analogous considerations as in (3.56), the condition (4.103) is equivalent to the identity (with $P_\pi \overset{\text{def}}{=} \sum_{x\in E} \pi(x)\, P_x$)

$$\underbrace{\theta_t \circ P_\pi}_{\substack{\uparrow \\ \text{image of } P_\pi \text{ under } \theta_t}} = P_\pi, \text{ for all } t \ge 0. \tag{4.106}$$

3) The characterizations (4.103) of a stationary distribution, and (4.104) of a reversible distribution, are not practical. We will soon see equivalent conditions, which are expressed in terms of the generator matrix $A$ from (4.78), and which can be checked in concrete examples. □

We will now provide a characterization of stationary distributions, which is easier to check than (4.103). It is a type of "infinitesimal version" of (4.103), which is expressed in terms of the generator matrix $A$, cf. (4.78).

**Theorem 4.22.** *Given a pure jump process (with no explosion), with generator matrix $(A_{x,y})_{x,y\in E}$, a probability $\pi$ on $E$ is a stationary distribution if and only if:*

$$\text{for all } y \in E, \sum_{x\ne y} \pi(x)\, A_{x,y} < \infty \text{ and } \sum_{x\in E} \pi(x)\, A_{x,y} = 0, \tag{4.107}$$

*(recall from (4.78) that $A_{x,x} = -\lambda(x)$, and $A_{x,y} = \lambda(x)\, q_{x,y} = \lambda_{x,y}$ for $x \ne y$).*

*Proof.*

- We first assume that $\pi$ is stationary.

  Then for $y \in E$, $t \geq 0$, by (4.103) and the forward integral equation (4.66):

(4.108)
$$\begin{aligned} \pi(y) &\overset{(4.103)}{=} \sum_{x \in E} \pi(x)\, r_{x,y}(t) \\ &\overset{(4.66)}{=} \pi(y)\, e^{-\lambda(y)t} + \int_0^t \sum_{x \in E, z \neq y} \pi(x)\, r_{x,z}(s)\, \lambda_{z,y}\, e^{-\lambda(y)(t-s)} ds \\ &\overset{(4.103)}{=} \pi(y)\, e^{-\lambda(y)t} + \int_0^t \sum_{z \neq y} \pi(z)\, \lambda_{z,y}\, e^{-\lambda(y)(t-s)} ds. \end{aligned}$$

As a result, setting $t - s = u$, we find

$$\pi(y) \underbrace{(1 - e^{-\lambda(y)t})}_{\lambda(y) \int_0^t e^{-\lambda(y)u} du} = \sum_{z \neq y} \pi(z) \underbrace{\lambda_{z,y}}_{A_{z,y}} \int_0^t e^{-\lambda(y)u} du,$$

from which we deduce that

(4.109)
$$\pi(y)\, \lambda(y) = \sum_{z \neq y} \pi(z)\; A_{z,y}\; (< \infty)$$

and this proves (4.107).

- We only prove the converse under a stronger assumption then (4.107), namely:

(4.110)
$$\sum_{x \in E} \pi(x)\, \lambda(x) < \infty \text{ and } \sum_{x \in E} \pi(x)\, A_{x,y} = 0 \text{ for all } y \in E.$$

(Since for $x \neq y$ $A_{x,y} = \lambda_{x,y} \leq \lambda(x)$, clearly (4.110) implies (4.107). It is in general a stronger assumption.)

From the backward equation (4.74), we find

$$\Big| \frac{d}{dt}\, r_{x,y}(t) \Big| = \Big| \sum_{z \in E} A_{x,z}\, r_{z,y}(t) \Big| \leq \sum_{z \in E} |A_{x,z}| = \underset{\substack{\uparrow \\ \pi\text{-integrable by (4.110)}}}{2\, \lambda(x)}.$$

We thus can exchange differentiation and summation as follows:

(4.111)
$$\begin{aligned} \frac{d}{dt} \Big( \sum_{x \in E} \pi(x)\, r_{x,y}(t) \Big) &= \sum_{x \in E} \pi(x)\, \frac{d}{dt}\, r_{x,y}(t) = \sum_{x \in E} \pi(x) \sum_{z \in E} A_{x,z}\, r_{z,y}(t) \\ &\underset{\text{and (4.110)}}{\overset{\text{Fubini}}{=}} \sum_{z \in E} \Big( \sum_{x \in E} \pi(x)\, A_{x,z} \Big)\, r_{z,y}(t) \overset{(4.110)}{=} 0. \end{aligned}$$

As a result we find that

$$\sum_{x \in E} \pi(x)\, r_{x,y}(t) = \sum_{x \in E} \pi(x)\, r_{x,y}(0) = \pi(y), \text{ for all } t \geq 0,\ y \in E,$$

and $\pi$ is a stationary distribution. $\square$

**Remark 4.23.** The assumption (4.110) means that:

$$C \overset{\text{def}}{=} \sum_{x \in E} \pi(x)\, \lambda(x) < \infty,$$

and the probability $\nu(x) = \frac{\lambda(x)\,\pi(x)}{C}$ is a stationary distribution for the discrete skeleton (i.e. the discrete time Markov chain $(\overline{X}_n)_{n \geq 0}$ attached to $(q_{x,y})_{x,y \in E}$).

One can however show that the weaker condition (4.107) also implies that $\pi$ is a stationary distribution of the pure jump process. One first observes that $\pi$ cannot put mass on transient states, see [13], p. 118-119, see also Corollary 4.28 below. So $\pi$ is supported by irreducible recurrent classes, and one can conclude using Theorem 3.5.5, p. 120 of the same reference, see otherwise [12], p. 393, and 398-401. $\square$

**Corollary 4.24.** *If a probability $\pi$ on $E$ is such that*

$$\pi(x)\, A_{x,y} = \pi(y)\, A_{y,x}, \text{ for all } y \neq x \text{ in } E, \tag{4.112}$$

*then $\pi$ is a stationary distribution of the pure jump process (with no explosion).*

*Proof.* For $y \in E$, one has

$$\sum_{x \neq y} \pi(x)\, A_{x,y} = \sum_{x \neq y} \pi(y)\, A_{y,x} = \pi(y)\, \lambda(y) < \infty,$$

and $\pi(y)\, \lambda(y) = -\pi(y)\, A_{y,y}$, so (4.107) holds true, and $\pi$ is a stationary distribution. $\square$

**Remark 4.25.** One can in fact show (see [13], p. 124-125), that:

(4.113) **under (4.112), $\boldsymbol{\pi}$ is a reversible distribution**.

