MARTINGALES and MARKOV CHAINS

Solved Exercises and Elements of Theory

Paolo Baldi
Laurent Mazliak
Pierre Priouret

CHAPMAN & HALL/CRC
A CRC Press Company
Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Baldi, Paolo, 1948-
[Martingales et chaînes de Markov. English]
Martingales and Markov chains : solved exercises and elements of theory / Paolo Baldi, Laurent Mazliak, Pierre Priouret.
p. ; cm.
Includes bibliographical references and index.
ISBN 1-58488-329-4 (alk. paper)
1. Martingales (Mathematics) 2. Martingales (Mathematics)—Problems, exercises, etc. 3. Markov processes. 4. Markov processes—Problems, exercises, etc. I. Mazliak, Laurent. II. Priouret. P. (Pierre), 1939- III. Title.

QA274.5 .B3513 2002
519.2'87—dc21 2002018810

International Standard Book Number 1-58488-329-4
Library of Congress Card Number 2002018810
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Contents

Preface

This is primarily a book of exercises with solutions about discrete time stochastic processes, more precisely on martingales and Markov chains with a countable state space. These are two of the most important instances of processes, and being familiar with them is an unavoidable step toward the study of more complex situations. We are therefore concerned with notions of stochastic processes such as are treated at the end of undergraduate or at the beginning of graduate studies. We resolved to give rather detailed solutions, mostly for the first exercises of each chapter.

In this book the reader can find:

1) A section of elementary exercises essentially requiring use of basic theory, followed by others in which the reader is prompted to more initiative.

2) A section of problems; often these come from examinations at the end of the undergraduate level (maîtrise) at the Université Pierre et Marie Curie, Paris 6. Often they open the perspective toward specific applications.

It seemed necessary to give a substantial recall of the elements of the theory in order to set precisely the context. The references to the theory are very frequent in the solution of the exercises; these are either to a theorem (whose reference is given by a chapter number followed by the statement number; "Theorem 4.25" suggests, therefore, to look at Theorem 25 of Chapter 4) or by the mark • followed by a number, pointing to the essential elements of the theory section.

It also appeared necessary, before starting the core of the subject, to write two short chapters recalling the basic notions about conditional expectations and stochastic processes. We do not claim to be thorough; the aim is only to state the essentials for the rest of the book. On the other hand we assume the reader is familiar with the basic notions of measure theory, integration, probability.

Exercises and problems are, in principle, sorted in increasing order of difficulty, but this rule, however subjective, is not unbreakable. We found it particularly advisable to put in a sequence some exercises having strong links between them but not necessarily of the same difficulty.

Since some of the readers might choose to study Markov chains without knowing about martingales, we have marked with the beacon Ⓜ those exercises that make use explicitly of martingales. The problems, conversely, very often need both notions and we did not deem necessary to give the same indication.

Paolo Baldi
Laurent Mazliak
Pierre Priouret

CHAPTER 1

Conditional Expectations

Introduction

•**1.1** Let us begin with a simple, albeit typical, case of conditioning. On a probability space $(\Omega, \mathscr{F}, \mathbf{P})$, let X be a bounded real random variable (r.v.) and T a r.v. with values in $E = \{t_1, t_2, \dots\}$ such that $\mathbf{P}(T = t_k) > 0$ for every $k \geq 1$. We define, for every Borel set $A \subset \mathbb{R}$ and $k \geq 1$,

$$n(t_k, A) = \mathbf{P}(X \in A \mid T = t_k) = \frac{\mathbf{P}(X \in A, T = t_k)}{\mathbf{P}(T = t_k)}.$$

For every fixed k, $A \to n(t_k, A)$ is a probability on $\mathbb{R}$ which is called the *conditional law of X given $T = t_k$*. The *conditional expectation* of X given $T = t_k$ is then defined as

$$\mathbf{E}(X \mid T = t_k) = \frac{1}{\mathbf{P}(T = t_k)} \int_{\{T = t_k\}} X \, d\mathbf{P} = \int x \, n(t_k, dx).$$

This is a key notion, as we shall see in the rest of this book. This explains why it is important to extend it to the case of a general r.v. T. This is the aim of this chapter: we are going to characterize $h(t_k) = \mathbf{E}(X \mid T = t_k)$ by a formula having a meaning in the general case, i.e., when the r.v. T takes its values in a general measurable space. We remark that, for every $B \subset E$,

$$\int_{\{T \in B\}} h(T) \, d\mathbf{P} = \sum_{t_k \in B} \mathbf{E}(X \mid T = t_k) \mathbf{P}(T = t_k) =$$

$$= \sum_{t_k \in B} \mathbf{E}(X 1_{\{T = t_k\}}) = \mathbf{E}(X 1_{\{T \in B\}}) = \int_{\{T \in B\}} X \, d\mathbf{P}.$$

The previous relation thus states that $h(T)$ is a $\sigma(T)$-measurable bounded r.v. such that its integral and the integral of X coincide on events belonging to $\sigma(T)$. This property characterizes the conditional expectation.

Definition and First Properties

•**1.2** Let $(\Omega, \mathscr{A}, \mathbf{P})$ be a probability space.

Theorem and Definition 1.1 *Let $\mathscr{B} \subset \mathscr{A}$ be a sub-σ-algebra and X a real integrable (resp. positive) r.v. Then there exists a real integrable (resp. positive) r.v. Y, $\mathscr{B}$-measurable and unique up to an equivalence such that*

$$\int_A Y \, d\mathbf{P} = \int_A X \, d\mathbf{P}, \quad \textit{for every } A \in \mathscr{B}. \tag{1.1}$$

Such a r.v. Y is called the conditional expectation of X given $\mathscr{B}$ and is written $\mathbf{E}(X \mid \mathscr{B})$ *or* $\mathbf{E}^{\mathscr{B}}(X)$.

•**1.3** We shall often make use of the following fact. Let $\mathscr{B} \subset \mathscr{A}$ be a sub-σ-algebra (thus, possibly, $\mathscr{B} = \mathscr{A}$). If X and Y are integrable (resp. positive) $\mathscr{B}$-measurable r.v.'s such that

$$\int_A X \, d\mathbf{P} \geq \int_A Y \, d\mathbf{P} \quad \text{for every } A \in \mathscr{B}, \tag{1.2}$$

then $X \geq Y$ a.s. This property already proves the uniqueness in Theorem–Definition 1.1.

•**1.4** The relation (1.1) can also be written $\mathbf{E}(YZ) = \mathbf{E}(XZ)$, where $Z = 1_A$. This equality extends to all finite linear combinations of indicators of sets in $\mathscr{B}$ and, by the usual techniques of integration theory, to more general classes of functions. More precisely, if X is a integrable (resp. positive) r.v. and $Y = \mathbf{E}(X \mid \mathscr{B})$, we have, for every bounded (resp. positive) $\mathscr{B}$-measurable r.v. Z,

$$\mathbf{E}(YZ) = \mathbf{E}(XZ). \tag{1.3}$$

•**1.5** We state now the main properties of the conditional expectation. In what follows $\mathscr{B}, \mathscr{C}$ are sub-σ-algebras of $\mathscr{A}$ and X. Y. X_n r.v.'s which are positive or satisfy suitable integrability conditions.

(i) $\mathbf{E}^{\mathscr{B}}(1) = 1$ a.s., $\mathbf{E}[\mathbf{E}^{\mathscr{B}}(X)] = \mathbf{E}(X)$.

(ii) $\mathbf{E}^{\mathscr{B}}(aX + bY) = a\mathbf{E}^{\mathscr{B}}(X) + b\mathbf{E}^{\mathscr{B}}(Y)$ a.s.

(iii) If $X \leq Y$, $\mathbf{E}^{\mathscr{B}}(X) \leq \mathbf{E}^{\mathscr{B}}(Y)$ a.s.

(iv) If f is a real convex function, $f(\mathbf{E}^{\mathscr{B}}(X)) \leq \mathbf{E}^{\mathscr{B}}(f(X))$ a.s. (Jensen's inequality). In particular $|\mathbf{E}^{\mathscr{B}}(X)|^p \leq \mathbf{E}^{\mathscr{B}}(|X|^p)$ a.s. for every $p \geq 1$.

(v) If Y is $\mathscr{B}$-measurable. $\mathbf{E}^{\mathscr{B}}(XY) = Y\mathbf{E}^{\mathscr{B}}(X)$ a.s., in particular $\mathbf{E}^{\mathscr{B}}(Y) = Y$ a.s.

(vi) If $\mathscr{C} \subset \mathscr{B}$. $\mathbf{E}^{\mathscr{C}}(X) = \mathbf{E}^{\mathscr{C}}[\mathbf{E}^{\mathscr{B}}(X)]$ a.s.

(vii) If $X_n \uparrow X$ a.s. (i.e.. $X_n \leq X_{n+1}$ a.s. for every n and $\lim_{n\to\infty} X_n = X$ a.s.), $\mathbf{E}^{\mathscr{B}}(X_n) \uparrow \mathbf{E}^{\mathscr{B}}(X)$ a.s. (Beppo Levi's theorem for conditional expectations).

(viii) If $X_n \geq 0$, $\mathbf{E}^{\mathscr{B}}(\underline{\lim}_{n\to\infty} X_n) \leq \underline{\lim}_{n\to\infty} \mathbf{E}^{\mathscr{B}}(X_n)$ a.s. (Fatou's lemma for conditional expectations).

(ix) If $X_n \to_{n\to\infty} X$ a.s. and, for every n, $|X_n| \leq Z \in \mathscr{L}^1$, then $\mathbf{E}^{\mathscr{B}}(X_n) \to_{n\to\infty} \mathbf{E}^{\mathscr{B}}(X)$ a.s. (Lebesgue's theorem for conditional expectations).

•**1.6** One can see immediately that, if $X = X'$ a.s., then $\mathbf{E}^{\mathscr{B}}(X) = \mathbf{E}^{\mathscr{B}}(X')$ a.s The conditional expectation is thus actually defined on equivalence classes of r.v.'s.

Having stressed this point, •1.5 (iv) states that the conditional expectation is an operator $L^p \to L^p$, $p \geq 1$. which is a contraction (recall that L^p is a space that is not formed of r.v.'s, but of equivalence classes of r.v.'s). The range of L^p through the operator $X \to \mathbf{E}^{\mathscr{B}}(X)$ will be written $L^p(\mathscr{B})$: it is the subspace of L^p that is formed by those equivalence classes of r.v.'s that contain at least one $\mathscr{B}$-measurable element.

•**1.7** The case L^2 deserves particular attention. Relation (1.3) means that, if $X \in L^2$ and $Y = \mathbf{E}^{\mathscr{B}}(X)$, then $X - Y \perp Z$ for every bounded $\mathscr{B}$-measurable Z. By density we have $X - Y \perp Z$ for every $Z \in L^2(\mathscr{B})$. Thus Y is the *orthogonal projection* of X on $L^2(\mathscr{B})$. Therefore, by a classical characterization of the orthogonal projection,

$\mathbf{E}[(X-Y)^2] = \inf_{Z\in L^2(\mathscr{B})} \mathbf{E}[(X-Z)^2]$ and, if $X \in L^2$, $\mathbf{E}^{\mathscr{B}}(X)$ is the best L^2 approximation of X by a $\mathscr{B}$-measurable r.v.

•1.8 Let T be a r.v. with values in the measurable space $(E, \mathscr{E})$ and $\sigma(T)$ the smallest σ-algebra on Ω making T measurable. We recall that a real r.v. Z is $\sigma(T)$-measurable if and only if there exists a measurable function h on $(E, \mathscr{E})$ such that $Z = h(T)$. Thus, if $\mathscr{B} = \sigma(T)$, there exists a function h, measurable on $(E, \mathscr{E})$, such that $\mathbf{E}(X \mid \mathscr{B}) = h(T)$ a.s. In this case the computation of the conditional expectation is reduced to the determination of h. One writes also $\mathbf{E}(X \mid T)$ instead of $\mathbf{E}(X \mid \sigma(T))$ and, with a suggestive notation,

$$h(t) = \mathbf{E}(X \mid T = t).$$

One should remark that h is only defined a.s. with respect to the law of T.

Thus, thanks to •1.4, for a integrable (resp. positive) r.v. X, $h(T) = \mathbf{E}(X \mid T)$ a.s. if and only if, for every bounded (resp. positive) measurable function g,

$$\mathbf{E}[h(T)g(T)] = \mathbf{E}[Xg(T)]. \tag{1.4}$$

If $X \in L^2$, the results of •1.7 allow characterizing the function h through the property

$$\mathbf{E}[(X-h(T))^2] = \inf\{\mathbf{E}[(X-g(T))^2];\ g \text{ such that } g(T) \in L^2\}.$$

•1.9 If $\mathscr{B} \subset \mathscr{A}$ is a sub-σ-algebra and $A \in \mathscr{A}$, let us set $\mathbf{P}^{\mathscr{B}}(A) = \mathbf{P}(A \mid \mathscr{B}) = \mathbf{E}(1_A \mid \mathscr{B})$. $\mathbf{P}^{\mathscr{B}}(A)$ is called *the conditional probability* of A given $\mathscr{B}$. It is worth pointing out that, thanks to •1.5 (ii) and (vii), it holds, for every sequence $(A_n)_{n\geq 1}$ of disjoint events,

$$\mathbf{P}^{\mathscr{B}}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mathbf{P}^{\mathscr{B}}(A_n) \text{ a.s.}$$

and also $\mathbf{P}^{\mathscr{B}}(A^c) = 1 - \mathbf{P}^{\mathscr{B}}(A)$ a.s. Notice however that $\mathbf{P}^{\mathscr{B}}(A)$ is a r.v. that is only defined up to a negligible set: because of this, this notion should be handled with a certain care.

•1.10 Let us state a first link between the theory developed so far and the Introduction. Let $(A_n)_{n\geq 1}$ be a partition of Ω with $A_n \in \mathscr{A}$ and $\mathscr{B} = \sigma(A_n, n \geq 1)$. We recall that a r.v. Y is $\mathscr{B}$-measurable if and only if $Y = \sum_{n\geq 1} a_n 1_{A_n}$, $a_n \in \mathbb{R}$. Let X be a positive r.v. and $Y = \mathbf{E}^{\mathscr{B}}(X)$. Then $Y = \sum_{n\geq 1} \alpha_n 1_{A_n}$ and, by the relation

$$\mathbf{E}(X 1_{A_n}) = \mathbf{E}(1_{A_n} Y) = \alpha_n \mathbf{P}(A_n),$$

we get $\alpha_n = \mathbf{P}(A_n)^{-1}\mathbf{E}(X1_{A_n})$ if $\mathbf{P}(A_n) > 0$, whereas α_n can be chosen arbitrarily if $\mathbf{P}(A_n) = 0$. In particular, if $\mathbf{P}(A_n) > 0$ for every n, we have, for every $A \in \mathscr{B}$,

$$\mathbf{P}^{\mathscr{B}}(A) = \sum_{n\geq 1} \frac{\mathbf{P}(A \cap A_n)}{\mathbf{P}(A_n)} 1_{A_n}.$$

•1.11 Let us recall that a r.v. X is *independent* of the σ-algebra $\mathscr{B}$ if X is independent of every $\mathscr{B}$-measurable r.v. Y. We have then, for every positive or bounded function f, $\mathbf{E}(f(X) \mid \mathscr{B}) = \mathbf{E}(f(X))$ a.s. Conversely we can prove, using the usual methods of integration theory, the following criterion:

(i) if, for every bounded function f, $\mathbf{E}^{\mathscr{B}}(f(X)) = constant$ a.s., then X and $\mathscr{B}$ are independent.

(ii) If X takes values in $\mathbb{R}^p$ and if, for every $t \in \mathbb{R}^p$, $\mathbf{E}^{\mathscr{B}}(e^{i\langle t, X\rangle}) = constant$ a.s., then X and $\mathscr{B}$ are independent.

Let us point out that, thanks to •1.5 (i), these constants are equal, respectively, to $\mathbf{E}(f(X))$ and $\mathbf{E}(e^{i\langle t, X\rangle})$.

In particular, if X and T are independent, we have $\mathbf{E}(f(X) \mid T) = \mathbf{E}(f(X))$ a.s. This relation admits the following very useful extension.

Lemma 1.2 *Let $\mathscr{B}$ be a sub-σ-algebra of $\mathscr{A}$, T a $\mathscr{B}$-measurable r.v. with values in $(E, \mathscr{E})$ and X a r.v. independent of $\mathscr{B}$ with values in $(F, \mathscr{F})$. Let h be a measurable function, positive, or such that $h(X, T)$ is integrable. Then*

$$\mathbf{E}[h(X, T) \mid \mathscr{B}] = H(T),$$

where $H(t) = \mathbf{E}(h(X, t))$.

The proof is very simple. We assume first that h is positive. Let Z be a positive $\mathscr{B}$-measurable r.v. Let us note μ_X and $\mu_{Z,T}$ the laws of X and (Z, T), respectively. It holds from one side

$$H(t) = \int h(x, t)\, d\mu_X(x)$$

and, from the other, the r.v.'s X and (Z, T) being independent,

$$\mathbf{E}(Zh(X, T)) = \int zh(x, t)\, d\mu_X(x)\, d\mu_{Z,T}(z, t) =$$

$$= \int z\Big(\int h(x, t)\, d\mu_X(x)\Big)\, d\mu_{Z,T}(z, t) = \int zH(t)\, d\mu_{Z,T}(z, t) = \mathbf{E}(ZH(T)).$$

Then it is easy to extend the proof to the case of an integrable function h.

If $\mathscr{B} = \sigma(T)$, the lemma states that, if X and T are independent, then

$$\mathbf{E}[h(X, T) \mid T = t] = \mathbf{E}[h(X, t)]. \tag{1.5}$$

Conditional Expectations and Conditional Laws

•1.12 We now establish the link between Theorem–Definition 1.1 and conditional laws. Let T be a r.v. with values in $(E, \mathscr{E})$, ν its law and X a r.v. with values in $(F, \mathscr{F})$. A family of probabilities on $(F, \mathscr{F})$, $(n(t, dx))_{t \in E}$, is a *conditional law of X given T* if,

(i) For every $A \in \mathscr{F}$, $t \to n(t, A)$ is $\mathscr{E}$-measurable.

(ii) For every $A \in \mathscr{F}$ and $B \in \mathscr{E}$,

$$\mathbf{P}(X \in A, T \in B) = \int_B n(t, A)\, \nu(dt).$$

We deduce that, if f and g are positive measurable functions,

$$E(f(X)g(T)) = \int_E \Big(\int_F f(x)\, n(t, dx)\Big) g(t)\, \nu(dt).$$

Thus, if we set $h(t) = \int_{\mathbb{R}} f(x)\, n(t, dx)$, then

$$\mathbf{E}(f(X)g(T)) = \int_E h(t)g(t)\, \nu(dt) = \mathbf{E}(h(T)g(T)),$$

which means exactly that

$$\mathbf{E}(f(X) \mid T) = \int_{\mathbb{R}} f(x)\, n(T, dx) \quad \text{a.s.}$$

In particular, for $X \geq 0$, it holds a.s.

$$\mathbf{E}(X \mid T) = \int_{\mathbb{R}} x\, n(T, dx).$$

Thus the conditional expectation appears to be the mean of the conditional law.

It is easy to see that the previous formula extends to the case $f(X)$ (resp. X) integrable.

Let us assume, in particular, that the couple (T, X) has a density h with respect to a product measure $\mu \otimes \rho$ on $E \times F$, μ and ρ being σ-finite. Let us set

$$\phi(t) = \int_{\mathbb{R}^q} h(t, x)\, \rho(dx)$$

the density of T and $Q = \{t; \phi(t) = 0\}$. Of course $\mathbf{P}(T \in Q) = 0$. We set

$$\bar{h}(x; t) = \begin{cases} \frac{h(t,x)}{\phi(t)} & \text{if } t \notin Q \\ \text{any arbitrary density} & \text{if } t \in Q. \end{cases} \tag{1.6}$$

It can be shown easily that $n(t, dx) = \bar{h}(x; t)\, \rho(dx)$ is a conditional law of X given $T = t$. Thus we have, for every positive Borel function f,

$$\mathbf{E}(f(X) \mid T) = \int_F f(x)\bar{h}(x; T)\, \rho(dx),$$

which allows computing explicitly the conditional expectation in many concrete situations.

Exercises

Exercise 1.1 Let X and Y be independent r.v.'s with a $B(1, p)$ law, i.e., Bernoulli with parameter p. We set $Z = 1_{\{X+Y=0\}}$ and $\mathscr{G} = \sigma(Z)$. Compute $\mathbf{E}(X \mid \mathscr{G})$ and $\mathbf{E}(Y \mid \mathscr{G})$. Are these r.v.'s still independent?

Exercise 1.2 Let X be a square integrable real r.v. on $(\Omega, \mathscr{F}, \mathbf{P})$ and $\mathscr{G} \subset \mathscr{F}$ a sub-σ-algebra. Show that, if we set $\mathrm{Var}(X \mid \mathscr{G}) = \mathbf{E}[(X - \mathbf{E}(X \mid \mathscr{G}))^2 \mid \mathscr{G}]$, then

$$\mathrm{Var}(X) = \mathbf{E}(\mathrm{Var}(X \mid \mathscr{G})) + \mathrm{Var}(\mathbf{E}(X \mid \mathscr{G})).$$

Exercise 1.3 Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability space and $\mathscr{G} \subset \mathscr{F}$ a sub-σ-algebra. For $A \in \mathscr{F}$, consider the event $B = \{\mathbf{E}(1_A \mid \mathscr{G}) = 0\}$.

Show that $B \subset A^c$ a.s.

Exercise 1.4 *1) Let X, Y, Z be r.v.'s with values in $(E, \mathscr{E})$ such that the couples (X, Z) and (Y, Z) have the same law (in particular X and Y have common law μ).*

1a) Show that, if f is a real positive function (resp. such that $f(X)$ is integrable), then

$$\mathrm{E}(f(X) \mid Z) = \mathrm{E}(f(Y) \mid Z) \quad \text{a.s.}$$

1b) Let g be a measurable application from $(E, \mathscr{E})$ to $\mathbb{R}$ that is positive (resp. such that $g(Z)$ is integrable), $h_1(X) = \mathrm{E}(g(Z) \mid X)$, $h_2(Y) = \mathrm{E}(g(Z) \mid Y)$. Show that $h_1 = h_2$ μ-a.s.

2) Let $T_1, \dots, T_n$ be real integrable r.v.'s, independent and having the same law. We set $T = T_1 + \dots + T_n$.

2a) Show that

$$\mathrm{E}(T_1 \mid T) = \frac{T}{n}.$$

2b) Compute $\mathrm{E}(T \mid T_1)$.

Exercise 1.5 (Conditional laws of Gaussian vectors) 1) Let X be a r.v. with values in $\mathbb{R}^m$, of the form

$$X = \phi(Y) + Z,$$

where Y and Z are independent. Show that the conditional law of X given $Y = y$ coincides with the law of $Z + \phi(y)$.

2) Let X, Y be Gaussian vectors taking values in $\mathbb{R}^k$ and $\mathbb{R}^p$, respectively. We assume that their joint law on $(\mathbb{R}^{k+p}, \mathscr{B}(\mathbb{R}^{k+p}))$ is Gaussian with mean and covariance matrix given, respectively, by

$$\begin{pmatrix} m_X \\ m_Y \end{pmatrix} \qquad \begin{pmatrix} R_X & R_{XY} \\ R_{YX} & R_Y \end{pmatrix}$$

where R_X, R_Y are the covariance matrices of X and Y, respectively, and where $R_{XY} = \mathrm{E}((X - \mathrm{E}(X))\, {}^t(Y - \mathrm{E}(Y))) = {}^tR_{YX}$ is the $k \times p$ matrix of the covariances between components of X and Y; we assume moreover that R_Y is positive defined (and thus invertible).

2a) Find a $k \times p$ matrix, A, such that the r.v.'s $X - AY$ and Y are independent.

2b) Show that the conditional law of X given Y is Gaussian with mean

$$\mathrm{E}(X \mid Y) = m_X + R_{XY} R_Y^{-1}(Y - m_Y) \tag{1.7}$$

and covariance matrix

$$R_X - R_{XY} R_Y^{-1} R_{YX}. \tag{1.8}$$

3) Let X be a signal, with normal law $N(0, 1)$. We assume that we cannot observe the value of X; we only know an observation $Y = X + W$, where W is a noise, independent of X and with law $N(0, \sigma^2)$.

3a) Given an observation $Y = y$, give an estimate of the value X of the signal.

3b) Let us assume $\sigma^2 = 0.1$ and that the value of the observation is $Y = 0.55$. What is the probability for the signal X to be in the interval $[\frac{1}{4}, \frac{3}{4}]$?

Exercise 1.6 In this exercise we give an application of the conditional laws of Gaussian vectors. The reader should first try to understand thoroughly Exercise 1.5.

A) Let $\sigma > 0$, $a \in \mathbb{R}^d$ and C a positive definite (symmetric) $d \times d$ matrix. Show that

$$\left(C + \sigma^2 a\, {}^t a\right)^{-1} = C^{-1} - \frac{C^{-1} a\, {}^t a C^{-1}}{\sigma^{-2} + \langle C^{-1} a, a \rangle}$$

(remark that ${}^t a C^{-1} a = \langle C^{-1}a, a\rangle$).

B) Let $(Z_n)_{n\geq 1}$ be a sequence of real independent r.v.'s such that Z_n has law $N(0, c_n^2)$, $c_n > 0$ and X a real r.v. with law $N(0, \sigma^2)$, $\sigma > 0$, independent of the sequence $(Z_n)_{n\geq 1}$. We set, for every $n \geq 1$,

$$Y_n = X + Z_n, \qquad \mathscr{G}_n = \sigma(Y_1, \dots, Y_n), \qquad \hat{X}_n = \mathrm{E}^{\mathscr{G}_n}(X).$$

Intuitively, Y_n is a noisy observation of X. We note $Y^n = (Y_1, \dots, Y_n)$.

B1) What is the law of the vector $(X, Y_1, \dots, Y_n)$?

B2) For every $y \in \mathbb{R}^n$, show that the conditional law of X given $Y^n = y$ is Gaussian and compute its mean and variance.

B3) Compute $\hat{X}_n$ and $\mathrm{E}[(X - \hat{X}_n)^2]$.

B4) Show that $\hat{X}_n \to_{n\to\infty} X$ in L^2 if and only if $\sum_{n\geq 1} c_n^{-2} = +\infty$.

Exercise 1.7 Let $T_1, T_2, \dots$ be independent r.v.'s having an exponential law with parameter λ. We set $T = T_1 + \dots + T_n$.

a) Determine the conditional law of T_1 given T and then compute $\mathrm{E}(T_1 \mid T)$.

b) Compute $\mathrm{E}(T_1^2 \mid T)$ and $\mathrm{E}(T_1 T_2 \mid T)$.

Exercise 1.8 Let $(X_n)_{n\geq 0}$ be a sequence of independent r.v.'s taking values in $\mathbb{R}^d$, all defined on the same probability space $(\Omega, \mathscr{F}, \mathrm{P})$. Let us define $S_0 = 0$, $S_n = X_1 + \dots + X_n$ and $\mathscr{F}_n = \sigma(S_k, k \leq n)$. Show that, for every bounded Borel function $f: \mathbb{R}^d \to \mathbb{R}$,

$$\mathrm{E}(f(S_n) \mid \mathscr{F}_{n-1}) = \mathrm{E}(f(S_n) \mid S_{n-1}) \tag{1.9}$$

and express this quantity in terms of the law μ_n of X_n.

Solutions

E1.1 Observe that the σ-algebra $\mathscr{G}$ is generated by the partition $A_0 = \{X + Y = 0\}$, $A_1 = \{X + Y \geq 1\}$. By •1.10 the r.v. $\mathrm{E}(X \mid \mathscr{G})$ takes on A_i, $i = 0, 1$, the value

$$\alpha_i = \frac{\mathrm{E}(X 1_{A_i})}{\mathrm{P}(A_i)}.$$

But $X = 0$ on A_0; hence $\mathrm{E}(X 1_{A_0}) = 0$ and $\alpha_0 = 0$. On the other hand $X 1_{A_1} = 1_{\{X=1\}} 1_{\{X+Y\geq 1\}} = 1_{\{X=1\}}$. Therefore

$$\alpha_1 = \frac{p}{\mathrm{P}(A_1)} = \frac{p}{(1 - (1-p)^2)}.$$

Finally

$$\mathrm{E}(X \mid \mathscr{G}) = \frac{p}{1 - (1-p)^2} 1_{\{X+Y\geq 1\}}.$$

By symmetry (the right-hand side is symmetric in X and Y) this formula also gives $\mathrm{E}(Y \mid \mathscr{G})$. Therefore $\mathrm{E}(X \mid \mathscr{G}) = \mathrm{E}(Y \mid \mathscr{G})$. As a nonconstant r.v. cannot be independent of itself; the two r.v.'s $\mathrm{E}(X \mid \mathscr{G})$ and $\mathrm{E}(Y \mid \mathscr{G})$ are not independent.

E1.2 One may write

$$X - \mathrm{E}(X) = X - \mathrm{E}(X \mid \mathscr{G}) + \mathrm{E}(X \mid \mathscr{G}) - \mathrm{E}(X).$$

Now the r.v. $\mathbf{E}(X \mid \mathscr{G}) - \mathbf{E}(X)$ is $\mathscr{G}$-measurable whereas $X - \mathbf{E}(X \mid \mathscr{G})$ is orthogonal to every $\mathscr{G}$-measurable r.v. (•1.7). Therefore

$$\mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{E}((X - \mathbf{E}(X \mid \mathscr{G}))^2) + \mathbf{E}((\mathbf{E}(X \mid \mathscr{G}) - \mathbf{E}(X))^2) =$$
$$= \mathbf{E}(\mathrm{Var}(X \mid \mathscr{G})) + \mathrm{Var}(\mathbf{E}(X \mid \mathscr{G})).$$

E1.3 By definition $\mathbf{E}(1_A \mid \mathscr{G})$ is a $\mathscr{G}$-measurable r.v. so that $B \in \mathscr{G}$. Therefore, by the definition of conditional expectation,

$$\mathbf{E}(1_A 1_B) = \mathbf{E}(\mathbf{E}(1_A \mid \mathscr{G}) 1_B). \tag{1.10}$$

As on B it holds $\mathbf{E}(1_A \mid \mathscr{G}) = 0$, (1.10) implies $\mathbf{E}(1_A 1_B) = 0$ and thus, since $1_A 1_B$ is positive, $1_{A \cap B} = 0$ a.s., which is equivalent to $B \subset A^c$ a.s.

E1.4 1a) For every measurable positive (resp. bounded) application ϕ from $(E, \mathscr{E})$ to $\mathbb{R}$,

$$\mathbf{E}[\phi(Z)\mathbf{E}(f(X) \mid Z)] = \mathbf{E}(\phi(Z)f(X)) = \mathbf{E}(\phi(Z)f(Y)) = \mathbf{E}[\phi(Z)\mathbf{E}(f(Y) \mid Z)],$$

which implies (using •1.3) that $\mathbf{E}(f(X) \mid Z) = \mathbf{E}(f(Y) \mid Z)$ a.s.

1b) For every measurable positive (resp. bounded) application ϕ from $(E, \mathscr{E})$ to $\mathbb{R}$,

$$\int \phi(x)h_1(x)\,d\mu(x) = \mathbf{E}(\phi(X)h_1(X)) = \mathbf{E}[\phi(X)\mathbf{E}(g(Z) \mid X)] = \mathbf{E}[\phi(X)g(Z)] =$$
$$= \mathbf{E}[\phi(Y)g(Z)] = \mathbf{E}[\phi(Y)\mathbf{E}(g(Z) \mid Y)] = \int \phi(y)h_2(y)\,d\mu(y)$$

which, again, implies $h_1 = h_2$, except at most on a μ-null set.

2a) Clearly the r.v.'s $(T_1, T), (T_2, T), \dots, (T_n, T)$ have the same law. Thanks to (1a), this produces

$$\mathbf{E}(T_1 \mid T) = \ldots = \mathbf{E}(T_n \mid T) \quad \text{a.s.}$$

Therefore

$$n\mathbf{E}(T_1 \mid T) = \mathbf{E}(T_1 \mid T) + \ldots + \mathbf{E}(T_n \mid T) = \mathbf{E}(T_1 + \ldots + T_n \mid T) =$$
$$= \mathbf{E}(T \mid T) = T,$$

where the last relation comes by •1.5 (v).

2b) If $Z = T_2 + \ldots + T_n$, the r.v.'s T_1 and Z are independent and $T = T_1 + Z$. Hence

$$\mathbf{E}(T \mid T_1) = \mathbf{E}(Z + T_1 \mid T_1) = \mathbf{E}(Z) + T_1 = (n-1)\mathbf{E}(T_1) + T_1,$$

where we used the properties •1.5 (ii), (v) and •1.11.

♦ This exercise gives an example of computation of a conditional expectation without previous determination of the conditional distribution (but using instead the definition and properties •1.5).

E1.5 1) If one denotes by $n(y, dx)$ the conditional distribution of X given $Y = y$, one knows that (•1.12), for every bounded Borel function f,

$$\int f(x)\,n(y, dx) = \mathbf{E}(f(X) \mid Y = y) = \mathbf{E}(f(\phi(Y) + Z) \mid Y = y).$$

By Lemma 1.2, the last quantity equals $\mathbf{E}(f(\phi(y)+Z))$. Therefore $n(y, dx)$ is the distribution of $\phi(y)+Z$. (To be precise, this last relation is true for almost every y, with respect to the distribution of Y...).

2a) Let $Z = X - AY$. The couple (X, Y) being Gaussian, this also holds for (Z, Y), which is obtained from (X, Y) through a linear transformation. Therefore, to prove the independence of Z and Y, it is sufficient to check that $\mathrm{Cov}(Z_i, Y_j) = 0$ for every $i = 1, \dots, k$, $j = 1, \dots, p$. Suppose first, in order to simplify the notations, that the means m_X and m_Y are equal to 0. The condition of noncorrelation between the components of Z and Y may therefore be written

$$0 = \mathbf{E}(Z\,{}^tY) = \mathbf{E}((X - AY)\,{}^tY) = \mathbf{E}(X\,{}^tY) - A\mathbf{E}(Y\,{}^tY) = R_{XY} - AR_Y.$$

The matrix A is therefore unique and given by $A = R_{XY}R_Y^{-1}$. If we do not suppose any more that the means are equal to 0, it is sufficient to repeat the same calculation with X and Y replaced by $X - m_X$ and $Y - m_Y$, respectively.

2b) One may write

$$X = (X - AY) + AY,$$

where the r.v.'s $X - AY$ and AY are independent. Thanks to 1) the conditional distribution of X given $Y = y$ is the same as the distribution of $X - AY + Ay$. Since $X - AY$ is Gaussian, this distribution is characterized by its mean

$$m_X - Am_Y + Ay = m_X - R_{XY}R_Y^{-1}(m_Y - y)$$

and its covariance matrix

$$\begin{aligned} C_{X-AY} &= R_X - R_{XY}\,{}^tA - AR_{YX} + AR_Y\,{}^tA = \\ &= R_X - R_{XY}R_Y^{-1}\,{}^tR_{XY} - R_{XY}R_Y^{-1}\,{}^tR_{XY} + R_{XY}R_Y^{-1}R_YR_Y^{-1}\,{}^tR_{XY} = \\ &= R_X - R_{XY}R_Y^{-1}\,{}^tR_{XY}, \end{aligned}$$

where we used the fact that R_Y is symmetric and the obvious relation $R_{YX} = {}^tR_{XY}$.

3a) We are looking for a function of the observation Y that is a good approximation of X. One knows (•1.7 and •1.8) that the r.v. $\phi(Y)$ that minimizes the discrepancy $\mathbf{E}((\phi(Y) - X)^2)$ is the conditional expectation $\phi(Y) = \mathbf{E}(X \mid Y)$. Therefore if one decides to measure the quality of the approximation of X by $\phi(Y)$ with this distance (other choices are also possible), the best approximation of the value of X is given by $\mathbf{E}(X \mid Y)$. From (1.7) and (1.8), with the values

$$m_X = m_Y = 0, \quad R_{XY} = 1, \quad R_Y = 1 + \sigma^2,$$

one obtains

$$\mathbf{E}(X \mid Y) = \frac{Y}{1+\sigma^2}.$$

3b) With the given values the conditional distribution of X given $Y = y$ is Gaussian with mean

$$\frac{0.55}{1.1} = \frac{1}{2}$$

and variance

$$1 - \frac{1}{1+\sigma^2} = \frac{\sigma^2}{1+\sigma^2} = \frac{1}{11}.$$

The probability that a r.v. with such a Gaussian distribution takes values in interval

$[\frac{1}{4}, \frac{3}{4}]$ is easily computed: it is equal to the probability that a $N(0, 1)$ r.v. takes its values in the interval

$$[-\tfrac{1}{4}\cdot\sqrt{11}, \tfrac{1}{4}\cdot\sqrt{11}] = [-0.83, 0.83],$$

which may be computed with the help of tables (or of specialized software), and one obtains the value 0.58.

E1.6 A) It is sufficient to compute

$$\begin{aligned}(C+\sigma^2 a\,{}^t a)\Big(C^{-1} - \frac{C^{-1}a\,{}^t a C^{-1}}{\sigma^{-2}+\langle C^{-1}a, a\rangle}\Big) &=\\ = I + \sigma^2 a\,{}^t a C^{-1} - \frac{1}{\sigma^{-2}+\langle C^{-1}a, a\rangle}\big(a\,{}^t a C^{-1} + \sigma^2 a\,{}^t a C^{-1} a\,{}^t a C^{-1}\big) &=\\ = I + \sigma^2 a\,{}^t a C^{-1} - \frac{1+\sigma^2\langle C^{-1}a, a\rangle}{\sigma^{-2}+\langle C^{-1}a, a\rangle} a\,{}^t a C^{-1} &= I.\end{aligned}$$

B1) The vector $(X, Y_1, \dots, Y_n)$ is obtained from $(X, Z_1, \dots, Z_n)$ through a linear transformation. As the r.v.'s $X, Z_1, \dots, Z_n$ are independent and each have a Gaussian distribution, $(X, Y_1, \dots, Y_n)$ is a Gaussian vector which is moreover centered. Let us look for its covariance matrix. If $e \in \mathbb{R}^n$ denotes the vector with components $1, \dots, 1$, then $\mathrm{Var}(X) = \sigma^2$, $\mathbf{E}(X\,{}^t Y^n) = \sigma^2\,{}^t e$ and $R_{Y^n} = \mathbf{E}(Y^n\,{}^t Y^n) = C_n + \sigma^2 e\,{}^t e$, where

$$C_n = \begin{pmatrix} c_1^2 & & & 0 \\ & c_2^2 & & \\ & & \ddots & \\ 0 & & & c_n^2 \end{pmatrix}$$

Thus the covariance matrix of $(X, Y_1, \dots, Y_n)$ is

$$\begin{pmatrix} \sigma^2 & \dots & \sigma^2 \\ \vdots & R_{Y^n} & \\ \sigma^2 & & \end{pmatrix}.$$

B2) By (1.7), (1.8) and A), the conditional distribution of X given $Y^n = y$ is Gaussian with mean

$$\sigma^2\,{}^t e R_{Y^n}^{-1} y = \sigma^2\Big(\langle C_n^{-1}y, e\rangle - \frac{\langle C_n^{-1}e, e\rangle\langle C_n^{-1}y, e\rangle}{\sigma^{-2}+\langle C_n^{-1}e, e\rangle}\Big) = \frac{\langle C_n^{-1}y, e\rangle}{\sigma^{-2}+\langle C_n^{-1}e, e\rangle}$$

and variance

$$\begin{aligned}\rho_n^2 \overset{\text{def}}{=} \sigma^2 - \sigma^4\,{}^t e\Big(C_n^{-1} - \frac{C_n^{-1}e\,{}^t e C_n^{-1}}{\sigma^{-2}+\langle C_n^{-1}e, e\rangle}\Big)e &=\\ = \sigma^2\Big(1 - \sigma^2\Big(\langle C_n^{-1}e, e\rangle - \frac{\langle C_n^{-1}e, e\rangle^2}{\sigma^{-2}+\langle C_n^{-1}e, e\rangle}\Big)\Big) &=\\ = \sigma^2\Big(1 - \frac{\langle C_n^{-1}e, e\rangle}{\sigma^{-2}+\langle C_n^{-1}e, e\rangle}\Big) = \frac{1}{\sigma^{-2}+\sum_{k=1}^n c_k^{-2}}.\end{aligned}$$

B3) Obviously

$$\hat{X}_n = \mathbf{E}^{\mathscr{G}_n}(X) = \frac{\langle C_n^{-1}Y, e\rangle}{\sigma^{-2} + \langle C_n^{-1}e, e\rangle} = \frac{\sum_{k=1}^n c_k^{-2} Y_k}{\sigma^{-2} + \sum_{k=1}^n c_k^{-2}},$$

and $\mathbf{E}[(X - \hat{X}_n)^2] = \mathbf{E}[(X - \hat{X}_n)^2 \,|\, \mathscr{G}_n] = \rho_n^2$.

B4) It is immediate, as $\rho_n^2 = \big(\sigma^{-2} + \sum_{k=1}^n c_k^{-2}\big)^{-1}$; hence $\mathbf{E}[(X - \hat{X}_n)^2] \to_{n\to\infty} 0$ if and only if $\sum_{k=1}^\infty c_k^{-2} = +\infty$.

E1.7 a) Let us compute the distribution of the couple (T_1, T). If we set $Z = T_2 + \ldots + T_n$, then the r.v. Z has a $\Gamma(n-1, \lambda)$ distribution, i.e., with density

$$g_Z(v) = \frac{\lambda^{n-1}}{(n-2)!} v^{n-2} e^{-\lambda v} 1_{\{v \ge 0\}}.$$

Moreover $T = T_1 + Z$. As the r.v.'s T_1 and Z are independent, they have a joint density $\bar{g}_2$ given by

$$\bar{g}_2(u, v) = \frac{\lambda^{n-1}}{(n-2)!} v^{n-2} e^{-\lambda(v+u)} 1_{\{u \ge 0\}} 1_{\{v \ge 0\}}.$$

As the couple (T_1, T) is the image of (T_1, Z) through the linear transformation associated to the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

by the change of variable theorem the density g_2 of (T_1, T) is given by

$$g_2(x, y) = \bar{g}_2\big(A^{-1}\tbinom{x}{y}\big) |\det A|^{-1}.$$

Now

$$A^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

$$A^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y - x \end{pmatrix};$$

hence

$$g_2(x, y) = \frac{\lambda^{n-1}}{(n-2)!} (y - x)^{n-2} e^{-\lambda y} 1_{\{0 \le x \le y\}}.$$

The conditional density of T_1 given $T = y$ is therefore

$$\bar{g}(x) = \frac{g_2(x, y)}{g_T(y)} = \frac{n-1}{y^{n-1}} (y - x)^{n-2} 1_{[0,y]}(x)$$

(remark: it does not depend on λ). $\mathbf{E}(T_1 \mid T)$ is the mean of this conditional distribution; integrating by parts or using the relation

$$\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\, du = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \qquad \alpha > 0, \beta > 0,$$

which holds for $\alpha, \beta > 0$, we have

$$\mathbf{E}(T_1 \mid T) = \frac{n-1}{T^{n-1}} \int_0^T u(T-u)^{n-2}\, du = \frac{T}{n}$$

(of course this is the same as the result obtained in Exercise 1.4, where no hypotheses on the distribution of the r.v.'s T_i, $1 \le i \le n$ were made).

♦ In particular note that $\mathbf{E}(T_1 \mid T) = \mathbf{E}(T_2 \mid T)$. T_1 and T_2 being independent, this gives another example of independent r.v.'s that are not independent any more after conditioning, a phenomenon already observed in Exercise 1.1.

b) Similarly

$$\mathbf{E}(T_1^2 \mid T) = \frac{n-1}{T^{n-1}} \int_0^T u^2 (T-u)^{n-2}\, du = \frac{2T^2}{n(n+1)}.$$

In order to compute $\mathbf{E}(T_1 T_2 \mid T)$, one can first compute the distribution of the triple (T_1, T_2, T). More quickly, just observe that

$$T^2 = \sum_{i=1}^n T_i^2 + \sum_{1 \le i \ne j \le n} T_i T_j. \tag{1.11}$$

Now, by symmetry, for every $i \ne j$ the r.v.'s $(T_i T_j, T)$ and $(T_1 T_2, T)$ have the same distribution and, thanks to Exercise 1.4 (a),

$$\mathbf{E}(T_i T_j \mid T) = \mathbf{E}(T_1 T_2 \mid T).$$

The same argument yields

$$\mathbf{E}(T_i^2 \mid T) = \mathbf{E}(T_1^2 \mid T) = \frac{2T^2}{n(n+1)}.$$

The relation (1.11) therefore gives

$$T^2 = \mathbf{E}(T^2 \mid T) = n \frac{2T^2}{n(n+1)} + (n^2 - n)\mathbf{E}(T_1 T_2 \mid T),$$

from which one gets

$$\mathbf{E}(T_1 T_2 \mid T) = \frac{T^2}{n(n+1)}.$$

E1.8 Observe that $S_n = X_n + S_{n-1}$. The r.v. S_{n-1} is $\mathscr{F}_{n-1}$-measurable whereas X_n is independent of $\mathscr{F}_{n-1}$. We are therefore in the situation of Lemma 1.2, which implies

$$\mathbf{E}(f(X_n + S_{n-1}) \mid \mathscr{F}_{n-1}) = \psi(S_{n-1}),$$

where the function ψ is defined by

$$\psi(x) = \mathbf{E}(f(X_n + x)) = \int f(y + x)\, d\mu_n(y).$$

Therefore

$$\mathbf{E}(f(S_n) \mid \mathscr{F}_{n-1}) = \psi(S_{n-1}) = \int f(y + S_{n-1})\, d\mu_n(y).$$

The right-hand side is a $\sigma(S_{n-1})$-measurable r.v. (as a function of S_{n-1}) and this immediately implies the relation (1.9); indeed as $\sigma(S_{n-1}) \subset \mathscr{F}_{n-1}$, by ▪1.5 (v) and (vi),

$$\begin{aligned}\mathbf{E}(f(S_n) \mid S_{n-1}) &= \mathbf{E}\big(\mathbf{E}(f(S_n) \mid \mathscr{F}_{n-1}) \mid \sigma(S_{n-1})\big) = \\ &= \mathbf{E}(\psi(S_{n-1}) \mid S_{n-1}) = \psi(S_{n-1}).\end{aligned}$$

♦ We shall see later the significance of the relation (1.9), which may also be expressed by saying that the sequence $(S_n)_{n\geq 0}$ satisfies the *Markov property*. Note also the role played by the very useful Lemma 1.2.