Note that (4.107) has the interpretation:

$$\underbrace{\sum_{x \neq y} \pi(x)\, A_{x,y}}_{\text{flow into } y} = \underbrace{\pi(y)\, \lambda(y)}_{\text{flow out of } y} \quad \text{"total balance equations"},$$

whereas (4.112) has the interpretation ($x \neq y$)

$$\underbrace{\pi(x)\, A_{x,y}}_{\text{flow from } x \text{ to } y} = \underbrace{\pi(y)\, A_{y,x}}_{\text{flow from } y \text{ to } x} \quad \text{"detailed balance equations"}.$$

$\square$

## 4.5 Stationary distributions and asymptotic behavior

The next result will be helpful (compare with (3.67)):

**Proposition 4.26.**

$$P_x[X_t = y] \underset{t\to\infty}{\longrightarrow} \frac{P_x[H_y < \infty]}{\lambda(y)\, E_y[\widetilde{H}_y]}, \textit{ for } x, y \in E. \tag{4.114}$$

*understood as* $= 1$*, when* $\lambda(y) = 0$*,*
*i.e. for* $y$ *an absorbing state*

*Proof.* We have $P_x[X_t = y] = P_x[H_y \le t, X_t = y]$.

As in the case of (4.69), we can use a discretization procedure to handle

$$P_x[H_y \le t, X_t = y] = P_x\big[H_y \le t, X_{(t-H_y(\omega))}\big(\theta_{H_y(\omega)})(\omega)\big) = y\big],$$

see (4.68), and find that

$$P_x[X_t = y] = E_x[H_y \le t, r_{X_{H_y},y}(t - H_y)] = E_x[H_y \le t,\, r_{y,y}(t - H_y)]. \tag{4.115}$$

In a similar way, if we replace $H_y$ with $\widetilde{H}_x$ (the return time to $x$), we can write

$$\begin{aligned} P_x[X_t = y] &= P_x[\widetilde{H}_x > t, X_t = y] + P_x[\widetilde{H}_x \le t,\, X_t = y] \\ &= P_x[\widetilde{H}_x > t, X_t = y] + E_x[\widetilde{H}_x \le t,\, r_{x,y}(t - \widetilde{H}_x)] \end{aligned}$$

(with a similar justification as for the derivation of (4.115)).

We thus see that $r_{x,y}(t)$ satisfies the renewal equation

$$r_{x,y}(t) = P_x[\widetilde{H}_x > t,\, X_t = y] + \int_0^t r_{x,y}(t-s)\, dF_x(s), \tag{4.116}$$

with $F_x(\cdot)$ the possibly defective distribution function (when $x$ is transient or $x$ absorbing), non-arithmetic with (4.16), (4.17), when $\lambda(x) > 0$,

$$F_x(u) = P_x[\widetilde{H}_x \le u], \text{ for } u \ge 0. \tag{4.117}$$

In particular, choosing $x = y$ in (4.116), we have a first term in the right-hand side $P_y[\widetilde{H}_y > t, X_t = y] = e^{-\lambda(y)t}$, which is direct Riemann integrable by (2.86), when $\lambda(y) > 0$, and

$$r_{y,y}(t) = e^{-\lambda(y)t} + \int_0^t r_{y,y}(t-s)\, dF_y(s), \text{ for all } t \ge 0. \tag{4.118}$$

We thus find that

- when $\lambda(y) = 0$, $r_{y,y}(t) \equiv 1$

- when $\lambda(y) > 0$, and $y$ is recurrent by Smith's key renewal theorem, cf. (2.88)

$$r_{y,y}(t) \underset{t\to\infty}{\longrightarrow} \frac{\displaystyle\int_0^\infty e^{-\lambda(y)u} du}{\displaystyle\int_0^\infty u\, dF_y(u)} = (\lambda(y)\, E_y[\widetilde{H}_y])^{-1}$$

(due to (4.16), (4.17), one sees that the non-arithmeticity condition of $F_y$ is satisfied)

- when $\lambda(y) > 0$, and $y$ is transient by (2.103) one finds that

$$r_{y,y}(t) \underset{t\to\infty}{\longrightarrow} 0\,.$$

Inserting these limit behaviors in (4.115), we find by dominated convergence that

$$P_x[X_t = y] \underset{t\to\infty}{\longrightarrow} \begin{cases} P_x[H_y < \infty]\ (\lambda(y)\, E_y[\widetilde{H}_y])^{-1}, \text{ if } \lambda(y) > 0, \\ P_x[H_y < \infty], \text{ if } \lambda(y) = 0, \end{cases}$$

and this proves (4.114).

$\square$

**Remark 4.27.**

1) If $E_x[\widetilde{H}_x] < \infty$, we see from (4.116) and the fact that $P_x[\widetilde{H}_x > t, X_t = y]$ is direct Riemann integrable, and $dF_x(\cdot)$ is non-arithmetic, that

(4.119) $$P_x[X_t = y] \underset{t\to\infty}{\overset{(2.88)}{\longrightarrow}} \frac{\int_0^\infty P_x[\widetilde{H}_x > t, X_t = y]\, dt}{E_x[\widetilde{H}_x]} \overset{\text{Fubini}}{=} \frac{E_x\big[\int_0^{\widetilde{H}_x} 1\{X_t = y\}\, dt\big]}{E_x[\widetilde{H}_x]}\,.$$

2) For pure jump processes with no explosion, unlike what happened for discrete time Markov chains, we do not need to consider limits in the Césaro sense, but we can directly take limits in $t \to \infty$, in (4.14) or (4.119) (compare with (3.67) and (3.87)).