CHAPTER 2

Stochastic Processes

General Facts

•**2.1** A stochastic process on a measurable space $(E, \mathscr{E})$ is a term

$$X = (\Omega, \mathscr{F}, (X_n)_{n\geq 0}, \mathbf{P})$$

where $(\Omega, \mathscr{F}, \mathbf{P})$ is a probability space and where, for every n, X_n is a r.v. with values in $(E, \mathscr{E})$.

Let us set $\mathscr{F}_n^0 = \sigma(X_0, X_1, \dots, X_n)$. We know that a function Φ on Ω is $\mathscr{F}_n^0$-measurable if and only if $\Phi = \phi(X_0, X_1, \dots, X_n)$ for a $\mathscr{E}^{\otimes(n+1)}$-measurable function ϕ. In particular for an event A to belong to $\mathscr{F}_n^0$ it means, intuitively, that one knows whether A has happened or not as soon as the values taken by $X_0, X_1, \dots, X_n$ are known. However, in many cases, what is known at time n is not just the values of $X_0, X_1, \dots, X_n$, since also other information may be available, the values taken by other processes, for instance. This explains the following definition

Definition 2.1 *We call filtered probability space a term*

$$(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$$

where $(\Omega, \mathscr{F}, \mathbf{P})$ is a probability space and $(\mathscr{F}_n)_{n\geq 0}$ is an increasing family of sub-σ-algebras of $\mathscr{F}$, called a filtration, on Ω.

$\mathscr{F}_n$ is called the σ-algebra of the *events prior to time n*.

Definition 2.2 *An adapted stochastic process on a measurable space $(E, \mathscr{E})$ is a term $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ where $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ is a filtered probability space and, for every n, X_n is a r.v. with values in $(E, \mathscr{E})$ and $\mathscr{F}_n$-measurable (this last property is also expressed by saying that $(X_n)_{n\geq 0}$ is adapted to the filtration $(\mathscr{F}_n)_{n\geq 0}$).*

•**2.2** We give now two examples of those processes that will be our main concern.

i) An adapted real stochastic process X is called a *martingale* if every r.v. X_n is integrable and, for every n,

$$\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = X_n \quad \text{a.s.}$$

It is a *supermartingale* if $\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) \leq X_n$ and a *submartingale* if $\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) \geq X_n$.

ii) A stochastic process X on $(E, \mathscr{E})$ is a *Markov chain* if, for every n and every $A \in \mathscr{E}$,

$$\mathbf{P}(X_{n+1} \in A \mid \mathscr{F}_n) = \mathbf{P}(X_{n+1} \in A \mid X_n) \quad \text{a.s.}$$

This means, in particular, that, if one knows the position X_n of the process at time n, in order to foresee its position X_{n+1} at time $n+1$, the additional knowledge of what happened before time n (and which is contained in $\mathscr{F}_n$) carries no useful additional information.

•**2.3** Let X be a process on $(E, \mathscr{E})$. We denote by μ_n the law of the r.v. $(X_0, X_1, \dots, X_n)$, which takes its values in $(E^{n+1}, \mathscr{E}^{\otimes(n+1)})$. The probabilities $(\mu_n)_{n\geq 0}$ are called the *finite dimensional distributions* of the process X. Since

$$\mu_n(A_0 \times A_1 \times \dots \times A_n) = \mathbf{P}(X_0 \in A_0, \dots, X_n \in A_n),$$

it holds

$$\mu_{n-1}(A_0 \times A_1 \times \dots \times A_{n-1}) = \mu_n(A_0 \times A_1 \times \dots \times A_{n-1} \times E). \tag{2.1}$$

Conversely, given for every $n \geq 0$ a probability μ_n on $(E^{n+1}, \mathscr{E}^{\otimes(n+1)})$ satisfying (2.1), does a process having $(\mu_n)_{n\geq 0}$ as its finite dimensional distributions exist? In order to answer this question, we introduce the *canonical space*

$$\begin{aligned} \Omega = E^{\mathbb{N}}, \quad \omega = (\omega_n)_{n\geq 0}, \quad X_n(\omega) = \omega_n, \\ \mathscr{F}_n = \sigma(X_k, k \leq n), \quad \mathscr{F} = \sigma(X_k, k \geq 0), \end{aligned} \tag{2.2}$$

which is going to play an important role in the chapter on Markov chains. We define a probability $\mathbf{P}_n$ on $(\Omega, \mathscr{F}_n)$ by setting, for $A \in \mathscr{E}^{\otimes(n+1)}$,

$$\mathbf{P}_n(A \times E \dots \times E \times \dots) = \mu_n(A)$$

and then defining a set function on $\bigcup_{n\geq 0} \mathscr{F}_n$ by

$$\mathbf{P}(A) = \mathbf{P}_n(A) \text{ if } A \in \mathscr{F}_n.$$

We must extend $\mathbf{P}$ to $\mathscr{F} = \sigma(\bigcup_{n=0}^{\infty} \mathscr{F}_n)$. It holds

Theorem 2.3 (Kolmogorov) *Let $(\mu_n)_{n\geq 0}$ be a family of probability laws, each μ_n being a probability on the product space $(E^{n+1}, \mathscr{E}^{\otimes(n+1)})$, satisfying (2.1). Then on the canonical space defined in (2.2) there exists a unique probability $\mathbf{P}$ such that the process $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ has $(\mu_n)_{n\geq 0}$ as its finite dimensional distributions.*

•**2.4** Let us recall the following result (which is called the *monotone class theorem*). Let $\mathscr{S}$ be a family of parts of Ω such that,

i) $\Omega \in \mathscr{S}$,

ii) if $A_n \in \mathscr{S}$ and $A_n \uparrow A$, then $A \in \mathscr{S}$,

iii) if $A, B \in \mathscr{S}$ and $B \subset A$, then $A \setminus B \in \mathscr{S}$.

Then, if $\mathscr{S}$ contains a class $\mathscr{C}$ that is stable with respect to finite intersections, $\mathscr{S}$ contains $\sigma(\mathscr{C})$.

This result has two important consequences.

Theorem 2.4 *Let $\mathbf{P}_1, \mathbf{P}_2$ be two probabilities on $(\Omega, \mathscr{F})$ such that $\mathbf{P}_1(A) = \mathbf{P}_2(A)$ for every $A \in \mathscr{C}$, where $\mathscr{C}$ is a class that is stable with respect to finite intersections. Then $\mathbf{P}_1(A) = \mathbf{P}_2(A)$ for every $A \in \sigma(\mathscr{C})$.*

Theorem 2.5 *Let $\mathscr{H}$ be a vector space of bounded real functions defined on Ω and $\mathscr{C}$ a class of subsets of Ω, which is stable with respect to finite intersections. Assume that*

i) $1 \in \mathscr{H}$;

ii) if $f_n \in \mathscr{H}$ and $0 \le f_n \uparrow f$ for a bounded function f, then $f \in \mathscr{H}$;
iii) $1_A \in \mathscr{H}$, for every $A \in \mathscr{C}$.
Then $\mathscr{H}$ contains all bounded $\sigma(\mathscr{C})$-measurable functions.

Both Theorems 2.4 and 2.5 are important technical tools. For instance Theorem 2.4 implies immediately the uniqueness of **P** in Kolmogorov's Theorem 2.3, since $\bigcup_{n=0}^{\infty} \mathscr{F}_n$ is a class that is stable with respect to finite intersections.

We shall make reference to Theorems 2.4 and 2.5 speaking of "monotone class arguments".

Stopping Times

•**2.5** Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge 0}, \mathbf{P})$ be a filtered probability space (Definition 2.1). We set $\mathscr{F}_\infty = \sigma(\bigcup_{n\ge 0} \mathscr{F}_n)$.

Definition 2.6 *(i) A stopping time (of the filtration $(\mathscr{F}_n)_{n\ge 0}$) is any application ν $\Omega \to \bar{\mathbb{N}}$ such that, for every $n \ge 0$,*

$$\{\nu \le n\} \in \mathscr{F}_n.$$

(ii) We call σ-algebra of the events prior to time ν and write $\mathscr{F}_\nu$, the σ-algebra

$$\mathscr{F}_\nu = \{ A \in \mathscr{F}_\infty;\ \text{for every } n \ge 0,\ A \cap \{\nu \le n\} \in \mathscr{F}_n\}.$$

Remark. In (i) and (ii), one can replace $\{\nu \le n\}$ and $A \cap \{\nu \le n\}$ by $\{\nu = n\}$ and $A \cap \{\nu = n\}$, respectively, since

$$\{\nu \le n\} = \bigcup_{k=0}^{n} \{\nu = k\}, \qquad \{\nu = n\} = \{\nu \le n\} \setminus \{\nu \le n-1\}.$$

Fundamental examples. Let X be a stochastic process with values in $(E, \mathscr{E})$ adapted to the filtration $(\mathscr{F}_n)_{n\ge 0}$. We set, for $A \in \mathscr{E}$,

$$\tau_A(\omega) = \inf\{n \ge 0;\ X_n(\omega) \in A\}, \tag{2.3}$$

with the understanding that $\inf \emptyset = +\infty$, which *we shall use throughout all* this text. Then τ_A is a stopping time since

$$\{\tau_A = n\} = \{X_0 \notin A, X_1 \notin A, \dots, X_{n-1} \notin A, X_n \in A\} \in \mathscr{F}_n;$$

τ_A is called the *passage* or hitting time in A. We shall also consider (beware of the difference ...)

$$\sigma_A(\omega) = \inf\{n \ge 1;\ X_n(\omega) \in A\}; \tag{2.4}$$

σ_A is a stopping time (same proof), which, of course, coincides with τ_A if $X_0(\omega) \notin A$. It is called the *return time* in A.

•**2.6** Let $(\mathscr{F}_n)_{n\ge 0}$ be a filtration and ν_1, ν_2 two stopping times for this filtration. We have then the following properties (elementary to check and that are the subject of Exercise 2.1).

i) $\nu_1 + \nu_2$, $\nu_1 \vee \nu_2$, $\nu_1 \wedge \nu_2$ are stopping times with respect to the same filtration.
ii) If $\nu_1 \le \nu_2$, then $\mathscr{F}_{\nu_1} \subset \mathscr{F}_{\nu_2}$.
iii) $\mathscr{F}_{\nu_1 \wedge \nu_2} = \mathscr{F}_{\nu_1} \cap \mathscr{F}_{\nu_2}$.

iv) $\{\nu_1 < \nu_2\}$ and $\{\nu_1 = \nu_2\}$ both belong to $\mathscr{F}_{\nu_1} \cap \mathscr{F}_{\nu_2}$.

•**2.7** Let X be an adapted process: we want to define the position of the process at time ν, i.e., $X_\nu(\omega) = X_{\nu(\omega)}(\omega)$. There is a problem when $\nu(\omega) = +\infty$. A possible way to handle this situation is to fix a r.v. X_∞, $\mathscr{F}_\infty$-measurable, and to set $X_\nu = X_\infty$ on $\{\nu = +\infty\}$. If E is a topological space (and $\mathscr{E}$ its Borel σ-algebra), there are two customary choices for X_∞.

i) The sequence $(X_n)_{n\geq 0}$ has a.s. a limit X as $n \to \infty$. One can then choose $X_\infty = X$.

ii) Otherwise one can add to the space E an isolated point ∂, sometimes referred to as the *cemetery*, and set $X_\infty = \partial$.

One can then define

$$X_\nu = X_n \text{ on } \{\nu = n\}, \quad n \in \bar{\mathbb{N}}.$$

Let us remark that the r.v. X_ν is $\mathscr{F}_\nu$-measurable since

$$\{X_\nu \in A\} \cap \{\nu = n\} = \{X_n \in A\} \cap \{\nu = n\} \in \mathscr{F}_n.$$

Finally the following result is very useful in order to compute the conditional expectation with respect to $\mathscr{F}_\nu$.

Proposition 2.7 *Let X be a positive or integrable r.v. and ν a stopping time. Let us set, for every $n \in \bar{\mathbb{N}}$, $X_n = \mathrm{E}(X \mid \mathscr{F}_n)$. Then on $\{\nu = n\}$, it holds $\mathrm{E}(X \mid \mathscr{F}_\nu) = \mathrm{E}(X \mid \mathscr{F}_n)$, i.e.,*

$$\mathrm{E}(X \mid \mathscr{F}_\nu) = X_\nu \quad a.s. \tag{2.5}$$

Let us check (2.5) for $X \geq 0$. If Z is ≥ 0 and $\mathscr{F}_\nu$-measurable, then it is easy to see that $Z1_{\{\nu=n\}}$ is $\mathscr{F}_n$-measurable and that

$$\mathrm{E}(ZX_\nu) = \sum_{n\in\bar{\mathbb{N}}} \mathrm{E}[ZX_n 1_{\{\nu=n\}}] = \sum_{n\in\bar{\mathbb{N}}} \mathrm{E}[ZX 1_{\{\nu=n\}}] = \mathrm{E}(ZX),$$

i.e., $\mathrm{E}(X \mid \mathscr{F}_\nu) = X_\nu$.

Exercises

Exercise 2.1 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ be a filtered probability space (note however that in this exercise $\mathbf{P}$ plays no role), τ and ν two stopping times of the filtration $(\mathscr{F}_n)_{n\geq 0}$, $\mathscr{F}_\tau$ (resp. $\mathscr{F}_\nu$) the σ-algebra of the events prior to τ (resp. ν). Show that

i) if $\tau \equiv p$, $p \in \mathbb{N}$, then $\mathscr{F}_\tau = \mathscr{F}_p$;

ii) $\tau \wedge \nu$, $\tau \vee \nu$, $\tau + \nu$ are stopping times;

iii) if $\tau \leq \nu$, then $\mathscr{F}_\tau \subset \mathscr{F}_\nu$;

iv) $\mathscr{F}_{\tau\wedge\nu} = \mathscr{F}_\tau \cap \mathscr{F}_\nu$;

v) $\{\tau < \nu\} \in \mathscr{F}_\tau \cap \mathscr{F}_\nu$, $\{\tau = \nu\} \in \mathscr{F}_\tau \cap \mathscr{F}_\nu$.

Exercise 2.2 Let E be a countable set, $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0})$ the canonical space (see (2.2)) of the applications from $\Omega = E^{\mathbb{N}}$ to E. Let $\theta: \Omega \to \Omega$ be *the translation operator*, defined by $\theta((\omega_n)_{n\geq 0}) = (\omega_{n+1})_{n\geq 0}$, and $(\theta_n)_{n\geq 0}$ the sequence of its iterates (i.e., θ_0 =identity and $\theta_n = \theta \circ \ldots \circ \theta$, n times).

1) Show that $X_n \circ \theta_m = X_{n+m}$ and $\theta_n^{-1}(\mathscr{F}_m) = \sigma(X_n, \ldots, X_{n+m})$.

2) Let τ and ν be two stopping times of the filtration $(\mathscr{F}_n)_{n\geq 0}$ (possibly $\tau = \nu$).

2a) Show that, for every positive integer k, $k + \tau \circ \theta_k$ is a stopping time.

2b) Show that $\rho = \nu + \tau \circ \theta_\nu$ is a stopping time (with the understanding $\rho = +\infty$ on $\{\nu = +\infty\}$). Show that, if moreover one assumes that ν and τ are finite, $X_\tau \circ \theta_\nu = X_\rho$.

2c) For $A \subset E$, let us note

$$\tau_A(\omega) = \inf\{n \geq 0;\, X_n(\omega) \in A\}, \qquad \sigma_A(\omega) = \inf\{n \geq 1;\, X_n(\omega) \in A\},$$

the hitting and return times in A. These are stopping times, as we saw in ∎2.5. Let σ be a stopping time. Show that

$$\begin{aligned} \sigma + \tau_A \circ \theta_\sigma &= \inf\{k \geq \sigma;\, X_k \in A\} \\ \sigma + \sigma_A \circ \theta_\sigma &= \inf\{k > \sigma;\, X_k \in A\} \end{aligned} \tag{2.6}$$

and that, if $A \subset B$, $\tau_A = \tau_B + \tau_A \circ \theta_{\tau_B}$.

Solutions

E2.1 i) Suppose $\tau \equiv p$. The event $\{\tau \leq n\}$ is therefore equal to $\emptyset$ if $p > n$ and to Ω if $p \leq n$. By consequence, as $\emptyset \in \mathscr{F}_n$ for every n,

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty,\, A \in \mathscr{F}_n \text{ for every } n \geq p\} = \bigcap_{n \geq p} \mathscr{F}_n = \mathscr{F}_p.$$

ii)Let $n \in \mathbb{N}$. One must show that the events $\{\tau \wedge \nu \leq n\}$, $\{\tau \vee \nu \leq n\}$ and $\{\tau + \nu \leq n\}$ belong to $\mathscr{F}_n$. Since

$$\{\tau \wedge \nu \leq n\} = \{\tau \leq n\} \cup \{\nu \leq n\}$$

and τ and ν are stopping times, $\{\tau \leq n\}$ and $\{\nu \leq n\}$ belong to $\mathscr{F}_n$; therefore $\{\tau \wedge \nu \leq n\} \in \mathscr{F}_n$. In the same way, $\{\tau \vee \nu \leq n\} = \{\tau \leq n\} \cap \{\nu \leq n\} \in \mathscr{F}_n$. At last,

$$\{\tau + \nu \leq n\} = \bigcup_{k=0}^{n} \{\tau \leq k\} \cap \{\nu \leq n - k\}.$$

Now, for $0 \leq k \leq n$, $\{\tau \leq k\} \in \mathscr{F}_k \subset \mathscr{F}_n$ and $\{\nu \leq n - k\} \in \mathscr{F}_{n-k} \subset \mathscr{F}_n$: hence $\{\tau + \nu \leq n\} \in \mathscr{F}_n$, which ends the proof of (i).

iii) Let $A \in \mathscr{F}_\tau$ and $n \in \mathbb{N}$. As $\{\nu \leq n\} \subset \{\tau \leq n\}$, $A \cap \{\nu \leq n\} = A \cap \{\tau \leq n\} \cap \{\nu \leq n\}$. But $A \cap \{\tau \leq n\} \in \mathscr{F}_n$ and $\{\nu \leq n\} \in \mathscr{F}_n$ (ν is a stopping time); therefore $A \cap \{\tau \leq n\} \cap \{\nu \leq n\} \in \mathscr{F}_n$ and $A \cap \{\nu \leq n\} \in \mathscr{F}_n$. Thus $A \in \mathscr{F}_\nu$.

iv) One already knows, by (ii) and (iii), that $\mathscr{F}_{\tau\wedge\nu} \subset \mathscr{F}_\tau \cap \mathscr{F}_\nu$. Let us prove the converse inclusion. Let $A \in \mathscr{F}_\tau \cap \mathscr{F}_\nu$, obviously $A \subset \mathscr{F}_\infty$. To prove that $A \in \mathscr{F}_{\tau\wedge\nu}$, one must show that, for every $k \geq 0$, $A \cap \{\tau \wedge \nu \leq k\} \in \mathscr{F}_k$. But one already knows that $A \cap \{\tau \leq k\} \in \mathscr{F}_k$ and $A \cap \{\nu \leq k\} \in \mathscr{F}_k$. Hence, recalling that $\{\tau \wedge \nu \leq k\} = \{\tau \leq k\} \cup \{\nu \leq k\}$,

$$A \cap \{\tau \wedge \nu \leq k\} = A \cap \big(\{\tau \leq k\} \cup \{\nu \leq k\}\big) = \big(A \cap \{\tau \leq k\}\big) \cup \big(A \cap \{\nu \leq k\}\big) \in \mathscr{F}_k.$$

v) Let $n \in \mathbb{N}$. It holds

$$\{\tau < \nu\} \cap \{\tau \leq n\} = \bigcup_{k=0}^{n} \{\tau = k\} \cap \{\nu > k\}.$$

But, for $0 \leq k \leq n$,

$$\{\tau = k\} = \{\tau \leq k\} \cap \{\tau \leq k-1\}^c \in \mathscr{F}_k \subset \mathscr{F}_n$$
$$\{\nu > k\} = \{\nu \leq k\}^c \in \mathscr{F}_k \subset \mathscr{F}_n.$$

Therefore, $\{\tau < \nu\} \cap \{\tau \leq n\} \in \mathscr{F}_n$ and $\{\tau < \nu\} \in \mathscr{F}_\tau$. In the same way,

$$\{\tau < \nu\} \cap \{\nu \leq n\} = \bigcup_{k=0}^{n} \{\nu = k\} \cap \{\tau < k\}$$

and as, for $0 \leq k \leq n$, $\{\nu = k\} \in \mathscr{F}_k \subset \mathscr{F}_n$ and $\{\tau < k\} = \{\tau \leq k-1\} \in \mathscr{F}_{k-1} \subset \mathscr{F}_n$, it also holds $\{\tau < \nu\} \cap \{\nu \leq n\} \in \mathscr{F}_n$ so that $\{\tau < \nu\} \in \mathscr{F}_\nu$. Finally,

$$\{\tau = \nu\} = \{\tau < \nu\}^c \cap \{\nu < \tau\}^c \in \mathscr{F}_\tau \cap \mathscr{F}_\nu.$$

E2.2 1) Let $\omega = (\omega_p)_{p\geq 0} \in \Omega$. By induction, $\theta_m(\omega) = (\omega_{p+m})_{p\geq 0}$. Therefore

$$X_n \circ \theta_m(\omega) = X_n((\omega_{p+m})_{p\geq 0}) = \omega_{n+m} = X_{n+m}(\omega).$$

It holds, for every $A \subset E$,

$$\theta_n^{-1}\{X_k \in A\} = \{X_k \circ \theta_n \in A\} = \{X_{k+n} \in A\}.$$

As $\mathscr{F}_n$ is generated by the events of the form

$$B = \{X_0 \in A_0\} \cap \{X_1 \in A_1\} \cap \ldots \cap \{X_m \in A_m\},$$

$\theta_n^{-1}(\mathscr{F}_m)$ is generated by the events

$$\begin{aligned}\theta_n^{-1}(B) = \theta_n^{-1}\{X_0 \in A_0\} \cap \theta_n^{-1}\{X_1 \in A_1\} \cap \ldots \cap \theta_n^{-1}\{X_m \in A_m\} =\\ = \{X_n \in A_0\} \cap \{X_{n+1} \in A_1\} \cap \ldots \cap \{X_{n+m} \in A_m\},\end{aligned}$$

which generate the σ-algebra $\sigma(X_n, \ldots, X_{n+m}) \subset \mathscr{F}_{n+m}$.

2a) One has

$$\{k + \tau \circ \theta_k = m\} = \{\tau \circ \theta_k = m - k\} = \theta_k^{-1}\{\tau = m - k\}.$$

But $\{\tau = m - k\} \in \mathscr{F}_{m-k}$ and, by 1), $\theta_k^{-1}\{\tau = m - k\} \in \mathscr{F}_m$.

2b) One has

$$\begin{aligned}\{\rho = m\} = \{\nu + \tau \circ \theta_\nu = m\} = \bigcup_{k=0}^{m} \{k + \tau \circ \theta_k = m, \nu = k\} =\\ = \bigcup_{k=0}^{m} \{k + \tau \circ \theta_k = m\} \cap \{\nu = k\}.\end{aligned}$$

The last event belongs to $\mathscr{F}_m$, thanks to (2a) and since ν is a stopping time.

One has $X_\tau \circ \theta_\nu(\omega) = X_{\tau(\omega)\circ\theta_\nu(\omega)}(\theta_\nu(\omega))$. Now, for every $k \geq 0$, $X_k(\theta_\nu(\omega)) = X_{k+\nu}(\omega)$ and

$$X_{\tau+\theta_\nu(\omega)}(\theta_{\nu(\omega)}(\omega)) = X_{\nu+\tau\circ\theta_\nu}(\omega).$$

2c) The relations are obviously true on $\{\sigma = +\infty\}$. On $\{\sigma = k\}$ it holds

$$\begin{aligned}\sigma + \tau_A \circ \theta_\sigma = k + \tau_A \circ \theta_k = k + \inf\{m, X_m \circ \theta_k \in A\} =\\ = k + \inf\{m, X_{m+k} \in A\}.\end{aligned} \tag{2.7}$$

But if $\inf\{m, X_{m+k} \in A\} = \ell$, this means that $X_k \notin A, \dots, X_{\ell-1} \notin A, X_\ell \in A$ and

$$\inf\{m, X_{m+k} \in A\} = \inf\{r \geq k, X_r \in A\} - k.$$

This, with (2.7), proves the relation we are looking for on $\{\sigma = k\}$ for every k. The proof for σ_A is identical.

If $A \subset B$, the relation $\tau_A = \tau_B + \tau_A \circ \theta_{\tau_B}$ is an immediate consequence of (2.6). Actually, as $X_k \notin A$ for every $k < \tau_B$,

$$\tau_A = \inf\{k, X_k \in A\} = \inf\{k \geq \tau_B, X_k \in A\} = \tau_B + \tau_A \circ \theta_{\tau_B}.$$

CHAPTER 3

Martingales

First Definitions

•3.1 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ be a fixed filtered probability space. We know that the family of sets $\bigcup_{n\geq 0} \mathscr{F}_n$ is not, in general, a σ-algebra. Let us define, thus, as in •2.5, $\mathscr{F}_\infty = \sigma(\bigcup_{n\geq 0} \mathscr{F}_n)$. $\mathscr{F}_\infty$ is a σ-algebra that is going to play an important role.

Definition 3.1 *A real adapted process $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale) if it is integrable and if, for every $n \in \mathbb{N}$,*

$$\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = X_n \quad \textit{a.s. (resp. } \leq, \geq).$$

Thus $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale) if and only if, for every $A \in \mathscr{F}_n$,

$$\int_A X_{n+1}\, d\mathbf{P} = \int_A X_n\, d\mathbf{P} \quad (\text{resp. } \leq, \geq).$$

Note that $(X_n)_{n\geq 0}$ is a supermartingale if and only if $(-X_n)_{n\geq 0}$ is a submartingale and $(X_n)_{n\geq 0}$ is a martingale if and only if $(X_n)_{n\geq 0}$ is both a supermartingale and a submartingale.

The definitions above still hold if "X_n integrable" is replaced by "X_n positive". We speak then of a positive martingale (resp. supermartingale, submartingale). It is understood that a martingale (resp. supermartingale, submartingale) without further specification is necessarily integrable.

Finally a process $X = (\Omega, \mathscr{F}, (X_n)_{n\geq 0}, \mathbf{P})$ (without specification of the filtration) is a martingale (resp. a supermartingale, a submartingale) if it is a martingale (resp. a supermartingale, a submartingale) with respect to the *natural filtration* $\mathscr{F}_n^0 = \sigma(X_0, \dots, X_n)$.

•3.2 Examples (i) Let $X \in L^1$ (resp. $X \geq 0$). Thanks to •1.5 (vi), $X_n = \mathbf{E}(X \mid \mathscr{F}_n)$ is a martingale (resp. a positive martingale, possibly nonintegrable).

(ii) Clearly a process $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale) if and only if $\mathbf{E}(X_{n+1} - X_n \mid \mathscr{F}_n) = 0$, (resp. $\leq, \geq$). Thus if $(Y_n)_{n\geq 0}$ is an adapted integrable process, $X_n = Y_0 + Y_1 + \ldots + Y_n$ defines a martingale (resp. : supermartingale, a submartingale) if and only if $\mathbf{E}(Y_{n+1} \mid \mathscr{F}_n) = 0$, (resp. $\leq 0, \geq 0$). In particular if $(Y_n)_{n\geq 1}$ is a sequence of independent integrable r.v.'s, then $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale) with respect to the filtration $\sigma(Y_0, Y_1, \dots, Y_n)$ if, for every $n \geq 1$, $\mathbf{E}(Y_n) = 0$ (resp. $\leq 0, \geq 0$).

First Properties

•**3.3** (i) If $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale), the sequence $(\mathrm{E}(X_n))_{n\geq 0}$ is constant (resp. decreasing, increasing). This follows immediately from •1.5 (i).

(ii) If $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale), for $m < n$ it holds $\mathrm{E}(X_n \mid \mathscr{F}_m) = X_m$ a.s. (resp. $\leq$, $\geq$). This follows from •1.5 (vi).

(iii) If $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are supermartingales, the same is true for $X_n + Y_n$ and $X_n \wedge Y_n$.

(iv) If $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale) and f is a real concave (resp. concave and increasing) function such that $f(X_n)$ is integrable for every $n \geq 0$, then $Y_n = f(X_n)$ is a supermartingale (this follows from Jensen's inequality •1.5 (v)). If we replace the assumption "$f(X_n)$ integrable" with "$f(X_n)$ positive", $(f(X_n))_n$ is a positive supermartingale.

In particular if $(X_n)_{n\geq 0}$ is a martingale, $(|X_n|)_{n\geq 0}$ and $(X_n^2)_{n\geq 0}$ are positive submartingales.

(v) If $(X_n)_{n\geq 0}$ is a martingale (resp. a supermartingale, a submartingale), this also holds for the stopped process $X_n^\nu = X_{n\wedge\nu}$, where ν is a stopping time of the filtration $(\mathscr{F}_n)_{n\geq 0}$. Actually, by the definition of a stopping time, $\{\nu \geq n+1\} = \{\nu \leq n\}^c \in \mathscr{F}_n$; hence

$$\begin{aligned}\mathrm{E}(X_{n+1}^\nu - X_n^\nu \mid \mathscr{F}_n) &= \mathrm{E}((X_{n+1} - X_n)1_{\{\nu\geq n+1\}} \mid \mathscr{F}_n) = \\ &= 1_{\{\nu\geq n+1\}}\mathrm{E}(X_{n+1} - X_n \mid \mathscr{F}_n) = 0.\end{aligned} \tag{3.1}$$

This result holds also for positive martingales, with slight changes in the proof.

•**3.4** (Doob's decomposition). An adapted process $(A_n)_{n\geq 0}$ is an *increasing predictable process* if $A_0 = 0$, $A_n \leq A_{n+1}$ for every $n \geq 0$ and A_{n+1} is $\mathscr{F}_n$-measurable. Let $(X_n)_{n\geq 0}$ be a submartingale. One defines

$$A_0 = 0, \quad A_{n+1} = A_n + \mathrm{E}(X_{n+1} - X_n \mid \mathscr{F}_n).$$

By construction, $(A_n)_{n\geq 0}$ is an integrable predictable increasing process and $M_n = X_n - A_n$ verifies $\mathrm{E}(M_{n+1} - M_n \mid \mathscr{F}_n) = 0$ (since A_{n+1} is $\mathscr{F}_n$-measurable!). This decomposition is moreover unique since, if $X_n = M_n' + A_n'$ is any such decomposition, then $A_0' = 0$ and

$$A_{n+1}' - A_n' = X_{n+1} - X_n - (M_{n+1}' - M_n')$$

from which, by conditioning with respect to $\mathscr{F}_n$, $A_{n+1}' - A_n' = \mathrm{E}(X_{n+1} - X_n \mid \mathscr{F}_n)$; thus $A_n' = A_n$ and $M_n' = M_n$. We have thus proved (recall that, by definition, $A_0 = 0$):

Theorem 3.2 *Every submartingale $(X_n)_{n\geq 0}$ can be written, uniquely, in the form $X_n = M_n + A_n$, where $(M_n)_{n\geq 0}$ is a martingale and $(A_n)_{n\geq 0}$ is an integrable increasing predictable process.*

Such an increasing process $(A_n)_{n\geq 0}$ is called the *compensator* of the submartingale $(X_n)_{n\geq 0}$.

The Stopping Theorem

•**3.5** Let $(X_n)_{n\geq 0}$ be a supermartingale. Thus, for $m < n$, $\mathrm{E}(X_n \mid \mathscr{F}_m) \leq X_m$ a.s. Is this property still true if one replaces m and n by stopping times? That is: if ν_1 and

ν_2 are two stopping times with $\nu_1 \le \nu_2$, does it hold $\mathbf{E}(X_{\nu_2} \mid \mathscr{F}_{\nu_1}) \le X_{\nu_1}$, i.e.,

$$\int_A X_{\nu_1}\, d\mathbf{P} \ge \int_A X_{\nu_2}\, d\mathbf{P} \quad \text{for every } A \in \mathscr{F}_{\nu_1}? \tag{3.2}$$

Let us assume $\nu_1 \le \nu_2 \le k \in \mathbb{N}$. Since $A \cap \{\nu_1 = j\} \in \mathscr{F}_j$ and X^{ν_2} is a supermartingale (•3.3 (v)),

$$\int_A X_{\nu_2}\, d\mathbf{P} = \sum_{j=0}^{k} \int_{A\cap\{\nu_1=j\}} X_{\nu_2\wedge k}\, d\mathbf{P} \le \sum_{j=0}^{k} \int_{A\cap\{\nu_1=j\}} X_{\nu_2\wedge j}\, d\mathbf{P}.$$

But $X_{\nu_2\wedge j} = j$ on $\{\nu_1 = j\}$ so that

$$\int_A X_{\nu_2}\, d\mathbf{P} \le \sum_{j=0}^{k} \int_{A\cap\{\nu_1=j\}} X_j\, d\mathbf{P} = \int_A X_{\nu_1}\, d\mathbf{P}.$$

We have proved

Theorem 3.3 *Let $(X_n)_{n\ge 0}$ be a positive or integrable martingale (resp. supermartingale, submartingale) and ν_1, ν_2 two bounded stopping times such that $\nu_1 \le \nu_2$; then $\mathbf{E}(X_{\nu_2} \mid \mathscr{F}_{\nu_1}) = X_{\nu_1}$ (resp. $\le$, $\ge$).*

By taking the expectations,

Corollary 3.4 *Let $(X_n)_{n\ge 0}$ be a positive or integrable martingale (resp. supermartingale, submartingale) and ν a bounded stopping time; then $\mathbf{E}(X_\nu) = \mathbf{E}(X_0)$ (resp $\le$, $\ge$).*

•3.6 In these statements, the assumption of boundedness for ν is essential and cannot be removed without additional hypotheses. Let us consider, for instance, the random walk $S_0 = 0$, $S_n = X_1 + X_2 + \ldots + X_n$ where $(X_n)_{n\ge 0}$ is a sequence of independent r.v.'s such that $\mathbf{P}(X_k = 1) = \mathbf{P}(X_k = -1) = \frac{1}{2}$ for every $k = 1, 2, \ldots$. We shall see (Exercise 3.13) that $\nu = \inf\{n;\ S_n = 1\}$ satisfies $\mathbf{P}(\nu < +\infty) = 1$. Besides $(S_n)_{n\ge 0}$ is a martingale (•3.2). Then $\mathbf{E}(S_0) = \mathbf{E}(S_n) = 0$ but also $S_\nu = 1$ a.s.: thus $\mathbf{E}(S_\nu) = 1$. The statement of Corollary 3.4 does not hold and obviously ν is not bounded; actually one can see easily that $\mathbf{P}(\nu > n) \ge 2^{-n} > 0$.

Maximal Inequalities

•3.7 These are inequalities concerning the r.v.'s

$$\max_{0\le k\le n} X_k, \quad \max_{0\le k\le n} |X_k|, \quad \sup_{n\ge 0} X_n, \quad \sup_{n\ge 0} |X_n|.$$

Let us start with two simple but typical examples.

Let $(X_n)_{n\ge 0}$ be a positive supermartingale and denote, for $a > 0$, by $\nu = \inf\{n \ge 0: X_n > a\}$ the passage time in $]a, +\infty[$. Then $\{\sup_{k\ge 0} X_k > a\} = \{\nu < +\infty\}$ and $X_\nu > a$ on $\{\nu < +\infty\}$; thus $X_{n\wedge\nu} \ge a 1_{\{\nu\le n\}}$. We deduce that $\mathbf{E}(X_0) \ge \mathbf{E}(X_{n\wedge\nu}) \ge a\mathbf{P}(\nu \le n)$ and, letting n go to ∞, $\mathbf{E}(X_0) \ge a\mathbf{P}(\nu < +\infty) = a\mathbf{P}(\sup_{k\ge 0} X_k > a)$.

Similarly, if $(X_n)_{n\ge 0}$ is a positive submartingale, one has (with the same stopping time ν) $\{\max_{0\le k\le n} X_k > a\} = \{\nu \le n\}$. For $k \le n$,

$$\int_{\{\nu=k\}} X_k \, d\mathbf{P} \le \int_{\{\nu=k\}} X_n \, d\mathbf{P},$$

which gives

$$\begin{aligned} a\mathbf{P}(\nu \le n) &\le \int_{\{\nu\le n\}} X_\nu \, d\mathbf{P} = \sum_{k=0}^{n} \int_{\{\nu=k\}} X_k \, d\mathbf{P} \le \\ &\le \sum_{k=0}^{n} \int_{\{\nu=k\}} X_n \, d\mathbf{P} = \int_{\{\nu\le n\}} X_n \, d\mathbf{P} \le \mathbf{E}(X_n). \end{aligned} \tag{3.3}$$

We thus proved

Theorem 3.5 *For $a > 0$ it holds*
(i) if $(X_n)_{n\ge 0}$ is a positive supermartingale, $\mathbf{P}(\sup_{k\ge 0} X_k > a) \le \frac{1}{a}\,\mathbf{E}(X_0)$;
(ii) if $(X_n)_{n\ge 0}$ is a positive submartingale,

$$\mathbf{P}\Big(\max_{0\le k\le n} X_k > a\Big) \le \frac{1}{a} \int_{\{\max_{0\le k\le n} X_k > a\}} X_n \, d\mathbf{P} \le \frac{1}{a}\,\mathbf{E}(X_n).$$

Starting from this theorem one can prove

Theorem 3.6 (Doob's inequality) *Let $p > 1$ and $(X_n)_{n\ge 0}$ a martingale (or a positive submartingale) such that* $\sup_{n\ge 0} \mathbf{E}|X_n|^p < +\infty$. *Then the r.v.* $\sup_{n\ge 0} |X_n|^p$ *is integrable and*

$$\Big\| \sup_{n\ge 0} |X_n| \Big\|_p \le \frac{p}{p-1} \sup_{n\ge 0} \|X_n\|_p .$$

Note that, if $p = 2$, the inequality becomes $\mathbf{E}[\sup_{n\ge 0} X_n^2] \le 4 \sup_{n\ge 0} \mathbf{E}(X_n^2)$.

The proof of Doob's inequality is actually simple. It is enough to prove the statement for a positive submartingale, as, if $(X_n)_{n\ge 0}$ is a martingale, $(|X_n|)_{n\ge 0}$ is a positive submartingale.

Note first that $\|X_n\|_p$ is increasing in n. Let $Y_m = \max_{k\le m} X_k$. By Theorem 3.5 ii),

$$a\mathbf{E}(1_{\{Y_m > a\}}) = a\mathbf{P}(Y_m > a) \le \mathbf{E}(X_m 1_{\{Y_m > a\}})$$

and, by multiplying by pa^{p-2} and integrating,

$$\int_0^{+\infty} pa^{p-1}\mathbf{E}(1_{\{Y_m > a\}}) \, da \le \int_0^{+\infty} pa^{p-2}\mathbf{E}(X_m 1_{\{Y_m > a\}}) \, da. \tag{3.4}$$

Since everything is positive, on one side one has

$$\int_0^{+\infty} pa^{p-1}\mathbf{E}(1_{\{Y_m > a\}}) \, da = \mathbf{E}\Big(\int_0^{+\infty} pa^{p-1} 1_{\{Y_m > a\}} \, da\Big) = \mathbf{E}(Y_m^p) = \|Y_m\|_p^p$$

and on the other one

$$\begin{aligned} \int_0^{+\infty} pa^{p-2}\mathbf{E}(X_m 1_{\{Y_m > a\}}) \, da &= \mathbf{E}\Big(X_m \int_0^{+\infty} pa^{p-2} 1_{\{Y_m > a\}} \, da\Big) = \\ &= \frac{p}{p-1}\,\mathbf{E}(X_m Y_m^{p-1}). \end{aligned}$$

Moreover, by Hölder's inequality for p and $q = \frac{p}{p-1}$,

$$\mathbf{E}(X_m Y_m^{p-1}) \le \|X_m\|_p \|Y_m^{p-1}\|_{\frac{p}{p-1}} = \|X_m\|_p \|Y_m\|_p^{p-1}$$

and, plugging this into (3.4), we get $\|Y_m\|_p^p \le \frac{p}{p-1}\|X_m\|_p \|Y_m\|_p^{p-1}$; dividing by $\|Y_m\|_p^{p-1}$ (which is finite) one gets $\|Y_m\|_p \le \frac{p}{p-1}\|X_m\|_p$. Finally just note that, as $m \to \infty$, $Y_m = \max_{k \le m} X_k \uparrow \sup_n X_n$.

If $\sup_{n\ge 0} \mathbf{E}|X_n|^p < +\infty$, $(X_n)_{n\ge 0}$ is said to be *bounded in* L^p.

Square Integrable Martingales

•**3.8** Let $(M_n)_{n\ge 0}$ be a martingale such that $\mathbf{E}(M_n^2) < +\infty$ for every $n \ge 0$. Note that $(M_n^2)_{n\ge 0}$ is a submartingale; hence the sequence $(\mathbf{E}(M_n^2))_{n\ge 0}$ is increasing. Also, for $p \ge 1$, it holds $\mathbf{E}^{\mathcal{F}_n}(M_n M_{n+p}) = M_n \mathbf{E}^{\mathcal{F}_n}(M_{n+p}) = M_n^2$ a.s., from which we derive the formulae

$$\mathbf{E}^{\mathcal{F}_n}((M_{n+p} - M_n)^2) = \mathbf{E}^{\mathcal{F}_n}(M_{n+p}^2 - M_n^2) \qquad a.s.$$
$$\mathbf{E}((M_{n+p} - M_n)^2) = \mathbf{E}(M_{n+p}^2 - M_n^2).$$

Usually one writes $(\langle M\rangle_n)_{n\ge 0}$ for the compensator $(A_n)_{n\ge 0}$ of the submartingale $(M_n^2)_{n\ge 0}$ (•3.4). This increasing process is characterized by $\langle M\rangle_0 = 0$, $\langle M\rangle_{n+1}$ is $\mathcal{F}_n$-measurable and

$$\mathbf{E}^{\mathcal{F}_n}((M_{n+p} - M_n)^2) = \mathbf{E}^{\mathcal{F}_n}(M_{n+p}^2 - M_n^2) = \langle M\rangle_{n+p} - \langle M\rangle_n \qquad \text{a.s.}$$

Then $U_n = M_n^2 - \langle M\rangle_n$ is a martingale and, in particular, if $M_0 = 0$, $\mathbf{E}(M_n^2) = \mathbf{E}(\langle M\rangle_n)$. The compensator $(\langle M\rangle_n)_{n\ge 0}$ is also called the *associated increasing process* of the square integrable martingale $(M_n)_{n\ge 0}$.

•**3.9** Let us assume now that the martingale $(M_n)_{n\ge 0}$ is bounded in L^2, and set $m^* = \sup_{n\ge 0} \mathbf{E}(M_n^2) < +\infty$. Actually, since $(\mathbf{E}(M_n^2))_{n\ge 0}$ is increasing, $\mathbf{E}(M_n^2) \uparrow m^* < +\infty$. Therefore $\mathbf{E}((M_{n+p} - M_n)^2) \le m^* - \mathbf{E}(M_n^2)$, which gives

$$\sup_{p\ge 0} \mathbf{E}(M_{n+p} - M_n)^2 \underset{n\to\infty}{\to} 0.$$

This shows that $(M_n)_{n\ge 0}$ converges in L^2, being a Cauchy sequence. Let us show that $(M_n)_{n\ge 0}$ converges a.s. Let us note $V_n = \sup_{i,j\ge n} |M_i - M_j|$; obviously $(V_n)_{n\ge 0}$ is decreasing and $V_n \downarrow V$. It suffices now to show that $V = 0$ a.s. since, then, $(M_n)_{n\ge 0}$ would be a.s. a Cauchy sequence and thus a.s. convergent. Applying Doob's inequality (Theorem 3.6) to the martingale $(M_i - M_n)_{i\ge n}$, one gets, for every $\rho > 0$,

$$\mathbf{P}(V_n > \rho) = \mathbf{P}\Big(\sup_{i,j\ge n} |M_i - M_j| > \rho\Big) \le \mathbf{P}\Big(\sup_{i\ge n} |M_i - M_n| > \frac{\rho}{2}\Big) \le$$
$$\le \frac{4}{\rho^2} \sup_{i\ge n} \mathbf{E}((M_i - M_n)^2) \underset{n\to\infty}{\to} 0.$$

This implies $\mathbf{P}(V > 0) = 0$ and

Theorem 3.7 *Let $(M_n)_{n\geq 0}$ be a martingale such that* $\sup_{n\geq 0} \mathbf{E}(M_n^2) < +\infty$. *Then $(M_n)_{n\geq 0}$ converges a.s. and in L^2.*

Convergence Theorems

•**3.10** One of the reasons for the importance of martingales is the convergence theorems; they guarantee that a martingale converges a.s., under a set of assumptions that are often easy to check. Already Theorem 3.7 states that a bounded martingale converges a.s.

Let $(X_n)_{n\geq 0}$ be a submartingale bounded in absolute value by some number K and let $X_n = M_n + A_n$ be its Doob decomposition (Theorem 3.2). Then $\mathbf{E}(A_n) = \mathbf{E}(X_n) - \mathbf{E}(M_0) \leq 2K$ so that $\mathbf{E}(A_\infty) \leq 2K$, which gives $A_\infty < +\infty$ a.s. We set $\tau_p = \inf\{n;\, A_{n+1} > p\}$; τ_p is a stopping time and $|M_{\tau_p\wedge n}| \leq |X_{\tau_p\wedge n} - A_{\tau_p\wedge n}| \leq K + p$, so that M^{τ_p} is a bounded martingale and converges a.s. by Theorem 3.7. We get then the a.s. convergence of $(X_n^{\tau_p})_{n\geq 0}$ and therefore that $(X_n)_{n\geq 0}$ converges on $\{\tau_p = +\infty\}$. Finally $(X_n)_{n\geq 0}$ converges a.s. since, A_∞ being a.s. finite, $\bigcup_{p=1}^\infty \{\tau_p = +\infty\} = \bigcup_{p=1}^\infty \{A_\infty < p\} = \{A_\infty < +\infty\} = \Omega$ a.s.