$\square$

**Corollary 4.28.** *Let $\pi$ be a stationary distribution of the pure jump process $(X_t)_{t\ge 0}$ with no explosion. Then for $y \in E$,*

(4.120) $\pi(y) > 0$ *implies that $y$ is positive recurrent for $(X_t)_{t\ge 0}$.*

*Proof.*

$$\pi(y) \overset{(4.103)}{=} \sum_{x\in E} \pi(x)\, r_{x,y}(t) = \sum_{x\in E} \pi(x)\, P_x[X_t = y] \xrightarrow[\substack{\text{if } y \text{ is transient or} \\ \text{null recurrent by (4.114)}}]{\text{dominated convergence}} 0\,.$$

Therefore $\pi(y) > 0 \Longrightarrow y$ positive recurrent. $\square$

**Example 4.29.** (Continuation of Remark 4.19)

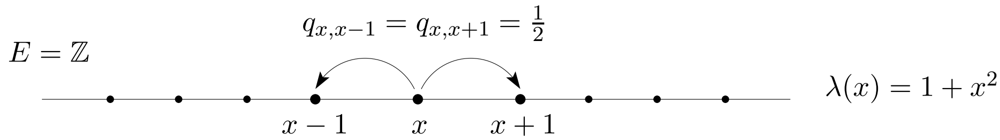



Fig. 4.3

The discrete skeleton $\overline{X}_n$, $n \geq 0$, is a simple random walk on $\mathbb{Z}$, which is recurrent. It follows that $\overline{P}_x$-a.s., $\sum_{n\geq 0} \lambda^{-1}(\overline{X}_n) = \infty$, and the non-explosion criterion (4.38) is fulfilled, and hence, the jump process with no explosion attached to $\lambda(\cdot)$ and $(q_{x,y})_{x,y\in\mathbb{Z}}$ is well-defined.

Define for $x \in \mathbb{Z}$

(4.121) $$\pi(x) = c(1+x^2)^{-1} \text{ where } c = 1/\Big(\sum_{z\in\mathbb{Z}} (1+z^2)^{-1}\Big),$$

so that $\pi$ is a probability.

Then, we have the identity

(4.122) $$\pi(x)\, A_{x,y} = \frac{c}{2}\, 1\{|x-y| = 1\} = \pi(y)\, A_{y,x}, \text{ for all } x \neq y \text{ in } \mathbb{Z}.$$

It follows that $\pi$ is a reversible, and hence stationary, distribution of $(X_t)_{t\geq 0}$. As a result of (4.120)

(4.123) $$\text{any site } x \in \mathbb{Z} \text{ is positive recurrent for } (X_t)_{t\geq 0},$$

but since $(\overline{X}_n)_{n\geq 0}$ is a simple random walk on $\mathbb{Z}$

(4.124) $$\text{any site } x \in \mathbb{Z} \text{ is null recurrent for } (\overline{X}_n)_{n\geq 0},$$ □

We carry on our discussion of **stationary distributions** and **asymptotic behavior**.

**Theorem 4.30.** *(compare with Theorem 3.28)*

*Consider a pure jump process (with no explosion) on $E$, which is irreducible (i.e. all $x, y$ in $E$ communicate). Then one has the equivalences*

(4.125) *some $x$ in $E$ is positive recurrent,*

(4.126) *all $x$ in $E$ are positive recurrent,*

(4.127) *there is a stationary distribution.*

*Furthermore, if one of the above equivalent conditions holds, then*

(4.128) $$\pi(x) = \frac{1}{\lambda(x)\, E_x[\widetilde{H}_x]} \; (> 0),\; x \in E, \text{ is the unique stationary distribution,}$$

*(if $\lambda(x) = 0$, then $E = \{x\}$ and $\lambda(x)\, E_x[\widetilde{H}_x]$ is to be understood as equal to 1).*

*Proof.* We assume $\lambda(x) > 0$, for all $x$ in $E$, otherwise due to irreducibility $E = \{x_0\}$ and the statements are obvious.

- $\underline{(4.125) \Longrightarrow (4.126) \text{ and } (4.125) \Longrightarrow (4.127)}$:

  We assume $x$ positive recurrent and define for all $y$ in $E$:

$$\pi(y) \stackrel{\text{def}}{=} \frac{E_x[\int_0^{\widetilde{H}_x} 1\{X_s = y\} ds]}{E_x[\widetilde{H}_x]} > 0, \text{ by irreducibility.} \tag{4.129}$$

  Then for $u \geq 0$ and $y$ in $E$, one has:

$$\begin{aligned} \pi(y) &\stackrel{(4.119)}{=} \lim_{t\to\infty} \underbrace{P_x[X_{t+u} = y]}_{r_{x,y}(t+u)} \stackrel{(4.12)}{=} \lim_{t\to\infty} \sum_{z\in E} r_{x,z}(t)\, r_{z,y}(u) \\ &\stackrel{\text{Fatou}}{\geq} \sum_{z\in E} \lim_{t\to\infty} r_{x,z}(t)\, r_{z,y}(u) \stackrel{(4.119)}{=} \sum_{z\in E} \pi(z)\, r_{z,y}(u). \end{aligned} \tag{4.130}$$

  Moreover,

$$\sum_{y\in E} \sum_{z\in E} \pi(z)\, r_{z,y}(u) = \sum_{z\in E} \pi(z) \underbrace{\sum_{y\in E} r_{z,y}(u)}_{=1} = \sum_{z\in E} \pi(z) \stackrel{(4.129)}{=} 1.$$

  So one has equality in (4.122) and

$$\pi(y) = \sum_{z\in E} \pi(z)\, r_{z,y}(u), \text{ for each } u \geq 0,\ y \in E: \quad \text{i.e. } \pi \text{ is stationary,} \tag{4.131}$$

  and (4.127) follows.

  Since $\pi(y) > 0$, for each $y \in E$, it follows from (4.120) that each $y$ in $E$ is positive recurrent, and (4.126) follows.

- $\underline{(4.126) \Longrightarrow (4.125)}$: is clear.

- $\underline{(4.127) \Longrightarrow (4.125)}$: follows from (4.120).