If $(X_n)_{n\geq 0}$ is a positive supermartingale, since the exponential function is convex and increasing, $Y_n = e^{-X_n}$ is a submartingale (•3.3) satisfying $0 \leq Y_n \leq 1$. Thus $(Y_n)_{n\geq 0}$ and $(X_n)_{n\geq 0}$ (with values in $\bar{\mathbb{R}}^+$) both converge a.s. Moreover by Fatou's lemma (•1.5)

$$\mathbf{E}(X_\infty \mid \mathscr{F}_k) = \mathbf{E}\Big(\varliminf_{n\to\infty} X_n \mid \mathscr{F}_k\Big) \leq \varliminf_{n\to\infty} \mathbf{E}(X_n \mid \mathscr{F}_k) \leq X_k \qquad \text{a.s.}$$

This implies that, for every $M \geq 0$, $\mathbf{E}(X_\infty 1_{\{X_0\leq M\}}) = \mathbf{E}(X_0 1_{\{X_0\leq M\}})$, so that $X_\infty < +\infty$ on $\{X_0 < +\infty\}$. We have proved

Theorem 3.8 *Let $(X_n)_{n\geq 0}$ be a positive supermartingale. Then it converges a.s. to a r.v. X_∞ and $X_n \geq \mathbf{E}(X_\infty \mid \mathscr{F}_n)$. Moreover, if $\mathbf{P}(X_0 < +\infty) = 1$, then $\mathbf{P}(X_\infty < +\infty) = 1$.*

Of course, a positive martingale is a positive supermartingale and the previous theorem gives the convergence also in this case.

Theorem 3.9 *Let $(X_n)_{n\geq 0}$ be a submartingale such that* $\sup_{n\geq 0} \mathbf{E}(|X_n|) < +\infty$. *Then $(X_n)_{n\geq 0}$ converges a.s.*

The proof simply amounts to showing that $(X_n)_{n\geq 0}$ can be written as the difference of a positive integrable martingale and a positive integrable supermartingale and then applying Theorem 3.8. Let $X_n^+ = M_n + A_n$ be the Doob decomposition of the submartingale $(X_n^+)_{n\geq 0}$. Then $\mathbf{E}(A_\infty) = \sup_{n\geq 0} \mathbf{E}(A_n) \leq \mathbf{E}|M_0| + \sup_{n\geq 0} \mathbf{E}|X_n| < +\infty$. If $Y_n = M_n + \mathbf{E}(A_\infty \mid \mathscr{F}_n)$, since $Y_n \geq X_n^+ \geq X_n$, $(Y_n)_{n\geq 0}$ is a positive martingale and $Z_n = Y_n - X_n$ is a positive supermartingale. Obviously $X_n = Y_n - Z_n$. Note that Theorem 3.8 states that $\mathbf{E}(Y_\infty) < +\infty$ and $\mathbf{E}(Z_\infty) < +\infty$, so that $Y_\infty < +\infty$ and $Z_\infty < +\infty$ a.s.

Let us point out that, if $(X_n)_{n\geq 0}$ is a submartingale, since $|x| = 2x^+ - x$, it holds $\mathbf{E}|X_n| = 2\mathbf{E}(X_n^+) - \mathbf{E}(X_n) \leq 2\mathbf{E}(X_n^+) - \mathbf{E}(X_0)$. Thus $\sup_{n\geq 0} \mathbf{E}|X_n| < +\infty$ if and only if $\sup_{n\geq 0}\mathbf{E}(X_n^+) < +\infty$. Similarly, for a (integrable) supermartingale, $\sup_{n\geq 0}\mathbf{E}|X_n| < +\infty$ if and only if $\sup_{n\geq 0}\mathbf{E}(X_n^-) < +\infty$.

•**3.11** We have already seen that a martingale that is bounded in L^2 converges a.s. and in L^2. More generally a martingale that is bounded in L^p, $p > 1$, converges a.s. and in L^p. Actually it is even bounded in L^1 and thus, thanks to Theorem 3.9, there exists a r.v. X_∞ such that $X_n \to_{n\to\infty} X_\infty$ a.s. Moreover, by Doob's inequality, the r.v. $X^* = \sup_{n\geq 0}|X_n|$ belongs to L^p. Thanks to the inequality $|X_n - X_\infty|^p \leq 2^p X^{*p}$, we are allowed to apply Lebesgue's theorem and get

$$\lim_{n\to\infty}\mathbf{E}(|X_n - X_\infty|^p) = 0.$$

Contrary to the case L^p, $p > 1$, the condition $\sup_{n\geq 0}\mathbf{E}|X_n| < +\infty$ does not imply the convergence in L^1, as we shall see in the exercises (Exercise 3.5, for instance).

•**3.12** Let us go back to our first example of a martingale, i.e., $X_n = \mathbf{E}^{\mathscr{F}_n}(X)$, $X \in L^1$ (see •3.2). This martingale is bounded in L^1, since

$$\mathbf{E}(|X_n|) = \mathbf{E}(|\mathbf{E}^{\mathscr{F}_n}(X)|) \leq \mathbf{E}(\mathbf{E}^{\mathscr{F}_n}(|X|)) = \mathbf{E}(|X|).$$

Therefore (Theorem 3.9) $X_n \to_{n\to\infty} X_\infty$ a.s. Let us study the convergence in L^1. Thanks to the decomposition $X = X^+ - X^-$, we can assume $X \geq 0$. Let $a > 0$. The Lebesgue theorem gives $\|X - X\wedge a\|_1 \to 0$ as $a \to +\infty$. We just proved that $\mathbf{E}(X\wedge a \mid \mathscr{F}_n)$ converges a.s. but, being bounded, $\mathbf{E}(X\wedge a \mid \mathscr{F}_n)$ also converges in L^1. It suffices then to write

$$\begin{aligned}
&\|\mathbf{E}(X \mid \mathscr{F}_n) - \mathbf{E}(X \mid \mathscr{F}_m)\|_1 \leq \\
&\leq \|\mathbf{E}(X \mid \mathscr{F}_n) - \mathbf{E}(X\wedge a \mid \mathscr{F}_n)\|_1 + \|\mathbf{E}(X\wedge a \mid \mathscr{F}_n) - \mathbf{E}(X\wedge a \mid \mathscr{F}_m)\|_1 + \\
&\qquad + \|\mathbf{E}(X\wedge a \mid \mathscr{F}_m) - \mathbf{E}(X \mid \mathscr{F}_m)\|_1 \leq \\
&\leq 2\|X - X\wedge a\|_1 + \|\mathbf{E}(X\wedge a \mid \mathscr{F}_n) - \mathbf{E}(X\wedge a \mid \mathscr{F}_m)\|_1
\end{aligned}$$

in order to derive that $(\mathbf{E}^{\mathscr{F}_n}(X))_{n\geq 0}$ is a Cauchy sequence in L^1 and thus that it converges in L^1. Let us identify its limit X_∞. Clearly X_∞ is $\mathscr{F}_\infty$-measurable. Moreover, by the definition of X_n,

$$\int_A X_n\,d\mathbf{P} = \int_A X\,d\mathbf{P}, \qquad \text{for every } A \in \mathscr{F}_n$$

from which, since $(X_n)_{n\geq 0}$ converges in L^1,

$$\int_A X_\infty\,d\mathbf{P} = \int_A X\,d\mathbf{P}, \qquad \text{for every } A \in \mathscr{F}_n. \tag{3.5}$$

It is easy to see, by a monotone class argument (•2.4), that (3.5) holds for every $A \in \mathscr{F}_\infty = \sigma(\bigcup_{n\geq 0}\mathscr{F}_n)$. This shows that $X_\infty = \mathbf{E}(X \mid \mathscr{F}_\infty)$. Summarizing,

Theorem 3.10 *Assume $X \in L^1$. The martingale $X_n = \mathbf{E}(X \mid \mathscr{F}_n)$ converges a.s. and in L^1 to $\mathbf{E}(X \mid \mathscr{F}_\infty)$.*

•**3.13** We just mention the following

Theorem 3.11 *Let $X \in L^1$ and $(\mathscr{G}_n)_{n\geq 0}$ be a decreasing sequence of σ-algebras with $\mathscr{G}_\infty = \bigcap_{n\geq 0}\mathscr{G}_n$. The sequence $X_n = \mathbf{E}(X \mid \mathscr{G}_n)$ then converges a.s. and in L^1 to $\mathbf{E}(X \mid \mathscr{G}_\infty)$.*

The sequence $(X_n)_{n\geq 0}$ of Theorem 3.11 is called an *inverse martingale*.

Regular Martingales

•3.14 A martingale of the form $X_n = \mathbf{E}(X \mid \mathscr{F}_n)$, $X \in L^1$, is said to be *regular*. We have just proved that such a martingale converges a.s. and in L^1 to $X_\infty = \mathbf{E}(X \mid \mathscr{F}_\infty)$. Conversely, let $(X_n)_{n\geq 0}$ be a martingale converging in L^1 to a r.v. X. Since $|X_n| \leq |X| + |X - X_n|$, $\sup_{n\geq 0} \mathbf{E}(|X_n|) < +\infty$. Thus (Theorem 3.9), $(X_n)_{n\geq 0}$ converges a.s. to X. We then have, for every $p \geq 0$ and for every $A \in \mathscr{F}_n$,

$$\int_A X_{n+p}\, d\mathbf{P} = \int_A X_n\, d\mathbf{P}$$

and, letting p go to $+\infty$, since X_{n+p} converges in L^1,

$$\int_A X\, d\mathbf{P} = \int_A X_n\, d\mathbf{P}, \qquad \text{for every } A \in \mathscr{F}_n$$

or equivalently, $X_n = \mathbf{E}(X \mid \mathscr{F}_n)$. In conclusion,

Theorem 3.12 *A martingale is regular if and only if it converges in L^1.*

•3.15 Let $(X_n)_{n\geq 0}$ be a regular martingale having the r.v. X_∞ as its limit. Then it is *closed* meaning that $(X_n)_{n\in\bar{\mathbb{N}}}$ is a martingale, i.e., verifies the definition for every $m \leq n \leq +\infty$. If ν is a stopping time with values $\bar{\mathbb{N}}$, then X_ν is defined by $X_\nu = X_\infty$ on $\{\nu = +\infty\}$. In this case, the stopping theorem takes an easy form:

Proposition 3.13 *Let $X_n = \mathbf{E}(X \mid \mathscr{F}_n)$, $X \in L^1$, be a regular martingale. Then,*
(i) if ν is a stopping time, X_ν is integrable and $X_\nu = \mathbf{E}(X \mid \mathscr{F}_\nu)$;
(ii) if $\nu_1 \leq \nu_2$ are stopping times, $\mathbf{E}(X_{\nu_2} \mid \mathscr{F}_{\nu_1}) = X_{\nu_1}$.

Exercises

Exercise 3.1 a) Let $X = (X_n)_{n\geq 0}$ be a supermartingale such that $\mathbf{E}(X_n)$ is constant. Show that $(X_n)_{n\geq 0}$ is a martingale.

b) Let $(X_n)_{n\geq 0}$ be an integrable process adapted to the filtration $(\mathscr{F}_n)_{n\geq 0}$. Show that $(X_n)_{n\geq 0}$ is a $(\mathscr{F}_n)_{n\geq 0}$-martingale if and only if $\mathbf{E}(X_\tau) = \mathbf{E}(X_0)$, for every bounded stopping time τ of the filtration $(\mathscr{F}_n)_{n\geq 0}$.

Exercise 3.2 Let $(X_n)_{n\geq 0}$, $(Y_n)_{n\geq 0}$ be two supermartingales (resp. martingales) on the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$. Let τ be a stopping time such that $X_\tau \leq Y_\tau$ (resp. $X_\tau = Y_\tau$) on the event $\{\tau < +\infty\}$. Let us define $Z_n = Y_n$ on $\{n < \tau\}$ and $Z_n = X_n$ on $\{n \geq \tau\}$. Prove that $(Z_n)_{n\geq 0}$ is a supermartingale (resp. a martingale) with respect to $(\mathscr{F}_n)_{n\geq 0}$.

Exercise 3.3 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ be a filtered probability space on which we consider two square integrable martingales $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$.

a) Show that $\mathbf{E}(X_m Y_n \mid \mathscr{F}_m) = X_m Y_m$ a.s. for every $m \leq n$.

b) Show that $\mathbf{E}(X_n Y_n) - \mathbf{E}(X_0 Y_0) = \sum_{k=1}^n \mathbf{E}((X_k - X_{k-1})(Y_k - Y_{k-1}))$.

c) Show that $\mathrm{Var}(X_n) = \mathrm{Var}(X_0) + \sum_{k=1}^n \mathrm{Var}(X_k - X_{k-1})$.

d) Show that the r.v.'s X_0, $X_k - X_{k-1}$, $k \geq 1$ are pairwise orthogonal.

Exercise 3.4 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ be a filtered probability space on which we consider a bounded martingale $(M_n)_{n\geq 0}$ ($|M_n| \leq K < +\infty$ for every $n \geq 0$). Let us define

$$X_n = \sum_{k=1}^{n} \frac{1}{k}(M_k - M_{k-1}).$$

Show that $(X_n)_{n\geq 1}$ is a $(\mathscr{F}_n)_{n\geq 1}$-martingale converging a.s. and in L^2.

Exercise 3.5 Let $(Y_n)_{n\geq 1}$ be a sequence of independent r.v's. with the same normal law $N(0, \sigma^2)$, with $\sigma > 0$. We set $\mathscr{F}_n = \sigma(Y_1, \ldots, Y_n)$ and $X_n = Y_1 + \ldots + Y_n$. Recall that

$$\mathbf{E}(e^{uY_1}) = e^{\frac{1}{2}u^2\sigma^2}. \tag{3.6}$$

1) Let $Z_n^u = \exp(uX_n - \frac{1}{2}nu^2\sigma^2)$. Show that, for every $u \in \mathbb{R}$, $(Z_n^u)_{n\geq 1}$ is a $(\mathscr{F}_n)_{n\geq 1}$-martingale.

2) Show that, for every $u \in \mathbb{R}$, $(Z_n^u)_{n\geq 1}$ converges a.s. to a finite r.v. Z_∞^u. Determine Z_∞^u. For what values of $u \in \mathbb{R}$ is $(Z_n^u)_{n\geq 1}$ a regular martingale?

Exercise 3.6 A process $(M_n)_{n\geq 0}$ is said to be *with independent increments* if, for every n, the r.v. $M_{n+1} - M_n$ is independent of the σ-algebra $\mathscr{F}_n = \sigma(M_0, \ldots, M_n)$.

1) Let $(M_n)_{n\geq 0}$ be a square integrable martingale with independent increments. We set $\sigma_0^2 = \mathrm{Var}(M_0)$ and, for $n \geq 1$, $\sigma_k^2 = \mathrm{Var}(M_k - M_{k-1})$.

1a) Show that $\mathrm{Var}(M_n) = \sum_{k=0}^{n} \sigma_k^2$.

1b) Let $(\langle M \rangle_n)_{n\geq 0}$ be the increasing process associated to $(M_n)_{n\geq 0}$ (see •3.8). Compute $\langle M \rangle_n$.

2) Let $(M_n)_{n\geq 0}$ be a Gaussian martingale (we recall that a process $(M_n)_{n\geq 0}$ is Gaussian if, for every n, the vector $(M_0, \ldots, M_n)$ is Gaussian).

2a) Show that $(M_n)_{n\geq 0}$ has independent increments.

2b) Show that, for every fixed $\theta \in \mathbb{R}$, the process

$$Z_n^\theta = e^{\theta M_n - \frac{1}{2}\theta^2 \langle M \rangle_n} \tag{3.7}$$

is a martingale. Does it converge a.s.?

Exercise 3.7 Let $(X_n)_{n\geq 0}$ be a sequence of r.v.'s with values in $[0, 1]$. We set $\mathscr{F}_n = \sigma(X_0, \ldots, X_n)$. Assume that $X_0 = a \in [0, 1]$ and

$$\mathbf{P}\Big(X_{n+1} = \frac{X_n}{2} \mid \mathscr{F}_n\Big) = 1 - X_n, \qquad \mathbf{P}\Big(X_{n+1} = \frac{1 + X_n}{2} \mid \mathscr{F}_n\Big) = X_n.$$

1) Show that $(X_n)_{n\geq 0}$ is a martingale converging a.s. and in L^2 to a r.v. Z.

2) Show that $\mathbf{E}((X_{n+1} - X_n)^2) = \frac{1}{4}\mathbf{E}(X_n(1 - X_n))$.

3) Compute $\mathbf{E}(Z(1 - Z))$. What is the law of Z?

Exercise 3.8 At time 1 an urn contains a white ball and a red ball. A ball is drawn and replaced with two balls having the same colour as the one that is drawn. This gives the new composition of the urn at time 2 and so on following the same procedure.

We denote by Y_n and $X_n = Y_n/(n + 1)$ the number and the proportion of white balls in the urn at time n, respectively. We set $\mathscr{F}_n = \sigma(Y_1, \ldots, Y_n)$.

1) Show that $(X_n)_{n\geq 1}$ is a martingale that converges a.s. to a r.v. U and that, for every $k \geq 1$, $\lim_{n\to\infty} \mathbf{E}(X_n^k) = \mathbf{E}(U^k)$.

2) Let $k \geq 1$ be fixed. We set, for $n \geq 1$,

$$Z_n = \frac{Y_n(Y_n+1)\dots(Y_n+k-1)}{(n+1)(n+2)\dots(n+k)}.$$

Show that $(Z_n)_{n\geq 1}$ is a $(\mathscr{F}_n)_{n\geq 0}$-martingale. What is its limit? Derive from this the value of $\mathrm{E}(U^k)$.

3) Let X be a real r.v. such that $|X| \leq M < +\infty$ a.s. Show that its characteristic function is given by the expansion in power series

$$\phi(t) = \sum_{k=0}^{\infty} \frac{\phi^{(k)}(0)}{k!} t^k \tag{3.8}$$

for every $t \in \mathbb{R}$.

4) What is the law of U?

Exercise 3.9 (A proof of the Kolmogorov 0-1 law by martingales) Let $(Y_n)_{n\geq 1}$ be a sequence of independent r.v.'s. Let us define

$$\mathscr{F}_n = \sigma(Y_1, \dots, Y_n), \qquad \mathscr{F}_\infty = \sigma\Big(\bigcup_{n\geq 1} \mathscr{F}_n\Big),$$

$$\mathscr{F}^n = \sigma(Y_n, Y_{n+1}, \dots), \qquad \mathscr{F}^\infty = \bigcap_{n\geq 1} \mathscr{F}^n.$$

1) Let $A \in \mathscr{F}^\infty$. Show that either $\mathrm{P}(A) = 0$ or $\mathrm{P}(A) = 1$. (Hint: $Z_n = \mathrm{E}^{\mathscr{F}_n}(1_A)$ is a martingale ...).

2) Show that if X is a real $\mathscr{F}^\infty$-measurable r.v., then X is constant a.s.

Exercise 3.10 (A proof of the strong law of large numbers via inverse martingales) Let $(Y_n)_{n\geq 1}$ be a sequence of real independent integrable r.v.'s, having the same law. We set $S_0 = 0$, $S_n = Y_1 + \dots + Y_n$ and $\mathscr{G}_n = \sigma(S_n, Y_{n+1}, Y_{n+2}, \dots)$.

1) Show that, for $1 \leq m \leq n$, $\mathrm{E}^{\mathscr{G}_n}(Y_m) = \mathrm{E}^{\mathscr{G}_n}(Y_1)$ a.s. and derive that $\mathrm{E}^{\mathscr{G}_n}(Y_1) = \frac{1}{n} S_n$ a.s.

2) Deduce that $(\frac{1}{n} S_n)_{n\geq 1}$ converges a.s. and in L^1 to a r.v. X.

3) Show that $X = \mathrm{E}(Y_1)$ a.s.

Exercise 3.11 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (M_n)_{n\geq 0}, \mathrm{P})$ be a square integrable martingale and denote by $(\langle M\rangle_n)_{n\geq 0}$ its associated increasing process (see •3.8). We set $\langle M\rangle_\infty = \lim_{n\to\infty} \uparrow \langle M\rangle_n$.

1) Let τ be a stopping time. Show that $(\langle M\rangle_{n\wedge\tau})_{n\geq 0}$ is the increasing process associated to the martingale $(M_n^\tau)_{n\geq 0}$ (recall that $M_n^\tau = M_{n\wedge\tau}$).

2) Let $a > 0$. Show that $\tau_a = \inf\{n;\ \langle M\rangle_{n+1} > a\}$ is a stopping time.

3) Show that the martingale $(M_n^{\tau_a})_{n\geq 0}$ converges a.s. and in L^2.

4) Show that, on $\{\langle M\rangle_\infty < +\infty\}$, $(M_n)_{n\geq 0}$ converges a.s.

Exercise 3.12 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (M_n)_{n\geq 0}, \mathrm{P})$ be a square integrable martingale. We note $A_n = \langle M\rangle_n$ the associated increasing process (see •3.8) and $A_\infty = \lim_{n\to\infty} \uparrow A_n$. We have seen in Exercise 3.11 that such a martingale converges a.s. on the event $\{A_\infty < +\infty\}$. In this exercise we are going to be more precise about what happens in $\{A_\infty = +\infty\}$. We set

$$X_0 = 0, \qquad X_n = \sum_{k=1}^{n} \frac{M_k - M_{k-1}}{1 + A_k}, \quad n \geq 1.$$

1) Show that $(X_n)_{n\geq 0}$ is a square integrable martingale and that

$$\mathbf{E}^{\mathscr{F}_{n-1}}[(X_n - X_{n-1})^2] \leq \frac{1}{1 + A_{n-1}} - \frac{1}{1 + A_n}.$$

Deduce that $\langle X \rangle_n \leq 1$ for every n and that $(X_n)_{n\geq 0}$ converges a.s.

2a) Let $(a_n)_{n\geq 0}$ be a sequence of strictly positive numbers such that $\lim_{n\to\infty} \uparrow a_n = +\infty$ and $(u_n)_{n\geq 0}$ be a sequence of real numbers converging to u. Show that

$$\frac{1}{a_n} \sum_{k=1}^{n} (a_k - a_{k-1}) u_k \underset{n\to\infty}{\to} u.$$

2b) (Kronecker's lemma) Let $(a_n)_{n\geq 1}$ be a sequence of strictly positive numbers such that $\lim_{n\to\infty} \uparrow a_n = +\infty$ and $(x_n)_{n\geq 0}$ be a sequence of real numbers. We set $s_n = x_1 + \ldots + x_n$. Show that if the series $\sum_{n=1}^{\infty} x_n/a_n$ converges, then

$$\lim_{n\to\infty} \frac{s_n}{a_n} = 0.$$

3) Show that, on $\{A_\infty = \infty\}$, $M_n/A_n \to_{n\to\infty} 0$ a.s.

Exercise 3.13 Let $(Z_n)_{n\geq 1}$ be a sequence of independent r.v.'s such that $\mathbf{P}(Z_i = 1) = \mathbf{P}(Z_i = -1) = \frac{1}{2}$ for $i \geq 1$. We set $S_0 = 0$, $S_n = Z_1 + \ldots + Z_n$, $\mathscr{F}_0 = \{\Omega, \emptyset\}$ and $\mathscr{F}_n = \sigma(Z_1, \ldots, Z_n)$. Let a be a strictly positive integer and $\tau = \inf\{n \geq 0; S_n = a\}$ the first passage time in a.

a) Show that, for $\theta \in \mathbb{R}$,

$$X_n^\theta = \frac{e^{\theta S_n}}{(\cosh \theta)^n}$$

is a $(\mathscr{F}_n)_{n\geq 0}$-martingale. Show that, if $\theta \geq 0$, $(X_{n\wedge\tau}^\theta)_{n\geq 0}$ is a bounded $(\mathscr{F}_n)_{n\geq 0}$-martingale.

b1) Show that, for every $\theta \geq 0$, $(X_{n\wedge\tau}^\theta)_{n\geq 0}$ converges a.s. and in L^2 to the r.v.

$$W^\theta = \frac{e^{\theta a}}{(\cosh \theta)^\tau} 1_{\{\tau < +\infty\}}. \tag{3.9}$$

b2) Show that $\mathbf{P}(\tau < +\infty) = 1$ and that, for every $\theta \geq 0$,

$$\mathbf{E}[(\cosh \theta)^{-\tau}] = e^{-\theta a}.$$

Exercise 3.14 Assume, as in Exercise 3.13, that $(Z_n)_{n\geq 1}$ is a sequence of independent r.v.'s such that $\mathbf{P}(Z_i = 1) = \mathbf{P}(Z_i = -1) = \frac{1}{2}$ for $i \geq 1$. We set $S_0 = 0$, $S_n = Z_1 + \ldots + Z_n$, $\mathscr{F}_0 = \{\Omega, \emptyset\}$ and $\mathscr{F}_n = \sigma(Z_1, \ldots, Z_n)$. Let a be a positive integer and λ a real number such that $0 < \lambda < \pi/(2a)$. Denote by $\tau = \inf\{n \geq 0; |S_n| = a\}$ the exit time from $]-a, a[$.

a) Show that $X_n = (\cos \lambda)^{-n} \cos(\lambda S_n)$ is a $(\mathscr{F}_n)_{n\geq 0}$-martingale.

b) Show that

$$1 = \mathbf{E}(X_{n\wedge\tau}) \geq \cos(\lambda a)\mathbf{E}((\cos \lambda)^{-n\wedge\tau}).$$

c) Deduce that $\mathbf{E}((\cos\lambda)^{-\tau}) \le (\cos(\lambda a))^{-1}$ and then that τ is a.s. finite and the martingale $(X_{n\wedge\tau})_{n\ge 0}$ is regular.

d) Compute $\mathbf{E}((\cos\lambda)^{-\tau})$. Does τ belong to L^p, $p \ge 1$?

Exercise 3.15 Let $(S_n)_{n\ge 0}$ be a simple random walk on $\mathbb{Z}$: $S_0 = 0$, $S_n = U_1 + \ldots + U_n$, where the r.v.'s U_i are independent and identically distributed such that $0 < \mathbf{P}(U_i = 1) = p < 1$, $\mathbf{P}(U_i = -1) = 1 - p = q$.

a) Let $Z_n = (\frac{q}{p})^{S_n}$. Show that $(Z_n)_{n\ge 0}$ is a positive martingale.

b) Derive by a maximal inequality applied to the martingale $(Z_n)_{n\ge 0}$ that

$$\mathbf{P}\Big(\sup_{n\ge 0} S_n \ge k\Big) \le \Big(\frac{p}{q}\Big)^k$$

and that, if $q > p$,

$$\mathbf{E}\Big(\sup_{n\ge 0} S_n\Big) \le \frac{p}{q-p}.$$

♦ We shall see in Problem 3.5 that actually, if $q > p$, in the two previous inequalities the equal sign holds and that the r.v. $\sup_{n\ge 0} S_n$ has a geometric law with parameter $1 - \frac{p}{q}$ (i.e., $\mathbf{P}(\sup_{n\ge 0} S_n = k) = (1 - \frac{p}{q})(\frac{p}{q})^k$).

Exercise 3.16 Let $(X_n)_{n\ge 1}$ be a sequence of real independent r.v.'s, all having normal distribution $N(m, \sigma^2)$ with $m < 0$. Let us set $S_0 = 0$, $S_n = X_1 + \ldots + X_n$ and

$$\mathscr{B}_n = \sigma(S_0, \ldots, S_n), \qquad W = \sup_{n\ge 0} S_n.$$

The aim of this exercise is to establish certain properties of the r.v. W.

1) Show that $\mathbf{P}(W < +\infty) = 1$.

2) Recall that, for every real λ, $\mathbf{E}(e^{\lambda X_1}) = e^{\frac{1}{2}\lambda^2\sigma^2} e^{\lambda m}$. Compute $\mathbf{E}(e^{\lambda S_{n+1}} \mid \mathscr{B}_n)$.

3) Show that there exists a unique $\lambda_0 > 0$ such that $(e^{\lambda_0 S_n})_{n\ge 0}$ is a martingale.

4) Show that, for every $a > 1$,

$$\mathbf{P}(e^{\lambda_0 W} > a) \le \frac{1}{a}$$

and that, for $t > 0$, $\mathbf{P}(W > t) \le e^{-\lambda_0 t}$.

5) Show that

$$\mathbf{E}(e^{\lambda W}) = 1 + \lambda \int_0^{+\infty} e^{\lambda t}\, \mathbf{P}(W > t)\, dt \tag{3.10}$$

and deduce that, for every $\lambda < \lambda_0$, $\mathbf{E}(e^{\lambda W}) < +\infty$. In particular W has moments of all orders.

Exercise 3.17 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge 0}, (X_n)_{n\ge 0}, \mathbf{P})$ be a submartingale bounded in L^1, i.e., such that $\sup_{n\ge 0} \mathbf{E}|X_n| < +\infty$.

1) Show that, for a fixed n, the sequence $(\mathbf{E}^{\mathscr{F}_n}(X_p^+))_{p\ge n}$ is increasing in p.

2) Let us set $M_n = \lim_{p\to\infty} \uparrow \mathbf{E}^{\mathscr{F}_n}(X_p^+)$. Show that $(M_n)_{n\ge 0}$ is a positive integrable martingale.

3) Let us set $Y_n = M_n - X_n$. Show that $(Y_n)_{n\ge 0}$ is a positive integrable supermartingale.

♦ We come to the conclusion that every submartingale that is bounded in L^1 can be written as the difference of a martingale and of a supermartingale, both positive and integrable (Krickeberg's decomposition).

Exercise 3.18 Let $(Y_n)_{n\geq 0}$ be a sequence of independent identically distributed r.v.'s such that $\mathbf{P}(Y_k = 1) = \mathbf{P}(Y_k = -1) = \frac{1}{2}$. We set $\mathscr{B}_0 = \{\emptyset, \Omega\}$, $\mathscr{B}_n = \sigma(Y_1, \dots, Y_n)$ and $S_0 = 0$, $S_n = Y_1 + \dots + Y_n$, $n \geq 1$. Let us consider the sign function

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

and the process defined by

$$M_0 = 0, \qquad M_n = \sum_{k=1}^{n} \operatorname{sign}(S_{k-1}) Y_k, \quad n = 1, 2, \dots .$$

1) What is the compensator (see •3.4) of the submartingale $(S_n^2)_{n\geq 0}$?

2) Show that $(M_n)_{n\geq 0}$ is a square integrable martingale and compute its associated increasing process (•3.8).

3) Compute the Doob decomposition of $(|S_n|)_{n\geq 0}$. Show that M_n is measurable with respect to the σ-algebra $\sigma(|S_1|, \dots, |S_n|)$.

Exercise 3.19 Let p, q be two probabilities on a countable set E such that $p \neq q$ and $q(x) > 0$ for every $x \in E$. Let $(X_n)_{n\geq 1}$ be a sequence of independent r.v.'s with values in E and having the same law q. Show that the sequence

$$Y_n = \prod_{k=1}^{n} \frac{p(X_k)}{q(X_k)}$$

is a positive martingale whose limit is 0 a.s. Is it regular? (Hint: compute the mean of $\sqrt{Y_n}$.)

Exercise 3.20 Let $(Y_n)_{n\geq 1}$ be a sequence of positive independent r.v'.s with expectation equal to 1 and $\mathscr{F}_0 = \{\emptyset, \Omega\}$, $\mathscr{F}_n = \sigma(Y_k; k \leq n)$. We set $X_0 = 1$ and $X_n = \prod_{k=1}^{n} Y_k$.

1) Show that $(X_n)_{n\geq 0}$ is a martingale with respect to the filtration $(\mathscr{F}_n)_{n\geq 0}$ and deduce that $(\sqrt{X_n})_{n\geq 0}$ is a supermartingale.

2) Assume that $\prod_{k=1}^{\infty} \mathbf{E}(\sqrt{Y_k}) = 0$. Investigate the convergence and the limit of $(\sqrt{X_n})_{n\geq 0}$, then of $(X_n)_{n\geq 0}$. Is the martingale $(X_n)_{n\geq 0}$ regular?

3) Assume that $\prod_{k=1}^{\infty} \mathbf{E}(\sqrt{Y_k}) > 0$. Show that $(\sqrt{X_n})_{n\geq 0}$ is a Cauchy sequence in L^2 and deduce that $(X_n)_{n\geq 0}$ is a regular martingale.

Exercise 3.21 The aim of this exercise is to establish a version of the Borel–Cantelli lemma for a family of not necessarily independent r.v.'s, which is often useful. We recall that, for a martingale $(M_n)_{n\geq 0}$, the three relationships

$$\sup_{n\geq 0} \mathbf{E}(|M_n|) < +\infty, \qquad \sup_{n\geq 0} \mathbf{E}(M_n^+) < +\infty, \qquad \sup_{n\geq 0} \mathbf{E}(M_n^-) < +\infty$$

are equivalent (•3.10). Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ be a filtered probability space and $(Y_n)_{n\geq 1}$ a sequence of positive, integrable, adapted r.v.'s (not necessarily independent).

1a) Let us set $X_0 = 0$, $X_n = Y_1 + \ldots + Y_n$. Show that $(X_n)_{n\geq 0}$ is a submartingale and determine its compensator $(A_n)_{n\geq 1}$ (•3.4).

1b) Show that, for every $a > 0$, $\tau_a = \inf\{n;\ A_{n+1} > a\}$ is a stopping time.

1c) Let us set $Z_n = X_n - A_n$. Show that, for every n, $Z^-_{n\wedge\tau_a} \leq a$. Deduce that $(Z_{n\wedge\tau_a})_{n\geq 0}$ converges a.s.

1d) Show that $\{\lim_{n\to\infty} \uparrow A_n < +\infty\} \subset \{\lim_{n\to\infty} \uparrow X_n < +\infty\}$ a.s. (hint: what happens on $\{\tau_a = +\infty\}$?).

2) Assume, furthermore, that $\sup_{n\geq 1} Y_n \in L^1$. Show that

$$\{\lim_{n\to\infty} \uparrow X_n < +\infty\} = \{\lim_{n\to\infty} \uparrow A_n < +\infty\} \quad \text{a.s.}$$

(hint: introduce the stopping time $\sigma_a = \inf\{n;\ X_n > a\}$ and consider $Z^+_{n\wedge\sigma_a}$).

3) Let $(B_n)_{n\geq 1}$ be a sequence of adapted events. Show that

$$\Big\{\sum_{n\geq 1} \mathbf{P}^{\mathscr{F}_{n-1}}(B_n) < +\infty\Big\} = \Big\{\sum_{n\geq 1} 1_{B_n} < +\infty\Big\} \qquad \text{a.s.}$$

Exercise 3.22 Let $(X_n)_{n\geq 1}$ be a sequence of real, square integrable, adapted r.v.'s, defined on $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, \mathbf{P})$ and $(\sigma_n^2)_{n\geq 0}$ a sequence of positive numbers. Assume that a.s., for every $n \geq 1$,

$$\mathbf{E}^{\mathscr{F}_{n-1}}(X_n) = 0, \qquad \mathbf{E}^{\mathscr{F}_{n-1}}(X_n^2) = \sigma_n^2 \text{ a.s.}$$

We set $S_0 = 0$, $A_0 = 0$ and, for $n \geq 1$, $S_n = X_1 + \ldots + X_n$, $A_n = \sigma_1^2 + \ldots + \sigma_n^2$, $V_n = S_n^2 - A_n$.

1) Show that $(S_n)_{n\geq 0}$ and $(V_n)_{n\geq 0}$ are martingales.

2) Show that, if $\sum_{k=1}^{\infty} \sigma_k^2 < +\infty$, $(S_n)_{n\geq 0}$ converges a.s. and in L^2.

3) Assume that $(S_n)_{n\geq 0}$ converges a.s. and that there exists a constant M such that, for every $n \geq 1$, $|X_n| \leq M$ a.s. For $a > 0$, we set $\tau_a = \inf\{n \geq 0;\ |S_n| > a\}$.

3a) Show that, for every n, $\mathbf{E}(S^2_{n\wedge\tau_a}) = \mathbf{E}(A_{n\wedge\tau_a})$.

3b) Show that there exists $a > 0$ such that $\mathbf{P}(\tau_a = +\infty) > 0$.

3c) Deduce that $\sum_{k=1}^{\infty} \sigma_k^2 < +\infty$.

♦ Let $(X_n)_{n\geq 1}$ be a sequence of real *centered*, independent, square integrable r.v.'s. It is clear that such a sequence satisfies the hypotheses of this exercise. Thus we have just met some classical results:

(i) if $\sum_{k=1}^{\infty} \mathrm{Var}(X_k) < +\infty$, then $\sum_{k=1}^{\infty} X_k$ converges a.s.,

(ii) if the r.v.'s X_k are uniformly bounded, then $\sum_{k=1}^{\infty} X_k$ converges a.s. if and only if $\sum_{k=1}^{\infty} \mathrm{Var}(X_k) < +\infty$.

Exercise 3.23 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (M_n)_{n\geq 0}, \mathbf{P})$ be a square integrable martingale and denote by $A_n = \langle M\rangle_n$ the associated increasing process (see •3.8). We set $\tau_a = \inf\{n \geq 0;\ A_{n+1} > a^2\}$.

1) Show that τ_a is a stopping time.

2) Show that $\mathbf{P}(\sup_{n\geq 0} |M_{n\wedge\tau_a}| > a) \leq a^{-2}\mathbf{E}(A_\infty \wedge a^2)$.

3) Show that

$$\mathbf{P}\Big(\sup_{n\geq 0} |M_n| > a\Big) \leq \mathbf{P}(A_\infty > a^2) + \mathbf{P}\Big(\sup_{n\geq 0} |M_{n\wedge\tau_a}| > a\Big). \tag{3.11}$$

4) Let X be a positive r.v. Show, using Fubini's theorem, the two relations

$$\int_0^\lambda \mathbf{P}(X > t)\,dt = \mathbf{E}(X \wedge \lambda) \text{ for every } \lambda \in [0, +\infty],$$
$$\int_0^\infty a^{-2}\mathbf{E}(X \wedge a^2)\,da = 2\mathbf{E}(\sqrt{X}).$$

5) Show that $\mathbf{E}(\sup_{n\geq 0} |M_n|) \leq 3\mathbf{E}(\sqrt{A_\infty})$ (hint: integrate (3.11) with respect to a from 0 to $+\infty$ ...).

6) Let $(Y_n)_{n\geq 1}$ be a sequence of centered, independent, square integrable and identically distributed r.v.'s. We set $S_0 = 0$, $\mathscr{F}_0 = \{\Omega, \emptyset\}$ and, for $n \geq 1$, $S_n = Y_1 + \ldots + Y_n$, $\mathscr{F}_n = \sigma(Y_1, \ldots, Y_n)$. Show that if τ is a stopping time such that $\mathbf{E}(\sqrt{\tau}) < +\infty$, then $\mathbf{E}(S_\tau) = 0$.

♦ It is interesting to compare the result of question (6) with part (A4) of Exercise 3.25 for $m = 0$. One can derive, furthermore, that the passage time τ of part (A4) of Exercise 3.25 is such that $\mathbf{E}(\sqrt{\tau}) = +\infty$.

Exercise 3.24 On a probability space $(\Omega, \mathscr{F}, \mathbf{P})$ let us consider the r.v's. X_n^i, $n = 1, 2, \ldots$, $i = 1, 2$, independent and Bernoulli $B(1, \frac{1}{2})$. We set $S_n^i = \sum_{k=1}^n X_k^i$, $\nu_i = \inf\{n;\ S_n^i = a\}$ where a is an integer ≥ 1 and set $\nu = \nu_1 \wedge \nu_2$.

1) Show that $\mathbf{P}(\nu_i < +\infty) = 1$, $i = 1, 2$.

2) We set, for $i = 1, 2$ and every $n \geq 0$,

$$M_n^i = 2S_n^i - n, \qquad M_n^{i,j} = (2S_n^i - n)(2S_n^j - n) - n\delta_{i,j}$$

where $\delta_{i,j} = 1$ if $i = j$ and $= 0$ otherwise. Show that $(M_n^i)_n$ and $(M_n^{i,j})_n$ are martingales with respect to the filtration

$$\mathscr{F}_n = \sigma(X_k^i, k \leq n, i = 1, 2).$$

3) Show that $\mathbf{E}(\nu) \leq 2a$.

4) Show that $\mathbf{E}(M_\nu^{i,j}) = 0$.

5) Show that $\mathbf{E}(|S_\nu^1 - S_\nu^2|) \leq \sqrt{a}$ (hint: consider the martingale $M_n^{1,1} - 2M_n^{1,2} + M_n^{2,2}$).

Exercise 3.25 (Wald identities) Let $(Y_n)_{n\geq 1}$ be a sequence of independent, integrable, identically distributed real r.v.'s. We set $m = \mathbf{E}(Y_1)$, $S_0 = 0$, $\mathscr{F}_0 = \{\Omega, \emptyset\}$ and, for $n \geq 1$, $S_n = Y_1 + \ldots + Y_n$, $\mathscr{F}_n = \sigma(Y_1, \ldots, Y_n)$. Let ν be an integrable stopping time.

A1) We set $X_n = S_n - nm$. Show that $(X_n)_{n\geq 0}$ is a martingale.

A2) Show that, for every n, $\mathbf{E}(S_{n\wedge\nu}) = m\mathbf{E}(n \wedge \nu)$.

A3) Show that S_ν is integrable and that $\mathbf{E}(S_\nu) = m\mathbf{E}(\nu)$ (hint: consider first the case $Y_n \geq 0$).

A4) Assume that, for every n, $\mathbf{P}(Y_n = -1) = \mathbf{P}(Y_n = 1) = \frac{1}{2}$ and $\tau = \inf\{n;\ S_n \geq a\}$, where $a \geq$ is an integer. In Exercise 3.13 we showed that $\tau < +\infty$ a.s. Show that τ is not integrable

B) Assume furthermore that $\mathbf{E}(Y_1^2) < +\infty$ and set $\sigma^2 = \mathrm{Var}(Y_1)$. Assume first that $m = 0$ and set $Z_n = S_n^2 - n\sigma^2$.

B1) Show that $(Z_n)_{n\geq 0}$ is a $(\mathscr{F}_n)_{n\geq 0}$-martingale.

B2) Show that, for every $j < k$,

$$\mathbf{E}[Y_j 1_{\{j\le\nu\}} Y_k 1_{\{k\le\nu\}}] = 0,$$

and then that $\mathbf{E}[\sum_{k=1}^{\infty} Y_k^2 1_{\{k\le\nu\}}] < +\infty$ (hint: use (A)).

B3) Show that $(S_{n\wedge\nu})_{n\ge0}$ is a Cauchy sequence in L^2. Deduce that $S_{n\wedge\nu} \to_{n\to\infty} S_\nu$ in L^2.

B4) Show that $\mathbf{E}(S_\nu^2) = \sigma^2\mathbf{E}(\nu)$.

B5) Let us drop now the assumption $m = 0$. Show that $\mathbf{E}((S_\nu - m\nu)^2) = \sigma^2\mathbf{E}(\nu)$.

Exercise 3.26 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge0}, \mathbf{P})$ be a filtered probability space and ν a finite measure on $\mathscr{F}_\infty = \sigma(\bigcup_{n=0}^{\infty}\mathscr{F}_n)$. Let us assume that, for every $n \ge 0$, $\mathbf{P}$ dominates ν on $\mathscr{F}_n$ and denote by M_n the Radon–Nikodym density: M_n is thus positive, $\mathscr{F}_n$-measurable and

$$\nu(A) = \int_A M_n\, d\mathbf{P}$$

for every $A \in \mathscr{F}_n$.

a) Prove that $(M_n)_{n\ge0}$ is a martingale.

b) Prove that $(M_n)_{n\ge0}$ converges a.s. to an integrable r.v. M_∞.

c) Prove that $\mathbf{P}$ dominates ν on $\mathscr{F}_\infty$ if and only if the martingale $(M_n)_n$ is regular and that, in this case, M_∞ is the corresponding Radon–Nikodym density.

d) Let us assume that the two measures ν and $\mathbf{P}$ are mutually orthogonal on $\mathscr{F}_\infty$, i.e., there exists $S \in \mathscr{F}_\infty$ such that $\mathbf{P}(S) = 1$ and $\nu(S) = 0$. Show that $M_\infty = 0$ a.s.

e) Let $\Omega = \mathbb{R}^{\mathbb{N}}$, $X_n(\omega) = \omega_n$, $\mathscr{F}_n = \sigma(X_k, k \le n)$ and let $\mathbf{P}$ be the probability on $(\Omega, \mathscr{F}_\infty)$, which is the product of the laws $N(0, 1)$. Thus with respect to $\mathbf{P}$ the r.v.'s X_n are independent and have law $N(0, 1)$. Let ν be the probability on $(\Omega, \mathscr{F}_\infty)$, which is the product of the laws $N(\mu_n, 1)$, where $(\mu_n)_{n\ge1}$ is a sequence of real numbers; with respect to ν the r.v.'s X_n are also independent and have law $N(\mu_n, 1)$.

e1) Prove that $\mathbf{P}$ and ν are equivalent on $\mathscr{F}_n$ for every n and compute the density of ν with respect to $\mathbf{P}$.

e2) Using the criterion of Exercise 3.20 (2) and of (3), give necessary and sufficient conditions on the sequence $(\mu_n)_{n\ge1}$ in order to ensure that $\mathbf{P}$ dominates ν on $\mathscr{F}_\infty$.

Exercise 3.27 Let $(X_n)_{n\ge0}$ be a martingale, defined on a space $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge0}, \mathbf{P})$ and ν a stopping time such that

$$\mathbf{P}(\nu < +\infty) = 1, \qquad \mathbf{E}(|X_\nu|) < +\infty, \qquad \int_{\{\nu>n\}} |X_n|\, d\mathbf{P} \underset{n\to\infty}{\to} 0.$$

1) Show that

$$\int_{\{\nu>n\}} |X_\nu|\, d\mathbf{P} \underset{n\to\infty}{\to} 0.$$

2) Show that $\mathbf{E}(|X_{\nu\wedge n} - X_\nu|) \to 0$ as $n \to \infty$.

3) Deduce that $\mathbf{E}(X_\nu) = \mathbf{E}(X_0)$.

Exercise 3.28 (A stopping theorem for nonbounded stopping times) Let us consider a supermartingale $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge0}, (X_n)_{n\ge0}, \mathbf{P})$. Assume that there exists a constant M such that, for every $n \ge 1$,

$$\mathbf{E}^{\mathscr{F}_{n-1}}(|X_n - X_{n-1}|) \le M \qquad \text{a.s.}$$

1) Show that, if $(V_n)_{n\geq 1}$ is a positive process such that, for every $n \geq 0$, V_n is $\mathscr{F}_{n-1}$-measurable,

$$\mathbf{E}\Big(\sum_{n=1}^{\infty} V_n |X_n - X_{n-1}|\Big) \leq M\mathbf{E}\Big(\sum_{n=1}^{\infty} V_n\Big).$$

2) Let ν be an *integrable* stopping time. Recall that $\mathbf{E}(\nu) = \sum_{n\geq 1} \mathbf{P}(\nu \geq n)$.

2a) Deduce from (1) that $\mathbf{E}(\sum_{n\geq 1} 1_{\{\nu\geq n\}}|X_n - X_{n-1}|) < +\infty$.

2b) What is the value of $\sum_{n\geq 1} 1_{\{\nu\geq n\}}(X_n - X_{n-1})$? Deduce that X_ν is integrable.

3) Show that $(X_{\nu\wedge p})_{p\geq 0}$ converges to X_ν in L^1 as $p \to \infty$.

4) Deduce that, if $\nu_1 \leq \nu_2$ are stopping times and ν_2 is integrable,

$$\mathbf{E}(X_{\nu_2} \mid \mathscr{F}_{\nu_1}) \leq X_{\nu_1}$$

(hint: one can use the fact that, if $A \in \mathscr{F}_{\nu_1}$, then $A \cap \{\nu_1 \leq k\} \in \mathscr{F}_{\nu_1\wedge k}$; why is this true? Have a look at Exercise 2.1 (iv)).

Problems

Problem 3.1 A family $(X_i)_{i\in I}$ of real r.v.'s on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$ is *equi-integrable* (or *uniformly integrable*) if

$$\sup_{i\in I} \int_{\{|X_i|>a\}} |X_i|\, d\mathbf{P} \underset{a\to+\infty}{\to} 0.$$

We now prove several properties of equi-integrable families of integrable r.v.'s and apply them to the theory of martingales.

1) Show that a finite family of integrable r.v.'s is equi-integrable.

2) Show that the family $(X_i)_{i\in I}$ is equi-integrable if and only if $\sup_{i\in I} \mathbf{E}|X_i| < +\infty$ and the property

(P) for every $\varepsilon > 0$ there exists $\alpha > 0$ such that $\mathbf{P}(A) < \alpha$ implies

$$\sup_{i\in I} \int_A |X_i|\, d\mathbf{P} < \varepsilon$$

holds.

3) Show that if $\sup_{i\in I} \mathbf{E}[g(|X_i|)] < +\infty$, g being an application from $\mathbb{R}^+$ in $\mathbb{R}^+$ satisfying $\lim_{t\to+\infty} \frac{1}{t}g(t) = +\infty$, then the family $(X_i)_{i\in I}$ is equi-integrable. Deduce that a family of r.v.'s that is bounded in $L^p(\Omega, \mathscr{F}, \mathbf{P})$, $p > 1$, is equi-integrable.

4) Let $(\mathscr{B}_i)_{i\in I}$ be the family of all sub-σ-algebras of $\mathscr{F}$.

4a) Let X be a r.v. of L^p, $p > 1$. Prove that the family $(\mathbf{E}^{\mathscr{B}_i}(X))_{i\in I}$ is equi-integrable.

4b) Go farther and prove that the family $(\mathbf{E}^{\mathscr{B}_i}(X))_{i\in I}$ is equi-integrable for every integrable r.v. X (recall that L^p, $p \geq 1$, is dense in L^1).

5) Let $(X_n)_{n\geq 0}$ be an equi-integrable sequence of real r.v.'s converging a.s. to a r.v. X. Let us set, for $a > 0$,

$$f_a(x) = -a 1_{\{x<-a\}} + x 1_{\{-a\leq x\leq a\}} + a 1_{\{x>a\}}.$$

5a) Show that $\mathbf{E}(|X|) < +\infty$.

5b) Show that, for every $a > 0$, $\|f_a(X_n) - f_a(X)\|_1 \to_{n\to\infty} 0$.

5c) Show that

$$\sup_{n\geq 0} \|X_n - f_a(X_n)\|_1 \underset{a\to+\infty}{\to} 0, \tag{3.12}$$

$$\|X - f_a(X)\|_1 \underset{a\to+\infty}{\to} 0. \tag{3.13}$$

5d) Show finally that, as $n \to \infty$, $(X_n)_{n\geq 0}$ converges to X in L^1.

5e) Show that, if there exists an integrable r.v. Z such that $|X_n| \leq Z$ for every $n \geq 0$, then $(X_n)_{n\geq 0}$ is equi-integrable. Remark that Lebesgue's theorem is a particular case of (5d).

6) Let $(X_n)_{n\geq 0}$ be a sequence of real r.v.'s converging in L^1 to a r.v. X. Show that, for every $A \in \mathscr{F}$,

$$\int_A |X_n|\, d\mathbf{P} \leq \int_A |X|\, d\mathbf{P} + \int_\Omega |X - X_n|\, d\mathbf{P}. \tag{3.14}$$

Deduce that the sequence is equi-integrable.

7) Show that a martingale is regular if and only if it is equi-integrable.

8) Let $(X_n)_{n\geq 0}$ be a martingale satisfying

$$\sup_{n\geq 0} \mathbf{E}[|X_n| \log^+ |X_n|] < +\infty. \tag{3.15}$$

Then $(X_n)_{n\geq 0}$ is regular.

♦ The most important result of this exercise is the following: a martingale $(X_n)_{n\geq 0}$ is regular if and only if it is equi-integrable, i.e., if it satisfies the condition

$$\sup_{n\geq 0} \int_{\{|X_n|>a\}} |X_n|\, d\mathbf{P} \underset{a\to+\infty}{\to} 0.$$

Problem 3.2 (The solution makes use of Problem 3.1) Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability space. Assume that the σ-algebra $\mathscr{F}$ is separable, i.e., that there exists a sequence $(F_n)_{n\geq 1}$ of events such that $\mathscr{F} = \sigma(F_n, n \geq 1)$. Let $\mathbf{Q}$ be a probability on $(\Omega, \mathscr{F})$ that is absolutely continuous with respect to $\mathbf{P}$, i.e., satisfying for every $A \in \mathscr{F}$ $\mathbf{Q}(A) = 0$ if $\mathbf{P}(A) = 0$.

1a) Let $(A_n)_{n\geq 1}$ be a sequence of events such that $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < +\infty$ and satisfying, for every n, $\mathbf{Q}(A_n) \geq \alpha > 0$. Let $A = \limsup_{n\to\infty} A_n$. Show that $\mathbf{P}(A) = 0$ and $\mathbf{Q}(A) \geq \alpha$.

1b) Deduce that, for every $\varepsilon > 0$, there exists $\eta > 0$ such that $\mathbf{P}(A) < \eta$ implies $\mathbf{Q}(A) < \varepsilon$.

2) Show that there exists a sequence $\mathscr{P}_n = (G_{n,k})_{1\leq k\leq r(n)}$ of finite partitions of Ω such that

$$\mathscr{F}_n \overset{\text{def}}{=} \sigma(F_1, \ldots, F_n) = \sigma(\mathscr{P}_n).$$

3) Let $I_n = \{k \in \{1, \ldots, r(n)\}, \mathbf{P}(G_{n,k}) > 0\}$. We set

$$X_n = \sum_{k\in I_n} \frac{\mathbf{Q}(G_{n,k})}{\mathbf{P}(G_{n,k})} 1_{G_{n,k}}.$$

Show that $(X_n)_{n\geq 1}$ is a $(\mathscr{F}_n)_{n\geq 0}$-martingale that converges a.s.

4) Show that the martingale $(X_n)_{n\ge 1}$ is equi-integrable (use (1b)) and conclude that it converges in L^1.

5) Show that, for every $A \in \mathscr{F}$, $\mathbf{Q}(A) = \int_A X\,d\mathbf{P}$, where $X = \lim_{n\to\infty} X_n$.

♦ This exercise produces a proof of the Radon–Nikodym theorem for a separable σ-algebra. It is possible, but rather technical, to extend this proof to any σ-algebra. Remark that the a.s. uniqueness of the density X comes from •1.3.

Problem 3.3 On the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge 0}, \mathbf{P})$, let us consider a submartingale $(X_n)_{n\ge 0}$ such that $X_0 = 0$. Let $(A_n)_{n\ge 0}$ be its compensator (see •3.4). We have already seen (in Exercise 3.11, for instance) how to derive information about the convergence of a martingale from $(A_n)_{n\ge 0}$. In this problem we go deeper into this question.

1) Assume that $X_n \ge 0$ for every n and set, for $a > 0$, $\sigma_a = \inf\{n \ge 0;\ A_{n+1} > a\}$.

1a) Show that σ_a is a stopping time.

1b) Show that $\mathbf{E}[X_{n\wedge\sigma_a}] \le a$.

1c) Deduce that $(X_n)_{n\ge 0}$ converges a.s. on $\{\sigma_a = +\infty\}$.

1d) Show that

$$\{A_\infty < +\infty\} \subset \{(X_n)_{n\ge 0} \text{ converges}\} \subset \left\{\sup_{n\ge 0} X_n < +\infty\right\} \quad \text{a.s.}$$

Let us set, for $a > 0$, $\tau_a = \inf\{n \ge 1;\ X_n > a\}$. $(X_n)_{n\ge 0}$ is said to be of class $\mathscr{C}^+$ if, for every $a > 0$,

$$\mathbf{E}[(\Delta_{\tau_a})^+ 1_{\{\tau_a < +\infty\}}] = C_a < +\infty$$

where $\Delta_n = X_n - X_{n-1}$, $n \ge 1$ (in particular, if $\mathbf{E}(\sup_{n\ge 1} |X_n - X_{n-1}|) < +\infty$, then $(X_n)_{n\ge 0}$ is of class $\mathscr{C}^+$).

2) Assume that $(X_n)_{n\ge 0}$ is of class $\mathscr{C}^+$ (we are not assuming $X_n \ge 0$ any more).

2a) Show that $\mathbf{E}[X^+_{\tau_a} 1_{\{\tau_a < +\infty\}}] \le a + C_a$ and then that $\mathbf{E}[X^+_{n\wedge\tau_a}] \le 2a + C_a$.

2b) Deduce that $\sup_{n\ge 0} \mathbf{E}(|X_{n\wedge\tau_a}|) < +\infty$, then that $(X_n)_{n\ge 0}$ converges a.s. on $\{\tau_a = +\infty\}$ (make use of the equality $|x| = 2x^+ - x$).

2c) Deduce that a.s. $\{(X_n)_{n\ge 0} \text{ converges}\} = \{\sup_{n\ge 0} X_n < +\infty\}$.

2d) Show that $\mathbf{E}[A_{n\wedge\tau_a}] = \mathbf{E}[X_{n\wedge\tau_a}] \le 2a + C_a$. Deduce that $\mathbf{E}[A_{\tau_a}] < +\infty$ and that, a.s., $\{\tau_a = +\infty\} \subset \{A_\infty < +\infty\}$.

2e) Show that a.s. $\{\sup_{n\ge 0} X_n < +\infty\} \subset \{A_\infty < +\infty\}$.

2f) Assume furthermore that $X_n \ge 0$ for every n. Show that

$$\{(X_n)_{n\ge 0} \text{ converges}\} = \left\{\sup_{n\ge 0} X_n < +\infty\right\} = \{A_\infty < +\infty\} \quad \text{a.s.}$$

3) Assume that $(X_n)_{n\ge 0}$ is a martingale satisfying the relation $\mathbf{E}(\sup_{n\ge 0} |\Delta_n|) < +\infty$. Show that

$$\Omega = \{(X_n)_{n\ge 0} \text{ converges}\} \cup \left\{\varlimsup_{n\to\infty} X_n = +\infty,\ \varliminf_{n\to\infty} X_n = -\infty\right\} \quad \text{a.s.}$$

4) Let $(B_n)_{n\ge 1}$ be a sequence of adapted events. Show that

$$\left\{\sum_{n=1}^{\infty} \mathbf{P}^{\mathscr{F}_{n-1}}(B_n) < +\infty\right\} = \left\{\sum_{n=1}^{\infty} 1_{B_n} < +\infty\right\} \quad \text{a.s.}$$

5) Let $(M_n)_{n\geq 0}$ be a square integrable martingale vanishing at 0 and let $(\langle M\rangle_n)_{n\geq 0}$ be the associated increasing process (see •3.8). We set $A_n = \langle M\rangle_n$.

5a) Show that the submartingale $X_n = (M_n+1)^2$ has $(A_n)_{n\geq 0}$ as its compensator.

5b) Deduce (one can make use of (1)) that

$$\{A_\infty < +\infty\} \subset \{(M_n)_{n\geq 0} \text{ converges}\} \quad \text{a.s.}$$

5c) Assume furthermore that $\mathbf{E}(\sup_{n\geq 0}|M_{n+1}-M_n|^2) < +\infty$. Show that

$$\{A_\infty < +\infty\} = \{(M_n)_{n\geq 0} \text{ converges}\} \quad \text{a.s.}$$
$$\{A_\infty = +\infty\} = \left\{\overline{\lim_{n\to\infty}}\, M_n = +\infty,\ \underline{\lim_{n\to\infty}}\, M_n = -\infty\right\} \quad \text{a.s.}$$

(a useful equality: $|y^2-x^2| \leq |y-x|^2 + 2|x||y-x|$).

Problem 3.4 In this problem we show, through probability arguments, that *a Lipschitz continuous function is the primitive of a bounded measurable function.*

Let $f:[0,1]\to\mathbb{R}$ be a Lipschitz continuous function, with Lipschitz constant $L>0$ and X a r.v. taking values in $]0,1[$ and uniformly distributed. We set

$$X_n = 2^{-n}[2^n X], \qquad Z_n = 2^n\left(f(X_n+\tfrac{1}{2^n}) - f(X_n)\right)$$

and $\mathscr{F}_n = \sigma(X_1,\ldots,X_n)$ ($[x]$ denotes the integer part of the real number x). It is useful to remark that $X_n = \frac{k}{2^n}$ if and only if $X\in[\frac{k}{2^n},\frac{k+1}{2^n}[$ and $\mathbf{P}(X_n = \frac{k}{2^n}) = \frac{1}{2^n}$, $0\leq k\leq 2^n-1$.

a) Investigate the convergence of the sequence $(X_n)_{n\geq 0}$.

b1) Prove the relationship $X_n = 2^{-n}[2^n X_{n+1}]$ and deduce that $\mathscr{F}_n = \sigma(X_n)$

b2) Show that

$$\bigcap_{n\geq 0}\sigma(X_n, X_{n+1},\ldots) = \sigma(X).$$

c) Determine the conditional law of X_{n+1} given $\mathscr{F}_n$. Deduce that $(Z_n)_{n\geq 0}$ is a bounded martingale.

We note Z_∞ its limit, a.s. and in L^1.

d) Show that there exists a Borel function g such that $Z_\infty = g(X)$ a.s.

e) Compute the conditional law of X given X_n and show that

$$Z_n = 2^n\int_{X_n}^{X_n+2^{-n}} g(u)\,du \qquad \text{a.s.}$$

f) Deduce that for every k, n, $0\leq k\leq 2^n-1$,

$$f(\tfrac{k}{2^n}+\tfrac{1}{2^n}) - f(\tfrac{k}{2^n}) = \int_{k2^{-n}}^{(k+1)2^{-n}} g(u)\,du$$

and conclude that, for every $x\in[0,1]$,

$$f(x)-f(0) = \int_0^x g(u)\,du. \tag{3.16}$$

Problem 3.5 Let $(Y_n)_{n\geq 1}$ be a sequence of r.v.'s with values in $\mathbb{Z}$, integrable, independent and having a common law μ. Assume that $\mathbf{E}(Y_i) = m < 0$ and $\mathbf{P}(Y_i = 1) > 0$,

$\mathbf{P}(Y_i \ge 2) = 0$. Let us set $X_0 = 0$, $X_n = Y_1 + \ldots + Y_n$ and

$$W = \sup_{n \ge 0} X_n.$$

The goal of this problem is to determine the law of W.

a) Prove that $W < +\infty$ a.s.

b1) Let X be a real r.v. We denote by $M(\lambda) = \mathbf{E}(e^{\lambda X})$ its Laplace transform (possibly $M(\lambda) = +\infty$) and set $\psi(\lambda) = \log M(\lambda)$. Prove that ψ is convex (hint: use Hölder's inequality).

b2) Let us set $M(\lambda) = \mathbf{E}(e^{\lambda Y_1})$ and $\psi(\lambda) = \log M(\lambda)$. Prove that $\psi(\lambda) < +\infty$ for every $\lambda \ge 0$. What is the value of $\psi'(0+)$? Prove that $\psi(\lambda) \to +\infty$ as $\lambda \to +\infty$ and that there exists a unique $\lambda_0 > 0$ such that $\psi(\lambda_0) = 0$.

c) Let λ_0 be as in (b2). Prove that $Z_n = e^{\lambda_0 X_n}$ is a martingale and remark that $\lim_{n\to\infty} Z_n = 0$ a.s.

d) Let $k \in \mathbb{N}$, $k \ge 1$ and $\tau_k = \inf\{n \ge 1;\ X_n \ge k\}$. Prove that

$$\lim_{n\to\infty} Z_{n\wedge\tau_k} = e^{\lambda_0 k}\, 1_{\{\tau_k < +\infty\}}. \tag{3.17}$$

e) Compute $\mathbf{P}(\tau_k < +\infty)$ and the law of W. Work out precisely this law if $\mathbf{P}(Y_i = 1) = p$, $\mathbf{P}(Y_i = -1) = q = 1 - p$, $p < \frac{1}{2}$.

Problem 3.6 In this problem we apply a convergence result for positive supermartingales to the study of a stochastic algorithm. Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge0}, \mathbf{P})$ be a filtered probability space.

A) (Robbins–Siegmund's lemma) Let us consider a sequence of positive adapted a.s. finite r.v.'s $(Z_n, \beta_n, \xi_n, \eta_n)_{n\ge0}$. We assume that, for every $n \ge 0$,

$$\mathbf{E}^{\mathscr{F}_n}(Z_{n+1}) \le (1 + \beta_n) Z_n + \xi_n - \eta_n \qquad \text{a.s.}$$

We investigate the a.s. convergence of the sequence $(Z_n)_{n\ge0}$.

A1) We set

$$\alpha_{-1} = 1, \qquad \alpha_n = \prod_{k=0}^{n} (1 + \beta_k)^{-1},$$
$$Z'_n = \alpha_{n-1} Z_n, \quad \xi'_n = \alpha_n \xi_n, \quad \eta'_n = \alpha_n \eta_n.$$

Show that

$$\mathbf{E}^{\mathscr{F}_n}(Z'_{n+1}) \le Z'_n + \xi'_n - \eta'_n$$

and that $U_n = Z'_n - \sum_{k=0}^{n-1} (\xi'_k - \eta'_k)$, $n \ge 1$, $U_0 = Z'_0$, is a supermartingale.

A2) Let $a > 0$ and consider the stopping time

$$\tau_a = \inf\left\{n \ge 0;\ \textstyle\sum_{k=0}^{n} (\xi'_k - \eta'_k) > a\right\}.$$

Show that, on $\{\tau_a = +\infty\}$, $(U_n)_{n\ge0}$ converges a.s. to a finite r.v.

A3) Let $\Gamma = \{\sum_{n=1}^{\infty} \beta_n < +\infty, \sum_{n=1}^{\infty} \xi_n < +\infty\}$.

i) Show that, on Γ, α_n converges to $\alpha_\infty > 0$ and that $\sum_{n=1}^{\infty} \xi'_n < +\infty$.

ii) Show that, on Γ, $(Z_n)_{n\ge0}$ converges a.s. to a finite r.v. Z_∞ and that $\sum_{n=1}^{\infty} \eta_n < +\infty$ a.s.

B) (The Robbins–Monro algorithms) We now consider an adapted sequence of random vectors $(X_n, Y_n)_{n\geq 0}$ with values in $\mathbb{R}^d \times \mathbb{R}^d$ and a sequence $(\gamma_n)_{n\geq 0}$ of positive adapted r.v.'s such that

$$X_{n+1} = X_n + \gamma_n Y_{n+1}.$$

Assume that there exist a continuous function $f: \mathbb{R}^d \to \mathbb{R}^d$, a constant $K < +\infty$ and $x_0 \in \mathbb{R}^d$ such that

$$\mathbf{E}^{\mathscr{F}_n}(Y_{n+1}) = f(X_n), \qquad \mathbf{E}^{\mathscr{F}_n}(|Y_{n+1}|^2) \leq K(1 + |X_n|^2) \quad \text{a.s.}$$

and that,

(i) $\sum_{n=1}^{\infty} \gamma_n = +\infty$, $\sum_{n=1}^{\infty} \gamma_n^2 < +\infty$ a.s.

(ii) $f(x_0) = 0$ and $\langle x - x_0, f(x)\rangle < 0$ for $x \neq x_0$.

B1) Let us set $Z_n = |X_n - x_0|^2$. Show that there exists a constant $\tilde{K}$ such that

$$\mathbf{E}^{\mathscr{F}_n}(Z_{n+1}) \leq (1 + \tilde{K}\gamma_n^2)Z_n + \tilde{K}\gamma_n^2 + 2\gamma_n\langle X_n - x_0, f(X_n)\rangle \quad \text{a.s.}$$

B2) Derive from (A) that $(Z_n)_{n\geq 0}$ converges a.s. to a r.v. Z and that the series $\sum_{n=1}^{\infty} \gamma_n \langle X_n - x_0, f(X_n)\rangle$ converges a.s.

B3) Show that $Z = 0$ a.s., i.e., that $X_n \to_{n\to\infty} x_0$ a.s.

Solutions

E3.1 a) If X is a supermartingale, then $\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) \leq X_n$ for $n \in \mathbb{N}$. Therefore the r.v. $U_n = X_n - \mathbf{E}(X_{n+1} \mid \mathscr{F}_n)$ is ≥ 0 a.s.; using the hypothesis, it has mean 0 as

$$\mathbf{E}(\mathbf{E}(X_{n+1} \mid \mathscr{F}_n)) = \mathbf{E}(X_{n+1}) = \mathbf{E}(X_n).$$

Thus $\mathbf{E}(U_n) = 0$ and $U_n = 0$ a.s. Therefore $\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = X_n$ a.s.

b) If X is a martingale the property follows from the optional sampling Theorem 3.3: $\mathbf{E}[X_\tau]$ is equal to $\mathbf{E}[X_0]$ for any bounded stopping time τ.

Conversely, in order to prove the martingale property, one has to show that, for every $A \in \mathscr{F}_n$,

$$\mathbf{E}(X_{n+1} 1_A) = \mathbf{E}(X_n 1_A).$$

The idea is to find two bounded stopping times τ_1, τ_2 such that the equality $\mathbf{E}[X_{\tau_1}] = \mathbf{E}[X_{\tau_2}]$ implies the previous martingale relation. Let us choose, for $A \in \mathscr{F}_n$,

$$\tau_1(\omega) = \begin{cases} n & \text{if } \omega \in A \\ n+1 & \text{if } \omega \in A^c \end{cases}$$

and $\tau_2 \equiv n + 1$; τ_1 is a stopping time: indeed

$$\{\tau_1 \leq k\} = \begin{cases} \emptyset & \text{if } k \leq n-1 \\ A & \text{if } k = n \\ \Omega & \text{if } k \geq n+1 \end{cases}$$

so that $\{\tau_1 \leq k\} \in \mathscr{F}_k$ in any case. Now $X_{\tau_1} = X_n 1_A + X_{n+1} 1_{A^c}$ and the relation $\mathbf{E}[X_{\tau_1}] = \mathbf{E}[X_{n+1}]$ implies

$$\begin{aligned} \mathbf{E}[X_n 1_A] + \mathbf{E}[X_{n+1} 1_{A^c}] &= \mathbf{E}[X_{\tau_1}] = \mathbf{E}[X_{n+1}] = \\ &= \mathbf{E}[X_{n+1} 1_A] + \mathbf{E}[X_{n+1} 1_{A^c}] \end{aligned}$$

from which the martingale property follows by subtraction.

E3.2 Since $|Z_n| \le |X_n| + |Y_n|$, Z_n is integrable. On $\{n+1=\tau\}$, it holds $Y_{n+1} = Y_\tau \ge X_\tau = X_{n+1}$; thus

$$\begin{aligned} Z_{n+1} &= Y_{n+1} 1_{\{n+1<\tau\}} + X_{n+1} 1_{\{n+1\ge\tau\}} = \\ &= Y_{n+1} 1_{\{n+1<\tau\}} + X_{n+1} 1_{\{n+1=\tau\}} + X_{n+1} 1_{\{n\ge\tau\}} \le \\ &\le Y_{n+1} 1_{\{n+1<\tau\}} + Y_{n+1} 1_{\{n+1=\tau\}} + X_{n+1} 1_{\{n\ge\tau\}} = Y_{n+1} 1_{\{n<\tau\}} + X_{n+1} 1_{\{n\ge\tau\}} \end{aligned}$$

Since the events $\{n \ge \tau\}$ and $\{n < \tau\} = \{n \ge \tau\}^c$ belong to $\mathscr{F}_n$, it holds a.s.

$$\begin{aligned} \mathrm{E}(Z_{n+1} \mid \mathscr{F}_n) &\le 1_{\{n<\tau\}} \mathrm{E}(Y_{n+1} \mid \mathscr{F}_n) + 1_{\{n\ge\tau\}} \mathrm{E}(X_{n+1} \mid \mathscr{F}_n) \le \\ &\le 1_{\{n<\tau\}} Y_n + 1_{\{n\ge\tau\}} X_n = Z_n. \end{aligned}$$

$(Z_n)_{n\ge0}$ is therefore a supermartingale. The other case under inquiry is easily treated, either by repeating the argument above with $=$ instead of $\le$, or by applying the result obtained so far twice, obtaining that $(Z_n)_{n\ge0}$ and $(-Z_n)_{n\ge0}$ are both supermartingales.

E3.3 a) As $m \le n$,

$$\mathrm{E}(X_m Y_n \mid \mathscr{F}_m) = X_m \mathrm{E}(Y_n \mid \mathscr{F}_m) = X_m Y_m \qquad \text{a.s.}$$

b) Taking the expectation in the relation proved in (a), due to the symmetric role of X_n and Y_n,

$$\begin{aligned} \mathrm{E}(X_{k-1} Y_k) &= \mathrm{E}(X_{k-1} Y_{k-1}) \\ \mathrm{E}(X_k Y_{k-1}) &= \mathrm{E}(X_{k-1} Y_{k-1}) \end{aligned}$$

from which

$$\mathrm{E}((X_k - X_{k-1})(Y_k - Y_{k-1})) = \mathrm{E}(X_k Y_k) - \mathrm{E}(X_{k-1} Y_{k-1}).$$

Summing up, one gets the requested relation.

c) Thanks to (b), for $X_n = Y_n$,

$$\mathrm{E}(X_n^2) = \mathrm{E}(X_0^2) + \sum_{k=1}^{n} \mathrm{E}((X_k - X_{k-1})^2).$$

It is sufficient now to remark that $\mathrm{Var}(X_n) = \mathrm{E}(X_n^2) - \mathrm{E}(X_n)^2$, that $\mathrm{E}(X_n)^2 = \mathrm{E}(X_0)^2$ and that, the r.v.'s $X_k - X_{k-1}, k = 1, \dots n$ being centered, $\mathrm{Var}(X_k - X_{k-1}) = \mathrm{E}((X_k - X_{k-1})^2)$.

d) For a fixed n and for every r.v. W, $\mathscr{F}_{n-1}$-measurable and square integrable,

$$\mathrm{E}(W(X_n - X_{n-1})) = \mathrm{E}(W \mathrm{E}(X_n - X_{n-1} \mid \mathscr{F}_{n-1})) = 0.$$

Therefore $X_n - X_{n-1}$ is, in particular, orthogonal to each of the r.v.'s $X_0, X_k - X_{k-1}$, $1 \le k < n$.

E3.4 Let us show that $(X_n)_{n\ge1}$ is a martingale:

$$\mathrm{E}(X_{n+1} \mid \mathscr{F}_n) = X_n + \frac{1}{n+1} \mathrm{E}(M_{n+1} - M_n \mid \mathscr{F}_n) = X_n \quad \text{a.s.}$$

since $\mathrm{E}(M_{n+1} \mid \mathscr{F}_n) = M_n$ a.s.

In order to prove the L^2 and a.s. convergence of the martingale $(X_n)_{n\geq 0}$, we know that it is sufficient to check the L^2-boundedness, i.e., $\sup_{n\geq 0} \mathrm{E}(X_n^2) < \infty$. Here,

$$X_n^2 = \sum_{k=1}^{n} \frac{1}{k^2}(M_k - M_{k-1})^2 + 2 \sum_{1\leq i<j\leq n} \frac{1}{ij}(M_j - M_{j-1})(M_i - M_{i-1}). \tag{3.18}$$

If $1 \leq i < j \leq n$,

$$\mathrm{E}((M_j - M_{j-1})(M_i - M_{i-1})) = \mathrm{E}((M_i - M_{i-1})\underbrace{\mathrm{E}(M_j - M_{j-1} \mid \mathscr{F}_{j-1})}_{=0}).$$

By (3.18),

$$\mathrm{E}(X_n^2) = \sum_{k=1}^{n} \frac{1}{k^2}\mathrm{E}((M_k - M_{k-1})^2). \tag{3.19}$$

But $(M_k - M_{k-1})^2 \leq (2K)^2 = 4K^2$, thus

$$\mathrm{E}(X_n^2) \leq 4K^2 \sum_{k=1}^{n} \frac{1}{k^2} \leq 4K^2 \sum_{k=1}^{\infty} \frac{1}{k^2} = C < +\infty$$

and the requested convergence result follows.

♦ In this exercise the hypothesis could have been weakened by assuming only that $(M_n)_{n\geq 0}$ is bounded in L^2.

E3.5 1) As Y_n and $\mathscr{F}_{n-1}$ are independent and using (3.6), one has a.s.

$$\begin{aligned}\mathrm{E}(Z_n^u \mid \mathscr{F}_{n-1}) &= \mathrm{E}(\exp(uX_n - \tfrac{1}{2}\sigma^2 nu^2) \mid \mathscr{F}_{n-1}) = \\ &= \exp(uX_{n-1} - \tfrac{1}{2}\sigma^2 nu^2)\mathrm{E}(\exp(uY_n) \mid \mathscr{F}_{n-1}) = \\ &= \exp(uX_{n-1} - \tfrac{1}{2}\sigma^2 nu^2)\mathrm{E}(\exp(uY_n)) = \\ &= \exp(uX_{n-1} - \tfrac{1}{2}\sigma^2 (n-1)u^2) = Z_{n-1}^u.\end{aligned}$$

2) The positive martingale $(Z_n^u)_{n\geq 1}$ converges a.s. and its limit Z_∞^u is $< +\infty$ a.s. (Theorem 3.8). By the law of large numbers, $\frac{1}{n}X_n \to_{n\to\infty} 0$ a.s. so that

$$\frac{1}{n}\left(uX_n - \frac{nu^2\sigma^2}{2}\right) \underset{n\to\infty}{\to} -\frac{u^2\sigma^2}{2}.$$

This limit being < 0 for $u \neq 0$ implies

$$uX_n - \frac{nu^2\sigma^2}{2} \underset{n\to\infty}{\to} -\infty;$$

hence, $Z_n^u \to_{n\to\infty} 0$. In this case the martingale cannot be regular, as $\mathrm{E}(Z_n^u) = 1$, so that the convergence cannot take place in L^1. If $u = 0$, $Z_n^u \equiv 1$ and the martingale is obviously regular.

E3.6 1a) One has $M_n = M_0 + \sum_{k=1}^n (M_k - M_{k-1})$ and, since the hypotheses imply that the r.v.'s $M_0, M_1 - M_0, \dots, M_n - M_{n-1}$ are independent,

$$\mathrm{Var}(M_n) = \mathrm{Var}(M_0) + \sum_{k=1}^{n} \mathrm{Var}(M_k - M_{k-1}) = \sum_{k=0}^{n} \sigma_k^2.$$

Let us remark however that this formula was already proved in Exercise 3.3 (c). In particular *it is not* a consequence of the independence of the increments and it remains true for every square integrable martingale.

1b) We know (•3.8) that the associated increasing process (•3.8) $\langle M\rangle$ is characterized by $\langle M\rangle_0 = 0$ and

$$\mathbf{E}((M_{k+1}-M_k)^2 \mid \mathscr{F}_k) = \langle M\rangle_{k+1} - \langle M\rangle_k \qquad \text{a.s.}$$

As $M_{k+1}-M_k$ is independent of $\mathscr{F}_k$, one has (•1.11)

$$\mathbf{E}((M_{k+1}-M_k)^2 \mid \mathscr{F}_k) = \mathbf{E}((M_{k+1}-M_k)^2) = \sigma_k^2$$

so that

$$\langle M\rangle_n = \sum_{k=1}^{n} \sigma_k^2.$$

2) Let us recall that, if two r.v.'s X and Y with values, respectively, in $\mathbb{R}^m$ and $\mathbb{R}^d$ are such that (X, Y) is Gaussian, then they are independent if and only if $\mathrm{Cov}(X_i, Y_j) = 0$ for every $i = 1, \dots, m$, $j = 1, \dots, d$, i.e., if and only if their components are uncorrelated.

2a) In order to prove that the r.v. $M_{n+1}-M_n$ is independent of $\mathscr{F}_n$, it suffices to show that it is independent of $(M_0, M_1, \dots, M_n)$. Being a linear transformation of $(M_0, M_1, \dots, M_n, M_{n+1})$, the vector $(M_0, M_1, \dots, M_n, M_{n+1}-M_n)$ is Gaussian. $M_{n+1}-M_n$ is centered so that it is enough to prove that

$$\mathbf{E}((M_{n+1}-M_n)M_k) = 0, \quad k = 0, 1, \dots, n.$$

But this is obviously true since

$$\mathbf{E}[(M_{n+1}-M_n)M_k] = \mathbf{E}[M_k \underbrace{\mathbf{E}^{\mathscr{F}_n}(M_{n+1}-M_n)}_{=0}].$$

2b) We have seen that $\langle M\rangle_n = \sum_{k=1}^{n} \sigma_k^2$, i.e., $\langle M\rangle_n - \langle M\rangle_{n-1} = \sigma_n^2$. Therefore a.s.

$$\begin{aligned}\mathbf{E}(\mathrm{e}^{\theta M_n - \frac{1}{2}\theta^2\langle M\rangle_n} \mid \mathscr{F}_{n-1}) &= \mathrm{e}^{\theta M_{n-1} - \frac{1}{2}\theta^2\langle M\rangle_n}\, \mathbf{E}(\mathrm{e}^{\theta(M_n - M_{n-1})} \mid \mathscr{F}_{n-1}) = \\ &= \mathrm{e}^{\theta M_{n-1} - \frac{1}{2}\theta^2\langle M\rangle_n}\, \mathbf{E}(\mathrm{e}^{\theta(M_n - M_{n-1})}).\end{aligned}$$

But it is known that, for a centered Gaussian r.v. X with variance σ^2, it holds $\mathbf{E}(\mathrm{e}^{\theta X}) = \mathrm{e}^{\frac{1}{2}\sigma^2\theta^2}$, so that

$$\mathbf{E}(\mathrm{e}^{\theta(M_n - M_{n-1})}) = \mathrm{e}^{\frac{1}{2}\theta^2\sigma_n^2} = \mathrm{e}^{\frac{1}{2}\theta^2(\langle M\rangle_n - \langle M\rangle_{n-1})}$$

from which we get

$$\mathbf{E}(\mathrm{e}^{\theta M_n - \frac{1}{2}\theta^2\langle M\rangle_n} \mid \mathscr{F}_{n-1}) = \mathrm{e}^{\theta M_{n-1} - \frac{1}{2}\theta^2\langle M\rangle_{n-1}}.$$

It remains to study the convergence of the martingale. As it is a positive martingale, one can use Theorem 3.8 and obtain that it converges a.s.

E3.7 1) For $f \geq 0$, it holds a.s.

$$\begin{aligned} &\mathbf{E}(f(X_{n+1}) \mid \mathscr{F}_n) = \\ &= \mathbf{E}\big(f(X_{n+1})1_{\{X_{n+1}=X_n/2\}} + f(X_{n+1})1_{\{X_{n+1}=(1+X_n)/2\}} \mid \mathscr{F}_n\big) = \\ &= \mathbf{E}\big(f(\tfrac{X_n}{2})1_{\{X_{n+1}=X_n/2\}} + f(\tfrac{X_n+1}{2})1_{\{X_{n+1}=(X_n+1)/2\}} \mid \mathscr{F}_n\big) = \\ &= f(\tfrac{X_n}{2})(1-X_n) + f(\tfrac{X_n+1}{2})X_n. \end{aligned} \tag{3.20}$$

In particular

$$\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = \frac{X_n}{2}(1-X_n) + \frac{X_n+1}{2}X_n = X_n \quad \text{a.s.}$$

The martingale $(X_n)_{n\geq 0}$ is bounded and converges to a r.v. Z, a.s. and in L^p for every $p \geq 1$.

2) It holds

$$\mathbf{E}((X_{n+1}-X_n)^2) = \mathbf{E}(\mathbf{E}((X_{n+1}-X_n)^2 \mid \mathscr{F}_n)). \tag{3.21}$$

But

$$\mathbf{E}((X_{n+1}-X_n)^2 \mid \mathscr{F}_n) = \mathbf{E}(X_{n+1}^2 - 2X_{n+1}X_n + X_n^2 \mid \mathscr{F}_n) \quad \text{a.s.} \tag{3.22}$$

If one chooses $f(x) = x^2$ in relation (3.20),

$$\mathbf{E}(X_{n+1}^2 \mid \mathscr{F}_n) = \Big(\frac{X_n}{2}\Big)^2(1-X_n) + \Big(\frac{X_n+1}{2}\Big)^2 X_n = \frac{X_n}{4}(1+3X_n) \quad \text{a.s.}$$

We replace in (3.22), taking into account that $\mathbf{E}(X_{n+1} \mid X_n) = X_n$:

$$\mathbf{E}((X_{n+1}-X_n)^2 \mid \mathscr{F}_n) = \frac{X_n}{4}(1+3X_n) - 2X_n^2 + X_n^2 = \frac{1}{4}X_n(1-X_n) \quad \text{a.s.}$$

3) Since $X_n \to_{n\to\infty} Z$ a.s.,

$$(X_{n+1}-X_n)^2 \underset{n\to\infty}{\to} 0 \quad \text{and} \quad X_n(1-X_n) \underset{n\to\infty}{\to} Z(1-Z) \quad \text{a.s.}$$

As these r.v.'s are bounded, the convergence takes place also in L^1. Therefore

$$\begin{aligned} &\mathbf{E}(X_n(1-X_n)) \underset{n\to\infty}{\to} \mathbf{E}(Z(1-Z)) \\ &\mathbf{E}((X_{n+1}-X_n)^2) \underset{n\to\infty}{\to} 0. \end{aligned}$$

But due to (2) $\mathbf{E}((X_{n+1}-X_n)^2) = \frac{1}{4}\mathbf{E}(X_n(1-X_n))$, from which we deduce $\mathbf{E}(Z(1-Z)) = 0$. One concludes that Z can only take the values 1 and 0 and has a Bernoulli distribution. It remains to compute $p = \mathbf{P}(Z=1)$. But

$$p = \mathbf{P}(Z=1) = \mathbf{E}(Z) = \lim_{n\to\infty} \mathbf{E}(X_n) = \mathbf{E}(X_0) = a.$$

E3.8 1) X_n represents the probability of choosing a white ball from the urn at time n. Thus $\mathbf{P}(Y_{n+1} = Y_n + 1 \mid \mathscr{F}_n) = X_n$ a.s., $\mathbf{P}(Y_{n+1} = Y_n \mid \mathscr{F}_n) = 1 - X_n$ a.s.,

$$\mathbf{E}(Y_{n+1} \mid \mathscr{F}_n) = (Y_n+1)X_n + Y_n(1-X_n) = X_n + Y_n = (n+2)X_n \quad \text{a.s.}$$

and

$$\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = \mathbf{E}\Big(\frac{1}{n+2}\,Y_{n+1} \,\Big|\, \mathscr{F}\Big) = \frac{1}{n+2}\,\mathbf{E}(Y_{n+1} \mid \mathscr{F}_n) = X_n \quad \text{a.s.}$$

Since $0 \le X_n \le 1$ for every $k \ge 0$, $(X_n)_{n\ge 1}$ is a bounded martingale and it converges a.s. to a r.v. U; moreover by Lebesgue's theorem $\lim_{n\to\infty} \mathbf{E}(X_n^k) = \mathbf{E}(U^k)$.

2) One has a.s.

$$\begin{aligned}\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n) &= \mathbf{E}(Z_{n+1} 1_{\{Y_{n+1}=Y_n\}} \mid \mathscr{F}_n) + \mathbf{E}(Z_{n+1} 1_{\{Y_{n+1}=Y_n+1\}} \mid \mathscr{F}_n) = \\ &= \frac{n+1-Y_n}{n+1} \, \frac{Y_n(Y_n+1)\dots(Y_n+k-1)}{(n+2)(n+3)\dots(n+k+1)} + \\ &\qquad + \frac{Y_n}{n+1} \, \frac{(Y_n+1)(Y_n+2)\dots(Y_n+k)}{(n+2)(n+3)\dots(n+k+1)} = \\ &= \frac{Y_n(Y_n+1)\dots(Y_n+k-1)[(n+1-Y_n)+(Y_n+k)]}{(n+1)(n+2)\dots(n+k+1)} = \\ &= \frac{Y_n(Y_n+1)\dots(Y_n+k-1)}{(n+1)(n+2)\dots(n+k)} = Z_n.\end{aligned}$$

Since $\frac{1}{n+1} Y_n \to_{n\to\infty} U$ a.s. then also $\frac{1}{n+r+1}(Y_n + r) \to_{n\to\infty} U$ a.s. for every fixed r, so that $Z_n \to_{n\to\infty} U^k$ a.s. and, as $0 \le Z_n \le 1$, $\lim_{n\to\infty} \mathbf{E}(Z_n) = \mathbf{E}(U^k)$. $(Z_n)_{n\ge 1}$ being a martingale,

$$\mathbf{E}(Z_n) = \mathbf{E}(Z_1) = \frac{1}{k+1} = \mathbf{E}(U^k).$$

3) We know that, for every t,

$$\left| \phi(t) - \sum_{k=0}^{n} \frac{\phi^{(k)}(0)}{k!} t^k \right| \le \sup_{\tau\in[0,t]} \frac{|\phi^{(n+1)}(\tau)|}{(n+1)!} |t|^{n+1}.$$

But

$$|\phi^{(n+1)}(\tau)| = |i^{n+1} \mathbf{E}(X^{n+1} \mathrm{e}^{i\tau X})| \le \mathbf{E}(|X^{n+1}|) \le M^{n+1}.$$

Therefore

$$\left| \phi(t) - \sum_{k=0}^{n} \frac{\phi^{(k)}(0)}{k!} t^k \right| \le \frac{|tM|^{n+1}}{(n+1)!}.$$

The right-hand side converges to 0 as $n \to \infty$ and this ends the proof of (3).

4) Denote by ϕ the characteristic function of U. Then $\phi^{(k)}(0) = i^k \mathbf{E}(U^k) = \frac{i^k}{k+1}$ and we may apply (3.8), as $0 \le U \le 1$; therefore

$$\phi(t) = \sum_{k=0}^{\infty} \frac{\phi^{(k)}(0)}{k!} t^k = \sum_{k=0}^{\infty} \frac{(it)^k}{(k+1)!} = \frac{\mathrm{e}^{it}-1}{it} = \int_0^1 \mathrm{e}^{itx}\, dx$$

and this shows that U has a distribution that is uniform on $[0, 1]$.

E3.9 1) $A \in \mathscr{F}^\infty$ implies $A \in \mathscr{F}^{n+1}$ for every n; therefore A is independent of $\mathscr{F}_n$ and a.s. $\mathbf{E}^{\mathscr{F}_n}(1_A) = \mathbf{E}(1_A) = \mathbf{P}(A)$. On the other hand (Theorem 3.8), $\mathbf{E}^{\mathscr{F}_n}(1_A) \to_{n\to\infty} \mathbf{E}^{\mathscr{F}_\infty}(1_A) = 1_A$ a.s. (as $\mathscr{F}^\infty \subset \mathscr{F}_\infty$!). So $\mathbf{P}(A)$ may only take the values 0 or 1.

2) Let $F(t) = \mathbf{P}(X \le t)$ be the distribution function of X. Thus F is right-continuous and increases from 0 to 1; but, thanks to (1), $F(t)$ may only take the values 0 or 1. Therefore there exists $a \in \mathbb{R}$ such that $F(t) = 1_{[a,+\infty[}$ and $X = a$ a.s.

E3.10 1) The r.v.'s Y_n being independent and having the same distribution, the law of

$(Y_1, \ldots, Y_n)$ is invariant by permutation; therefore, for $1 \le m \le n$, the distributions of $(Y_1, S_n, Y_{n+1}, Y_{n+2}, \ldots)$ and $(Y_m, S_n, Y_{n+1}, Y_{n+2}, \ldots)$ coincide, so that $\mathrm{E}^{\mathscr{G}_n}(Y_m) = \mathrm{E}^{\mathscr{G}_n}(Y_1)$ a.s. Therefore

$$\mathrm{E}^{\mathscr{G}_n}(Y_1) = \frac{1}{n}\sum_{m=1}^{n}\mathrm{E}^{\mathscr{G}_n}(Y_m) = \mathrm{E}^{\mathscr{G}_n}\Big(\frac{1}{n}\sum_{m=1}^{n}Y_m\Big) = \frac{S_n}{n} \quad \text{a.s.}$$

since $S_n = \sum_{m=1}^{n} Y_m$ is $\mathscr{G}_n$-measurable.