- $\underline{(4.128)}$:

  If one of (4.125), (4.126), (4.127) holds and $\pi$ is a stationary distribution, then

$$\pi(y) = \sum_{x\in E} \pi(x)\, P_x[X_t = y] \underset{t\to\infty}{\longrightarrow} (\lambda(y)\, E_y[\widetilde{H}_y])^{-1}, \tag{4.132}$$

  using (4.114) and dominated convergence.

$\square$

**Remark 4.31.** Under the equivalent conditions (4.125) - (4.127), the unique stationary distribution can also be expressed as

$$\pi(y) = \frac{E_x[\int_0^{\widetilde{H}_x} 1\{X_s = y\} ds]}{E_x[\widetilde{H}_x]}, \text{ for } y \in E, \text{ with } x \in E \text{ an arbitrary in } E, \tag{4.133}$$

as seen in (4.129). □

The large time behavior of the transition probability is described by

**Theorem 4.32.** *Consider a pure jump process with no explosion on $E$ and $y \in E$.*

- *If $y$ is transient or null recurrent, then*

$$P_x[X_t = y] \underset{t\to\infty}{\longrightarrow} 0, \text{ for } x \in E. \tag{4.134}$$

*If the pure jump process is irreducible and positive recurrent, then*

$$P_x[X_t = y] \to \pi(y), \text{ for } x \in E, \text{ with } \pi \text{ the unique stationary distribution.} \tag{4.135}$$

*Proof.* This follows from (4.114) and (4.128). □

# 5 Brownian motion

Brownian motion is definitely a fundamental example of stochastic processes with continuous trajectory. It is at the same time an example of Gaussian process, of Markov process, and of continuous martingale. This leads to a very rich mathematical theory. The present chapter only contains a brief introduction to Brownian motion. We refer for instance to [8] for a detailed exposition.

**Definition 5.1.** $(\Omega, \mathcal{A}, P)$, *a probability space. A stochastic process* $B_t(\omega), t \geq 0$, *is called (standard) Brownian motion if:*

(5.1) $\quad B_0 = 0,$

(5.2) *for any* $0 = t_0 < t_1 < \cdots < t_n$, *the random variables* $B_{t_i} - B_{t_{i-1}}$, $1 \leq i \leq n$, *are independent ("independent increments", cf. (1.1)),*

(5.3) *for* $t > 0$, $s \geq 0$, $A \in \mathcal{B}(\mathbb{R})$: $P[B_{t+s} - B_s \in A] = \displaystyle\int_A \frac{1}{\sqrt{2\pi\, t}} \exp\Big\{ -\frac{x^2}{2t}\Big\}\, dx,$ *("stationary Gaussian increments"),*

(5.4) *with probability* $1$, $t \to B_t(\omega)$ *is a continuous function from* $\mathbb{R}_+$ *into* $\mathbb{R}$.

### Sketch of construction:

There are several possible methods, we explain a construction which highlights the link with the simple random walk. Consider

$$X_n, n \geq 1, \text{ on some } (\Omega, \mathcal{F}, Q) \text{ which are i.i.d., with } \mathrm{var}(X_n) = 1,\ E[X_n] = 0,\ n \geq 1, \tag{5.5}$$

$$S_n = \begin{cases} X_1 + \cdots + X_n, n \geq 1, \\ 0,\ n = 0, \end{cases} \tag{5.6}$$

the "random walk" based on the increments $(X_i)_{i \geq 1}$.

We then introduce the polygonal interpolation:

$$S_t = \text{the polygonal line interpolating } S_n, \text{ for } n \geq 0, \tag{5.7}$$

$$B_t^{(n)} = \frac{1}{n}\, S_{n^2 t},\ t \geq 0, \text{ the rescaled (in space and time) trajectory.} \tag{5.8}$$

Note that from the central limit theorem for any $t > 0$

$$\frac{1}{n}\, S_{[n^2t]} = \underbrace{\frac{\sqrt{[n^2t]}}{n}}_{\substack{\downarrow\ n \to \infty \\ \sqrt{t}}} \quad \frac{S_{[n^2t]}}{[n^2t]^{1/2}} \xrightarrow[n \to \infty]{\text{in law}} \underset{\substack{\uparrow \\ \text{centered Gaussian} \\ \text{law with variance } t}}{\mu_t} \tag{5.9}$$

and in fact for $t_0 = t_0 < t_1 \cdots < t_k$, using the vector valued version of the central limit theorem, cf. [4], p. 151:

$$(5.10)\qquad \Big(\frac{1}{n}\, S_{[n^2 t_1]},\ \frac{1}{n}\, \big(S_{[n^2 t_2]} - S_{[n^2 t_1]}\big), \dots,\ \frac{1}{n}\, \big(S_{[n^2 t_k]} - S_{[n^2 t_{k-1}]}\big)\Big) \underset{n\to\infty}{\overset{\text{in law}}{\longrightarrow}} \mu_{t_1} \otimes \mu_{t_2 - t_1} \otimes \cdots \otimes \mu_{t_k - t_{k-1}}$$

(in other words, for $f$ bounded continuous on $\mathbb{R}^d$:

$$\int f\Big(\frac{1}{n}\, S_{[n^2 t_1]}, \dots, \frac{1}{n}\, \big(S_{[n^2 t_k]} - S_{[n^2 t_{k-1}]}\big)\Big)\, dQ \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^k} f\, d\mu_{t_1} \otimes \cdots \otimes d\mu_{t_k - t_{k-1}}. \quad )$$

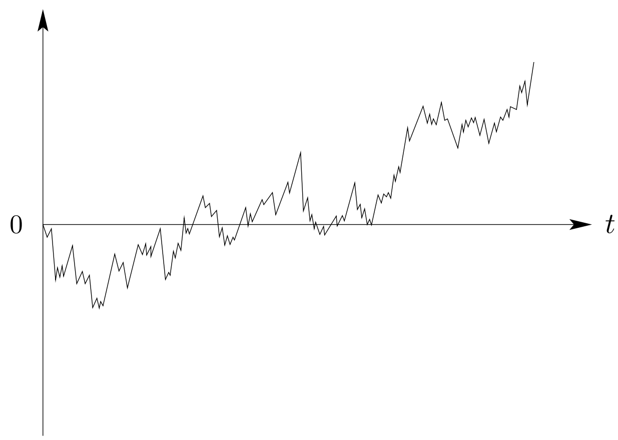



Fig. 5.1: The polygonal line $S_t$

The difference of the random vector in (5.10) and $(B_{t_1}^{(n)}, B_{t_2}^{(n)} - B_{t_1}^{(n)}, \dots, B_{t_k}^{(n)} - B_{t_{k-1}}^{(n)})$, tends to 0 in $Q$ probability, and from this one can deduce that

$$(5.11)\qquad (B_{t_1}^{(n)}, B_{t_2}^{(n)} - B_{t_1}^{(n)}, \dots, B_{t_k}^{(n)} - B_{t_{k-1}}^{(n)}) \underset{n\to\infty}{\overset{\text{in law}}{\longrightarrow}} \mu_{t_1} \otimes \mu_{t_2 - t_1} \otimes \cdots \otimes \mu_{t_k - t_{k-1}}$$
for $k \ge 1$, and $t_0 = 0 < t_1 < \cdots < t_k$ arbitrary.

From this fact, one sees that **all finite dimensional distributions** $(B_{t_1}^{(n)}, B_{t_2}^{(n)}, \dots, B_{t_k}^{(n)}$ **converge in law** as $n \to \infty$, and the limit corresponds to centered independent Gaussian increments with respective variances $t_i - t_{i-1}$, $1 \le i \le k$.

But a much stronger type of convergence takes place usually referred to as "Invariance Principle" (of Donsker). One can view $(B_t^{(n)})_{t \ge 0}$ as a random variable on $(\Omega, \mathcal{F}, Q)$ with values into the canonical space

$$(5.12)\qquad \Omega_{\text{can}} = C([0,\infty), \mathbb{R}) \text{ space of continuous functions } [0,\infty) \to \mathbb{R},$$

endowed with the $\sigma$-algebra:

$$(5.13)\qquad \mathcal{F}^{\text{can}} = \sigma\text{-algebra generated by all maps } \omega \to \omega(s),\ s \ge 0.$$

The canonical space $(\Omega_{\text{can}}, \mathcal{F}^{\text{can}})$ also has the canonical process

$$X_t(\omega) = \omega(t), \text{ for } t \geq 0, \ \omega \in \Omega_{\text{can}}. \tag{5.14}$$

One endows $\Omega_{\text{can}}$ with the topology of uniform convergence on compact intervals, for instance metrized with the distance:

$$d(\omega, \omega') = \sum_{n \geq 1} \frac{1}{2^n} \big( \sup_{0 \leq s \leq n} |\omega(s) - \omega'(s)|\big) \wedge 1. \tag{5.15}$$

One then shows the **invariance principle** (we refer to [8], p. 70 for a proof)

$$\begin{array}{l} (B_t^{(n)})_{t \geq 0} \text{ converges in law to a probability } W \text{ on } (\Omega_{\text{can}}, \mathcal{F}_{\text{can}}), \\ (\text{i.e. } \int_\Omega f(B_\cdot^n)\, dQ \to \int_{\Omega_{\text{can}}} f(\omega)\, dW, \text{ as } n \to \infty, \text{ for any bounded} \\ \text{continuous } f\colon \Omega_{\text{can}} \to \mathbb{R}), \text{ and } (X_t)_{t \geq 0} \text{ under the probability } W \\ \text{is a Brownian motion.} \end{array} \tag{5.16}$$

There is at most one measure on $(\Omega_{\text{can}}, \mathcal{F}^{\text{can}})$ for which $(X_t)_{t \geq 0}$ is a Brownian motion (indeed all finite dimensional distributions $(X_0, X_{t_1}, \dots, X_{t_k})$ are then specified by (5.1), (5.2), (5.3) and we can then apply Dynkin's lemma). This unique measure $W$ on $(\Omega_{\text{can}}, \mathcal{F}_{\text{can}})$ is then called **Wiener measure**.

Note that the map $\omega \in \Omega_{\text{can}} \overset{\Phi_{t_0,\dots,t_k}}{\longrightarrow} (X_{t_0}(\omega), \dots, X_{t_k}(\omega)) \in \mathbb{R}^{k+1}$ is continuous, so from (5.16) we recover that

$$\int g \circ \Phi_{t_0,\dots,t_k}(B_\cdot^{(n)})\, dQ \underset{n \to \infty}{\longrightarrow} \int_{\Omega_{\text{can}}} g \circ \Phi_{t_0,\dots,t_k} dW \text{ for } g \text{ continuous bounded in } \mathbb{R}^{k+1},$$

i.e. all finite dimensional distributions of $(B_t^{(n)})_{t \geq 0}$ converge in law to the corresponding finite $d$-dimensional distribution of $(B_t)_{t \geq 0}$, (see below (5.11)).

But the invariance principle (5.16) contains more information. For instance:

$$\omega \in \Omega_{\text{can}} \longrightarrow \sup_{0 \leq s \leq 1} \omega(s) \in \mathbb{R}$$

is also a continuous map and by the same argument we see that for bounded continuous $f\colon \mathbb{R} \to \mathbb{R}$

$$\int f\Big(\frac{1}{n} \sup_{0 \leq k \leq n^2} S_k\Big)\, dQ \underset{n \to \infty}{\longrightarrow} \int_{\Omega_{\text{can}}} \big(\sup_{0 \leq s \leq 1} X_s(\omega)\big)\, dW(\omega) \tag{5.17}$$

(that is, $\frac{1}{n} \sup_{0 \leq k \leq n^2} S_k$ converges in law to $\sup_{0 \leq s \leq 1} X_s$).

So (5.16) also contains information about trajectorial behavior. □

It is convenient to discuss Brownian motion with the help of the canonical model introduced above $(\Omega_{\text{can}}, \mathcal{F}_{\text{can}}, W, (X_t)_{t \geq 0})$ cf. (5.12) - (5.14), (5.16).