2) One remarks that the σ-algebras $\mathscr{G}_n$ are decreasing in n. By •3.13, as $Y_1 \in L^1$, $(\mathrm{E}^{\mathscr{G}_n}(Y_1))_{n\ge 0}$ converges a.s. and in L^1 to $X = \mathrm{E}^{\mathscr{G}_\infty}(Y_1)$, where $\mathscr{G}_\infty = \bigcap_{n\ge 0}\mathscr{G}_n$.

3) The r.v. $X = \lim_{n\to\infty}\frac{1}{n}S_n$ is, for every p, $\sigma(X_p, X_{p+1}, \ldots)$-measurable, since the limit does not depend on the first n terms. Therefore by Kolmogorov's 0-1 law (see also Exercise 3.9), X is constant, so that $X = a$ a.s. for $a = \mathrm{E}(X)$. But, as the convergence takes place in L^1, $\mathrm{E}(X) = \lim_{n\to\infty}\mathrm{E}(\frac{1}{n}S_n) = \mathrm{E}(Y_1)$.

We have shown that $(\frac{1}{n}S_n)_n$ converges to $\mathrm{E}(Y_1)$ a.s. and in L^1.

E3.11 1) The process $(\langle M\rangle_{n\wedge\tau})_{n\ge 0}$ is equal to 0 at time 0; it is increasing and predictable. Indeed, $\langle M\rangle_{n\wedge\tau} = \sum_{k=0}^{n-1}\langle M\rangle_k 1_{\{\tau=k\}} + \langle M\rangle_n 1_{\{\tau\le n-1\}^c}$ and all these terms are clearly $\mathscr{F}_{n-1}$-measurable and positive. Finally $N_n = M_n^2 - \langle M\rangle_n$ is, by hypothesis, a martingale and this also holds for $N_{n\wedge\tau} = M^2_{n\wedge\tau} - \langle M\rangle_{n\wedge\tau}$, which shows that $\langle M^\tau\rangle_n = \langle M\rangle_{n\wedge\tau}$.

2) Indeed $\{\tau_a = n\} = \{\langle M\rangle_1 \le a, \ldots, \langle M\rangle_n \le a, \langle M\rangle_{n+1} > a\} \in \mathscr{F}_n$, as $\langle M\rangle_{n+1}$ is $\mathscr{F}_n$-measurable (the main point here is that the process $(\langle M\rangle_n)_{n\ge 0}$ is predictable).

3) Due to (1), $\mathrm{E}[(M_n^{\tau_a})^2] = \mathrm{E}[\langle M\rangle_{n\wedge\tau_a}] \le a$ by the definition of τ_a. The martingale $(M_n^{\tau_a})_{n\ge 0}$ is bounded in L^2 and therefore converges a.s. and in L^2 (Theorem 3.7).

4) By (3) $(M_n)_{n\ge 0}$ converges a.s. on $\{\tau_a = +\infty\}$, as the two sequences $(M_n(\omega))_{n\ge 0}$ and $(M_n^{\tau_a}(\omega))_{n\ge 0}$ coincide if $\tau_a(\omega) = +\infty$; but $\{\langle M\rangle_\infty < +\infty\} = \bigcup_a\{\tau_a = +\infty\}$.

E3.12 1) As A_n is $\mathscr{F}_{n-1}$-measurable, it holds a.s.

$$\begin{aligned}\mathrm{E}^{\mathscr{F}_{n-1}}(X_n - X_{n-1}) &= \mathrm{E}^{\mathscr{F}_{n-1}}\Big(\frac{M_n - M_{n-1}}{1+A_n}\Big) = \\ &= \frac{1}{1+A_n}\mathrm{E}^{\mathscr{F}_{n-1}}(M_n - M_{n-1}) = 0,\end{aligned}$$

i.e., $(X_n)_{n\ge 0}$ is a square integrable martingale, as $(1+A_n)^{-1} \le 1$. In the same way, one has a.s.

$$\begin{aligned}\mathrm{E}^{\mathscr{F}_{n-1}}[(X_n - X_{n-1})^2] &= \frac{1}{(1+A_n)^2}\mathrm{E}^{\mathscr{F}_{n-1}}[(M_n - M_{n-1})^2] = \\ &= \frac{A_n - A_{n-1}}{(1+A_n)^2} \le \frac{1}{1+A_{n-1}} - \frac{1}{1+A_n}.\end{aligned}$$

But $\langle X\rangle_n - \langle X\rangle_{n-1} = \mathrm{E}^{\mathscr{F}_{n-1}}[(X_n - X_{n-1})^2]$, so that $\langle X\rangle_n \le 1 - (1+A_n)^{-1} \le 1$. By Exercise 3.11 we deduce that $(X_n)_{n\ge 0}$ converges a.s

2a) Let $\varepsilon > 0$ and p such that $\sup_{n\ge p}|u_n - u| < \varepsilon$ and

$$\rho_n = \frac{1}{a_n}\sum_{k=1}^{n}(a_k - a_{k-1})u_k - u.$$

Then, for $n > p$,

$$\rho_n = \frac{1}{a_n}\sum_{k=1}^{p}(a_k - a_{k-1})u_k + \frac{1}{a_n}\sum_{k=p+1}^{n}(a_k - a_{k-1})(u_k - u) + \frac{a_n - a_p}{a_n}u - u$$

and

$$|\rho_n| \le \left|\frac{1}{a_n}\sum_{k=1}^{p}(a_k - a_{k-1})u_k\right| + \varepsilon\frac{a_n - a_p}{a_n} + \frac{|a_p|}{|a_n|}|u|$$

from which $\overline{\lim}_{n\to\infty}|\rho_n| \le \varepsilon$.

2b) Let us set $u_n = \sum_{k=1}^{n} x_k/a_k$ so that $x_k/a_k = u_k - u_{k-1}$ and $u_n \to_{n\to\infty} u = \sum_{k=1}^{\infty} x_k/a_k$. Therefore, by (2a),

$$\frac{s_n}{a_n} = \frac{1}{a_n}\sum_{k=1}^{n} a_k(u_k - u_{k-1}) = u_n - \frac{1}{a_n}\sum_{k=1}^{n} u_{k-1}(a_k - a_{k-1}) \underset{n\to\infty}{\to} u - u = 0.$$

3) By the a.s. convergence of $(X_n)_{n\ge 0}$ and Kronecker's lemma applied to $a_n = 1 + A_n$ and $x_n = M_n - M_{n-1}$, we conclude that on $\{A_\infty = \infty\}$ it holds $M_n/(1 + A_n) \to_{n\to\infty} 0$ a.s.; but, on $\{A_\infty = \infty\}$, it holds $A_n > 0$ for sufficiently large n and

$$\frac{M_n}{A_n} = \frac{M_n}{1 + A_n}\,\frac{1 + A_n}{A_n} \underset{n\to\infty}{\to} 0 \qquad \text{a.s.}$$

E3.13 a) Note that X_n^θ is positive and integrable, as $S_n \le n$. Then a.s.

$$\begin{aligned} \mathrm{E}^{\mathscr{F}_n}(X_{n+1}^\theta) &= \frac{1}{(\cosh\theta)^{n+1}}\mathrm{E}^{\mathscr{F}_n}(\mathrm{e}^{\theta S_{n+1}}) = \\ &= \frac{1}{(\cosh\theta)^{n+1}}\mathrm{E}^{\mathscr{F}_n}(\mathrm{e}^{\theta S_n}\,\mathrm{e}^{\theta Z_{n+1}}) = \frac{\mathrm{e}^{\theta S_n}}{(\cosh\theta)^{n+1}}\mathrm{E}^{\mathscr{F}_n}(\mathrm{e}^{\theta Z_{n+1}}). \end{aligned} \tag{3.23}$$

As Z_{n+1} is independent of $\mathscr{F}_n$,

$$\mathrm{E}^{\mathscr{F}_n}(\mathrm{e}^{\theta Z_{n+1}}) = \mathrm{E}(\mathrm{e}^{\theta Z_{n+1}}) = \frac{1}{2}(\mathrm{e}^{\theta} + \mathrm{e}^{-\theta}) = \cosh\theta \quad \text{a.s.}$$

so that (3.23) becomes

$$\mathrm{E}^{\mathscr{F}_n}(X_{n+1}^\theta) = \frac{\mathrm{e}^{\theta S_n}}{(\cosh\theta)^n} = X_n^\theta \quad \text{a.s.}$$

and $(X_n^\theta)_{n\ge 0}$ is a martingale; by the optional sampling Theorem 3.3 this is also the case for $(X_{n\wedge\tau}^\theta)_{n\ge 0}$. The latter is moreover bounded. Indeed due to the definition of τ and as S_n cannot cross the level a without taking the value a, $S_{n\wedge\tau} \le a$ (this being true even on $\{\tau = +\infty\}$). Therefore, $\cosh\theta$ being always ≥ 1,

$$0 \le X_{n\wedge\tau}^\theta \le \mathrm{e}^{\theta a}.$$

b1) Let $\theta > 0$. $(X_{n\wedge\tau}^\theta)_{n\ge 0}$ is a martingale that is bounded in L^2. Therefore it converges in L^2 (and thus in L^1) and a.s. to a r.v. W^θ. On $\{\tau < \infty\}$ $W^\theta = \lim_{n\to\infty} X_{n\wedge\tau}^\theta = X_\tau^\theta = \mathrm{e}^{\theta a}(\cosh\theta)^{-\tau}$; on the other hand $W^\theta = 0$ on $\{\tau = \infty\}$, since in this case $S_n \le a$ for every n whereas the denominator tends to $+\infty$. Therefore (3.9) is proved.

b2) One has $W^\theta \to_{\theta\to 0} 1_{\{\tau<+\infty\}}$ and as, for $\theta \le 1$,

$$W^\theta = \frac{e^{\theta a}}{(\cosh\theta)^\tau} 1_{\{\tau<+\infty\}} \le e^a,$$

one may apply Lebesgue's theorem and get

$$\mathbf{E}(W^\theta) \underset{\theta\to 0}{\to} \mathbf{P}(\tau < +\infty).$$

The martingale $(X^\theta_{n\wedge\tau})_{n\ge 0}$ being regular, for every $\theta > 0$ and $n \ge 0$, it holds $1 = \mathbf{E}(X^\theta_0) = \mathbf{E}(X^\theta_{n\wedge\tau}) = \mathbf{E}(W^\theta)$, which implies $\mathbf{P}(\tau < +\infty) = 1$. Therefore $W^\theta = e^{\theta a}(\cosh\theta)^{-\tau}$ a.s. whence $e^{-\theta a} = \mathbf{E}((\cosh\theta)^{-\tau})$.

E3.14 a) Z_{n+1} being independent of $\mathscr{F}_n$, one has a.s.

$$\mathbf{E}(\sin(\lambda Z_{n+1}) \mid \mathscr{F}_n) = \mathbf{E}(\sin(\lambda Z_{n+1})) = 0$$
$$\mathbf{E}(\cos(\lambda Z_{n+1}) \mid \mathscr{F}_n) = \mathbf{E}(\cos(\lambda Z_{n+1})) = \cos\lambda.$$

Therefore a.s.

$$\begin{aligned}\mathbf{E}^{\mathscr{F}_n}(X_{n+1} \mid \mathscr{F}_n) &= (\cos\lambda)^{-(n+1)}\mathbf{E}(\cos(\lambda(S_n + Z_{n+1}))) = \\ &= (\cos\lambda)^{-(n+1)}\mathbf{E}^{\mathscr{F}_n}\big(\cos(\lambda S_n)\cos(\lambda Z_{n+1}) - \sin(\lambda S_n)\sin(\lambda Z_{n+1})\big) = \\ &= (\cos\lambda)^{-(n+1)}\cos(\lambda S_n)\mathbf{E}(\cos(\lambda Z_{n+1})) = (\cos\lambda)^{-n}\cos(\lambda S_n) = X_n.\end{aligned}$$

b) $n\wedge\tau$ being a bounded stopping time, $\mathbf{E}(X_{n\wedge\tau}) = \mathbf{E}(X_0) = 1$. On the other hand

$$\mathbf{E}(X_{n\wedge\tau}) = \mathbf{E}\big((\cos\lambda)^{-n\wedge\tau}\cos(\lambda S_{n\wedge\tau})\big) \ge \cos(\lambda a)\mathbf{E}((\cos\lambda)^{-n\wedge\tau}). \tag{3.24}$$

as $|S_{n\wedge\tau}| \le a$ and $\lambda S_{n\wedge\tau} \in [-\lambda a, \lambda a] \subset]-\frac{\pi}{2}, \frac{\pi}{2}[$.

c) Since $\cos(\lambda a) > 0$, one gets from (b) that $\mathbf{E}((\cos\lambda)^{-n\wedge\tau}) \le \cos(\lambda a)^{-1}$. As $n \to \infty$, $(\cos\lambda)^{-n\wedge\tau} \uparrow (\cos\lambda)^{-\tau}$ and, by the monotone convergence theorem,

$$\mathbf{E}((\cos\lambda)^{-\tau}) \le \cos(\lambda a)^{-1} \tag{3.25}$$

and therefore, as $\cos\lambda < 1$ due to the assumptions made on λ, $\mathbf{P}(\tau = +\infty) = 0$. Thus $|S_{n\wedge\tau}| \to_{n\to\infty} |S_\tau| = a$ a.s. and

$$X_{n\wedge\tau} = (\cos\lambda)^{-n\wedge\tau}\cos(\lambda S_{n\wedge\tau}) \underset{n\to\infty}{\to} (\cos\lambda)^{-\tau}\cos(\lambda a) = X_\tau. \tag{3.26}$$

As

$$|X_{n\wedge\tau}| = |(\cos\lambda)^{-n\wedge\tau}\cos(\lambda S_{n\wedge\tau})| \le (\cos\lambda)^{-\tau},$$

thanks to (3.25) one may apply Lebesgue's theorem and deduce that $X_{n\wedge\tau} \to_{n\to\infty} X_\tau$ in L^1. Since it converges in L^1, the martingale is regular (see Theorem 3.12).

d) Thanks to (c)

$$1 = \mathbf{E}(X_\tau) = \cos(\lambda a)\mathbf{E}((\cos\lambda)^{-\tau})$$

so that $\mathbf{E}((\cos\lambda)^{-\tau}) = (\cos\lambda a)^{-1}$. If $\eta = \log((\cos\lambda)^{-1})$, then $\eta > 0$ and $\mathbf{E}(e^{\eta\tau}) = \mathbf{E}((\cos\lambda)^{-\tau}) < +\infty$; thus, for every $p \ge 1$, $\mathbf{E}(\tau^p) < +\infty$ since for every p there exists $\alpha(p)$ such that $\tau^p \le \alpha(p)\,e^{\eta\tau}$.

♦ In Exercises 3.13 and 3.14 it has been proved that, for the simple symmetric random walk, the two stopping times

$$\tau_1 = \inf\{n \geq 0, S_n = a\}$$
$$\tau_2 = \inf\{n \geq 0, |S_n| = a\}$$

are both a.s. finite. But the first one is not integrable (see the observation at the end of Exercise 3.23 and Exercise 3.25 (A4)) whereas the second one has finite moments of all orders (and even some exponential moments).

E3.15 a) Let us set $\mathscr{F}_n = \sigma(U_1, \dots, U_n)$. Observe that $Z_{n+1} = Z_n(\frac{q}{p})^{U_{n+1}}$. As U_{n+1} is independent of $\mathscr{F}_n$ and Z_n is $\mathscr{F}_n$-measurable,

$$\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n) = Z_n \mathbf{E}\big((\tfrac{q}{p})^{U_{n+1}}\big) =$$
$$= Z_n\big[\tfrac{q}{p}\mathbf{P}(U_{n+1} = 1) + \tfrac{p}{q}\mathbf{P}(U_{n+1} = -1)\big] = Z_n[\tfrac{q}{p}\cdot p + \tfrac{p}{q}\cdot q] = Z_n.$$

$(Z_n)_{n\geq 0}$ is therefore a martingale, obviously positive.

b) Note that the first inequality is obvious (and not interesting!) if $p > q$. Suppose therefore $p < q$. Then $\sup_{n\geq 0} S_n \geq k$ if and only if

$$\sup_{n\geq 0} Z_n = \sup_{n\geq 0}\Big(\frac{q}{p}\Big)^{S_n} \geq \Big(\frac{q}{p}\Big)^k.$$

By the maximal inequality (Theorem 3.5) applied to the positive martingale $(Z_n)_{n\geq 0}$,

$$\mathbf{P}\Big(\sup_{n\geq 0} S_n \geq k\Big) = \mathbf{P}\Big(\sup_{n\geq 0} Z_n \geq \Big(\frac{q}{p}\Big)^k\Big) \leq \Big(\frac{p}{q}\Big)^k \mathbf{E}(Z_0) = \Big(\frac{p}{q}\Big)^k.$$

By the relation $\mathbf{E}(X) = \sum_{k\geq 1}\mathbf{P}(X \geq k)$, which holds true for every positive integer valued r.v. X,

$$\mathbf{E}\Big(\sup_{n\geq 0} S_n\Big) \leq \sum_{k\geq 1}\mathbf{P}\Big(\sup_{n\geq 0} S_n \geq k\Big) \leq \sum_{k\geq 1}\Big(\frac{p}{q}\Big)^k = \frac{\frac{p}{q}}{1-\frac{p}{q}} = \frac{p}{q-p}.$$

♦ We proved in Theorem 3.5 that, if $(X_n)_{n\geq 0}$ is a positive supermartingale, then, for every $a > 0$, $\mathbf{P}(\sup_{k\geq 0} X_k > a) \leq \frac{1}{a}\mathbf{E}(X_0)$. But, for every ε such that $0 \leq \varepsilon \leq a$,

$$\mathbf{P}\Big(\sup_{k\geq 0} X_k \geq a\Big) \leq \mathbf{P}\Big(\sup_{k\geq 0} X_k > a - \varepsilon\Big) \leq \frac{1}{a-\varepsilon}\mathbf{E}(X_0).$$

Hence also $\mathbf{P}(\sup_{k\geq 0} X_k \geq a) \leq \frac{1}{a}\mathbf{E}(X_0)$.

E3.16 1) By the strong law of large numbers,

$$\lim_{n\to\infty}\frac{S_n}{n} = m < 0 \qquad \text{a.s.}$$

so that $\lim_{n\to\infty} S_n = -\infty$, which implies $W = \sup_{n\geq 0} S_n < +\infty$ a.s.

2) As X_{n+1} is independent of $\mathscr{B}_n$, one has a.s.

$$\mathbf{E}(e^{\lambda S_{n+1}} \mid \mathscr{B}_n) = e^{\lambda S_n}\mathbf{E}(e^{\lambda X_{n+1}} \mid \mathscr{B}_n) = e^{\lambda S_n}\mathbf{E}(e^{\lambda X_{n+1}}) = e^{\lambda S_n} e^{\lambda^2\sigma^2/2} e^{\lambda m}.$$

3) From the previous question, such a λ_0 should satisfy

$$\frac{\lambda_0^2\sigma^2}{2} + \lambda_0 m = 0.$$

This equation has the roots $\lambda_0 = 0$, $\lambda_0 = -2m/\sigma^2 > 0$. The latter is therefore the requested value of λ_0.

4) One has $e^{\lambda_0 W} = \sup_{n\geq 0} e^{\lambda_0 S_n}$. Thanks to one of the maximal inequalities (Theorem 3.5 (i)),

$$\mathbf{P}(e^{\lambda_0 W} > a) = \mathbf{P}\Big(\sup_{n\geq 0} e^{\lambda_0 S_n} > a\Big) \leq \frac{1}{a}\,\mathbf{E}(e^{\lambda_0 S_0}) = \frac{1}{a}$$

so that, by choosing $a = e^{\lambda_0 t}$,

$$\mathbf{P}(W > t) = \mathbf{P}(e^{\lambda_0 W} > e^{\lambda_0 t}) \leq e^{-\lambda_0 t}.$$

5) For $\lambda < \lambda_0$,

$$\mathbf{E}(e^{\lambda W}) = 1 + \lambda\mathbf{E}\Big(\int_0^W e^{\lambda t}\,dt\Big) = 1 + \lambda\int_0^{+\infty} \mathbf{E}(1_{[0,W[}(t))\,e^{\lambda t}\,dt =$$
$$= 1 + \lambda\int_0^{+\infty} e^{\lambda t}\,\mathbf{P}(W > t)\,dt \leq 1 + \lambda\int_0^{+\infty} e^{-(\lambda_0-\lambda)t}\,dt \leq$$
$$\leq 1 + \frac{\lambda}{\lambda_0 - \lambda} < +\infty.$$

As, for a fixed $\lambda > 0$, there exists a constant c_p such that, for any $x > 0$, $x^p \leq c_p \cdot e^{\lambda x}$, $\mathbf{E}(W^p) \leq c_p\mathbf{E}(e^{\lambda W}) < +\infty$ for every $p \geq 0$.

E3.17 1) The function $x \to x^+$ being convex and increasing, $(X_n^+)_{n\geq 0}$ is a submartingale. Therefore, for $p \geq n$,

$$\mathbf{E}^{\mathcal{F}_n}(X_p^+) \leq \mathbf{E}^{\mathcal{F}_n}[\mathbf{E}^{\mathcal{F}_p}(X_{p+1}^+)] = \mathbf{E}^{\mathcal{F}_n}(X_{p+1}^+).$$

2) M_n is $\mathcal{F}_n$-measurable and positive. By definition

$$M_n \geq \mathbf{E}^{\mathcal{F}_n}(X_n^+) = X_n^+ \geq X_n.$$

Moreover $(M_n)_{n\geq 0}$ is bounded in L^1 as, by the monotone convergence theorem,

$$\mathbf{E}(M_n) = \mathbf{E}\Big[\lim_{p\to\infty}\uparrow \mathbf{E}^{\mathcal{F}_n}(X_p^+)\Big] = \lim_{p\to\infty}\uparrow \mathbf{E}[\mathbf{E}^{\mathcal{F}_n}(X_p^+)] =$$
$$= \lim_{p\to\infty}\uparrow \mathbf{E}(X_p^+) < +\infty.$$

Finally $(M_n)_{n\geq 0}$ is a martingale since (using •1.5 (vii))

$$\mathbf{E}^{\mathcal{F}_n}(M_{n+1}) = \mathbf{E}^{\mathcal{F}_n}\Big[\lim_{p\to\infty}\uparrow \mathbf{E}^{\mathcal{F}_{n+1}}(X_p^+)\Big] =$$
$$= \lim_{p\to\infty}\uparrow \mathbf{E}^{\mathcal{F}_n}(\mathbf{E}^{\mathcal{F}_{n+1}}(X_p^+)) = \lim_{p\to\infty}\uparrow \mathbf{E}^{\mathcal{F}_n}(X_p^+) = M_n.$$

3) We have seen that $M_n \geq X_n$; therefore, $Y_n = M_n - X_n$ is positive. It is integrable as the difference of two integrable r.v.'s and a submartingale, since $(M_n)_{n\geq 0}$ and $(-X_n)_{n\geq 0}$ are submartingales.

♦ We have only used the assumption $\sup_{n\geq 0}\mathbf{E}(X_n^+) < +\infty$ but, for a submartingale, this is equivalent to $\sup_{n\geq 0}\mathbf{E}|X_n| < +\infty$, as pointed out at the end of •3.10.

♦ This result can also be deduced from Doob's decomposition, as in the proof of Theorem 3.9.

E3.18 1) Let $(V_n)_{n\geq 0}$ be the compensator of $(S_n^2)_{n\geq 0}$. As Y_{n+1} is independent of $\mathscr{B}_n$, $\mathbf{E}^{\mathscr{B}_n}(Y_{n+1}) = \mathbf{E}(Y_{n+1}) = 0$ and

$$V_{n+1} - V_n = \mathbf{E}^{\mathscr{B}_n}(S_{n+1}^2 - S_n^2) = \mathbf{E}^{\mathscr{B}_n}(Y_{n+1}^2 + 2Y_{n+1}S_n) =$$
$$= \mathbf{E}^{\mathscr{B}_n}(Y_{n+1}^2) + 2S_n\mathbf{E}^{\mathscr{B}_n}(Y_{n+1}) = \mathbf{E}(Y_{n+1}^2) + 2S_n\mathbf{E}(Y_{n+1}) = 1.$$

Therefore $V_n = n$.

2) One easily gets

$$\mathbf{E}^{\mathscr{B}_n}(M_{n+1} - M_n) = \mathbf{E}^{\mathscr{B}_n}(\operatorname{sign}(S_n)Y_{n+1}) = \operatorname{sign}(S_n)\mathbf{E}^{\mathscr{B}_n}(Y_{n+1}) = 0 \qquad \text{a.s.}$$

Therefore $(M_n)_{n\geq 0}$ is a martingale. It is obviously square integrable and its associated increasing process (•3.8) satisfies

$$A_{n+1} - A_n = \mathbf{E}^{\mathscr{B}_n}(M_{n+1}^2 - M_n^2) =$$
$$= \mathbf{E}^{\mathscr{B}_n}(\operatorname{sign}(S_n)^2 Y_{n+1}^2 + 2M_n \operatorname{sign}(S_n)Y_{n+1}) =$$
$$= \operatorname{sign}(S_n)^2 \underbrace{\mathbf{E}^{\mathscr{B}_n}(Y_{n+1}^2)}_{=1} + 2M_n \operatorname{sign}(S_n) \underbrace{\mathbf{E}^{\mathscr{B}_n}(Y_{n+1})}_{=0} = \operatorname{sign}(S_n)^2 = 1_{\{S_n\neq 0\}}$$

from which

$$A_n = \sum_{k=0}^{n-1} 1_{\{S_k\neq 0\}}.$$

3) Let $|S_n| = N_n + A'_n$ be the Doob decomposition of the submartingale $(|S_n|)_{n\geq 0}$. On $\{S_n > 0\}$ it holds $S_{n+1} \geq 0$ and $|S_{n+1}| - |S_n| = S_{n+1} - S_n = Y_{n+1}$ so that

$$\mathbf{E}^{\mathscr{B}_n}((|S_{n+1}| - |S_n|)1_{\{S_n>0\}}) = 1_{\{S_n>0\}}\mathbf{E}^{\mathscr{B}_n}(Y_{n+1}) = 0.$$

In the same way, on $\{S_n < 0\}$, $|S_{n+1}| - |S_n| = S_n - S_{n+1} = -Y_{n+1}$ and $\mathbf{E}^{\mathscr{B}_n}((|S_{n+1}| - |S_n|)1_{\{S_n<0\}}) = 0$. Therefore

$$A'_{n+1} - A'_n = E^{\mathscr{B}_n}(|Y_{n+1}|1_{\{S_n=0\}}) = 1_{\{S_n=0\}}$$

and, for $n \geq 1$,

$$A'_n = \sum_{k=0}^{n-1} 1_{\{S_k=0\}}.$$

Thus

$$N_{n+1} - N_n = |S_{n+1}| - |S_n| - 1_{\{S_n=0\}} = Y_{n+1}1_{\{S_n>0\}} - Y_{n+1}1_{\{S_n<0\}} =$$
$$= \operatorname{sign}(S_n)Y_{n+1}$$

so that $N_n = M_n$. Thus $M_n = |S_n| - \sum_{k=0}^{n-1} 1_{\{|S_k|=0\}}$ and M_n is measurable with respect to $\sigma(|S_1|, \dots, |S_k|)$.

E3.19 As usual we set $\mathscr{F}_n = \sigma(X_1, \dots, X_n)$. Since X_{n+1} is independent of $\mathscr{F}_n$,

$$\mathbf{E}(Y_{n+1} \mid \mathscr{F}_n) = \prod_{k=1}^{n} \frac{p(X_k)}{q(X_k)}\mathbf{E}\Big(\frac{p(X_{n+1})}{q(X_{n+1})} \Bigm| \mathscr{F}_n\Big) = Y_n\mathbf{E}\Big(\frac{p(X_{n+1})}{q(X_{n+1})}\Big).$$

It is therefore sufficient to show that the rightmost expectation is equal to 1; but

$$\mathbf{E}\Big(\frac{p(X_{n+1})}{q(X_{n+1})}\Big) = \sum_{x\in E} \frac{p(x)}{q(x)} q(x) = \sum_{x\in E} p(x) = 1. \tag{3.27}$$

The positive martingale $(Y_n)_{n\geq 0}$ converges to a r.v. $Y_\infty \geq 0$ a.s. In order to identify the limit, let us compute the mean of $\sqrt{Y_n}$. It holds

$$\mathbf{E}(\sqrt{Y_n}) = \prod_{k=1}^{n} \mathbf{E}\Big(\sqrt{\frac{p(X_k)}{q(X_k)}}\Big) = \alpha^n,$$

where we have set

$$\alpha = \mathbf{E}\Big(\sqrt{\frac{p(X_1)}{q(X_1)}}\Big).$$

If $\alpha < 1$ therefore $\mathbf{E}(\sqrt{Y_n}) \to_{n\to\infty} 0$; as $\sqrt{Y_n} \to_{n\to\infty} \sqrt{Y_\infty}$ a.s.,

$$\mathbf{E}(\sqrt{Y_\infty}) \leq \underline{\lim}_{n\to\infty} \mathbf{E}(\sqrt{Y_n}) = 0$$

and $Y_\infty = 0$ a.s. It remains to prove that $\alpha < 1$. By Schwarz inequality and using (3.27),

$$\alpha = \mathbf{E}\Big(\sqrt{\frac{p(X_1)}{q(X_1)}}\Big) = \mathbf{E}\Big(1 \cdot \sqrt{\frac{p(X_1)}{q(X_1)}}\Big) \leq \sqrt{\mathbf{E}(1)} \cdot \sqrt{\mathbf{E}\Big(\frac{p(X_1)}{q(X_1)}\Big)} = 1. \tag{3.28}$$

But we also know that this inequality is *strict*, unless the two vectors 1 and $\sqrt{p/q} = (p(x)^{1/2}q(x)^{-1/2})_{x\in E}$ are linearly dependent, i.e., unless there exists $\lambda \in \mathbb{R}$ such that

$$\sqrt{\frac{p(x)}{q(x)}} = \lambda$$

for every $x \in E$. But this implies $p = q$. Therefore $\alpha < 1$. Note that the martingale $(Y_n)_{n\geq 0}$ is not regular since $\mathbf{E}(Y_n) = 1$ whereas $\mathbf{E}(Y_\infty) = 0$; thus it cannot converge in L^1.

E3.20 1) Since Y_{n+1} is independent of $\mathscr{F}_n$ and $\mathbf{E}(Y_{n+1}) = 1$,

$$\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = \mathbf{E}(Y_1 \dots Y_{n+1} \mid \mathscr{F}_n) = Y_1 \dots Y_n \mathbf{E}(Y_{n+1} \mid \mathscr{F}_n) = X_n.$$

$(X_n)_{n\geq 0}$ is therefore a martingale and $(\sqrt{X_n})_{n\geq 0}$ is a supermartingale, as the function $x \to \sqrt{x}$ is concave and $\mathbf{E}(\sqrt{X_n}) \leq \sqrt{\mathbf{E}(X_n)} = 1$.

2) One has

$$\mathbf{E}(\sqrt{X_n}) = \prod_{k=1}^{n} \mathbf{E}(\sqrt{Y_n}) \underset{n\to\infty}{\to} 0.$$

But the positive supermartingale $(\sqrt{X_n})_{n\geq 0}$ converges a.s. to a r.v. $Z \geq 0$. By Fatou's lemma

$$\mathbf{E}(Z) = \mathbf{E}\Big(\lim_{n\to\infty} \sqrt{X_n}\Big) \leq \underline{\lim}_{n\to\infty} \mathbf{E}(\sqrt{X_n}) = 0.$$

Therefore $Z = 0$ and $X_n \to_{n\to\infty} 0$ a.s. If $(X_n)_{n\geq 0}$ were regular, the convergence would also take place in L^1 and one would get $\mathbf{E}(X_n) \to_{n\to\infty} 0$, in contradiction with the fact that $\mathbf{E}(X_n) = 1$ for every n.

3) The hypothesis $\prod_{k=1}^{\infty} \mathbf{E}(\sqrt{Y_k}) > 0$ implies $\prod_{k=n}^{\infty} \mathbf{E}(\sqrt{Y_k}) \to_{n\to\infty} 1$. Again $\mathbf{E}(\sqrt{Y_k}) \le \sqrt{\mathbf{E}(Y_k)} = 1$, by Jensen's inequality.

If $n \ge m$, $\sqrt{X_n X_m} = Y_1 \dots Y_m \sqrt{Y_{m+1} \dots Y_n}$, from which

$$\mathbf{E}(\sqrt{X_n X_m}) = \mathbf{E}(\sqrt{Y_{m+1}}) \dots \mathbf{E}(\sqrt{Y_n}).$$

Therefore

$$\mathbf{E}\big[(\sqrt{X_n} - \sqrt{X_m})^2\big] = \mathbf{E}(X_n + X_m - 2\sqrt{X_n X_m}) = 2 - 2\prod_{k=m+1}^{n} \mathbf{E}(\sqrt{Y_k}) \underset{n,m\to\infty}{\to} 0$$

and $(\sqrt{X_n})_{n\ge 0}$ is a Cauchy sequence in L^2. It therefore converges in L^2. Let us study the L^1 convergence of $(X_n)_{n\ge 0}$:

$$\begin{aligned}\|X_n - X_m\|_1 &= \mathbf{E}(|X_n - X_m|) = \mathbf{E}(|\sqrt{X_n} - \sqrt{X_m}|(\sqrt{X_n} + \sqrt{X_m})) \le \\ &\le \|\sqrt{X_n} - \sqrt{X_m}\|_2 \|\sqrt{X_n} + \sqrt{X_m}\|_2 \le 2\|\sqrt{X_n} - \sqrt{X_m}\|_2.\end{aligned}$$

Therefore $(X_n)_{n\ge 0}$ is a Cauchy sequence in L^1, which implies that it is a regular martingale.

E3.21 1a) $(X_n)_{n\ge 0}$ is a submartingale as the r.v.'s Y_i are positive (see •3.2). Its compensator is determined by $A_0 = 0$ and

$$A_{n+1} = A_n + \mathbf{E}(X_{n+1} - X_n \mid \mathscr{F}_n) = A_n + \mathbf{E}(Y_{n+1} \mid \mathscr{F}_n)$$

so that $A_n = \mathbf{E}(Y_1) + \mathbf{E}(Y_2 \mid \mathscr{F}_1) + \dots + \mathbf{E}(Y_n \mid \mathscr{F}_{n-1})$.

1b) $\{\tau_a = n\} = \{A_1 \le a, \dots, A_n \le a, A_{n+1} > a\} \in \mathscr{F}_n$ so that τ_a is a stopping time. We use here the fact that the increasing process is predictable (i.e., A_{n+1} is $\mathscr{F}_n$-measurable for every n).

1c) By Doob's decomposition $(Z_n)_n$ is a martingale and $-Z_n = A_n - X_n \le A_n$. Hence $-Z_{n\wedge\tau_a} \le A_{n\wedge\tau_a} \le a$ and $Z^-_{n\wedge\tau_a} = (-Z_{n\wedge\tau_a}) \vee 0 \le a$. This implies that $\sup_{n\ge 0} \mathbf{E}|Z^-_{n\wedge\tau_a}| < +\infty$ and therefore the martingale $(Z_{n\wedge\tau_a})_{n\ge 0}$ converges a.s. to a finite r.v.

1d) On $\{\tau_a = +\infty\}$, $(Z_n)_{n\ge 0}$ converges a.s. since, for $\omega \in \{\tau_a = +\infty\}$, the two sequences $(Z_n(\omega))_{n\ge 0}$ and $(Z_{n\wedge\tau_a}(\omega))_{n\ge 0}$ coincide; moreover, still on $\{\tau_a = +\infty\}$, $\sup_{n\ge 0} A_n(\omega) \le a$, so that

$$\lim_{n\to\infty} \uparrow X_n \le a + \sup_{n\ge 0} Z_n < +\infty \qquad \text{a.s.}$$

It is enough now to remark that $\{\lim_{n\to\infty} \uparrow A_n < +\infty\} = \bigcup_{a>0}\{\tau_a = +\infty\}$.

2) Observe that $X_{\sigma_a} = X_{\sigma_a - 1} + Y_{\sigma_a} \le a + \sup_{n\ge 1} Y_n$ (of course $\sigma_a \ge 1$) and, as $Z_n \le X_n$,

$$Z^+_{n\wedge\sigma_a} \le X_{n\wedge\sigma_a} \le a + \sup_{n\ge 1} Y_n$$

so that $\sup_{n\ge 0} \mathbf{E}|Z^+_{n\wedge\sigma_a}| < +\infty$ and the martingale $(Z_{n\wedge\sigma_a})_{n\ge 0}$ converges a.s. to a finite r.v. On $\{\sigma_a = +\infty\}$, it also holds

$$\lim_{n\to\infty} \uparrow A_n \le a + \sup_{n\ge 0}(-Z_n) < +\infty \quad \text{a.s.}$$

as $A_n = -Z_n + X_n$. But

$$\left\{\lim_{n\to\infty} \uparrow X_n < +\infty\right\} \subset \bigcup_{a>0}\{\sigma_a = +\infty\} \subset \left\{\lim_{n\to\infty} \uparrow A_n < +\infty\right\}$$

and, thanks to (1d), one obtains the requested inequality.

3) It is sufficient to apply (2) to $Y_n = 1_{A_n}$.

E3.22 1) As $\mathbf{E}(X_n^2) < +\infty$, S_n and V_n are integrable. One has

$$\mathbf{E}^{\mathscr{F}_n}(S_{n+1} - S_n) = \mathbf{E}^{\mathscr{F}_n}(X_{n+1}) = 0 \quad \text{a.s.},$$

i.e., $(S_n)_{n\geq 0}$ is a martingale. In the same way, a.s.,

$$\begin{aligned}\mathbf{E}^{\mathscr{F}_n}(V_{n+1} - V_n) &= \mathbf{E}^{\mathscr{F}_n}(S_{n+1}^2 - S_n^2) - \sigma_{n+1}^2 = \\ &= \mathbf{E}^{\mathscr{F}_n}(X_{n+1}(2S_n + X_{n+1})) - \sigma_{n+1}^2 = \\ &= 2S_n\mathbf{E}^{\mathscr{F}_n}(X_{n+1}) + \mathbf{E}^{\mathscr{F}_n}(X_{n+1}^2) - \sigma_{n+1}^2 = 0\end{aligned}$$

(in other words we have proved that $A_n = \langle S\rangle_n$ with the notations of •3.8).

2) Thanks to (1), $\mathbf{E}(V_n) = 0$, i.e., $\mathbf{E}(S_n^2) = A_n$; thus

$$\sup_{n\geq 0}\mathbf{E}(S_n^2) = \sum_{k=1}^{\infty}\sigma_k^2 < +\infty.$$

$(S_n)_{n\geq 0}$ is a thus a bounded martingale in L^2 and it converges a.s. and in L^2.

3a) It is an immediate consequence of the optional sampling Theorem 3.3: $0 = \mathbf{E}(V_{n\wedge\tau_a}) = \mathbf{E}(S_{n\wedge\tau_a}^2 - A_{n\wedge\tau_a})$ so that $\mathbf{E}(S_{n\wedge\tau_a}^2) = \mathbf{E}(A_{n\wedge\tau_a})$ ($n \wedge \tau_a$ is a bounded stopping time).

3b) The sequence $(S_n)_{n\geq 0}$ converges a.s.; it is therefore a.s. bounded and

$$\Omega = \bigcup_{k\geq 0}\{\tau_k = +\infty\} = \lim_{k\to\infty} \uparrow \{\tau_k = +\infty\} \quad \text{a.s.},$$

which implies that $\mathbf{P}(\tau_k = +\infty) > 0$ for k sufficiently large.

3c) Let us choose a such that $\mathbf{P}(\tau_a = +\infty) > 0$. On $\{n < \tau_a\}$ it holds $|S_{n\wedge\tau_a}| \leq a$ whereas, on $\{n \wedge \tau_a = \tau_a\}$,

$$|S_{n\wedge\tau_a}| = |S_{\tau_a}| \leq |S_{\tau_a - 1}| + |X_{\tau_a}| \leq a + M.$$

Therefore $\mathbf{E}(A_{n\wedge\tau_a}) \leq (a + M)^2$ and

$$(a + M)^2 \geq \mathbf{E}(A_{n\wedge\tau_a} 1_{\{\tau_a=+\infty\}}) = \mathbf{E}(A_n 1_{\{\tau_a=+\infty\}}) = \mathbf{P}(\tau_a = +\infty)\sum_{k=1}^{n}\sigma_k^2,$$

which implies the convergence of the series $\sum_{k=1}^{\infty}\sigma_k^2$.

E3.23 1) Indeed $\{\tau_a = n\} = \{A_1 \leq a^2, \ldots, A_n \leq a^2, A_{n+1} > a^2\} \in \mathscr{F}_n$ (the process $(A_n)_{n\geq 0}$ is predictable).

2) For every $p \geq 1$, $(M_{n\wedge\tau_a}^2)_{n\geq 0}$ being a positive submartingale and thanks to Theorem 3.5 (ii),

$$\mathbf{P}\Big(\sup_{n\leq p}|M_{n\wedge\tau_a}| > a\Big) \leq a^{-2}\mathbf{E}(M_{p\wedge\tau_a}^2) = a^{-2}\mathbf{E}(A_{p\wedge\tau_a}) \leq a^{-2}\mathbf{E}(A_{\tau_a}).$$

Hence $\mathbf{P}(\sup_{n\geq 0}|M_{n\wedge\tau_a}| > a) \leq a^{-2}\mathbf{E}(A_{\tau_a})$. But on $\{\tau_a < +\infty\}$ $A_{\tau_a} \leq a^2$ whereas $A_\infty \leq a^2$ on $\{\tau_a = +\infty\}$; therefore in any case $A_{\tau_a} \leq A_\infty \wedge a^2$.

3) Indeed, as $\tau_a = +\infty$ on $\{A_\infty \leq a^2\}$, $\{A_\infty \leq a^2\} \cap \{\sup_{n\geq 0}|M_n| > a\} \subset \{\sup_{n\geq 0}|M_{n\wedge\tau_a}| > a\}$ and

$$\Big\{\sup_{n\geq 0}|M_n| > a\Big\} \subset \{A_\infty > a^2\} \cup \Big(\{A_\infty \leq a^2\} \cap \Big\{\sup_{n\geq 0}|M_n| > a\Big\}\Big) \subset$$
$$\subset \{A_\infty > a^2\} \cup \Big\{\sup_{n\geq 0}|M_{n\wedge\tau_a}| > a\Big\}.$$

4) We use Fubini's theorem many times:

$$\int_0^\lambda \mathbf{P}(X > t)\,dt = \int_0^\lambda \mathbf{E}(1_{\{X>t\}})\,dt = \mathbf{E}\Big[\int_0^\lambda 1_{\{X>t\}}\,dt\Big] =$$
$$= \mathbf{E}\Big[\int_0^{X\wedge\lambda} dt\Big] = \mathbf{E}(X\wedge\lambda)$$

and

$$\int_0^\infty a^{-2}\mathbf{E}(X\wedge a^2)\,da = \int_0^\infty a^{-2}\,da \int_0^{a^2} \mathbf{P}(X > t)\,dt =$$
$$= \int_0^\infty \mathbf{P}(X > t)\,dt \int_{\sqrt{t}}^\infty a^{-2}\,da = \int_0^\infty \mathbf{E}(1_{\{X>t\}})t^{-1/2}\,dt =$$
$$= \mathbf{E}\Big[\int_0^{+\infty} 1_{\{X>t\}}t^{-1/2}\,dt\Big] = \mathbf{E}\Big[\int_0^X t^{-1/2}\,dt\Big] = 2\mathbf{E}(\sqrt{X}).$$

5) One has

$$\mathbf{E}\Big(\sup_{n\geq 0}|M_n|\Big) = \int_0^\infty \mathbf{P}\Big(\sup_{n\geq 0}|M_n| > a\Big)\,da \leq$$
$$\leq \int_0^\infty \mathbf{P}(A_\infty > a^2)\,da + \int_0^\infty \mathbf{P}\Big(\sup_{n\geq 0}|M_{n\wedge\tau_a}| > a\Big)\,da.$$

On one hand

$$\int_0^\infty \mathbf{P}(A_\infty > a^2)\,da = \int_0^\infty \mathbf{P}(\sqrt{A_\infty} > a)\,da = \mathbf{E}(\sqrt{A_\infty})$$

and on the other hand, thanks to (2),

$$\mathbf{P}\Big(\sup_{n\geq 0}|M_{n\wedge\tau_a}| > a\Big) \leq a^{-2}\mathbf{E}(A_\infty \wedge a^2).$$

Thanks to (4),

$$\int_0^\infty a^{-2}\mathbf{E}(A_\infty\wedge a^2)\,da = 2\mathbf{E}(\sqrt{A_\infty}),$$

which gives the inequality we were looking for.

6) It is easy to see that $(S_{n\wedge\tau})_{n\geq 0}$ is a square integrable martingale with associated increasing process $A_n = \sigma^2(n\wedge\tau)$ where $\sigma^2 = \mathbf{E}(Y_1^2)$; in particular $A_\infty = \sigma^2\tau$.

We apply the previous results: if $Z = \sup_{n\geq 0}|S_{n\wedge\tau}|$; then $\mathbf{E}(Z) \leq 3\sigma\mathbf{E}(\sqrt{\tau}) < +\infty$. Obviously the condition $\mathbf{E}(\sqrt{\tau}) < +\infty$ implies $\mathbf{P}(\tau < +\infty) = 1$ and

$S_{n\wedge\tau} \to S_\tau$ a.s. as $n \to \infty$; but $|S_{n\wedge\tau}| \le Z \in L^1$ for every n, so that, by the dominated convergence theorem and the optional sampling theorem, $\mathbf{E}(S_\tau) = \lim_{n\to\infty} \mathbf{E}(S_{n\wedge\tau}) = \mathbf{E}(S_0) = 0$.