## 5.1 Brownian motion as a Gaussian process

**Proposition 5.2.** $(X_t)_{t\geq 0}$ *(under $W$) is a centered Gaussian process (i.e. all the finite linear combinations $\sum_i \alpha_i X_{t_i}$ are centered Gaussian variables) with covariance function:*

$$\operatorname{cov}(X_s, X_t) = s \wedge t\,. \tag{5.18}$$

*This uniquely characterizes $W$.*

*Proof.*

- $(X_t)_{t>0}$ is a Gaussian process under $W$:

  Indeed, for $0 = t_0 < \cdots < t_n$, the variables $Y_j = X_{t_j} - X_{t_{j-1}}$, $1 \leq j \leq n$ are independent centered normal variables. But $Z = \sum \alpha_j X_{t_j}$ can be expressed as a linear combination of the $Y_j$; $Z = \sum_{j=1}^n \beta_j Y_t$, hence for $t \in \mathbb{R}$

$$\begin{aligned}\varphi_Z(t) = E^W[\exp\{it\,Z\}] \;&=\; E^W\Big[\exp\Big\{i\,t\sum_{j=1}^n \beta_j\,Y_j\Big\}\Big] \;\overset{\text{independence}}{=}\; \prod_{j=1}^n E^W[\exp\{i\,\beta_j\,Y_j\}] \\ &\overset{\text{centered normal}}{=}\; \exp\Big\{-\sum_{j=1}^n \frac{t^2}{2}\,\beta_j^2\,E^W[Y_j^2]\Big\},\end{aligned} \tag{5.19}$$

  so that $Z$ is a centered Gaussian variable.

- The covariance function of $(X_t)_{t\geq 0}$:
  Pick $0 \leq s \leq t$

$$\begin{aligned}E^W[X_s\,X_t] &= E^W[X_s(X_s + X_t - X_s)] \\ &= \underbrace{E^W[X_s^2]}_{\overset{(5.3)}{=}\, s} \;+\; \underbrace{E^W[X_s(X_t - X_s)]}_{\overset{(5.2)}{=}\, E[X_s]\,E[X_t - X_s] \overset{(5.3)}{=} 0} = s\end{aligned}$$

  and (5.18) follows.

- $W$ is uniquely determined by the above mentioned properties:
  for $0 \leq t_1 \leq \cdots < t_n$ and $\alpha \in \mathbb{R}^n$, $U = (X_{t_1}, \ldots, X_{t_n})$,

$$\begin{aligned}E^W[\exp\{i\,\alpha.U\}] \;&=\; E^W\Big[\exp\Big\{i\sum_1^n \alpha_j\,X_{t_j}\Big\}\Big] \\ &\overset{\substack{\text{centered}\\ \text{Gaussian}}}{=}\; \exp\Big\{-\frac{1}{2}\,E^W\Big[\Big(\sum_1^n \alpha_j\,X_{t_j}\Big)^2\Big]\Big\} \\ &=\; \exp\Big\{-\frac{1}{2}\sum_{j,k=1}^n \alpha_j\,\alpha_k\,E^W[X_{t_j}\,X_{t_k}]\Big\} \\ &\overset{(5.18)}{=} \exp\Big\{-\frac{1}{2}\sum_{j,k=1}^n \alpha_j\,\alpha_k\,t_j \wedge t_k\Big\}.\end{aligned} \tag{5.20}$$

So the characteristic function of $U$ is uniquely determined and hence the law of $U$ on $\mathbb{R}^n$ is uniquely determined of [4], p. 150, this for arbitrary $t_1 < \cdots < t_n$. Using Dynkin's lemma, we see that $W$ is uniquely determined. $\square$

**Example 5.3.** (scaling property)

(5.21) $$\text{for } \lambda > 0,\ (\lambda\, X_{t/\lambda^2})_{t\geq 0} \text{ is a Brownian motion.}$$

Indeed, it is clearly a centered Gaussian process with continuous trajectories and covariance function:

$$E[\lambda\, X_{t/\lambda/2}\ \lambda\, X_{s/\lambda^2}] = \lambda^2 \Big\{ \Big(\frac{t}{\lambda^2}\Big) \wedge \Big(\frac{s}{\lambda^2}\Big) \Big\} = t \wedge s$$

and the claim follows by noting that the law of $(\lambda\, X_{t/\lambda^2})_{t\geq 0}$ on $\Omega_{\text{can}}$ must now coincide with the probability $W$ (i.e. the Wiener measure). $\square$

## 5.2 Brownian motion as a Markov process

We introduce the filtration:

(5.22) $$\mathcal{F}_t^{\text{can}} = \sigma(X_s, s \leq t),\ t \geq 0.$$

**Proposition 5.4.** *(simple Markov property)*

*For $f$ bounded measurable: $\mathbb{R} \to \mathbb{R}$, $t \geq 0$, $h > 0$,*

(5.23) $$E^W[f(X_{t+h}) \,|\, \mathcal{F}_t] \overset{W\text{-}a.s.}{=} (R_h\, f)(X_t)$$

*where the semigroup $(R_u)_{u\geq 0}$ acting on bounded measurable functions is defined by*

(5.24) $$(R_u\, f)(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi u}} \displaystyle\int_{\mathbb{R}} f(x+z)\ \exp\Big\{-\frac{z^2}{2u}\Big\}\, dz,\ u > 0, \\ f(x),\ u = 0, \end{cases}$$

*and (5.23) together with $W[X = 0] = 1$, uniquely determines $W$.*

*Proof.* To prove (5.23), it suffices to show that for $t_0 = 0 < t_1 < \cdots < t_n = t$, and

$$G = \prod_{j=0}^{n} g_j(X_{t_j}),\ \text{with } g_0, \ldots, g_n \text{ bounded measurable functions,}$$

one has:

(5.25) $$E^W[G\, f(X_{t+h})] = E^W[G(R_h\, f)(X_t)],$$

indeed, with Dynkin's lemma, it then follows that for $A \in \mathcal{F}_t^{\mathrm{can}}$:

$$E^W[A, f(X_{t+h})] = E^W[A, (R_h\, f)(X_t)]$$

and this yields (5.23).

To check (5.25) observe that:

$X_{t+h} - X_t$ is $N(0,h)$ distributed and independent of $X_{t_j} - X_{t_{j-1}}$, $1 \le j \le n$, and $X_{t_0} = 0$.