E3.24 1) By the law of large numbers $\frac{1}{n} S_n^i \to_{n\to\infty} \frac{1}{2}$ a.s. so that $S_n^i \to_{n\to\infty} +\infty$; the stopping times ν_1, ν_2, ν are therefore finite a.s.

2) Let $Y_n^i = 2X_n^i - 1$; thus $M_n^i = \sum_{k=1}^n Y_k^i$ and $M_n^{i,j} = M_n^i M_n^j - n\delta_{ij}$. Then $\mathbf{E}(Y_n^i) = 0$, $\mathbf{E}((Y_n^i)^2) = 1$ and

$$\mathbf{E}(M_{n+1}^i - M_n^i \mid \mathscr{F}_n) = \mathbf{E}(Y_{n+1}^i \mid \mathscr{F}_n) = 0 \quad \text{a.s.},$$

so that $(M_n^i)_{n\ge 0}$, $i = 1, 2$, are martingales. In the same way a.s.

$$\begin{aligned}\mathbf{E}((M_{n+1}^i)^2 - (M_n^i)^2 \mid \mathscr{F}_n) &= \mathbf{E}((2M_n^i + Y_{n+1}^i)Y_{n+1}^i \mid \mathscr{F}_n) = \\ &= 2M_n^i \underbrace{\mathbf{E}(Y_{n+1}^i \mid \mathscr{F}_n)}_{=0} + \mathbf{E}((Y_{n+1}^i)^2) = 1\end{aligned}$$

from which one deduces that $M_n^{i,i} = (M_n^i)^2 - n$ is a martingale. At last a.s.

$$\begin{aligned}&\mathbf{E}(M_{n+1}^1 M_{n+1}^2 - M_n^1 M_n^2 \mid \mathscr{F}_n) = \\ &= \mathbf{E}(Y_{n+1}^1 M_n^2 + Y_{n+1}^2 M_n^1 + Y_{n+1}^1 Y_{n+1}^2 \mid \mathscr{F}_n) = 0;\end{aligned}$$

hence, $M_n^{1,2} = M_n^1 M_n^2$ is also a martingale. One should remember that it would be always necessary to check that these r.v.'s are integrable. Here this is obvious, as $0 \le S_n^i \le n$.

3) By the optional sampling Theorem 3.3 applied to the bounded stopping time $\nu \wedge n$, $\mathbf{E}(2S_{\nu\wedge n}^i) = \mathbf{E}(\nu \wedge n)$. As $0 \le S_{\nu\wedge n}^i \le a$, the dominated convergence theorem gives $\mathbf{E}(S_{\nu\wedge n}^i) \to_{n\to\infty} \mathbf{E}(S_\nu^i) \le a$; on the other hand $\mathbf{E}(\nu \wedge n) \uparrow_{n\to\infty} \mathbf{E}(\nu)$ by the monotone convergence theorem and $\mathbf{E}(\nu) = 2\mathbf{E}(S_\nu^i) \le 2a$.

4) If it was possible to apply the optional sampling Theorem 3.3 to the stopping time ν (this is not the case since it is not bounded), one would get

$$\mathbf{E}(M_\nu^{i,j}) = \mathbf{E}(M_0^{i,j}) = 0.$$

Therefore we try, as in the previous question, to approximate ν with the stopping times $\nu \wedge n$, which are bounded. It holds then $\mathbf{E}(M_{\nu\wedge n}^{i,j}) = 0$, i.e., for $i = j$,

$$\mathbf{E}((2S_{\nu\wedge n}^i - \nu \wedge n)^2) = \mathbf{E}(\nu \wedge n). \tag{3.29}$$

Note first that $\sup_{n\ge 0} |S_{n\wedge\nu}^i| \le a$. Let us show that $\mathbf{E}(\nu^2) < +\infty$. One can write $\nu \wedge n = (\nu \wedge n - 2S_{\nu\wedge n}^i) + 2S_{\nu\wedge n}^i$ and, since $\mathbf{E}(\nu) \le 2a$ as seen in (3),

$$\begin{aligned}\mathbf{E}((\nu \wedge n)^2) &\le 2\mathbf{E}((2S_{\nu\wedge n}^i)^2) + 2\mathbf{E}((\nu \wedge n - 2S_{\nu\wedge n}^i)^2) \le \\ &\le 8a^2 + 2\mathbf{E}(\nu \wedge n) \le 8a^2 + 4a.\end{aligned}$$

By the monotone convergence theorem, $\mathbf{E}(\nu^2) < +\infty$. Therefore it is easy to take the limit in (3.29), which gives

$$\mathbf{E}((2S_\nu^i - \nu)^2) = \mathbf{E}(\nu)$$

and $\mathbf{E}(M_\nu^{i,i}) = 0$. In order to prove that $\mathbf{E}(M_\nu^{1,2}) = 0$, the argument is the same: first, by Theorem 3.3,

$$0 = \mathbf{E}(M^{1,2}_{\nu\wedge n}) = \mathbf{E}((2S^1_{\nu\wedge n} - \nu\wedge n)(2S^2_{\nu\wedge n} - \nu\wedge n))$$

and one can easily take the limit, as we already know that $2S^i_{\nu\wedge n} - \nu\wedge n \to_{n\to\infty} 2S^i_\nu - \nu$ in L^2.

5) It is immediate that

$$M^{1,1}_n - 2M^{1,2}_n + M^{2,2}_n = 4(S^1_n)^2 + 4(S^2_n)^2 - 8S^1_n S^2_n - 2n = 4(S^1_n - S^2_n)^2 - 2n.$$

This and (3) give

$$4\mathbf{E}((S^1_\nu - S^2_\nu)^2) = \mathbf{E}(M^{1,1}_\nu - 2M^{1,2}_\nu + M^{2,2}_\nu) + 2\mathbf{E}(\nu) = 2\mathbf{E}(\nu) \le 2a.$$

Hence $\mathbf{E}(|S^1_\nu - S^2_\nu|)^2 \le \mathbf{E}((S^1_\nu - S^2_\nu)^2) \le a$ which allows to conclude.

E3.25 Let us first remark that $\mathbf{P}(\nu < +\infty) = 1$, ν being integrable.

A1) As Y_{n+1} is independent of $\mathscr{F}_n$, $\mathbf{E}^{\mathscr{F}_n}(Y_{n+1}) = \mathbf{E}(Y_{n+1}) = m$ a.s. and therefore $\mathbf{E}^{\mathscr{F}_n}(S_{n+1} - S_n) = \mathbf{E}^{\mathscr{F}_n}(Y_{n+1}) - m = 0$ a.s.

A2) $(X_{n\wedge\nu})_{n\ge1}$ is a martingale thanks to (A1). Therefore $\mathbf{E}(X_{n\wedge\nu}) = \mathbf{E}(X_0)$, i.e., $\mathbf{E}(S_{n\wedge\nu}) = m\mathbf{E}(n\wedge\nu)$.

A3) Assume $Y_n \ge 0$. The sequences $(S_n)_{n\ge0}$ and $(S_{n\wedge\nu})_{n\ge0}$ are positive and increasing, so that $\mathbf{E}(S_{n\wedge\nu}) \uparrow \mathbf{E}(S_\nu)$; for the same reason $\mathbf{E}(n\wedge\nu) \uparrow \mathbf{E}(\nu)$. This implies, thanks to (A2), $\mathbf{E}(S_\nu) = m\mathbf{E}(\nu) < +\infty$. As for the general case, it is sufficient to set

$$Y^{(1)}_n = Y^+_n, \quad S^{(1)}_n = Y^{(1)}_1 + \ldots + Y^{(1)}_n,$$
$$Y^{(2)}_n = Y^-_n, \quad S^{(2)}_n = Y^{(2)}_1 + \ldots + Y^{(2)}_n.$$

If $m_1 = \mathbf{E}(Y^{(1)}_n)$, $m_2 = \mathbf{E}(Y^{(2)}_n)$ (so that $m = m_1 - m_2$), we have just seen that

$$\mathbf{E}(S^{(1)}_\nu) = m_1\mathbf{E}(\nu), \qquad \mathbf{E}(S^{(2)}_\nu) = m_2\mathbf{E}(\nu).$$

Thus, by subtraction, all quantities appearing in the expression being finite,

$$\mathbf{E}(S_\nu) = \mathbf{E}(S^{(1)}_\nu) - \mathbf{E}(S^{(2)}_\nu) = (m_1 - m_2)\mathbf{E}(\nu) = m\mathbf{E}(\nu).$$

A4) The process $(S_n)_{n\ge1}$ can take, on $\mathbb{Z}$, just one step to the right or to the left; it is clear then that $S_\tau = a$ a.s. on $\{\tau < +\infty\}$. One should therefore have $\mathbf{E}(S_\tau) = a$ whereas, if τ were integrable, by (A3)

$$\mathbf{E}(S_\tau) = \mathbf{E}(\tau)\mathbf{E}(Y_1) = 0$$

(this nonintegrability property of τ is further investigated in Exercise 3.23).

B1) As Y_{n+1} is independent of $\mathscr{F}_n$, $\mathbf{E}^{\mathscr{F}_n}(Y^2_{n+1}) = \mathbf{E}(Y^2_{n+1}) = \sigma^2$ a.s. and therefore it holds a.s.

$$\mathbf{E}^{\mathscr{F}_n}(Z_{n+1} - Z_n) = \mathbf{E}^{\mathscr{F}_n}(S^2_{n+1} - S^2_n) - \sigma^2 =$$
$$\mathbf{E}^{\mathscr{F}_n}[(2S_n + Y_{n+1})Y_{n+1}] - \sigma^2 = 2S_n \underbrace{\mathbf{E}^{\mathscr{F}_n}(Y_{n+1})}_{=0} + \mathbf{E}^{\mathscr{F}_n}(Y^2_{n+1}) - \sigma^2 = 0.$$

B2) Note first that $\{k \le \nu\} = \{\nu \le k-1\}^c \in \mathscr{F}_{k-1}$. If $j < k$, the r.v.'s Y_j, $1_{\{j\le\nu\}}$ and $1_{\{k\le\nu\}}$ are thus $\mathscr{F}_{k-1}$-measurable and therefore

$$E[Y_j 1_{\{j\le\nu\}} Y_k 1_{\{k\le\nu\}}] = E[Y_j 1_{\{j\le\nu\}} 1_{\{k\le\nu\}} E^{\mathscr{F}_{k-1}}(Y_k)] = 0.$$

Let $R_n = \sum_{k=1}^{n} Y_k^2$. By (A), R_ν is integrable and it is sufficient to observe that

$$R_\nu = \sum_{k=1}^{\nu} Y_k^2 = \sum_{k=1}^{\infty} Y_k^2 1_{\{k\le\nu\}}.$$

B3) The series $\sum_{k=1}^{\infty} \mathbf{E}(Y_k^2 1_{\{k\le\nu\}})$ being convergent, it holds, for $q < p$,

$$\mathbf{E}(S_{p\wedge\nu} - S_{q\wedge\nu})^2 = \mathbf{E}\Big[\sum_{k=q+1}^{p} Y_k 1_{\{k\le\nu\}}\Big]^2 = \sum_{k=q+1}^{p} \mathbf{E}(Y_k^2 1_{\{k\le\nu\}}) \underset{p,q\to+\infty}{\to} 0.$$

$(S_{n\wedge\nu})_{n\ge0}$ is a Cauchy sequence in L^2 and converges in L^2. But $S_{n\wedge\nu} \to_{n\to\infty} S_\nu$ a.s., hence the result.

B4) By (A), $\mathbf{E}(S_{n\wedge\nu}^2) = \sigma^2\mathbf{E}(n\wedge\nu)$. By (B3) $\mathbf{E}(S_{n\wedge\nu}^2) \to_{n\to\infty} \mathbf{E}(S_\nu^2)$. Since $\mathbf{E}(n\wedge\nu) \uparrow \mathbf{E}(\nu)$, one gets $\mathbf{E}(S_\nu^2) = \sigma^2\mathbf{E}(\nu)$.

B5) It is sufficient to apply the previous arguments to the r.v.'s $\bar{Y}_i = Y_i - m$.

E3.26 a) Since $\Omega \in \mathscr{F}_n$ and $M_n \ge 0$,

$$\nu(\Omega) = \int_\Omega M_n\, d\mathbf{P} = \mathbf{E}(M_n).$$

$\nu(\Omega)$ being finite by hypothesis, M_n is integrable with respect to $\mathbf{P}$. On the other hand, let $A \in \mathscr{F}_n$; it holds $\nu(A) = \mathbf{E}(M_n 1_A)$. Moreover, since also 1_A is $\mathscr{F}_{n+1}$-measurable, $\nu(A) = \mathbf{E}(M_{n+1} 1_A)$, from which we get $\mathbf{E}(M_n 1_A) = \mathbf{E}(M_{n+1} 1_A)$ for every $A \in \mathscr{F}_n$. Thus $M_n = \mathbf{E}(M_{n+1} \mid \mathscr{F}_n)$.

b) Since $(M_n)_{n\ge0}$ is a positive martingale, it converges a.s. to a positive r.v. M_∞. We remark that M_∞ is integrable since, by Fatou's lemma,

$$\mathbf{E}(M_\infty) = \mathbf{E}\Big(\lim_{n\to\infty} M_n\Big) \le \varliminf_{n\to\infty} \mathbf{E}(M_n) = \nu(\Omega) < \infty.$$

c) Let us assume that $\mathbf{P}$ dominates ν on $\mathscr{F}_\infty$. By the Radon–Nikodym theorem, there exists a $\mathscr{F}_\infty$-measurable r.v. Y such that

$$\nu(A) = \mathbf{E}(Y 1_A), \qquad \text{for every } A \in \mathscr{F}_\infty. \tag{3.30}$$

If $A \in \mathscr{F}_n \subset \mathscr{F}_\infty$ then

$$\mathbf{E}(M_n 1_A) = \mathbf{E}(Y 1_A). \tag{3.31}$$

This gives $M_n = \mathbf{E}(Y \mid \mathscr{F}_n)$; hence, $(M_n)_{n\ge0}$ is regular. On the other hand, by Theorem 3.10, $M_n = \mathbf{E}(Y \mid \mathscr{F}_n) \to_{n\to\infty} \mathbf{E}(Y \mid \mathscr{F}_\infty) = Y$ a.s.; thus $M_\infty = Y$ $\mathbf{P}$-a.s.

Conversely, if the martingale $(M_n)_{n\ge0}$ is regular, then there exists a r.v. $Y \in L^1(\Omega, \mathscr{F}_\infty, \mathbf{P})$ such that $M_n = \mathbf{E}(Y \mid \mathscr{F}_n)$. It holds then, for every $A \in \mathscr{F}_n$,

$$\nu(A) = \mathbf{E}[1_A M_n] = \mathbf{E}[1_A Y].$$

The two measures ν and $A \to \mathbf{E}[1_A Y]$ coincide on $\bigcup_{n\ge0} \mathscr{F}_n$. This is a class that is stable with respect to finite intersections and that generates $\mathscr{F}_\infty$. By Theorem 2.4 thus $\nu(A) = \mathbf{E}[1_A Y]$ for every $A \in \mathscr{F}_\infty$.

d) For every $A \in \mathscr{F}_n$ and $k \ge n$, $\mathbf{E}(M_k 1_A) = \mathbf{E}(M_n 1_A)$. Thus, by Fatou's lemma,

$$\mathbf{E}(M_\infty 1_A) \le \varliminf_{k\to\infty} \mathbf{E}(M_k 1_A) = \mathbf{E}(M_n 1_A) = \nu(A).$$

This is true for every $A \in \bigcup_{n=1}^{\infty} \mathscr{F}_n$ and also for every $A \in \mathscr{F}_\infty$ by monotone class arguments. Thus, for $A = S$, since $1_S = 1$ **P**-a.s.,

$$\mathbf{E}(M_\infty) = \mathbf{E}(M_\infty 1_S) \le \nu(S) = 0.$$

Since $M_\infty \ge 0$ we get $M_\infty = 0$ **P**-a.s.

e1) Let us denote by f_n the quotient of the densities $N(\mu_n, 1)$ and $N(0,1)$, i.e.,

$$f_n(z) = e^{-\frac{1}{2}\mu_n^2 + z\mu_n}.$$

The r.v.'s X_k, $k \le n$ are independent with respect to both **P** and ν. Let us denote by γ_n the centered Gaussian probability on $\mathbb{R}^n$ with covariance matrix equal to the identity. Thus γ_n is the law of the vector $(X_1, \dots, X_n)$ with respect to **P**. Then, for every bounded Borel function $\Phi : \mathbb{R}^n \to \mathbb{R}$,

$$\begin{aligned}\mathbf{E}^\nu[\Phi(X_1, \dots, X_n)] &= \int_{\mathbb{R}^n} \Phi(x_1, \dots, x_n) f_1(x_1) \dots f_n(x_n)\, d\gamma_n(x) = \\ &= \mathbf{E}^{\mathbf{P}}[\Phi(X_1, \dots, X_n) f_1(X_1) \dots f_n(X_n)].\end{aligned}$$

This implies that, for every n,

$$M_n = \prod_{k=1}^{n} f_k(X_k)$$

is a density of ν with respect to **P** on $\mathscr{F}_n$.

e2) By Exercise 3.20 (2) and (3), the martingale $(M_n)_{n\ge 1}$ is regular if and only if

$$\prod_{k=1}^{\infty} \mathbf{E}\big(\sqrt{f_k(X_k)}\big) > 0. \tag{3.32}$$

But

$$\begin{aligned}\mathbf{E}\big(\sqrt{f_k(X_k)}\big) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sqrt{f_k(x)}\, e^{-x^2/2}\, dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x^2 - x\mu_n + \frac{1}{2}\mu_n^2)}\, dx = \\ &= \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{8}\mu_n^2} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x - \frac{1}{2}\mu_n)^2}\, dx = e^{-\frac{1}{8}\mu_n^2}.\end{aligned}$$

Thus the infinite product in (3.32) is equal to

$$\exp\Big(-\frac{1}{8}\sum_{k=1}^{\infty} \mu_n^2\Big)$$

and is > 0 if and only if $\sum_{k=1}^{\infty} \mu_n^2 < +\infty$. In this case the martingale $(M_n)_{n\ge 1}$ is regular and **P** dominates ν on $\mathscr{F}_\infty$. If conversely the series diverges, by Exercise 3.20 (2) it holds $M_n \to_{n\to\infty} 0$ and the martingale is not regular; by (c), **P** does not dominate ν on $\mathscr{F}_\infty$.

E3.27 1) Since $\nu < +\infty$ and $|X_\nu| < +\infty$ a.s., $|X_\nu| 1_{\{\nu > n\}} \to 0$ a.s. as $n \to \infty$;

since $|X_\nu|$ is supposed to be integrable and $|X_\nu|1_{\{\nu>n\}} \le |X_\nu|$, Lebesgue's theorem gives $\mathbf{E}(|X_\nu|1_{\{\nu>n\}}) \to_{n\to\infty} 0$.

2) It holds

$$\mathbf{E}(|X_{\nu\wedge n} - X_\nu|) = \mathbf{E}(|X_n - X_\nu|1_{\{\nu>n\}}) \le \mathbf{E}(|X_n|1_{\{\nu>n\}}) + \mathbf{E}(|X_\nu|1_{\{\nu>n\}})$$

and the right-hand side converges to 0 as $n \to \infty$, thanks to the assumptions and to (1).

3) As $\nu\wedge n$ is a bounded stopping time, by the optional sampling theorem $\mathbf{E}(X_{\nu\wedge n}) = \mathbf{E}(X_0)$. Therefore

$$|\mathbf{E}(X_\nu) - \mathbf{E}(X_0)| = |\mathbf{E}(X_\nu) - \mathbf{E}(X_{\nu\wedge n})| \le \mathbf{E}(|X_\nu - X_{\nu\wedge n}|)$$

and, by (2), the rightmost term tends to 0 as $n \to \infty$.

E3.28 1) Since V_n is $\mathscr{F}_{n-1}$-measurable,

$$\mathbf{E}\Big(\sum_{n=1}^{\infty} V_n|X_n - X_{n-1}|\Big) = \sum_{n=1}^{\infty}\mathbf{E}(V_n\mathbf{E}^{\mathscr{F}_{n-1}}(|X_n - X_{n-1}|)) \le M\mathbf{E}\Big(\sum_{n=1}^{\infty} V_n\Big).$$

2a) Since $\{\nu \ge n\} = \{\nu \le n-1\}^c \in \mathscr{F}_{n-1}$, one can apply (1) to $V_n = 1_{\{\nu\ge n\}}$ and get

$$\mathbf{E}\Big(\sum_{n=1}^{\infty} 1_{\{\nu\ge n\}}|X_n - X_{n-1}|\Big) \le M\mathbf{E}\Big(\sum_{n=1}^{\infty} 1_{\{\nu\ge n\}}\Big) = M\mathbf{E}(\nu) < +\infty.$$

2b) It holds

$$\sum_{n=1}^{\infty} 1_{\{\nu\ge n\}}(X_n - X_{n-1}) = \sum_{n=1}^{\nu}(X_n - X_{n-1}) = X_\nu - X_0.$$

Therefore, by (2a),

$$\mathbf{E}(|X_\nu|) \le \mathbf{E}(X_0) + M\mathbf{E}(\nu) < +\infty.$$

3) Because of (2b),

$$X_\nu - X_{\nu\wedge p} = \sum_{n=1}^{\infty} V_n(X_n - X_{n-1})$$

with $V_n = 1_{\{\nu\ge n\}} - 1_{\{\nu\wedge p\ge n\}} = 1_{\{\nu\wedge p<n\le\nu\}}$. It is immediate to check that V_n is $\mathscr{F}_{n-1}$-measurable so that, by (1),

$$\mathbf{E}(|X_\nu - X_{\nu\wedge p}|) \le M\mathbf{E}\Big(\sum_{n=1}^{\infty} V_n\Big) = \mathbf{E}(\nu - \nu\wedge p) \underset{p\to\infty}{\to} 0$$

by the dominated convergence theorem.

4) Let $A \in \mathscr{F}_{\nu_1}$. Then $A \cap \{\nu_1 \le k\} \in \mathscr{F}_{\nu_1\wedge k}$ since for every $n \ge 0$

$$A\cap\{\nu_1 \le k\}\cap\{\nu_1\wedge k \le n\} = A\cap\{\nu_1 \le k\wedge n\} \in \mathscr{F}_{k\wedge n} \subset \mathscr{F}_n.$$

The optional sampling Theorem 3.3 implies that

$$\int_{A\cap\{\nu_1\le p\}} X_{\nu_2\wedge p}\,d\mathbf{P} \le \int_{A\cap\{\nu_1\le p\}} X_{\nu_1\wedge p}\,d\mathbf{P}.$$

Thus,

$$\int_A (X_{\nu_2} - X_{\nu_1})\,d\mathbf{P} = \int_{A\cap\{\nu_1\le p\}} (X_{\nu_2} - X_{\nu_1})\,d\mathbf{P} + \int_{A\cap\{\nu_1>p\}} (X_{\nu_2} - X_{\nu_1})\,d\mathbf{P} =$$
$$= \int_{A\cap\{\nu_1\le p\}} (X_{\nu_2\wedge p} - X_{\nu_1\wedge p})\,d\mathbf{P} + \int_{A\cap\{\nu_1\le p\}} (X_{\nu_2} - X_{\nu_2\wedge p})\,d\mathbf{P} +$$
$$+ \int_{A\cap\{\nu_1>p\}} (X_{\nu_2} - X_{\nu_1})\,d\mathbf{P} \le$$
$$\le \mathbf{E}(|X_{\nu_2} - X_{\nu_2\wedge p}|) + \mathbf{E}(|X_{\nu_2} - X_{\nu_1}| 1_{\{\nu_1>p\}}).$$

But the last two terms both tend to 0 as $p \to \infty$; the first one thanks to (3) whereas, for the second one, just apply Lebesgue's theorem, noting that X_{ν_1} and X_{ν_2} are integrable and ν_1 is finite a.s.

P3.1 1) It is sufficient to consider the case of a family consisting of just one r.v. $X \in L^1$. Let us show that $\mathbf{E}(|X| 1_{\{|X|>a\}}) \to_{a\to+\infty} 0$. Indeed $X \in L^1$ implies $|X| < +\infty$ a.s. and therefore $|X| 1_{\{|X|>a\}} \to 0$ a.s. as $a \to +\infty$; since $|X| 1_{\{|X|>a\}} \le |X| \in L^1$, one can apply Lebesgue's theorem.

2) Suppose the family $(X_i)_{i\in I}$ equi-integrable. Then, for every $A \in \mathscr{F}$,

$$\begin{aligned}\int_A |X_i|\,d\mathbf{P} &= \int_A |X_i| 1_{\{|X_i|\le a\}}\,d\mathbf{P} + \int_A |X_i| 1_{\{|X_i|>a\}}\,d\mathbf{P} \le \\ &\le a\mathbf{P}(A) + \int |X_i| 1_{\{|X_i|>a\}}\,d\mathbf{P}.\end{aligned} \tag{3.33}$$

First, if $A = \Omega$, one gets $\sup_{i\in I} \mathbf{E}|X_i| < +\infty$. Let us prove property (P). For a fixed $\varepsilon > 0$ one chooses first a so that $\sup_{i\in I} \mathbf{E}(|X_i| 1_{\{|X_i|>a\}}) < \frac{\varepsilon}{2}$ and then $\alpha = \varepsilon(2a)^{-1}$; (3.33) implies that $\int_A |X_i|\,d\mathbf{P} < \varepsilon$.

Conversely, let $M = \sup_{i\in I} \mathbf{E}|X_i|$. It holds (Markov inequality) $\mathbf{P}(|X_i| > a) \le a^{-1}\mathbf{E}|X_i|$ so that, for every $i \in I$, $\mathbf{P}(|X_i| > a) \le M/a$ and it is sufficient to use (P), choosing a large enough so that $M/a \le \alpha$.

3) Let $M = \sup_{i\in I} \mathbf{E}[g(|X_i|)]$. By hypothesis there exists a_0 such that, if $t > a_0$, $\frac{1}{t}g(t) > M\varepsilon^{-1}$ so that $t < \varepsilon M^{-1} g(t)$. Then, for every $a > a_0$ and every $i \in I$,

$$\int_{\{|X_i|>a\}} |X_i|\,d\mathbf{P} \le \varepsilon M^{-1} \int_{\{|X_i|>a\}} g(|X_i|)\,d\mathbf{P} \le \varepsilon M^{-1}\mathbf{E}[g(|X_i|)] \le \varepsilon.$$

In order to show that the sets that are bounded in L^p, $p > 1$ are equi-integrable, it is sufficient to apply this criterion to the function $g(t) = t^p$ (the condition $p > 1$ is essential, of course).

4a) If $X \in L^p$, $p > 1$, then, by Jensen inequality,

$$\mathbf{E}(|\mathbf{E}^{\mathscr{B}_i}(X)|^p) \le \mathbf{E}(\mathbf{E}^{\mathscr{B}_i}(|X|^p)) = \mathbf{E}(|X|^p).$$

The family $(\mathbf{E}^{\mathscr{B}_i}(X))_{i\in I}$ is therefore bounded in L^p and equi-integrable, thanks to (3).

4b) If X is integrable, again by Jensen inequality, $\sup_{i\in I} \mathbf{E}(|\mathbf{E}^{\mathscr{B}_i}(X)|) \le \mathbf{E}(|X|) < +\infty$. The first of the conditions of (2) is thus satisfied. Let us prove that the family $(\mathbf{E}^{\mathscr{B}_i}(X))_{i\in I}$ satisfies property (P). Let $Z \in L^2$ such that $\|X - Z\|_1 \le \frac{\varepsilon}{2}$. Let $\alpha > 0$

such that, if $\mathbf{P}(A) \leq \alpha$, then, according to (2) and (4a),

$$\int_A |\mathbf{E}^{\mathscr{B}_i}(Z)|\, d\mathbf{P} \leq \frac{\varepsilon}{2}$$

for every sub-σ-algebra $\mathscr{B}_i$. Then

$$\int_A |\mathbf{E}^{\mathscr{B}_i}(X)|\, d\mathbf{P} \leq \int_A |\mathbf{E}^{\mathscr{B}_i}(X) - \mathbf{E}^{\mathscr{B}_i}(Z)|\, d\mathbf{P} + \int_A |\mathbf{E}^{\mathscr{B}_i}(Z)|\, d\mathbf{P} \leq$$
$$\leq \mathbf{E}(|X - Z|) + \frac{\varepsilon}{2} \leq \varepsilon.$$

Property (P) is thus satisfied, which allows to conclude.

5a) Thanks to (2), $\sup_{n\geq 0} \mathbf{E}|X_n| < +\infty$. Moreover, by Fatou's lemma

$$\mathbf{E}|X| = \mathbf{E}\Big(\lim_{n\to\infty} |X_n|\Big) \leq \varliminf_{n\to\infty} \mathbf{E}|X_n| \leq \sup_{n\geq 0} \mathbf{E}|X_n| < +\infty$$

and therefore $X \in L^1$.

5b) $|f_a(X_n) - f_a(X)| \leq 2a$ and $|f_a(X_n) - f_a(X)| \to 0$ a.s. since f_a is continuous; therefore, by Lebesgue's theorem, $\mathbf{E}|f_a(X_n) - f_a(X)| \to 0$.

5c) Since the family $(X_n)_{n\geq 0}$ is equi-integrable,

$$\sup_{n\geq 0} \|X_n - f_a(X_n)\|_1 \leq \sup_{n\geq 0} \int_{\{|X_n|>a\}} |X_n|\, d\mathbf{P} \underset{a\to+\infty}{\longrightarrow} 0.$$

In the same way, $\|X - f_a(X)\|_1 \leq \mathbf{E}(|X| 1_{\{|X|>a\}}) \to_{a\to+\infty} 0$ by (1), since $X \in L^1$.

5d) Obviously

$$\|X_n - X\|_1 \leq \|X_n - f_a(X_n)\|_1 + \|f_a(X_n) - f_a(X)\|_1 + \|X - f_a(X)\|_1.$$

Let $\varepsilon > 0$. Choose first, thanks to (5c), a number $a > 0$ such that, for every n, $\|X_n - f_a(X_n)\|_1 + \|X - f_a(X)\|_1 < \frac{\varepsilon}{2}$ and then, thanks to (5b), an index n_0 such that, for $n \geq n_0$, $\|f_a(X_n) - f_a(X)\|_1 < \frac{\varepsilon}{2}$; this implies $\|X_n - X\|_1 < \varepsilon$ for $n \geq n_0$.

5e) It is sufficient to remark that

$$\int_{\{Z>a\}} Z\, d\mathbf{P} \geq \int_{\{|X_n|>a\}} |X_n|\, d\mathbf{P}$$

and then to use (1).

6) One has

$$\int_A |X_n|\, d\mathbf{P} \leq \int_A |X|\, d\mathbf{P} + \int_A |X_n - X|\, d\mathbf{P} \leq \int_A |X|\, d\mathbf{P} + \int_\Omega |X_n - X|\, d\mathbf{P}$$

whence (3.14). Since the condition $\sup_{n\geq 0} \mathbf{E}|X_n| < +\infty$ is immediate, thanks to the L^1 convergence of $(X_n)_{n\geq 0}$, it is sufficient to check (P) and to apply (2). Let $\varepsilon > 0$. By (2) there exists $\beta > 0$ such that $\int_A |X|\, d\mathbf{P} < \frac{\varepsilon}{2}$ if $\mathbf{P}(A) < \beta$; hence, thanks to (3.14), $\int_A |X_n|\, d\mathbf{P} < \varepsilon$ if $n \geq n_0$ and $\mathbf{P}(A) < \beta$. Still by (2), for every $k \leq n_0$ there exists α_k such that $\int_A |X_k|\, d\mathbf{P} < \varepsilon$ if $\mathbf{P}(A) < \alpha_k$. It is then sufficient to choose $\alpha = \inf\{\beta, \alpha_0, \ldots, \alpha_{n_0}\}$: for such a choice of α it holds $\int_A |X_n|\, d\mathbf{P} < \varepsilon$ for every n and the family $(X_n)_{n\geq 0}$ satisfies condition (P).

7) Suppose that $(X_n)_{n\geq 0}$ is a regular martingale. Therefore it converges in L^1 and, by (6), it is equi-integrable.

Suppose that the martingale $(X_n)_{n\geq 0}$ is equi-integrable. By (2) $\sup_{n\geq 0}\mathbf{E}(|X_n|) < +\infty$ and therefore it converges a.s. and also, by (5), in L^1. It is therefore regular.

8) Condition (3.15), applied to the function $g(t) = t\log^+ t$, implies, thanks to criterion (3), that the martingale is equi-integrable and therefore, by (7), regular.

P3.2 1a) By the Borel–Cantelli lemma, $\mathbf{P}(A) = 0$. Also, by Fatou's lemma,

$$\mathbf{Q}(A) = \mathbf{E}(1_A) = \mathbf{E}\Big(\overline{\lim_{n\to\infty}}\, 1_{A_n}\Big) = 1 - \mathbf{E}\Big(\underline{\lim_{n\to\infty}}\, 1_{A_n^c}\Big) \geq$$
$$\geq 1 - \underline{\lim_{n\to\infty}}\, \mathbf{E}(1_{A_n^c}) \geq 1 - (1-\alpha) = \alpha.$$

1b) Otherwise there would exist an $\alpha > 0$ and, for every n, an event A_n satisfying $\mathbf{P}(A_n) < 2^{-n}$ and $\mathbf{Q}(A_n) \geq \alpha$; this, due to (1a), would contradict the assumption of absolute continuity.

2) Let $\mathscr{C}$ be the (finite) family of intersections of events F_i and F_j^c, $i, j \leq n$. One chooses for $G_{n,k}$ the atoms of $\mathscr{C}$, i.e., the sets in $\mathscr{C}$ that do not contain a proper subset belonging to $\mathscr{C}$.

3) It follows, by the definition of X_n, that $\mathbf{E}(X_n 1_A) = \mathbf{Q}(A)$ for every $A \in \mathscr{F}_n$. But it also holds $A \in \mathscr{F}_{n+1}$ and therefore $\mathbf{E}(X_{n+1}1_A) = \mathbf{Q}(A)$, which proves that $(X_n)_{n\geq 1}$ is a martingale. As $X_n \geq 0$, this martingale converges a.s. to a r.v. X.

4) On one hand $\mathbf{E}(X_n 1_{\{X_n > a\}}) = \mathbf{Q}(X_n > a)$ and, on the other hand, $\mathbf{P}(X_n > a) \leq a^{-1}\mathbf{E}(X_n) \leq a^{-1}$. Let $\varepsilon > 0$; if $a > \eta^{-1}$, η being given by (1b), one has $\mathbf{Q}(X_n > a) < \varepsilon$, which proves that $(X_n)_{n\geq 1}$ is equi-integrable. It is therefore a regular martingale (Problem 3.1) and it converges in L^1.

5) For $A \in \mathscr{F}_n$, for every p, $\int_A X_n\, d\mathbf{P} = \int_A X_{n+p}\, d\mathbf{P}$. As $X_{n+p} \to_{p\to\infty} X$ in L^1,

$$\mathbf{Q}(A) = \int_A X_n\, d\mathbf{P} = \int_A X\, d\mathbf{P}.$$

Let us define $\mathbf{Q}'(A) = \int_A X\, d\mathbf{P}$. $\mathbf{Q}'$ is a probability measure as $\mathbf{E}(X) = 1$. Both probabilities $\mathbf{Q}$ and $\mathbf{Q}'$ coincide on $\bigcup_n \mathscr{F}_n$, which is a class that is stable by finite intersections. Therefore they coincide on $\mathscr{F} = \sigma(\bigcup_n \mathscr{F}_n)$.

P3.3 1a) $(A_n)_{n\geq 0}$ being increasing, $\{\sigma_a \leq n\} = \{A_{n+1} > a\} \in \mathscr{F}_n$ (the compensator is predictable).

1b) Let $(M_n)_{n\geq 0}$ be the martingale $(X_n - A_n)_{n\geq 0}$. By the optional sampling Theorem 3.3, $\mathbf{E}[M_{n\wedge\sigma_a}] = 0$; hence, $\mathbf{E}[X_{n\wedge\sigma_a}] = \mathbf{E}[A_{n\wedge\sigma_a}] \leq a$.

1c) We have seen that $\sup_{n\geq 0}\mathbf{E}[X_{n\wedge\sigma_a}] < +\infty$. As $X_n \geq 0$, this implies that the submartingale $(X_{n\wedge\sigma_a})_{n\geq 0}$ is bounded in L^1. It converges a.s. and therefore $(X_n)_{n\geq 0}$ converges a.s. on $\{\sigma_a = +\infty\}$.

1d) One has $\{A_\infty \leq a\} = \{\sigma_a = +\infty\} \subset \{X_n \text{ converges}\}$. It is sufficient to observe that $\{A_\infty < +\infty\} = \bigcup_{a>0}\{A_\infty \leq a\}$ and that $\{X_n \text{ converges}\} \subset \{\sup_{n\geq 0} X_n < +\infty\}$.

2a) On $\{\tau_a < +\infty\}$,

$$X_{\tau_a}^+ = (X_{\tau_a - 1} + \Delta_{\tau_a})^+ \leq X_{\tau_a-1}^+ + \Delta_{\tau_a}^+ \leq a + \Delta_{\tau_a}^+;$$

hence, $\mathbf{E}[X_{\tau_a}^+ 1_{\{\tau_a < +\infty\}}] \leq a + C_a$. Therefore, as $X_n \leq a$ on $\{\tau_a > n\}$,

$$\mathbf{E}[X_{n\wedge\tau_a}^+] = \mathbf{E}[X_{\tau_a}^+ 1_{\{\tau_a \leq n\}}] + \mathbf{E}[X_n^+ 1_{\{\tau_a > n\}}] \leq (a + C_a) + a.$$

2b)

$$\mathbf{E}|X_{n\wedge\tau_a}| = 2\mathbf{E}[X^+_{n\wedge\tau_a}] - \mathbf{E}[X_{n\wedge\tau_a}] \leq 4a + 2C_a$$

as, by the optional sampling Theorem 3.3, $\mathbf{E}[X_{n\wedge\tau_a}] \geq \mathbf{E}(X_0) = 0$. By Theorem 3.9, $(X_{n\wedge\tau_a})_{n\geq 0}$ converges a.s.; therefore $(X_n)_{n\geq 0}$ converges a.s. on $\{\tau_a = +\infty\}$.

2c) Obviously

$$\{X_n \text{ converges}\} \subset \left\{\sup_{n\geq 0} X_n < +\infty\right\}.$$

Conversely $\{\sup_{n\geq 0} X_n < +\infty\} = \bigcup_{a>0}\{\tau_a = +\infty\}$; hence, according to (2b),

$$\left\{\sup_{n\geq 0} X_n < +\infty\right\} \subset \{X_n \text{ converges}\}.$$

2d) One has

$$\mathbf{E}[A_{n\wedge\tau_a}] = \mathbf{E}[X_{n\wedge\tau_a}] \leq \mathbf{E}[X^+_{n\wedge\tau_a}] \leq 2a + C_a.$$

Therefore

$$\mathbf{E}[A_{\tau_a}] = \lim_{n\to\infty} \uparrow \mathbf{E}[A_{n\wedge\tau_a}] \leq 2a + C_a;$$

hence, $A_{\tau_a} < +\infty$ and $A_\infty < +\infty$ on $\{\tau_a = +\infty\}$ a.s.

2e) It is sufficient to observe that $\{\sup_{n\geq 0} X_n < +\infty\} = \bigcup_{a>0}\{\tau_a = +\infty\} \subset \{A_\infty < +\infty\}$.

2f) This follows from (1) and (2e).

3) The hypothesis implies that $(X_n)_{n\geq 0}$ and $(-X_n)_{n\geq 0}$ are in the class $\mathscr{C}^+$. One next applies (2c): on $[X_n$ does not converge], one has a.s. both $\sup_{n\geq 0} X_n = +\infty$ and $\sup_{n\geq 0}(-X_n) = -\inf_{n\geq 0} X_n = +\infty$, hence the result.

4) It is sufficient to observe that $X_n = \sum_{k=1}^n 1_{B_k}$ is a positive submartingale in the class $\mathscr{C}^+$, as its increments are bounded by 1. Its compensator is $A_n = \sum_{k=1}^n \mathbf{P}^{\mathscr{F}_{k-1}}(B_k)$ and one may then apply (2f).

5a) It is straightforward as (•3.8)

$$\mathbf{E}^{\mathscr{F}_n}[X_{n+1} - X_n] = \mathbf{E}^{\mathscr{F}_n}[(M_{n+1} - M_n)^2] = A_{n+1} - A_n.$$

5b) It follows from (1) that, on $\{A_\infty < +\infty\}$, $(M_n^2)_{n\geq 0}$ and $(M_n + 1^2)_{n\geq 0}$ converge a.s.; hence, by subtraction, also $(M_n)_{n\geq 0}$.

5c) One has a fortiori $\mathbf{E}(\sup_{n\geq 0}|M_{n+1} - M_n|) < +\infty$. By (3), either $(M_n)_{n\geq 0}$ converges a.s., or simultaneously $\overline{\lim}_{n\to\infty} M_n = +\infty$ and $\underline{\lim}_{n\to\infty} M_n = -\infty$. If one shows that $(M_n^2)_{n\geq 0}$ is in the class $\mathscr{C}^+$, one gets, because of (3) and (2), that

$$\{A_\infty < +\infty\} = \{M_n \text{ converges}\}$$

and the requested result. Let $\tau_a = \inf\{n \geq 1; M_n^2 > a\}$. Thanks to the inequality $|y^2 - x^2| \leq |y - x|^2 + 2|x||y - x|$, that, on $\{\tau_a < +\infty\}$,

$$\begin{aligned}|M^2_{\tau_a} - M^2_{\tau_a-1}| &\leq |M_{\tau_a} - M_{\tau_a-1}|^2 + 2|M_{\tau_a-1}||M_{\tau_a} - M_{\tau_a-1}| \leq \\ &\leq \Delta^2_{\tau_a} + 2a^{1/2}|\Delta_{\tau_a}|;\end{aligned}$$

hence,

$$\mathbf{E}(|M^2_{\tau_a} - M^2_{\tau_a-1}|1_{\{\tau_a<+\infty\}}) \leq \mathbf{E}(\Delta^2_{\tau_a}1_{\{\tau_a<+\infty\}}) + 2a^{1/2}\mathbf{E}(\Delta^2_{\tau_a}1_{\{\tau_a<+\infty\}})^{1/2} \leq$$

$$\leq \mathbf{E}\Big(\sup_{n\geq 0}\Delta_n^2\Big) + 2a^{1/2}\mathbf{E}\Big(\sup_{n\geq 0}\Delta_n^2\Big)^{1/2} < +\infty$$

and $(M_n^2)_{n\geq 0}$ actually belongs to the class $\mathscr{C}^+$.

P3.4 a) As it has been remarked, $X_n = \frac{k}{2^n}$ if and only if $X \in [\frac{k}{2^n}, \frac{k+1}{2^n}[$, so that $|X_n - X| \leq \frac{1}{2^n}$ and $X_n(\omega) \to_{n\to\infty} X(\omega)$ for every $\omega \in \Omega$.

b1) If $X_n = \frac{k}{2^n}$ then either $X \in [\frac{k}{2^n}, \frac{2k+1}{2^{n+1}}[$ or $X \in [\frac{2k+1}{2^{n+1}}, \frac{k+1}{2^n}[$. In the first case $X_{n+1} = \frac{k}{2^n}$ whereas in the second one $X_{n+1} = \frac{2k+1}{2^{n+1}}$. It is immediate that the relationship we are after is satisfied in both cases. This proves that $\sigma(X_n) \subset \sigma(X_{n+1})$, which implies that X_k is $\sigma(X_n)$-measurable for every $k \leq n$ and $\mathscr{F}_n = \sigma(X_n)$.

b2) X is the limit of the sequence $(X_n)_{n\geq 0}$ and, for each fixed $k \geq 0$, $(X_n)_{n\geq k}$ is a sequence of $\sigma(X_k, X_{k+1}, \ldots)$-measurable r.v.'s; therefore X is $\sigma(X_k, X_{k+1}, \ldots)$-measurable and, this being true for every k, X is measurable with respect to the tail σ-algebra $\bigcap_{n\geq 0}\sigma(X_n, X_{n+1}, \ldots)$.

Conversely, each of the r.v.'s X_n is $\sigma(X)$-measurable, since they can be expressed as functions of X. Therefore $\sigma(X_n, X_{n+1}, \ldots) \subset \sigma(X)$ from which we get the inclusion $\bigcap_{n\geq 0}\sigma(X_n, X_{n+1}, \ldots) \subset \sigma(X)$ and the equality of the σ-algebras.

c) Thanks to (b1) we must compute the conditional distribution of X_{n+1} given X_n. By the same argument as in (b1), if $X_n = \frac{k}{2^n}$ then either $X_{n+1} = \frac{k}{2^n}$ or $X_{n+1} = \frac{2k+1}{2^{n+1}}$ and both the events $\{X_{n+1} = \frac{k}{2^n}\}$ and $\{X_{n+1} = \frac{2k+1}{2^{n+1}}\}$ are contained in $\{X_n = \frac{k}{2^n}\}$. Thus

$$\begin{cases} \mathbf{P}(X_{n+1} = \frac{k}{2^n} \mid X_n = \frac{k}{2^n}) = \dfrac{\mathbf{P}(X_{n+1} = \frac{k}{2^n})}{\mathbf{P}(X_n = \frac{k}{2^n})} = \dfrac{1}{2} \\ \mathbf{P}(X_{n+1} = \frac{2k+1}{2^{n+1}} \mid X_n = \frac{k}{2^n}) = \dfrac{\mathbf{P}(X_{n+1} = \frac{2k+1}{2^{n+1}})}{\mathbf{P}(X_n = \frac{k}{2^n})} = \dfrac{1}{2} \end{cases} \tag{3.34}$$

whereas $\mathbf{P}(X_{n+1} = \frac{j}{2^{n+1}}) = 0$ for $j \neq 2k, 2k+1$. Taking the mean with respect to the conditional distribution and remarking that $\frac{2k+1}{2^{n+1}} = \frac{k}{2^n} + \frac{1}{2^{n+1}}$,

$$\begin{cases} \mathbf{E}(f(X_{n+1}) \mid X_n) = \frac{1}{2}(f(X_n) + f(X_n + \frac{1}{2^{n+1}})) \\ \mathbf{E}(f(X_{n+1} + \frac{1}{2^{n+1}}) \mid X_n) = \frac{1}{2}(f(X_n + \frac{1}{2^{n+1}}) + f(X_n + \frac{2}{2^{n+1}})). \end{cases} \tag{3.35}$$

We may now determine $\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n) = \mathbf{E}(Z_{n+1} \mid X_n)$:

$$\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n) = 2^{n+1}\cdot\frac{1}{2}(f(X_n + \tfrac{1}{2^n}) - f(X_n)) = Z_n.$$

$(Z_n)_{n\geq 0}$ is therefore a martingale, obviously bounded since the Lipschitz property implies $|f(X_n + \frac{1}{2^n}) - f(X_n)| \leq L2^{-n}$ and $|Z_n| \leq L$. Hence $(Z_n)_{n\geq 0}$ converges a.s. and in L^1.

d) Note that Z_n is $\sigma(X_n)$-measurable (as a function of X_n) and therefore measurable with respect to $\sigma(X_n, X_{n+1}, \ldots)$. The limit Z_∞ is thus $\sigma(X_n, X_{n+1}, \ldots)$-measurable for every n. Thanks to the equality of σ-algebras proved in (b), Z_∞ is $\sigma(X)$-measurable and there exists a Borel function g such that $Z_\infty = g(X)$. Actually this argument is not rigorous, as Z_∞ is *only the a.s.* limit of $(Z_n)_{n\geq 0}$ and negligible sets may not belong to $\sigma(X)$. Just note however that the r.v. $\overline{\lim}_{n\to\infty} Z_n$ is $\sigma(X)$-measurable and $Z_\infty = \overline{\lim}_{n\to\infty} Z_n$ a.s.

e) Let us determine the distribution of X given X_n. It is immediate that if $0 \le k \le 2^n - 1$ and $\frac{k}{2^n} \le a < b \le \frac{k+1}{2^n}$, then $\{a \le X < b\} \subset \{X_n = \frac{k}{2^n}\}$ and

$$\mathrm{P}(a \le X < b \mid X_n = \tfrac{k}{2^n}) = \frac{\mathrm{P}(a \le X < b)}{\mathrm{P}(X_n = \frac{k}{2^n})} = 2^n(b-a).$$

Hence the conditional distribution of X given X_n is the uniform distribution on the interval $[X_n, X_n + \frac{1}{2^n}[$. The martingale $(Z_n)_{n\ge0}$ is regular with limit Z_∞ and

$$\begin{aligned} Z_n = \mathrm{E}(Z_\infty \mid \mathscr{F}_n) = \mathrm{E}(g(X) \mid \mathscr{F}_n) = \mathrm{E}(g(X) \mid \sigma(X_n)) = \\ = 2^n \int_{X_n}^{X_n+2^{-n}} g(u)\,du. \end{aligned} \tag{3.36}$$

f) Since, for $0 \le k \le 2^n - 1$, $\mathrm{P}(X_n = \frac{k}{2^n}) = 2^{-n} > 0$, just note that on the event $\{X_n = \frac{k}{2^n}\}$, the equality (3.36) becomes

$$f(\tfrac{k+1}{2^n}) - f(\tfrac{k}{2^n}) = \int_{k2^{-n}}^{(k+1)2^{-n}} g(u)\,du$$

and, adding up,

$$f(\tfrac{k+1}{2^n}) - f(0) = \int_0^{(k+1)2^{-n}} g(u)\,du.$$

Using the fact that the dyadic numbers are dense in $[0, 1]$ and the continuity of f, it follows that (3.16) is true for every $x \in [0, 1]$.

P3.5 a) This point has already been worked out elsewhere (see (1) of Exercise 3.16, for instance). It is a consequence of the strong law of large numbers: it holds $\frac{1}{n} X_n \to m < 0$ as $n \to \infty$ so that $X_n \to_{n\to\infty} -\infty$ a.s. Thus $\sup_{n\ge0} X_n < +\infty$.

b1) Thanks to Hölder's inequality, for every $\alpha \ge 0$ and $\beta \ge 0$ such that $\alpha + \beta = 1$,

$$M(\alpha s + \beta t) = \mathrm{E}((\mathrm{e}^{sX})^\alpha (\mathrm{e}^{tX})^\beta) \le \mathrm{E}(\mathrm{e}^{sX})^\alpha \mathrm{E}(\mathrm{e}^{tX})^\beta = M(s)^\alpha M(t)^\beta,$$

which proves that ψ is convex.

b2) As $\lambda \ge 0$, $\mathrm{e}^{\lambda Y_1} \le \mathrm{e}^\lambda$, since $Y_1 \le 1$ a.s. Also $M(\lambda) < +\infty$ for $\lambda \ge 0$. Moreover $M(\lambda) \ge \mathrm{e}^\lambda \mathrm{P}(Y_1 = 1)$, which gives $\lim_{\lambda\to+\infty} \psi(\lambda) = +\infty$.

It is well known that $M'(0+) = \mathrm{E}(Y_i) = m$ (it is easy to check that derivation under the integral sign is allowed); hence

$$\psi'(0+) = \frac{M'(0+)}{M(0)} = m < 0$$

since $M(0) = 1$. ψ is continuous, vanishes at 0 with a right derivative that is strictly negative and converges to $+\infty$ as $\lambda \to +\infty$. It has necessarily another zero, λ_0, which is strictly positive. Thanks to the convexity of ψ this zero is unique.

c) We have already seen this computation many times: if $\mathscr{F}_n = \sigma(Y_1, \dots, Y_n)$, since Z_{n-1} is $\mathscr{F}_{n-1}$-measurable and Y_n is independent of $\mathscr{F}_{n-1}$,

$$\begin{aligned} \mathrm{E}(Z_n \mid \mathscr{F}_{n-1}) = \mathrm{E}(Z_{n-1}\,\mathrm{e}^{\lambda_0 Y_n} \mid \mathscr{F}_{n-1}) = Z_{n-1}\mathrm{E}(\mathrm{e}^{\lambda_0 Y_n}) = \\ = \mathrm{e}^{\psi(\lambda_0)}\, Z_{n-1} = Z_{n-1}. \end{aligned}$$

We remarked already that $X_n \to_{n\to\infty} -\infty$ a.s.; thus $\lim_{n\to\infty} Z_n = 0$ a.s.

d) This is almost immediate since

$$\begin{cases} \lim_{n\to\infty} Z_{n\wedge\tau_k} = \lim_{n\to\infty} Z_n = 0 & \text{on } \{\tau_k = +\infty\} \\ \lim_{n\to\infty} Z_{n\wedge\tau_k} = Z_{\tau_k} = e^{\lambda_0 k} & \text{on } \{\tau_k < +\infty\}. \end{cases}$$

We use here the assumptions made on the law of the r.v.'s $(Y_n)_{n\geq 1}$, which imply that X_n takes at most one step to the right and necessarily $X_{\tau_k} = k$.

e) The stopped martingale $(Z_{n\wedge\tau_k})_{n\geq 1}$ is bounded (it takes values between 0 and $e^{\lambda_0 k}$) and we can apply Lebesgue's theorem in (3.17), so that

$$1 = \mathbf{E}(Z_0) = \lim_{n\to\infty} \mathbf{E}(Z_{n\wedge\tau_k}) = e^{\lambda_0 k}\, \mathbf{P}(\tau_k < +\infty).$$

Thus $\mathbf{P}(\tau_k < +\infty) = e^{-\lambda_0 k}$. Since obviously $\mathbf{P}(\tau_k < +\infty) = \mathbf{P}(W \geq k)$, W follows a geometric law with parameter $p = 1 - e^{-\lambda_0}$.

With the given law for Y_n, the Laplace transform is immediately computed:

$$M(\lambda) = q\, e^{-\lambda} + p\, e^{\lambda}.$$

We must now determine the value $\lambda_0 > 0$ such that $M(\lambda_0) = 1$. This reduces to the equation of the second degree

$$p\, e^{2\lambda} - e^{\lambda} + q = 0.$$

Its two roots are $e^{\lambda} = 1$ (obviously) and $e^{\lambda} = \frac{q}{p}$. Thus $\lambda_0 = \log\frac{q}{p}$ In this case W has a geometric law with parameter $1 - e^{-\lambda_0} = 1 - \frac{p}{q}$.

P3.6 A1) Observe that $\alpha_n(1+\beta_n) = \alpha_{n-1}$. Therefore, a.s.,

$$\begin{aligned} \mathbf{E}^{\mathscr{F}_n}(Z'_{n+1}) &= \mathbf{E}^{\mathscr{F}_n}(\alpha_n Z_{n+1}) = \alpha_n \mathbf{E}^{\mathscr{F}_n}(Z_{n+1}) \leq \\ &\leq \alpha_n(1+\beta_n)Z_n + \alpha_n\xi_n - \alpha_n\eta_n = \alpha_{n-1}Z_n + \xi'_n - \eta'_n \end{aligned}$$

and also

$$\begin{aligned} \mathbf{E}^{\mathscr{F}_n}(U_{n+1}) &= \mathbf{E}^{\mathscr{F}_n}(Z'_{n+1}) - \sum_{k=0}^{n}(\xi'_k - \eta'_k) \leq \\ &\leq Z'_n + \xi'_n - \eta'_n - \sum_{k=0}^{n}(\xi'_k - \eta'_k) = U_n, \end{aligned}$$

i.e., $(U_n)_{n\geq 0}$ is a supermartingale.

A2) The process $(U_{n\wedge\tau_a})_{n\geq 0}$ is a supermartingale minorized by $-a$ and converges a.s., as $(U_{n\wedge\tau_a} + a)_{n\geq 0}$ is a positive supermartingale (Theorem 3.8). This implies that $(U_n)_{n\geq 1}$ converges a.s. on $\{\tau_a = +\infty\}$, as, on this event. the two sequences $(U_n)_{n\geq 1}$ and $(U_{n\wedge\tau_a})_{n\geq 1}$ coincide.

A3i) If $\sum_{n=1}^{\infty}\beta_n < +\infty$, then $\sum_{n=1}^{\infty}\log(1+\beta_n) < +\infty$ and

$$\alpha_n \underset{n\to\infty}{\to} \prod_{n=0}^{\infty}(1+\beta_n)^{-1} = \alpha_\infty \in]0, +\infty[.$$

One easily deduces that, on Γ, $\sum_{n=0}^{\infty}\xi'_n < +\infty$.

A3ii) As

$$\sum_{k=0}^{n} \eta'_k \le U_{n+1} + \sum_{k=0}^{n} \xi'_k,$$

one has $\sum_{n=0}^{\infty} \eta'_n < +\infty$ a.s. on $\Gamma \cap \{\tau_a = +\infty\}$. This implies the a.s. convergence of the sequences $(Z'_n)_{n\ge 0}$, $(Z_n)_{n\ge 0}$ and of $\sum_{n=0}^{\infty} \eta_n$ on $\Gamma \cap \{\tau_a = +\infty\}$. Now, on Γ, $\sup_{n\ge 0} \sum_{k=0}^{n}(\xi'_k - \eta'_k) < +\infty$, which implies that

$$\Gamma = \bigcup_{n\ge 0} \Gamma \cap \{\tau_n = +\infty\}$$

and the requested result follows.

B1) It holds

$$\begin{aligned} Z_{n+1} &= \langle X_n - x_0 + \gamma_n Y_{n+1}, X_n - x_0 + \gamma_n Y_{n+1} \rangle = \\ &= Z_n + \gamma_n^2 |Y_{n+1}|^2 + 2\gamma_n \langle X_n - x_0, Y_{n+1}\rangle. \end{aligned}$$

There exists $\bar{K}$ such that

$$\mathbf{E}^{\mathcal{F}_n}(|Y_{n+1}|^2) \le K(1 + |X_n|^2) \le \bar{K}(1 + Z_n) \qquad \text{a.s.}$$

and

$$\mathbf{E}^{\mathcal{F}_n}(Z_{n+1}) \le (1 + \bar{K}\gamma_n^2) Z_n + \bar{K}\gamma_n^2 + 2\gamma_n \langle X_n - x_0, f(X_n)\rangle \qquad \text{a.s.}$$

B2) By hypothesis $\gamma_n \langle X_n - x_0, f(X_n)\rangle \le 0$. One can therefore apply the results of (A), with $\xi_n = \beta_n = \bar{K}\gamma_n^2$ and $\eta_n = -2\gamma_n\langle X_n - x_0, f(X_n)\rangle$. As $\sum_{n=1}^{\infty} \gamma_n^2 < +\infty$ a.s., $\Gamma = \Omega$ a.s. and both sequences $(Z_n)_{n\ge 0}$ and $\sum_{k=1}^{n} -\gamma_k \langle X_k - x_0, f(X_k)\rangle$ converge a.s.

B3) Let ω be such that $Z(\omega) = \lim_{n\to\infty} Z_n(\omega) \ne 0$. There exists $0 < a < b$ and $N(\omega)$ such that, for every $n \ge N(\omega)$, $a \le |X_n(\omega) - x_0| \le b$. As f is continuous, there exists $c > 0$ such that $-\langle x - x_0, f(x)\rangle \ge c$ for $a \le |x - x_0| \le b$. Thus, for every $n \ge N(\omega)$,

$$-\langle X_n(\omega) - x_0, f(X_n(\omega))\rangle \ge c,$$

which implies that

$$\sum_{n=1}^{\infty} -\gamma_n \langle X_n(\omega) - x_0, f(X_n(\omega))\rangle = +\infty.$$

Therefore, by (B2), $\mathbf{P}(Z \ne 0) = 0$.

CHAPTER 4

Markov Chains

Transition Matrices, Markov Chains

•**4.1** Throughout this chapter we denote by E a countable set (possibly finite).

Definition 4.1 *A transition matrix on E is a family of real numbers $(P(x, y))_{x,y\in E}$ such that, for every $x, y \in E$,*

$$P(x, y) \geq 0, \qquad \sum_{y\in E} P(x, y) = 1. \tag{4.1}$$

Thus, for every $x \in E$, $P(x, \cdot)$ is a probability on E.

Definition 4.2 *Let μ be a probability and P a transition matrix on E. We call a (homogeneous) Markov chain with initial law μ and transition matrix P a stochastic process $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ with values in E such that*

i) $\mathbf{P}(X_0 \in A) = \mu(A)$, *for every $A \subset E$*

ii) $\mathbf{P}(X_{n+1} \in A \mid \mathscr{F}_n) = P(X_n, A)$ a.s., *for every $A \subset E$ and $n \geq 0$.*

It is possible to define *time inhomogeneous* Markov chains, for which (ii) of Definition 4.2 is replaced by

$$\mathbf{P}(X_{n+1} \in A \mid \mathscr{F}_n) = P_n(X_n, A)$$

where P_n is a transition matrix possibly depending on n. In the sequel we only take into account homogeneous Markov chains and the word "homogeneous" will always be understood.

Intuitively the quantity $P(x, y)$ represents the probability for the chain, which is in x at time n, to switch to y at time $n + 1$.

Formula (ii) of Definition 4.2 is called the *Markov property* (of which we shall see other, more powerful, forms). It implies the relation

$$\mathbf{P}(X_{n+1} \in A \mid \mathscr{F}_n) = \mathbf{P}(X_{n+1} \in A \mid X_n)$$

whose intuitive meaning has already been described in •2.2.

If $\mathscr{F}_n = \mathscr{F}_n^0 = \sigma(X_0, X_1, \dots, X_n)$ (the *natural filtration* of the process), the previous definition takes a more elementary form. Actually, since E is countable, the σ-algebra $\mathscr{F}_n^0$ is generated by the partition of Ω, formed by the events $\{X_0 = a_0, X_1 = a_1, \dots, X_n = a_n\}$, $(a_0, a_1, \dots, a_n) \in E^{n+1}$, and (ii) of Definition 4.2 can be replaced by

$$\mathbf{P}(X_{n+1} = b \mid X_n = a_n, X_{n-1} = a_{n-1}, \dots, X_0 = a_0) = P(a_n, b) \tag{4.2}$$

for every $n \geq 0$ and every $(a_0, a_1, \dots, a_n, b) \in E^{n+2}$ such that $\mathbf{P}(X_n = a_n, X_{n-1} = a_{n-1}, \dots, X_0 = a_0) > 0$.

•**4.2** Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ be a Markov chain with initial law μ and transition matrix P. The following computation shows that the finite distributions of X (•2.3) are entirely determined by μ and P. Actually, it holds

$$\begin{aligned}\mathbf{P}(X_0 = a_0, \dots, X_{n-1} = a_{n-1}, X_n = a_n) &= \\ = \mathbf{E}(1_{\{X_0=a_0,\dots,X_{n-1}=a_{n-1},X_n=a_n\}}) &= \\ = \mathbf{E}[1_{\{X_0=a_0,\dots,X_{n-1}=a_{n-1}\}}\mathbf{P}^{\mathscr{F}_{n-1}}(X_n = a_n)] &= \\ = \mathbf{P}(X_0 = a_0, \dots, X_{n-1} = a_{n-1})P(a_{n-1}, a_n).\end{aligned} \tag{4.3}$$

We can derive the following result, part (ii) being an immediate consequence of (4.2).

Theorem 4.3 *(i) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ be a Markov chain with values in E, with initial distribution μ and transition matrix P. It holds, for every $(a_0, \dots, a_n) \in E^{n+1}$,*

$$\begin{aligned}\mathbf{P}(X_0 = a_0, X_1 = a_1, \dots, X_{n-1} = a_{n-1}, X_n = a_n) &= \\ = \mu(a_0)P(a_0, a_1)\dots P(a_{n-1}, a_n).\end{aligned} \tag{4.4}$$

(ii) Conversely, if a process $X = (\Omega, \mathscr{F}, (X_n)_{n\geq 0}, \mathbf{P})$ with values in E satisfies (4.4) for a probability μ and a transition matrix P, X is a Markov chain with initial law μ and transition matrix P with respect to the natural filtration $(\mathscr{F}_n^0)_{n\geq 0}$.

•**4.3** We denote by $\mathscr{E}^+$ the set of all applications $E \to \bar{\mathbb{R}}^+$ and $\mathscr{M}^+$ the set of all positive measures on E. Obviously, E being countable, these measures can be identified with the family $\mu = (\mu(x))_{x\in E}$. Besides, given a σ-algebra $\mathscr{F}$ on Ω, we shall denote by $\mathscr{F}^+$ (resp. $b\mathscr{F}$) the measurable positive (resp. real and bounded) functions.

•**4.4** A positive matrix Q on E is a family $(Q(x, y))_{x,y\in E}$ where $Q(x, y) \in \bar{\mathbb{R}}^+$ (the value $+\infty$ is thus admissible). Given two such matrices Q and R, we note QR the product matrix, defined, as usual, by

$$QR(x, y) = \sum_{z\in E} Q(x, z)R(z, y)$$

(recall that, in $\bar{\mathbb{R}}^+$, $+\infty \cdot 0 = 0$). We shall write $Q^0 = I$, where $I(x, y) = 1_{\{x=y\}}$ and, for $n \geq 1$,

$$Q^n = \underbrace{Q\dots Q}_{n \text{ times}}.$$

If $f \in \mathscr{E}^+$, we write

$$Qf(x) = \sum_{y\in E} Q(x, y)f(y)$$

and then $f \to Qf$ is an operator $\mathscr{E}^+ \to \mathscr{E}^+$ such that $Q(f + g) = Qf + Qg$, $Q(\alpha f) = \alpha Qf$ for $f, g \in \mathscr{E}^+, \alpha \geq 0$; the composition of these operators corresponds to the product of the matrices, i.e., $R(Qf)(x) = (RQ)f(x)$. If $\mu \in \mathscr{M}^+$, we set

$$\mu Q(y) = \sum_{x\in E} \mu(x)Q(x, y)$$

and $\mu \to \mu Q$ is an operator from $\mathscr{M}^+$ to $\mathscr{M}^+$; we still have $(\mu Q)R(y) = \mu(QR)(y)$.

Moreover, if we define, for $f \in \mathscr{E}^+$ and $\mu \in \mathscr{M}^+$, the natural duality

$$\langle \mu, f \rangle = \int f \, d\mu = \sum_{x \in E} \mu(x) f(x),$$

then

$$\langle \mu, Qf \rangle = \langle \mu Q, f \rangle.$$

Finally, for $A \subset E$, we set

$$Q(x, A) = \sum_{y \in A} Q(x, y) = Q 1_A(x).$$

•4.5 The positive matrices such that, for every $x \in E$, $A \to Q(x, A)$ is a probability on E, are exactly the transition matrices. It is immediately checked that a positive matrix P is a transition matrix if and only if $P1 = 1$; in this case, if E is finite, the matrix P *always has* 1 *as an eigenvalue*. The product of two transition matrices is still a transition matrix and, if P is a transition matrix and μ a probability, μP is a probability.

If E is finite, we can identify it with $\{1, 2, \dots, m\}$. We then identify $f \in \mathscr{E}^+$ with the column vector with components $f(1), f(2), \dots, f(m)$ and Pf is the column vector which is the product of the matrix P by the column vector f. Similarly $\mu \in \mathscr{M}^+$ is identified with the line vector with components $\mu(1), \mu(2), \dots, \mu(m)$ and μP is the line vector which is the product of the line vector μ with the matrix P.

•4.6 Let us now derive some consequences of (4.4). If f is an application $E^{n+1} \to \mathbb{R}$, positive or bounded, it holds (it suffices to sum)

$$\begin{aligned} &\mathbf{E}[f(X_0, \dots, X_n)] = \\ &= \sum_{a_0, \dots, a_n \in E} f(a_0, \dots, a_n) \mu(a_0) P(a_0, a_1) \dots P(a_{n-1}, a_n). \end{aligned} \tag{4.5}$$

If in (4.5) we choose $f(x_0, \dots, x_n) = 1_{A_0}(x_0) \dots 1_{A_n}(x_n)$, $A_i \subset E$, then we get

$$\begin{aligned} &\mathbf{P}(X_0 \in A_0, \dots, X_n \in A_n) = \\ &= \sum_{a_0 \in A_0, \dots, a_n \in A_n} \mu(a_0) P(a_0, a_1) \dots P(a_{n-1}, a_n). \end{aligned} \tag{4.6}$$

Finally, if in (4.6) we choose $A_i = \{a_i\}$ for some of the indices and $A_j = E$ for the others, we have, for $l_1, \dots, l_k \in \mathbb{N}$,

$$\begin{aligned} &\mathbf{P}(X_0 = a_0, X_{l_1} = a_1, \dots, X_{l_1 + \dots + l_k} = a_k) = \\ &= \mu(a_0) P^{l_1}(a_0, a_1) \dots P^{l_k}(a_{k-1}, a_k). \end{aligned} \tag{4.7}$$

Thus for every $(a_0, \dots, a_n) \in E^{n+1}$ such that $\mathbf{P}(X_0 = a_0, \dots, X_n = a_n) > 0$,

$$\begin{aligned} &\mathbf{P}(X_{n+1} = b_1, \dots, X_{n+k} = b_k \mid X_0 = a_0, \dots, X_n = a_n) = \\ &= P(a_n, b_1) \dots P(b_{k-1}, b_k) \end{aligned} \tag{4.8}$$

and also

$$\mathbf{P}(X_{n+k} = b \mid X_0 = a_0, \dots, X_n = a_n) = P^k(a_n, b).$$

Construction and Existence

•**4.7** Let us fix a probability μ and a transition matrix P on E. Does a Markov chain, with values in E, with initial law μ and transition matrix P exist? In order to answer this question, we make use of Theorem 2.3 (Kolmogorov's theorem). Actually formula (4.4) determines the finite distributions of the process; we define a probability μ_n on E^{n+1} by

$$\mu_n(a_0, \dots, a_n) = \mu(a_0)P(a_0, a_1) \dots P(a_{n-1}, a_n).$$

This is a probability thanks to (4.1) (it suffices to sum first for $a_n \in E$, then for $a_{n-1} \in E$ and so on). We must check that these probabilities are compatible, i.e., satisfy (2.1). But, in this case, this condition can be written, for every $a_0, \dots, a_{n-1}$,

$$\mu_{n-1}(a_0, \dots, a_{n-1}) = \mu_n(\{a_0\} \times \dots \times \{a_{n-1}\} \times E)$$

and this equality is an immediate consequence of (4.1) (just sum over all possible values of $a_n \in E$). Thus there exists a unique probability $\mathbf{P}$ on the canonical space

$$\begin{aligned} &\Omega = E^{\mathbb{N}}, \quad \omega = (\omega_n)_{n\geq 0}, \quad X_n(\omega) = \omega_n, \\ &\mathscr{F}_n = \sigma(X_k, k \leq n), \quad \mathscr{F} = \sigma(X_k, k \geq 0) \end{aligned} \tag{4.9}$$

such that $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ is a process satisfying

$$\begin{aligned} \mathbf{P}(X_0 = a_0, \dots, X_n = a_n) &= \mu_n(a_0, \dots, a_n) = \\ &= \mu(a_0)P(a_0, a_1) \dots P(a_{n-1}, a_n). \end{aligned}$$

Because of Theorem 4.3 (ii), this process is a Markov chain with initial law μ and transition matrix P.

•**4.8** We denote by $\mathbf{P}_\mu$ the probability above, relative to the initial law μ. If $\mu = \delta_x$, δ_x being the Dirac mass at x defined by $\delta_x(y) = 1_{\{y=x\}}$, we write $\mathbf{P}_x$ instead of $\mathbf{P}_{\delta_x}$. Thanks to (4.4), for every event A of the form $\{X_0 = a_0, \dots, X_n \in a_n\}$,

$$\mathbf{P}_\mu(A) = \sum_{x\in E} \mu(x)\mathbf{P}_x(A). \tag{4.10}$$

It is easy to show, using Theorem 2.5, that this relation holds true for every event $A \in \mathscr{F}$.

Definition 4.4 *Given a transition matrix P on E, a canonical Markov chain with transition matrix P is a term*

$$(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$$

where $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0})$ is the canonical space defined in (4.9) and where, for every x, $\mathbf{P}_x$ is a probability on $(\Omega, \mathscr{F})$ satisfying, for every $a_0, \dots, a_n$,

$$\mathbf{P}_x(X_0 = a_0, \dots, X_n = a_n) = 1_{\{x\}}(a_0)P(a_0, a_1) \dots P(a_{n-1}, a_n). \tag{4.11}$$

The following formulae hold. For every $n \geq 0$, $f \in \mathscr{E}^+$, $\mu \in \mathscr{M}^+$,

$$\mathbf{P}_x(X_n = y) = P^n(x, y), \quad \mathbf{E}_x(f(X_n)) = P^n f(x), \tag{4.12}$$

$$\mathbf{E}_\mu(f(X_n)) = \sum_{x\in E} \mu(x)P^n f(x) = \langle \mu, P^n f \rangle = \langle \mu P^n, f \rangle. \tag{4.13}$$

In particular (4.13) shows that, if μ is the law of X_0, μP^n is the law of X_n.

Computations on the Canonical Chain

•**4.9** The main advantage of the canonical space is the use of the translation operators, which greatly simplify some computations.

On the canonical space (4.9), we define an application $\theta: \Omega \to \Omega$ through $\omega = (\omega_n)_{n\geq 0} \to \theta(\omega) = (\omega_{n+1})_{n\geq 0}$. This application is called the *translation operator* on Ω. We set then

$$\theta_0 = \text{Identity},\ \ \theta_1 = \theta,\ \ \theta_n = \theta_{n-1} \circ \theta = \theta \circ \theta_{n-1}. \tag{4.14}$$

It is easy to check that θ_p is a measurable application from $(\Omega, \mathscr{F})$ to $(\Omega, \sigma(X_k,\ k \geq p))$. Moreover, the operators θ_p are characterized, for every $n \geq 0$, by

$$X_n \circ \theta_p = X_{n+p}. \tag{4.15}$$

Let τ be a stopping time. We set

$$\Omega_\tau = \{\tau < +\infty\} \tag{4.16}$$

and define θ_τ, on Ω_τ, by

$$\theta_\tau = \theta_n \quad \text{on } \{\tau = n\}. \tag{4.17}$$

Then for every $n \geq 0$, on Ω_τ,

$$X_n \circ \theta_\tau = X_{n+\tau}.$$

•**4.10** The relation (4.8) can be written

$$\mathbf{P}_\mu(X_{n+1} = b_1, \ldots, X_{n+k} = b_k \mid \mathscr{F}_n) = P(X_n, b_1) \ldots P(b_{k-1}, b_k) = \\ = \mathbf{P}_{X_n}(X_1 = b_1, \ldots, X_k = b_k).$$

Thus, if $\Phi = 1_{\{X_0=b_0, X_1=b_1,\ldots,X_k=b_k\}}$, then a.s.

$$\mathbf{E}_\mu(\Phi \circ \theta_n \mid \mathscr{F}_n) = 1_{\{X_0=b_0\}}\mathbf{P}_\mu(X_{n+1} = b_1, \ldots, X_{n+k} = b_k \mid \mathscr{F}_n) = \\ 1_{\{X_0=b_0\}}\mathbf{P}_{X_n}(X_1 = b_1, \ldots, X_k = b_k) = \mathbf{E}_{X_n}(\Phi).$$

Using Theorem 2.5,

Theorem 4.5 (Markov Property) *Let X be a canonical Markov chain. One has, for every $n \geq 0$, every $\Phi \in b\mathscr{F}$ or $\mathscr{F}^+$ and every initial law μ,*

$$\mathbf{E}_\mu(\Phi \circ \theta_n \mid \mathscr{F}_n) = \mathbf{E}_{X_n}(\Phi) \qquad \text{a.s.}$$

or equivalently, for every $n \geq 0$, every $\Phi \in b\mathscr{F}$, every $\Psi \in b\mathscr{F}_n$ (resp. every $\Phi \in \mathscr{F}^+$ every $\Psi \in \mathscr{F}_n^+$) and every initial law μ,

$$\mathbf{E}_\mu[\Psi\Phi \circ \theta_n] = \mathbf{E}_\mu[\Psi\mathbf{E}_{X_n}(\Phi)]. \tag{4.18}$$

•**4.11** It is easy to extend this result to stopping times. Actually, thanks to (4.17) and to Proposition 2.7, $\mathbf{P}_\mu$-a.s. on $\{\tau = n\}$ it holds

$$\mathbf{E}_\mu(\Phi \circ \theta_\tau \mid \mathscr{F}_\tau) = \mathbf{E}_\mu(\Phi \circ \theta_n \mid \mathscr{F}_n)$$

from which one gets immediately:

Theorem 4.6 (Strong Markov property) *Let X be a canonical Markov chain and τ a stopping time. It holds, for every $n \geq 0$, every $\Phi \in b\mathscr{F}$ (or $\mathscr{F}^+$) and every initial law μ,*

$$\mathbf{E}_\mu(1_{\{\tau<+\infty\}}\Phi \circ \theta_\tau \mid \mathscr{F}_\tau) = 1_{\{\tau<+\infty\}}\mathbf{E}_{X_\tau}(\Phi) \qquad \text{a.s.,}$$

or, equivalently, for every $n \geq 0$, every $\Phi \in b\mathscr{F}$, every $\Psi \in b\mathscr{F}_\tau$ (resp. every $\Phi \in \mathscr{F}^+$, every $\Psi \in \mathscr{F}_\tau^+$) and every initial law μ,

$$\mathbf{E}_\mu[1_{\{\tau<+\infty\}}\Psi\Phi \circ \theta_\tau] = \mathbf{E}_\mu[1_{\{\tau<+\infty\}}\Psi\mathbf{E}_{X_\tau}(\Phi)]. \tag{4.19}$$

•**4.12** It is often more convenient to make computations on the canonical chain, using Theorems 4.6 and 4.5.

However a Markov chain is not, in general, defined on the canonical space. How can this be fixed? Starting from its transition matrix P, one can construct the associated canonical chain and make the computations on that one.

More precisely, let $X' = (\Omega', \mathscr{F}', (\mathscr{F}'_n)_{n\geq 0}, (X'_n)_{n\geq 0}, \mathbf{P}')$ be a Markov chain with initial law μ and transition matrix P (Definition 4.2) and let us consider the canonical chain $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ with transition matrix P (Definition 4.4); then, for every $(a_0, \ldots, a_n) \in E^{n+1}$,

$$\begin{aligned}\mathbf{P}'(X'_0 = a_0 \ldots, X'_n = a_n) &= \mu(a_0)P(a_0, a_1)\ldots P(a_{n-1}, a_n) = \\ &= \mathbf{P}_\mu(X_0 = a_0 \ldots, X_n = a_n);\end{aligned}$$

hence, by a monotone class argument (•2.4), for every $\Phi \in b\mathscr{F}$ or $\mathscr{F}^+$, it holds, noting $\mathbf{E}'$ the expectation with respect to $\mathbf{P}'$,

$$\mathbf{E}'[\Phi(X'_0, X'_1, \ldots, X'_n, \ldots)] = \mathbf{E}_\mu[\Phi(X_0, X_1, \ldots, X_n, \ldots)].$$

The two processes X and X' have thus same law. For instance, if one needs to compute $\mathbf{P}'(X'$ visits $A)$, it suffices to compute $\mathbf{P}_\mu(X$ visits $A)$, which has the same value.

Potential Operators

•**4.13** Let P be a transition matrix and X the associated canonical chain.

Definition 4.7 *We call potential matrix of P (or of X) the matrix*

$$U = I + P + \ldots + P^n + \ldots = \sum_{k\geq 0} P^k.$$

If $f \in \mathscr{E}^+$, recalling (4.12), it holds

$$Uf(x) = \sum_{k\geq 0} P^k f(x) = \sum_{k\geq 0} \mathbf{E}_x(f(X_k)) = \mathbf{E}_x\Big[\sum_{k\geq 0} f(X_k)\Big]. \tag{4.20}$$

In particular, if $f = 1_A$, $A \subset E$,

$$U(x, A) = U1_A(x) = \mathbf{E}_x\Big[\sum_{k\geq 0} 1_A(X_k)\Big] = \mathbf{E}_x(N_A), \tag{4.21}$$

where

$$N_A = \sum_{k\geq 0} 1_A(X_k). \tag{4.22}$$

We see that $N_A(\omega)$ is equal to the number of the indices (or of the times) $n \geq 0$ such that $X_n(\omega) \in A$; thus $N_A(\omega)$ is the number of visits of the path ω to the set A. For $A = \{y\}$,

$$U(x, y) = \mathbf{E}_x(N_y). \tag{4.23}$$

Thus $U(x, y)$ is the mean number of visits in y of the chain starting at x.

•4.14 Let $f \in \mathcal{E}^+$: the function $v = Uf$ is called the *potential* of f. Since U satisfies

$$U = I + P + \dots + P^n + \dots = I + P(I + P + \dots + P^n + \dots) = I + PU,$$

it holds

$$v = f + Pv. \tag{4.24}$$

It is quite easy to characterize v among the solutions of this equation. Actually let $w \in \mathcal{E}^+$ be such that $w = f + Pw$. Obviously $w \geq f$; let us assume $w \geq \sum_{k=0}^n P^k f$, then

$$w = f + Pw \geq f + P\sum_{k=0}^{n} P^k f = \sum_{k=0}^{n+1} P^k f.$$

By induction $w \geq \sum_{k=0}^n P^k f$ for every n and, taking the limit, $w \geq v$. Since we know already that Uf is a solution of (4.24),

Proposition 4.8 *If $f \in \mathcal{E}^+$, $v = Uf$ is the smallest positive solution of the equation $v = f + Pv$.*

This result applied to $f = 1_{\{y\}}$ gives $U(x, y) = U1_{\{y\}}(x)$ and we get

Corollary 4.9 *For every $y \in E$, $u_y(x) = U(x, y)$ is the smallest positive solution of the equation*

$$u(x) = \begin{cases} 1 + Pu(x) & \text{if } x = y \\ Pu(x) & \text{if } x \neq y. \end{cases}$$

Remark. Speaking of positive solution we mean a solution with values in $[0, +\infty]$. In particular, $u \equiv +\infty$ for every x is a positive solution.

•4.15 The following result highlights the link between martingales and Markov chains. A function $f \in \mathcal{E}^+$ is said to be *excessive* (or P-excessive if there is a need to specify the transition matrix) if $f \geq Pf$; it is said to be *invariant* (or *harmonic*) if $f = Pf$. Let us point out that, for $g \in \mathcal{E}^+$, the potential of g, Ug, is an excessive function since $Ug = g + PUg \geq PUg$. Similarly, a measure μ is said to be *excessive* if $\mu P \leq \mu$, *invariant* (or *stationary*) if $\mu P = \mu$.

Proposition 4.10 *If $f \in \mathcal{E}^+$ is an excessive function (resp. invariant), then $Y_n = f(X_n)$ is a positive supermartingale (resp. martingale) for every law $\mathbf{P}_\mu$.*

Actually, thanks to Theorem 4.5,

$$\mathbf{E}_\mu^{\mathcal{F}_n}(Y_{n+1}) = \mathbf{E}_\mu^{\mathcal{F}_n}(f(X_1) \circ \theta_n) = \mathbf{E}_{X_n}(f(X_1)) = Pf(X_n) \leq f(X_n) = Y_n \quad \text{a.s.}$$

if f is excessive whereas the equality sign holds if it is invariant.

Passage Problems

A typical problem with Markov chains is the following: if $F \subset E$ and $x \in E \setminus F$, what is the probability of visiting F for the chain starting at x (i.e., what is the value of $\mathbf{P}_x(\tau_F < +\infty)$)? *And what is the law of X_{τ_F} (hitting law in F)? In this paragraph we develop the tools needed in order to tackle these problems.*

•**4.16** Let P be a transition matrix, X the associated canonical chain and τ a stopping time. Let us define, for $f \in \mathcal{E}^+$,

$$P_\tau f(x) = \mathbf{E}_x(1_{\{\tau<+\infty\}} f(X_\tau)). \tag{4.25}$$

P_τ is the operator associated to the (positive) matrix

$$P_\tau(x, y) = \mathbf{P}_x(\tau < +\infty, X_\tau = y).$$

Let us remark that

$$P_\tau 1(x) = \mathbf{P}_x(\tau < +\infty).$$

Let σ be a second stopping time (possibly $\sigma = \tau$!). It is easy to show that the r.v. ρ defined by $\rho = \tau + \sigma \circ \theta_\tau$ on $\{\tau < +\infty\}$ and $\rho = +\infty$ on $\{\tau = +\infty\}$ is also a stopping time and that,

$$\begin{aligned} &\{\rho < +\infty\} = \{\tau < +\infty\} \cap \{\sigma \circ \theta_\tau < +\infty\}, \\ &X_\rho = X_\sigma \circ \theta_\tau \quad \text{on } \{\rho < +\infty\}. \end{aligned} \tag{4.26}$$

For a better understanding of the stopping time ρ, one can think of the case $\tau = \tau_A$, $\sigma = \tau_B$ where $A, B \subset E$. A moment of reflection shows that $\tau_A + \tau_B \circ \theta_{\tau_A}$ is the first visiting time of the process in B after visiting A. Look at Exercise 2.2 for more information. It holds then

Proposition 4.11 *If σ and τ are two stopping times and $\rho = \tau + \sigma \circ \theta_\tau$, then $P_\rho = P_\tau P_\sigma$.*

Actually, using (4.26) and the strong Markov property, it holds for $f \in \mathcal{E}^+$

$$\begin{aligned} P_\tau P_\sigma f(x) &= \mathbf{E}_x[1_{\{\tau<+\infty\}} P_\sigma f(X_\tau)] = \mathbf{E}_x[1_{\{\tau<+\infty\}} \mathbf{E}_{X_\tau}(1_{\{\sigma<+\infty\}} f(X_\sigma))] = \\ &= \mathbf{E}_x[1_{\{\tau<+\infty\}} 1_{\{\sigma\circ\theta_\tau<+\infty\}} f(X_\sigma \circ \tau)] = \mathbf{E}_x[1_{\{\rho<+\infty\}} f(X_\rho)] = P_\rho f(x). \end{aligned}$$

•**4.17** Let $F \subset E$ and consider

$$\tau_F = \inf\{n \geq 0; X_n \in F\}, \qquad \sigma_F = \inf\{n \geq 1; X_n \in F\}, \tag{4.27}$$

which are, respectively, the hitting and return times in F. These are the stopping times that were introduced in •2.5. We shall write more simply P_F instead of P_{τ_F}. It holds then

$$P_F(x, y) = \mathbf{P}_x(\tau < +\infty, X_{\tau_F} = y), \quad P_F f(x) = \mathbf{E}_x(1_{\{\tau_F<+\infty\}} f(X_{\tau_F}))$$

and also

$$\begin{aligned} &\mathbf{P}_x(\tau_F = 0) = 1 \text{ if } x \in F, && \mathbf{P}_x(\tau_F = \sigma_F) = 1 \text{ if } x \in F^c, \\ &\sigma_F = 1 + \tau_F \circ \theta_1, && X_{\sigma_F} = X_{\tau_F} \circ \theta_1 \text{ on } \{\sigma_F < +\infty\} \end{aligned}$$

and finally (Proposition 4.11) $P_{\sigma_F} = P P_F$.

Let $g \in \mathcal{E}^+$ and $u(x) = P_F g(x)$. If $x \in F$, then $u(x) = g(x)$ since $\tau_F = 0$ $\mathbf{P}_x$-a.s. Conversely if $x \in F^c$, then $\tau_F = \sigma_F$ $\mathbf{P}_x$-a.s. and

$$u(x) = \mathbf{E}_x(1_{\{\sigma_F<+\infty\}} g(X_{\sigma_F})) = P_{\sigma_F} g(x) = P(P_F g)(x) = Pu(x).$$

In summary, u is a solution of

$$u(x) = \begin{cases} g(x) & \text{if } x \in F \\ Pu(x) & \text{if } x \notin F \end{cases} \tag{4.28}$$

and it can be proved that

Theorem 4.12 *For $g \in \mathcal{E}^+$, $u(x) = P_F g(x)$ is the smallest positive solution of (4.28).*

There are two particular cases that deserve a closer look. If one chooses $g = 1$,

Corollary 4.13 *Let $u(x) = \mathbf{P}_x(\tau_F < +\infty)$. Then u is the smallest positive solution of*

$$u(x) = \begin{cases} 1 & \text{if } x \in F \\ Pu(x) & \text{if } x \notin F. \end{cases}$$

Conversely, if $g(x) = 1_{\{y\}}(x)$, then

Corollary 4.14 *For every $y \in F$, $u_y(x) = P_F(x, y)$ is the smallest positive solution of*

$$u(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \in F \setminus \{y\} \\ Pu(x) & \text{if } x \notin F. \end{cases}$$

•**4.18** Similarly, it is possible to characterize $v(x) = \mathbf{E}_x(\tau_F)$ (mean time of passage in F). Actually if $x \in F$, $v(x) = 0$ and, if $x \in F^c$, using the Markov property (Theorem 4.5),

$$\begin{aligned} v(x) &= \mathbf{E}_x(\tau_F) = \mathbf{E}_x(\sigma_F) = 1 + \mathbf{E}_x(\tau_F \circ \theta_1) = \\ &= 1 + \mathbf{E}_x(\mathbf{E}_{X_1}(\tau_F)) = 1 + \mathbf{E}_x(v(X_1)) = 1 + Pv(x) \end{aligned}$$

and it is possible to prove:

Corollary 4.15 *Let $v(x) = \mathbf{E}_x(\tau_F)$. Then v is the smallest positive solution of*

$$v(x) = \begin{cases} 0 & \text{if } x \in F \\ 1 + Pv(x) & \text{if } x \notin F. \end{cases} \tag{4.29}$$

Recurrence, Transience

In this section, X is a canonical Markov chain with transition matrix P and potential matrix U. Thus

$$P(x, y) = \mathbf{P}_x(X_1 = y), \quad U(x, y) = \mathbf{E}_x(N_y), \quad N_x = \sum_{n \geq 0} 1_{\{x\}}(X_n).$$

•**4.19** Let us consider

$$\tau_x = \inf\{n \geq 0,\ X_n = x\}, \quad \sigma_x = \inf\{n \geq 1,\ X_n = x\}, \tag{4.30}$$

respectively, the hitting and return times in x; let us define a sequence of stopping times by $\sigma_x^1 = \sigma_x$ and

$$\sigma_x^n = \begin{cases} \sigma_x^{n-1} + \sigma_x \circ \theta_{\sigma_x^{n-1}} & \text{if } \sigma_x^{n-1} < +\infty \\ \sigma_x^n = +\infty & \text{otherwise.} \end{cases} \tag{4.31}$$

The sequence $(\sigma_x^n)_{n \geq 1}$ is formed by the subsequent times of passage in x of the chain, the time 0 possibly excepted (see Exercise 2.2 for details). The following easy properties hold:

$$N_x = 1_{\{x\}}(X_0) + \sum_{n \geq 1} 1_{\{\sigma_x^n < +\infty\}}. \tag{4.32}$$

$$X_{\sigma_x^n} = x \quad \text{on } \{\sigma_x^n < +\infty\}. \tag{4.33}$$

By the strong Markov property (Theorem 4.6),

$$\begin{aligned}\mathbf{P}_x(\sigma_x^n < +\infty) &= \mathbf{P}_x(\sigma_x^{n-1} + \sigma_x \circ \theta_{\sigma_x^{n-1}} < +\infty) = \\ &= \mathbf{P}_x(\sigma_x^{n-1} < +\infty, \sigma_x \circ \theta_{\sigma_x^{n-1}} < +\infty) = \\ &= \mathbf{E}_x\left[1_{\{\sigma_x^{n-1}<+\infty\}} 1_{\{\sigma_x<+\infty\}} \circ \theta_{\sigma_x^{n-1}}\right] = \\ = \mathbf{E}_x\left[1_{\{\sigma_x^{n-1}<+\infty\}} \mathbf{P}_{X_{\sigma_x^{n-1}}}(\sigma_x < +\infty)\right] &= \mathbf{P}_x(\sigma_x^{n-1} < +\infty)\mathbf{P}_x(\sigma_x < +\infty)\end{aligned}$$

from which we derive the important relation

$$\mathbf{P}_x(\sigma_x^n < +\infty) = \mathbf{P}_x(\sigma_x < +\infty)^n. \tag{4.34}$$

Two cases are then possible.