As a result we have

$$E^W[G\, f(X_{t+h})] = E^W[\underset{\uparrow}{G}\, f(\underset{\uparrow}{X_t} + \overbrace{X_{t+h} - X_t}^{\swarrow\ \text{independent, } N(0,h) \text{ distributed}})] \tag{5.26}$$

depends on $X_{t_0} = 0$, $X_{t_j} - X_{t_{j-1}}$, $1 \le j \le n$.

integrating out the $X_{t+h} - X_t$ variable:

$$\begin{aligned} &= E^W\Big[G \int_{\mathbb{R}} f(X_t + z)\ \exp\Big\{-\frac{z^2}{2h}\Big\}\ \frac{dz}{\sqrt{2\pi\, h}}\Big] \\ &= E^W[G(R_h\, f)(X_t)] \end{aligned}$$

which proves (5.25).

The semigroup property $R_{u+v} = R_u\, R_v$ follows from the fact that

$$\begin{aligned} &(R_u\, f)(x) = E[f(x + \overset{N(0,u)\text{ variable}}{\overset{\swarrow}{Z}})], \text{ so that} \\ &(R_u\, R_v\, f)(x) = E[f(x + \underbrace{\underset{\underset{N(0,u)}{\uparrow}}{Z} + \underset{\underset{N(0,v)}{\uparrow}}{Y}}_{\text{independent}})] = E[f(x + \underset{\underset{N(0,u+v)}{\uparrow}}{T})] \end{aligned}$$

Finally, let us check that:

- $W$ is uniquely determined by (5.23) and $W[X = 0] = 1$:

Indeed, using induction, these properties completely specify the finite dimensional distributions $(X_{t_0}, X_{t_1}, \dots, X_{t_n})$, for $t_0 = 0 < t, < \dots < t_n$ arbitrary. The claim now follows with the help of Dynkin's lemma. $\square$

## 5.3 Brownian motion as a continuous martingale

The heart of the matter lies in the following statement.

**Theorem 5.5.** *(Paul Lévy)*

(5.27) *$X_t$, $t \geq 0$, and $X_t^2 - t$, $t \geq 0$ are continuous $(\mathcal{F}_t^{\mathrm{can}})$-martingales under $W$, and $W$-a.s., $X_0 = 0$.*

*This uniquely characterizes $W$.*

*Proof of the easy part*: Clearly, $X_t$ and $X_t^2 - t$, are continuous in the $t$ variable, and for fixed $t \geq 0$, $\mathcal{F}_t^{\mathrm{can}}$-measurable. Also they are both integrable. Then for $s \leq t$:

$$(5.28) \qquad \begin{aligned} E^W[X_t \,|\, \mathcal{F}_s^{\mathrm{can}}] &= E^W[X_t - X_s + \overset{\mathcal{F}_s^{\mathrm{can}}\text{-meas.}}{X_s} \,|\, \mathcal{F}_s^{\mathrm{can}}] \\ &= E^W[X_t - X_s \,|\, \mathcal{F}_s^{\mathrm{can}}] + X_s. \end{aligned}$$

From the independence of increments, $W[X_0 = 0] = 1$, and Dynkin's lemma, we see that $X_t - X_s$ is independent of $\mathcal{F}_s^{\mathrm{can}}$. Hence:

$$(5.29) \qquad E^W[X_t - X_s \,|\, \mathcal{F}_s^{\mathrm{can}}] = E^W[X_t - X_s] \overset{(5.3)}{=} 0$$

Coming back to (5.28), we obtain

$$(5.30) \qquad E^W[X_t \,|\mathcal{F}_s^{\mathrm{can}}] = X_s, \text{ for all } s \leq t,$$

and $(X_t)_{t \geq 0}$ is an $(\mathcal{F}_t^{\mathrm{can}})$-martingale.

Similarly for $s \leq t$:

$$(5.31) \qquad \begin{aligned} &E^W[X_t^2 - t \,|\, \mathcal{F}_s^{\mathrm{can}}] = E^W[(X_s + X_t - X_s)^2 - t \,|\, \mathcal{F}_s^{\mathrm{can}}] \\ &= E^W[X_s^2 + 2X_s(X_t - X_s) + (X_t - X_s)^2 - s - (t-s) \,|\, \mathcal{F}^{\mathrm{can}}] \\ &= X_s^2 - s + 2E^W[X_s(X_t - X_s) \,|\, \mathcal{F}_s^{\mathrm{can}}] + E^W[(X_t - X_s)^2 \,|\, \mathcal{F}_s^{\mathrm{can}}] - (t-s). \end{aligned}$$

Note that $X_s$ is $\mathcal{F}_s^{\mathrm{can}}$ measurable and

$$(5.32) \qquad E^W[X_s(X_t - X_s) \,|\mathcal{F}_s^{\mathrm{can}}] = X_s \, E^W[X_t - X_s \,|\, \mathcal{F}_s^{\mathrm{can}}] \overset{(5.29)}{=} 0$$

and using the remark above (5.29), $X_t - X_s$ is independent of $\mathcal{F}_s^{\mathrm{can}}$. Hence

$$(5.33) \qquad E^W[(X_t - X_s)^2 \,|\, \mathcal{F}_s^{\mathrm{can}} = E^W[(X_t - X_s)^2] \overset{(5.3)}{=} t - s\,.$$

Inserting (5.32) and (5.33) into (5.31), we obtain:

$$E^W[X_t^2 - t \,|\, \mathcal{F}_s^{\text{can}}] = X_s^2 - s, \text{ for } s \le t.$$

Hence $(X_t^2 - t)_{t \ge 0}$ is an $(\mathcal{F}_t^{\text{can}})$-martingale.

The statement that (5.27) implies that $(X_t)_{t \ge 0}$ is a standard Brownian motion has a proof which goes beyond the scope of these notes (but can be found on p. 157 of [8]). □

**Remark 5.6.**

1) The martingale point of view turns out to be crucial in the development of "stochastic integrals" and stochastic calculus. Paul Lévy's theorem offers a powerful characterization of Brownian motion.

2) Closely related to the above theorem is the fact that for $t > 0$, the "quadratic variation" of $X_.$ in $[0, t]$ does not vanish and:

$$W\text{-a.s.}, \sum_{k=0}^{2^n - 1} \Big(X_{\frac{(k+1)}{2^n}\,t} - X_{\frac{k}{2^n}\,t}\Big)^2 \underset{n \to \infty}{\longrightarrow} t\,. \tag{5.34}$$

To see (5.34), set:

$$a_n = \sum_{k=0}^{2^n - 1} b_{k,n}, \text{ with } b_{k,n} = \Big(X_{\frac{(k+1)}{2^n}\,t} - X_{\frac{k}{2^n}\,t}\Big)^2 - \frac{t}{2^n}\,, \tag{5.35}$$

then $b_{k,n}$, for fixed $n$ are i.i.d. centered variables.

$$\begin{aligned} E^W[a_n^2] &\overset{\substack{\text{i.i.d.}\\ \text{centered}}}{=} 2^n\, E^W[b_{0,n}^2] = 2^n\, E\Big[\Big(X^2_{\frac{t}{2^n}} - \frac{t}{2^n}\Big)^2\Big] \\ &\underset{(5.21)}{\overset{\text{scaling}}{=}} 2^n \times \Big(\frac{t}{2^n}\Big)^2\, E[(X_1^2 - 1)^2] \le \text{const.}\, \frac{t^2}{2^n}\,, \end{aligned} \tag{5.36}$$

in particular $E^W[\sum_{n \ge 0} a_n] < \infty$, so that

$$W\text{-a.s.}, \; a_n \underset{n \to \infty}{\longrightarrow} 0, \tag{5.37}$$

which proves (5.34).

The property (5.34) is in particular an obstruction to the possibility that with positive $W$-probability the trajectory $s \in [0, t] \to X_s$ has bounded variation (because then the limit in (5.34) would be equal to 0 since $X_.$ is continuous as well, and not equal to $t$). One thus already has an indication of the fact that typically under Wiener measure, the trajectory $s \to X_s(\omega)$ is very rough!

□

# References

[1] R.N. Bhattacharya and E.C. Waymire. *Stochastic processes with applications.* Classics in Applied Mathematics, 61 (SIAM), Philadelphia, 2009.

[2] L. Breiman. *Probability.* Addison-Wesley, Reading, 1968.

[3] K.L. Chung. *Markov chains with stationary transition probabilities.* Springer, New York, second edition, 1967.

[4] R. Durrett. *Probability: Theory and Examples.* Duxbury Press, Belmont, second edition, 2004.

[5] R. Durrett. *Essentials of stochastic processes.* Springer Texts in Statistics, second edition. Springer, New York, 2012.

[6] W. Feller. *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, New York, 1971.

[7] E. Hewitt and K. Stromberg. *Real and abstract analysis.* Graduate Texts in Mathematics, No. 25, Springer-Verlag, New York-Heidelberg, 1975.

[8] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus.* Springer, Berlin, 1988.

[9] S. Karlin and H.M. Taylor. *A first course in stochastic processes.* Academic Press, New York-London, 1975.

[10] S. Karlin and H.M. Taylor. *A second course in stochastic processes.* Academic Press, New York-London, 1981.

[11] M. Lefebvre. *Applied stochastic processes.* Springer, New York, 2007.

[12] S.I. Resnick. *Adventures in Stochastic Processes.* Birkhäuser, Boston, 4th edition, 2005.

[13] J.R. Norris. *Markov chains.* Cambridge Series in Statistical and Probabilistic Mathematics, 2. Cambridge University Press, Cambridge, 1998.

[14] A.S. Sznitman. *Probability Theory.* Lecture Notes of Fall Semester 2016.

# Index

age process, 18, 47, 56
arithmetic, 43

backward integral equation, 95
birth and death process, 93
Brownian motion, 111
    as a continuous martingale, 117
    as a Gaussian process, 114
    Markov property, 115
    scaling property, 115

Chapman-Kolmogorov equation, 62
communicating states, 72, 102
counting process, 8

detailed balance equations, 105
directly Riemann-integrable, 52
discrete skeleton, 93, 101

Ehrenfest model, 65
entrance time, 68
excess process, 18, 47, 56
explosion, 87, 88

forward integral equation, 95

Galton-Watson chain, 65
generator, 16

hitting time, 68

independent increments, 111
inhomogeneous Poisson process, 25
    by variable thinning, 29
    via time change, 25, 29
invariance principle, 113
irreducible Markov chain, 72, 78
    aperiodic, 81

jump rate, 87
jump transition probability, 87

Kolmogorov equation, 95, 97
    backward, 97
    forward, 97

Markov chain, 61, 63, 83
    continuous time, 83
    discrete time, 61
    initial distribution, 63
    period, 81
    transition probability, 61
Markov property, 15, 67
    simple, 67, 115
    strong, 67

null recurrent, 73, 102

Poisson process, 9, 15, 16, 19, 22, 24, 88
    inhomogeneous, 24
    marking, 19
    stationary, 16
    superposition, 22
    thinning, 20
positive recurrent, 73, 77, 78, 102
pure jump process (with no explosion), 83, 103
    discrete skeleton, 101
    generator, 99

Q-matrix, 99

recurrent state, 68, 81, 101
reflected random walk, 65
renewal equation, 45, 47, 50–52, 57
    asymptotic behavior, 52
    existence and uniqueness, 51
    with defect, 57
renewal function, 33, 46
renewal process, 33, 39, 40, 57, 69
    stationary, 40
    with defect, 57
    with delay, 39, 47
renewal reward process, 37
renewal theorem, 35, 43, 54
    Blackwell, 43
    elementary, 43
    Smith, 54
renewal with delay, 38, 47

reversible distribution, 74, 103

semigroup, 16
    generator, 16
simple Markov property, 84
simple random walk, 64
state space, 61
stationary distribution, 74, 77, 78, 80, 103
stochastic process, 7
    counting process, 8
    independent increments, 7
    stationary increments, 7

thinning, 20
total balance equations, 105
transient state, 68, 101
transition probability, 61, 84

Wald identity, 36
Wiener measure, 113