(i) $\mathbf{P}_x(\sigma_x < +\infty) = 1$. Then, for every $n \geq 0$, $\mathbf{P}_x(\sigma_x^n < +\infty) = 1$ and, thanks to (4.32), $N_x = +\infty$ $\mathbf{P}_x$-a.s. and $U(x, x) = \mathbf{E}_x(N_x) = +\infty$.

(ii) $\mathbf{P}_x(\sigma_x < +\infty) = a < 1$. Then again (4.32) gives

$$\begin{aligned}U(x, x) = \mathbf{E}_x(N_x) &= 1 + \sum_{n\geq 1} \mathbf{P}_x(\sigma_x^n < +\infty) = \\ &= 1 + \sum_{n\geq 1} a^n = (1-a)^{-1} < +\infty,\end{aligned}$$

which implies $\mathbf{P}_x(N_x < +\infty) = 1$. We proved:

Theorem 4.16 *Let $x \in E$.*

(i) If $U(x, x) = +\infty$, then $\mathbf{P}_x(\sigma_x < +\infty) = 1$ and $\mathbf{P}_x(N_x = +\infty) = 1$. The state x is called recurrent.

(ii) If $U(x, x) < +\infty$, then $\mathbf{P}_x(\sigma_x < +\infty) < 1$ and $\mathbf{P}_x(N_x < +\infty) = 1$. The state x is called transient.

The terms employed are clear: if x is recurrent, then starting at x the chain visits x infinitely many times. Conversely, if x is transient, the chain only visits x finitely many times; thus there exists a certain time after which the chain never visits x any more.

•**4.20** The quantities $U(x, y)$ and $\mathbf{P}_x(\sigma_y < +\infty)$ are related according to the following relations.

Proposition 4.17 *If $x \neq y$ then $U(x, y) = \mathbf{P}_x(\sigma_y < +\infty)U(y, y)$.*

Actually, by Theorem 4.6, since $X_{\sigma_y} = y$ on $\{\sigma_y < +\infty\}$ and $N_y = 1_{\{\sigma_y<+\infty\}} N_y \circ \theta_{\sigma_y}$,

$$\begin{aligned}U(x, y) = \mathbf{E}_x[N_y] &= \mathbf{E}_x[1_{\{\sigma_y<+\infty\}} N_y \circ \theta_{\sigma_y}] = \mathbf{E}_x[1_{\{\sigma_y<+\infty\}} \mathbf{E}_{X_{\sigma_y}}(N_y)] = \\ &= \mathbf{E}_x[1_{\{\sigma_y<+\infty\}} \mathbf{E}_y(N_y)] = \mathbf{P}_x(\sigma_y < +\infty)U(y, y).\end{aligned}$$

Remark 1. Proposition 4.17 shows that the function $x \to U(x, y)$ attains its maximum at y. This is the *maximum principle*.

Remark 2. If E is finite, there exists at least one recurrent state. Actually $\sum_{y\in E} N_y = N_E = +\infty$ a.s., from which

$$\sum_{y\in E} U(x, y) = \mathbf{E}_x\Big(\sum_{y\in E} N_y\Big) = \mathbf{E}_x(N_E) = +\infty.$$

Hence there exists $y \in E$ such that $U(x, y) = +\infty$. But (Proposition 4.17) $U(y, y) \geq U(x, y) = +\infty$ and y is recurrent.

Remark 3. If y is transient, then $U(x, y) < +\infty$ for every $x \in E$. Thus $P^n(x, y)$ is the general term of a summable series and, necessarily, $P^n(x, y) \to_{n\to\infty} 0$.

•**4.21** We have studied the returns to a state x; let us now look at the communications between states.

We say that x *leads* to y, noted $x \rightsquigarrow y$, if $\mathbf{P}_x(\tau_y < +\infty) > 0$. Let us remark that $x \rightsquigarrow x$ and that, for $y \neq x$, $\mathbf{P}_x(\tau_y < +\infty) = \mathbf{P}_x(\sigma_y < +\infty)$; thus, since $\{\sigma_y < +\infty\} = \bigcup_{n\geq 1}\{X_n = y\}$, for every $x, y \in E$, $x \neq y$, the following three conditions are equivalent:

$$\begin{aligned}&\text{a) } \mathbf{P}_x(\sigma_y < +\infty) > 0,\\ &\text{b) there exists } n \geq 1 \text{ such that } P^n(x, y) > 0,\\ &\text{c) } U(x, y) > 0.\end{aligned} \tag{4.35}$$

An important point is that if $x \rightsquigarrow y$ and $y \rightsquigarrow z$, then $x \rightsquigarrow z$ since

$$\begin{aligned}\mathbf{P}_x(\tau_z < +\infty) &\geq \mathbf{P}_x(\tau_y + \tau_z \circ \theta_{\tau_y} < +\infty) = \\ = \mathbf{E}_x[1_{\{\tau_y<+\infty\}} 1_{\{\tau_z<+\infty\}} \circ \theta_{\tau_y}] &= \mathbf{E}_x[1_{\{\tau_y<+\infty\}}\mathbf{P}_{X_{\tau_y}}(\tau_z < +\infty)] = \\ &= \mathbf{P}_x(\tau_y < +\infty)\mathbf{P}_y(\tau_z < +\infty) > 0.\end{aligned}$$

Assume that x is recurrent and that $x \rightsquigarrow y$; then y satisfies the following properties.

(i) Thanks to (4.35), there exists $n \geq 1$ such that $P^n(x, y) > 0$. This gives, by Proposition 4.17,

$$\begin{aligned}U(y, y) \geq U(x, y) &= \sum_{k\geq 0} P^k(x, y) \geq \sum_{k\geq 0} P^{n+k}(x, y) \geq \\ &\geq \sum_{k\geq 0} P^k(x, x)P^n(x, y) = U(x, x)P^n(x, y) = +\infty.\end{aligned}$$

Thus $U(y, y) = +\infty$ and y is recurrent.

(ii) Since x is recurrent, we know that $\mathbf{P}_x$-a.s. the chain visits x infinitely many times; thus, on $\{\sigma_y < +\infty\}$, it holds $\sigma_x \circ \theta_{\sigma_y} < +\infty$ $\mathbf{P}_x$-a.s. and, since $x \rightsquigarrow y$,

$$\begin{aligned}0 < \mathbf{P}_x(\sigma_y < +\infty) &= \mathbf{P}_x(\sigma_y < +\infty, \sigma_x \circ \theta_{\sigma_y} < +\infty) = \\ &= \mathbf{P}_x(\sigma_y < +\infty)\mathbf{P}_y(\sigma_x < +\infty),\end{aligned}$$

which gives $\mathbf{P}_y(\sigma_x < +\infty) = 1$ and $y \rightsquigarrow x$. By symmetry, $\mathbf{P}_x(\sigma_y < +\infty) = 1$.

(iii) More generally

$$\begin{aligned}\mathbf{P}_x(\sigma_y^n < +\infty) &= \mathbf{P}_x(\sigma_y^{n-1} < +\infty, \sigma_y \circ \theta_{\sigma_y^{n-1}} < +\infty) = \\ = \mathbf{P}_x(\sigma_y^{n-1} < +\infty)\mathbf{P}_y(\sigma_y < +\infty) &= \mathbf{P}_x(\sigma_y^{n-1} < +\infty) = \\ &= \ldots = \mathbf{P}_x(\sigma_y < +\infty) = 1,\end{aligned}$$

from which we get $\mathbf{P}_x(N_y = +\infty) = 1$ and, by symmetry, $\mathbf{P}_y(N_x = +\infty) = 1$.

Remark. By (ii) we derive the important criterion: *If there exists y such that $x \rightsquigarrow y$ and $y \not\rightsquigarrow x$, then x is transient.*

We say that x and y communicate if $x \rightsquigarrow y$ and $y \rightsquigarrow x$ both hold and we shall denote this property $x \leftrightsquigarrow y$; $\leftrightsquigarrow$ is clearly an equivalence relation on E. Let $(C_i, i \in I)$ be the partition of E associated to this equivalence relation (i.e., if $x \in C_i$, C_i is the set of all states that communicate with x). By the previous considerations it comes that, if a class C_i contains a recurrent state, all the states in C_i are recurrent and the states in C_i do not lead to any state in C_i^c. We have shown:

Proposition 4.18 *There exists a partition $(C_i, i \in I)$ of the set of recurrent states of E such that,*

(i) if $x, y \in C_i$, $U(x, y) = +\infty$ and $\mathbf{P}_x(N_y = +\infty) = 1$;
(ii) If $x \in C_i$ and $y \in C_j$, $i \neq j$, $U(x, y) = 0$ and $\mathbf{P}_x(N_y = 0) = 1$.

•4.22 A subset F of E is said to be *closed* if, for every $x \in F$, it holds $\mathbf{P}_x(X_n \in F$ for every $n \geq 0) = 1$. If $F = \{x\}$ is closed, the state x is said to be *absorbing*. If E is the only closed subset (i.e., if all states communicate), the chain is said to be *irreducible*. It is clear that a chain is irreducible if and only if, for every $x, y \in E$, $U(x, y) > 0$. A transition matrix is said to be *irreducible* if the associated chain is irreducible. It is easy to show:

Proposition 4.19 *F is closed if and only if, for every $x \in F$, $P(x, F) = 1$ or if and only if, for every $x \in F$ and every $y \notin F$, $U(x, y) = 0$.*

It follows from •4.21 that, for an irreducible chain, all states are either recurrent and the chain is said to be *recurrent*, or transient and the chain is said to be *transient*. In the first case, it holds, for every $x, y \in E$, $U(x, y) = +\infty$; in the second one for every $x, y \in E$, $0 < U(x, y) < +\infty$. Of course the second case cannot happen unless E is infinite (see Remark 2 of •4.20): *an irreducible chain on a finite set of states is always recurrent.*

•4.23 We end this paragraph with an important remark: if X is an irreducible Markov chain and X has an invariant probability λ (•4.15), then X is recurrent.

Actually

$$\lambda(y) = \lambda P^n(y) = \sum_{x \in E} \lambda(x) P^n(x, y). \tag{4.36}$$

Assume that the chain is transient. Then $U(x, y) < +\infty$ for every $x, y \in E$ and $P^n(x, y) \to_{n \to \infty} 0$. As $P^n(x, y) \leq 1$, we can apply Lebesgue's theorem and take the limit as $n \to \infty$ in (4.36). We get $\lambda(y) = 0$ for every $y \in E$, which is impossible, since λ is a probability.

Recurrent Irreducible Chains

•4.24 Let us consider a canonical irreducible recurrent Markov chain X. Then, for every $x, y \in E$, $U(x, y) = +\infty$. We shall write $Z_n \to_{n \to \infty} Z$ a.s. for $Z_n \to_{n \to \infty} Z$, $\mathbf{P}_\mu$-a.s. *for every* initial law μ.

A first property is:

Proposition 4.20 *If X is recurrent irreducible and if f is an excessive function (•4.15), then f is constant.*

Actually $(f(X_n))_{n \geq 0}$ is a positive supermartingale and thus converges a.s. (Theorem

3.8) to a r.v. Z. Let $x, y \in E, x \neq y$. It holds $X_n = x$ for infinitely many n and $X_n = y$ for infinitely many n, thus $Z = f(x) = f(y)$.

Proposition 4.20 provides, in particular, a criterion of transience: if there exists a nonconstant excessive function, then the chain is necessarily transient (but this is just a sufficient condition).

•4.25 We now investigate the invariant measures (•4.15). One of the main interests of these measures is related to the ergodic theorems, which we shall see later. Let us start with two remarks.

(i) Assume that μ is an invariant probability for X. We know that, if μ is the law of X_0, μP^n is the law of X_n. Thus, under $\mathbf{P}_\mu$, the laws of the X_n's are constant and equal to μ. More generally it holds, for every $r, n \in \mathbb{N}$, for f positive on E^{n+1} and setting $\Phi = f(X_0, X_1, \dots, X_n)$, $\phi(x) = \mathbf{E}_x(\Phi)$,

$$\begin{aligned}\mathbf{E}_\mu[f(X_r, X_{r+1}, \dots, X_{r+n})] &= \mathbf{E}_\mu(\Phi \circ \theta_r) = \mathbf{E}_\mu[\mathbf{E}_{X_r}(\Phi)] = \mathbf{E}_\mu(\phi(X_r)) = \\ &= \mathbf{E}_\mu(\phi(X_0)) = \mathbf{E}_\mu[\mathbf{E}_{X_0}(\Phi)] = \mathbf{E}_\mu(\Phi) = \mathbf{E}_\mu[f(X_0, X_1, \dots, X_n)].\end{aligned}$$

i.e., under $\mathbf{P}_\mu$, the law of the r.v.'s

$$(X_r, X_{r+1}, \dots, X_{r+n}) \quad \text{and} \quad (X_0, X_1, \dots, X_n)$$

does not depend on r. Thus, under $\mathbf{P}_\mu$, the process $(X_n)_{n\geq 0}$ is stationary.

(ii) Assume that μ is an invariant measure such that $0 < \mu(x) < +\infty$ for every $x \in E$. Let us define (the *dual chain*)

$$\hat{P}(x, y) = \frac{\mu(y)}{\mu(x)} P(y, x).$$

$\hat{P}$ is a transition matrix since

$$\sum_{y\in E} \hat{P}(x, y) = \mu^{-1}(x) \sum_{y\in E} \mu(y) P(y, x) = \mu^{-1}(x)\mu(x) = 1.$$

It is immediate to check that $\hat{P}^2(x, y) = \mu^{-1}(x)\mu(y) P^2(y, x)$ and, more generally, that $\hat{P}^n(x, y) = \mu^{-1}(x)\mu(y) P^n(y, x)$. Thus

$$\hat{U}(x, y) = \mu^{-1}(x)\mu(y) U(y, x).$$

From this $\hat{U}(x, y) = +\infty$ and the dual chain is recurrent irreducible too.

•4.26 Actually a recurrent irreducible chain always has an invariant measure (possibly not finite).

Theorem 4.21 *Let X be a recurrent irreducible chain. Then an invariant measure μ for X exists and is unique up to a multiplicative constant. Furthermore, for every $y \in E$, $0 < \mu(y) < +\infty$. Finally, for $x \in E$, the invariant measure μ_x such that $\mu_x(x) = 1$ is given by*

$$\mu_x(y) = \mathbf{E}_x\Big[\sum_{k=0}^{\sigma_x - 1} 1_{\{y\}}(X_k)\Big] \tag{4.37}$$

where σ_x denotes the return time in x (see (4.30)).

The proof of this theorem is instructive. Let us fix $x \in E$ and define

$$\mu'_x(y) = \mathbf{E}_x\Big[\sum_{k=1}^{\sigma_x} 1_{\{y\}}(X_k)\Big]. \tag{4.38}$$

It holds

$$\mu_x(x) = \mu'_x(x) = 1, \quad \mu_x(E) = \langle \mu_x, 1\rangle = \mathbf{E}_x(\sigma_x) \tag{4.39}$$

and, for $f \in \mathscr{E}^+$, since $X_{\sigma_x} = x = X_0$ $\mathbf{P}_x$-a.s.,

$$\langle \mu_x, f\rangle = \mathbf{E}_x\Big[\sum_{k=0}^{\sigma_x-1} f(X_k)\Big] = \mathbf{E}_x\Big[\sum_{k=1}^{\sigma_x} f(X_k)\Big] = \langle \mu'_x, f\rangle, \tag{4.40}$$

which gives $\mu_x = \mu'_x$. Then:

(i) The measure μ_x is invariant. Actually, for $f \in \mathscr{E}^+$, thanks to (4.40),

$$\langle \mu_x P, f\rangle = \langle \mu_x, Pf\rangle = \mathbf{E}_x\Big[\sum_{k=0}^{\sigma_x-1} Pf(X_k)\Big] =$$

$$= \mathbf{E}_x\Big[\sum_{k=0}^{\sigma_x-1} \mathbf{E}_{X_k}(f(X_1))\Big] = \sum_{k\geq 0} \mathbf{E}_x[1_{\{k<\sigma_x\}}\mathbf{E}_{X_k}(f(X_1))].$$

Since $\{k < \sigma_x\} \in \mathscr{F}_k$, thanks to (4.18),

$$\langle \mu_x P, f\rangle = \sum_{k\geq 0} \mathbf{E}_x\Big[1_{\{k<\sigma_x\}} f(X_{k+1})\Big] = \mathbf{E}_x\Big[\sum_{k=1}^{\sigma_x} f(X_k)\Big] = \langle \mu'_x, f\rangle = \langle \mu_x, f\rangle.$$

(ii) For every $y \in E$, $0 < \mu_x(y) < +\infty$. This comes from:

Lemma 4.22 *Let P be an irreducible transition matrix and λ a measure such that $\lambda \geq \lambda P$. If, for some $x \in E$, $\lambda(x) = 0$ (resp. $\lambda(x) < +\infty$), then $\lambda(y) = 0$ (resp. $\lambda(y) < +\infty$) for every y.*

Actually let $x, y \in E$, $x \neq y$. There exists $n \geq 1$ such that $P^n(y, x) > 0$. Since $\lambda \geq \lambda P^n$,

$$\lambda(x) \geq \sum_{z\in E} \lambda(z) P^n(z, x) \geq \lambda(y) P^n(y, x).$$

Thus $\lambda(x) = 0$ implies $\lambda(y) = 0$ and $\lambda(x) < +\infty$ implies $\lambda(y) < +\infty$. ∎

(iii) We still have to investigate the uniqueness. Let λ be an excessive measure (possibly invariant) such that $\lambda(y) < +\infty$ for every $y \in E$. Let us set $f(y) = \lambda(y)/\mu_x(y)$. Then f is an excessive function for the chain $\hat{X}$ introduced in •4.25 (ii) with $\mu = \mu_x$. Actually

$$\hat{P} f(y) = \sum_{z\in E} \hat{P}(y, z)\mu_x^{-1}(z)\lambda(z) = \sum_{z\in E} \mu_x^{-1}(y)\mu_x(z)P(z, y)\mu_x^{-1}(z)\lambda(z) =$$

$$= \mu_x^{-1}(y)\lambda P(z) \leq \mu_x^{-1}(y)\lambda(y) = f(y).$$

Since $\hat{X}$ is irreducible recurrent, this implies that f is constant (Proposition 4.20).

•4.27 The following dichotomy holds: for every invariant measure μ either $\mu(E) =$

$+\infty$ or $\mu(E) < +\infty$. In the first case, thanks to (4.39), $\mathbf{E}_x(\sigma_x) = +\infty$; in the second one $\mathbf{E}_x(\sigma_x) < +\infty$ (and thus not only the return time is finite a.s., but it is even integrable). Finally, in the second case, a unique invariant probability exists, namely $\pi(y) = \mu_x(y)/\mu_x(E) = \mu_x(y)/\mathbf{E}_x(\sigma_x)$ and, in particular, $\pi(x) = \mu_x(x)/\mathbf{E}_x(\sigma_x) = 1/\mathbf{E}_x(\sigma_x)$. In summary:

Theorem 4.23 *Let X be an irreducible recurrent chain with invariant measure μ. Two situations may arise.*

(i) $\mu(E) = +\infty$. Then, for every $x \in E$, $\mathbf{E}_x(\sigma_x) = +\infty$. The chain is called null recurrent.

(ii) $\mu(E) < +\infty$. Then, for every $x \in E$, $\mathbf{E}_x(\sigma_x) < +\infty$. The chain is called positive recurrent. In this case there exists a unique invariant probability, given by

$$\pi(x) = \frac{1}{\mathbf{E}_x(\sigma_x)}. \tag{4.41}$$

•**4.28** Consider an irreducible recurrent Markov chain X. The name *ergodic theorems* refers to theorems concerning the behaviour, as $n \to \infty$, of quantities of the type

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \quad \text{or} \quad \frac{\sum_{k=0}^{n-1} f(X_k)}{\sum_{k=0}^{n-1} g(X_k)}$$

for functions f, g belonging to classes to be specified. For a fixed $x \in E$, let $\mu = \mu_x$ be the invariant measure taking the value 1 at x and $\sigma^p = \sigma_x^p$ the p-th return time in x. We denote by λ an arbitrary initial law. For $f \in L^1(\mu)$, i.e., satisfying $\sum_{x \in E} |f(x)|\mu(x) < +\infty$, let us set

$$Z_0 = \sum_{k=0}^{\sigma^1 - 1} f(X_k),$$

$$Z_p = Z_0 \circ \theta_{\sigma^p} = \sum_{k=0}^{\sigma^1 - 1} f(X_k) \circ \theta_{\sigma^p} = \sum_{k=\sigma^p}^{\sigma^{p+1}-1} f(X_k).$$

By Theorem 4.21, $\mathbf{E}_x(Z_0) = \langle \mu, f \rangle$ and, by the Markov property,

$$\mathbf{E}_\lambda(Z_p) = \mathbf{E}_\lambda[\mathbf{E}_{X_{\sigma^p}}(Z_0)] = \mathbf{E}_x(Z_0) = \langle \mu, f \rangle.$$

With respect to the law $\mathbf{P}_\lambda$, the r.v.'s $Z_1, \ldots, Z_p, \ldots$ are independent, identically distributed and have an expectation that is equal to $\langle \mu, f \rangle$. The care of checking these properties, nevertheless intuitive, is left to the reader. Actually the law of Z_p depends neither on the position of the chain at time σ^p, i.e., on x, nor on its behaviour before σ^p (one can find some hints in Exercise 4.40). By the strong law of large numbers,

$$\frac{1}{n}\sum_{k=0}^{\sigma^n} f(X_k) = \frac{Z_0}{n} + \frac{1}{n}(Z_1 + \ldots + Z_{n-1}) + \frac{f(x)}{n} \underset{n \to \infty}{\to} \langle \mu, f \rangle \ \mathbf{P}_\lambda\text{-a.s.}$$

Let us associate to $m \in \mathbb{N}$ the unique integer $\nu(m)$ such that $\sigma^{\nu(m)} \le m < \sigma^{\nu(m)+1}$;

then $\nu(m) \to_{m\to\infty} +\infty$ and

$$\nu(m) \le \sum_{k=0}^{m} 1_{\{x\}}(X_k) = \nu(m) + 1_{\{x\}}(X_0) \le \nu(m) + 1.$$

If $f \ge 0$, then

$$\frac{\nu(m)}{\nu(m)+1}\,\frac{\sum_{k=0}^{\sigma^{\nu(m)}} f(X_k)}{\nu(m)} \le \frac{\sum_{k=0}^{m} f(X_k)}{\sum_{k=0}^{m} 1_{\{x\}}(X_k)} \le \frac{\nu(m)+1}{\nu(m)}\,\frac{\sum_{k=0}^{\sigma^{\nu(m)+1}} f(X_k)}{\nu(m)+1}$$

and, since both the leftmost and rightmost terms tend to $\langle \mu, f\rangle$ $\mathbf{P}_\lambda$-a.s. as $m \to \infty$,

$$\frac{\sum_{k=0}^{m} f(X_k)}{\sum_{k=0}^{m} 1_{\{x\}}(X_k)} \underset{m\to\infty}{\to} \langle \mu, f\rangle \qquad \mathbf{P}_\lambda\text{-a.s.} \tag{4.42}$$

Writing $f = f^+ - f^-$, (4.42) follows for $f \in L^1(\mu)$. Finally, if $f, g \in L^1(\mu)$ and $\langle \mu, g\rangle \ne 0$, one gets (4.42) for f and g and it is enough to take the quotient. We proved:

Theorem 4.24 *Let X be a recurrent irreducible chain with invariant measure μ. Let $f, g \in L^1(\mu)$ with $\langle \mu, g\rangle \ne 0$. then*

$$\frac{\sum_{k=0}^{n} f(X_k)}{\sum_{k=0}^{n} g(X_k)} \underset{n\to\infty}{\to} \frac{\langle \mu, f\rangle}{\langle \mu, g\rangle} \quad a.s.$$

•**4.29** Let us point out some consequences of Theorem 4.24. First if we choose $f = 1_{\{y\}}$ and $g = 1_{\{x\}}$, then

$$\frac{\text{number of visits of } X \text{ in } y \text{ before time } n}{\text{number of visits of } X \text{ in } x \text{ before time } n} \underset{n\to\infty}{\to} \frac{\mu(y)}{\mu(x)} \text{ a.s.}$$

Moreover, if $\mu(E) < +\infty$ (i.e., the chain is positive recurrent), it is possible to choose $g = 1$ and $\mu = \pi$, the invariant probability, in Theorem 4.24, so that

Theorem 4.25 *Let X be a positive recurrent irreducible chain, with invariant probability π. Then, for every $f \in L^1(\pi)$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \underset{n\to\infty}{\to} \langle \pi, f\rangle \quad a.s.$$

Finally

Proposition 4.26 *(i) If X is a positive recurrent irreducible chain, with invariant probability π. then. for every $x \in E$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} 1_{\{x\}}(X_k) \underset{n\to\infty}{\to} \pi(x) \quad a.s.$$

(ii) If X is a null recurrent irreducible chain, then, for every $x \in E$,

$$\frac{1}{n}\sum_{k=0}^{n-1} 1_{\{x\}}(X_k) \underset{n\to\infty}{\to} 0 \quad a.s.$$

Only the second point needs to be proved. Let us choose $\mu = \mu_x$, so that $\mu(x) = 1$. For a finite subset $F \subset E$, by Theorem 4.24,

$$\frac{\sum_{k=0}^{n} 1_{\{x\}}(X_k)}{n+1} \leq \frac{\sum_{k=0}^{n} 1_{\{x\}}(X_k)}{\sum_{k=0}^{n} 1_F(X_k)} \underset{n\to\infty}{\to} \frac{\mu(x)}{\mu(F)} = \frac{1}{\mu(F)} \quad \text{a.s.};$$

hence,

$$\varlimsup_{n\to\infty} \frac{\sum_{k=0}^{n} 1_{\{x\}}(X_k)}{n+1} \leq \frac{1}{\mu(F)} \quad \text{a.s.}$$

for every finite $F \subset E$. If $\mu(E) = +\infty$ the result is thus proved, $\mu(F)$ being arbitrarily large.

Remark. Since the fractions appearing in Proposition 4.26 are bounded by 1, one can take the expectation $\mathbf{E}_y$ and apply Lebesgue's theorem. Since $\mathbf{E}_y(1_{\{x\}}(X_k)) = P^k(y, x)$, we get

(i) if X is positive recurrent with invariant probability π,

$$\frac{1}{n}\sum_{k=0}^{n-1} P^k(y, x) \underset{n\to\infty}{\to} \pi(x);$$

(ii) if X is null recurrent

$$\frac{1}{n}\sum_{k=0}^{n-1} P^k(y, x) \underset{n\to\infty}{\to} 0.$$

Periodicity

The remark above shows that, for a positive recurrent chain with invariant probability π, $P^n(y, x) \to_{n\to\infty} \pi(x)$ in the sense of Cesaro. It is natural to ask whether this property holds in the usual sense. In general this is not the case as the following example shows.

•**4.30** Let us consider the (deterministic!) chain with values in $E = \{1, 2, 3\}$ associated to the transition matrix P where $P(1, 2) = P(2, 3) = P(3, 1) = 1$. It holds

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad P^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad P^3 = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and, more generally, $P^{3n} = I$, $P^{3n+1} = P$, $P^{3n+2} = P^2$. One sees immediately that the sequence $P^n(1, 1)$ is of the form $1, 0, 0, 1, 0, 0, 1, \ldots$, the sequence $P^n(1, 2)$ the form $0, 1, 0, 0, 1, 0, 0, 1, \ldots$; thus the sequence $(P^n(x, y))_{n\geq 0}$ does not converge. However $(P^{3n}(1, 1))_{n\geq 0}$ converges to 1, $(P^{3n+1}(1, 1))_{n\geq 0}$ and $(P^{3n+2}(1, 1))_{n\geq 0}$ both converge to 0. There is a periodicity phenomenon.

•**4.31** Let P be a transition matrix. For $a \in E$, let us set

$$I(a) = \{n > 0;\ P^n(a, a) > 0\}.$$

It is easy to check that, if P is irreducible, then $I(a)$ in nonempty.

Definition 4.27 *If $I(a) \neq \emptyset$, the greatest common divisor (g.c.d.) of the set $I(a)$ is called the period of a, noted d_a.*

Proposition 4.28 *If X is an irreducible chain, all states have the same period. The chain is called aperiodic if its period is equal to 1.*

•**4.32** The following results hold.

Theorem 4.29 *Let X be an irreducible, aperiodic, positive recurrent chain with invariant probability π. Then $P^n(x, y) \to_{n\to\infty} \pi(y)$, for every $x, y \in E$.*

Theorem 4.30 *Let X be an irreducible positive recurrent chain, with invariant probability π, with period d. Then, for every $x, y \in E$, there exists r, $0 \le r < d$, with $r = 0$ if $x = y$, such that, for every s, $0 \le s < d$,*

$$\lim_{n\to\infty} P^{nd+r}(x, y) = d\pi(y), \qquad P^{nd+s}(x, y) = 0 \quad \text{for } s \neq r.$$

Conversely for a null recurrent chain, the following general result holds, regardless of aperiodicity:

Theorem 4.31 *Let X be an irreducible null recurrent chain Then $P^n(x, y) \to_{n\to\infty} 0$, for every $x, y \in E$.*

•**4.33** The aperiodicity of a chain is a question that one is often led to investigate. The two following criteria are often useful.

First if a state $a \in E$ is such that $P(a, a) > 0$, then it is necessarily aperiodic. Actually the set $I(a)$ of •4.31 contains the number 1 and thus its g.c.d. is 1 (one can even see that $I(a) = \mathbb{N}$). Thus, if the chain is irreducible and there exists at least one state $a \in E$ such that $P(a, a) > 0$, then it is aperiodic, because of Proposition 4.28.

Moreover if there exists an index m such that $P^m(x, y) > 0$ for every $x, y \in E$, then the chain is aperiodic. Actually it is immediate to check that it also holds $P^{m+1}(x, y) > 0$ for every $x, y \in E$; thus for every $x \in E$ the set $I(x)$ defined in •4.31 contains two consecutive numbers and its g.c.d. must be 1. This criterion is only sufficient and there are examples of aperiodic chains for which it is not satisfied But if E is finite this criterion is also necessary.

•**4.34** The importance of the invariant probability of a chain (whenever it exists) appears fully in the statements of Theorems 4.24, 4.29 and 4.30. One is thus often led to compute it, which means the resolution of the linear system

$$\mu = \mu P \tag{4.43}$$

with the condition $\sum_{x\in E} \mu(x) = 1$. There are two situations in which the computation of μ becomes simpler.

i) If E is finite, a transition matrix P on E is called *bistochastic* if $\sum_{x\in E} P(x, y) = 1$ for every $y \in E$, i.e., if the sum of its columns is also equal to 1. In this case, it is immediate to check that the vector μ having all its components equal to $|E|^{-1}$ is a probability on E which solves (4.43). For a bistochastic chain the uniform law on E is thus invariant.

ii) A probability μ on E is called *reversible* if, for every $x, y \in E$,

$$\mu(x) P(x, y) = \mu(y) P(y, x). \tag{4.44}$$

It is immediate to check that a reversible probability is stationary:

$$\sum_{x\in E} \mu(x) P(x, y) = \sum_{x\in E} \mu(y) P(y, x) = \mu(y) \sum_{x\in E} P(y, x) = \mu(y)$$

since the sum of the rows of P is constant and equal to 1. (4.44) is the *detailed balance equation*. It is in general easier to solve the latter than the equation of stationarity (4.43); this suggests that, even if the converse is not true (a probability can be invariant without being reversible), one might first try to solve equation (4.44).

The existence of a reversible probability gives useful information on P. For instance, if E is assumed to be finite, equation (4.44) is equivalent to assert that P is self-adjoint with respect to the scalar product

$$\langle f, g\rangle_\mu = \sum_{x\in E} f(x)g(x)\mu(x).$$

Thus P must have a real spectrum and is diagonalizable.

Exercises

Exercise 4.1 a) Let E be a countable set, $(S, \mathscr{S})$ a measurable space, $(Y_n)_{n\geq 1}$ a sequence of independent r.v.'s identically distributed with values in $(S, \mathscr{S})$. We define a sequence $(X_n)_{n\geq 0}$ by $X_0 = x \in E$ and $X_{n+1} = \Phi(X_n, Y_{n+1})$, where $\Phi : E \times S \to E$ is a fixed measurable application. Show that $(X_n)_{n\geq 0}$ is a Markov chain and determine its transition matrix.

b) Let μ be a probability and P a transition matrix on $E = \{0, 1, \dots, m\}$ or $E = \mathbb{N}$. Let us set $t_0 = 0$, $t_1 = \mu(0)$ and

$$t_n = \sum_{k=0}^{n-1} \mu(k), \qquad n \in E$$

and, for every $i \in E$, $s_0^{(i)} = 0$, $s_1^{(i)} = P(i, 0)$ and

$$s_n^{(i)} = \sum_{k=0}^{n-1} P(i, k).$$

Let us define also

$$\begin{aligned} \psi(x) &= k, \quad \text{if } x \in [t_k, t_{k+1}[, \qquad && \psi(1) = 1 \\ g(i, x) &= k, \quad \text{if } x \in [s_k^{(i)}, s_{k+1}^{(i)}[, \qquad && g(1) = 1. \end{aligned}$$

Let $(U_n)_{n\geq 0}$ be a sequence of independent r.v.'s having uniform law on $[0, 1]$. We define

$$X_0 = \psi(U_0), \quad X_{n+1} = g(X_n, U_{n+1}).$$

Prove that $(X_n)_{n\geq 0}$ is a Markov chain with transition matrix P and starting distribution μ.

♦ Point (b) of this exercise gives an alternate proof of the existence of a Markov chain with a given transition matrix and initial distribution. This proof moreover suggests a way of simulating a chain, which is often of practical interest.

Exercise 4.2 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathrm{P}_x)_{x\in E})$ be a canonical Markov chain on a countable set E with transition matrix Q. Let $F \subset E$ and $\tau = \inf\{n \geq 0; X_n \in F\}$ be the passage time in F. We set $Y_n = X_{n\wedge\tau}$.

Show that $Y = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (Y_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ is a Markov chain and determine its transition matrix.

Exercise 4.3 Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on the countable set E with transition matrix P. Let $F \subset E$ and τ_F be the passage time in F.

a) Let $f: E \to \mathbb{R}^+$ be a function such that $f(x) \geq Pf(x)$ (resp. $=$) if $x \in F^c$. Show that, for every $x \in E$, $(f(X_{n\wedge\tau_F}))_{n\geq 0}$ is a positive $\mathbf{P}_x$-supermartingale (resp. martingale) with respect to the filtration $(\mathscr{F}_n)_{n\geq 0}$.

b) Complete the proof of Theorem 4.12, showing that $P_F g$ is the smallest positive solution of (4.28).

Exercise 4.4 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in\mathbb{Z}})$ be a canonical Markov chain on $\mathbb{Z}$ with transition matrix Q. Assume that $\sum_{y\in\mathbb{Z}} y^2 Q(x, y) < +\infty$ for every $x \in \mathbb{Z}$. Let us define $b(x) = \mathbf{E}_x(X_1)$, $a(x) = \mathrm{Var}_x(X_1) = \mathbf{E}_x((X_1 - b(x))^2)$.

1) Express $b(x)$ and $a(x)$ in terms of the matrix Q.

2) Show that

$$\mathbf{E}_x(X_{n+1}) = \mathbf{E}_x(b(X_n)) \quad \text{and} \quad \mathrm{Var}_x(X_{n+1}) = \mathrm{Var}_x(b(X_n)) + \mathbf{E}_x(a(X_n)).$$

Exercise 4.5 Let E be a countable set and $\mathscr{H}$ the set of all bounded applications $E \to \mathbb{R}$. Let $(X_n)_{n\geq 0}$ be a Markov chain on E with transition matrix P and consider the natural filtration $(\mathscr{F}_n)_{n\geq 0}$ of $(X_n)_{n\geq 0}$ $(\mathscr{F}_n = \sigma(X_0, X_1, \dots, X_n))$.

Show that there exists an operator $A : \mathscr{H} \to \mathscr{H}$ such that, for every $f \in \mathscr{H}$,

$$f(X_n) - \sum_{k=0}^{n-1} Af(X_k)$$

is a $(\mathscr{F}_n)_{n\geq 0}$-martingale.

Exercise 4.6 Let $(\Omega, \mathscr{F}, (\mathscr{F})_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on the countable set E with potential matrix U. Assume that, for every $x, y \in E$, $U(x, y) < +\infty$. Let f be a positive function on E. The aim of this exercise is to find a formula expressing the moment of order 2 of N_y (•4.13) as a function of the potential matrix.

1) Show that $\mathbf{E}_x[\sum_{n\geq 0} f(X_n) \sum_{p\geq n} f(X_p)] = U(f\cdot Uf)(x)$.

2) Express $\mathbf{E}_x[(\sum_{n\geq 0} f(X_n))^2]$ in terms of the operator U.

3) Express $\mathbf{E}_x(N_y^2)$ in terms of $U(x, y)$ and $U(y, y)$.

Exercise 4.7 a) On $E = \{0, 1, \dots, n\}$ let us consider the Markov chain with transition matrix P given, for $0 \leq x \leq n-1$, by

$$P(x, y) = \begin{cases} p & \text{if } y = x+1 \\ 1-p & \text{if } y = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$ and with the state n absorbing. Classify the states of this chain. Let τ be the passage time in n. What is the value of $\mathbf{E}_x(\tau)$, $0 \leq x \leq n$? (Hint: recall Corollary 4.15).

b) A coin gives head with probability p (and tail with probability $q = 1 - p$). It is thrown successively several times. Let n be a fixed integer > 0. Show that, with probability 1, n consecutive heads are obtained with a finite numbers of tosses. How many tosses are necessary on average?

Ⓜ **Exercise 4.8** In order to model the evolution of the genetic configurations in a population, one is led to consider the following Markov chain. Let P be the transition matrix on $E = \{0, 1, \dots, N\}$ defined by

$$P(i, j) = \binom{N}{j}\left(\frac{i}{N}\right)^j\left(1 - \frac{i}{N}\right)^{N-j}.$$

In other words, for $i \in E$ fixed, $P(i, \cdot)$ is a binomial law $B(N, \frac{i}{N})$ (in the interpretation introduced above, N represents the dimension of the population and i the number of individuals in the population bearing a certain genetic character).

a) What is the value of $P(0, j)$? And $P(N, j)$? Is this chain irreducible? Which states are recurrent?

b) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge 0}, (X_n)_{n\ge 0}, (\mathrm{P}_i)_{i\in E})$ be the canonical Markov chain with transition matrix P. Show that, for every $i \in E$, X is a P_i-martingale and that the limit

$$\lim_{n\to\infty} X_n = X_\infty$$

exists P_i-a.s. Determine the law of X_∞.

Ⓜ **Exercise 4.9** Let Q be an irreducible transition matrix on a finite set E. Two players A and B choose each a different element of E, x_A and x_B. Let us consider the Markov chain on E with transition matrix Q and starting law that is uniform on E. The game is won by the player whose state is attained first. We want to compute the probabilities p_A and p_B of success for each of the players.

In order to model this game we note $E = \{1, 2, \dots, N\}$, $N - 1 = x_A$, $N = x_B$ and consider $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge 0}, (X_n)_{n\ge 0}, (\mathrm{P}_k)_{k\in E})$, the canonical Markov chain with transition matrix Q. Let $F = \{N - 1, N\} = \{x_A, x_B\}$ and $\tau = \tau_F = \inf\{n \ge 0; X_n \in F\}$ the passage time in F. We write

$$Q = \begin{pmatrix} C & D \\ M & N \end{pmatrix}$$

where C is the $(N-2) \times (N-2)$ matrix formed by the numbers $Q(i, j)$, $1 \le i, j \le N - 2$.

1) Let $y \in \mathbb{R}^{N-2}$ be a vector such that $y = Cy$ and define $h : E \to \mathbb{R}$ by $h(i) = y_i$, $i \in E \setminus F$, $h(N-1) = h(N) = 0$. Show that $Qh(i) = h(i)$ for every $i \in E \setminus F$ and then that $(h(X_{n\wedge\tau}))_{n\ge 0}$ is a P_k-martingale for every $k \in E$. Deduce that $y = 0$ and that $I - C$ is invertible.

2) Let $\bar p_j \in \mathbb{R}^{N-2}$, $j = N - 1, N$, the vectors with components $\mathrm{P}_i(X_\tau = j)$, $i = 1, \dots, N-2$ and $e_1 \in \mathbb{R}^2$ (resp. e_2) the vector with components $1, 0$ (resp. $0, 1$). Show that

$$\bar p_{N-1} = (I - C)^{-1} De_1 \quad \text{and} \quad \bar p_N = (I - C)^{-1} De_2.$$

3) Compute p_{N-1} and p_N in terms of C and D.

Exercise 4.10 a) On $E = \{1, \dots, 10\}$ consider the transition matrix P below, where $\cdot$ denotes 0 and $*$ a strictly positive number. Which states are recurrent? Transient?

Determine the closed classes.

$$
\begin{array}{c}
\\
1\\ 2\\ 3\\ 4\\ 5\\ 6\\ 7\\ 8\\ 9\\ 10
\end{array}
\begin{array}{c}
\begin{array}{cccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10
\end{array}\\
\begin{pmatrix}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * & \cdot & * & \cdot \\
\cdot & * & \cdot & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & * & \cdot & \cdot & * & \cdot & \cdot \\
\cdot & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & * & * & * & * & \cdot & \cdot & \cdot & \cdot & * \\
* & \cdot & \cdot & \cdot & * & \cdot & \cdot & \cdot & \cdot & \cdot \\
* & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * & \cdot \\
\cdot & \cdot & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & * & \cdot & * & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & *
\end{pmatrix}
\end{array}
$$

b) More generally let P be a transition matrix on the countable set E. Show that $x \rightsquigarrow y$ if and only if there exists $n \in \mathbb{N}$ and states $h_1, \dots, h_{n-1}$ such that

$$P(x, h_1) > 0,\ P(h_1, h_2) > 0, \dots, P(h_{n-1}, y) > 0 \tag{4.45}$$

($x, h_1, \dots, h_{n-1}, y$ is called a *path* of positive probability of length n). Show that, furthermore, one can assume that the states $x, h_1, \dots, h_{n-1}, y$ are distinct.

c) Let P, Q be two transition matrices on E. Assume that

$$P(x, y) > 0 \quad \Rightarrow \quad Q(x, y) > 0 \qquad \text{for every } x, y \in E. \tag{4.46}$$

Show that if $x \rightsquigarrow y$ for P, then $x \rightsquigarrow y$ for Q. Show that the same result holds if (4.46) is weakened into

$$P(x, y) > 0 \quad \Rightarrow \quad Q(x, y) > 0 \qquad \text{for every } x, y \in E, x \neq y. \tag{4.47}$$

d) Assume that (4.47) holds. Show that, if P is irreducible, the same is true for Q. Does the aperiodicity of P imply the aperiodicity of Q? What if (4.46) was true?

♦ This exercise shows that the communications between states of a Markov chain depend uniquely on the presence in certain positions of the transition matrix of elements > 0 (and not on their values).

Exercise 4.11 (A recurrence criterion) Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n \geq 0}, (X_n)_{n \geq 0}, (\mathbf{P}_x)_{x \in E})$ be a canonical Markov chain on the countable set E with transition matrix P. Assume that there exists a state $x_0 \in E$ such that

a) x_0 leads to every other state $x \in E$,

b) $\mathbf{P}_x(\tau_{x_0} < +\infty) = 1$ for every $x \in E$.

Show that the chain is recurrent irreducible.

Exercise 4.12 Let X be a canonical Markov chain on the countable set E, with transition matrix P. Let $A \subset E$ and τ_A be the passage time in A. Assume that there exist $n \geq 1$ and $\alpha > 0$ such that, for every $x \in A^c$, $P^n(x, A) \geq \alpha$.

a) Show that, for every $k \in \mathbb{N}$ and $x \in E$, $\mathbf{P}_x(\tau_A > kn) \leq (1 - \alpha)^k$. Deduce that $\mathbf{E}_x(\tau_A) \leq \frac{n}{\alpha}$. In particular $\tau_A < +\infty$ $\mathbf{P}_x$-a.s.

b) Show that, for every $u > 0$, $\mathbf{P}_x(\tau_A > u) \leq (1 - \alpha)^{-1}(1 - \alpha)^{u/n}$. Deduce that, for every $\rho < -\frac{1}{n}\log(1 - \alpha)$, $\mathbf{E}_x(e^{\rho\tau_A}) < +\infty$. (Hint: show first that $\mathbf{E}_x(e^{\rho\tau_A}) = 1 + \rho\int_0^{+\infty} e^{\rho u}\, \mathbf{P}_x(\tau_A > u)\, du$.)

Exercise 4.13 Classify, according to the values of the numbers $p_k, q_k, k \geq 0$, the states of the Markov chain on $\mathbb{N}$ with transition matrix

$$\begin{array}{lll} Q(0,0) = \alpha, & Q(0,1) = 1-\alpha, & 0 < \alpha < 1 \\ Q(1,2) = \beta, & Q(1,3) = 1-\beta, & 0 < \beta < 1 \\ Q(k,1) = p_k, & Q(k,k+2) = q_k = 1 - p_k, & 0 < p_k < 1, k \geq 2. \end{array}$$

Ⓜ **Exercise 4.14** Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_k)_{k\in\mathbb{N}})$ be the canonical chain on $\mathbb{N}$ with transition matrix P defined by

$$P(0,0) = 1, \qquad P(k,m) = e^{-k}\frac{k^m}{m!}, k \geq 1, m \geq 0$$

($P(k, \cdot)$ is thus a Poisson law with parameter k).

1) Which states are recurrent?

2) Show that the identity function on $\mathbb{N}$, $f(k) = k$, is P-harmonic.

3) Show that, for every $k \in \mathbb{N}$, the sequence $(X_n)_{n\geq 0}$ converges to 0 $\mathbf{P}_k$-a.s.

Exercise 4.15 (The Ehrenfest model of gas diffusion) m balls, numbered $1, \dots, m$, are distributed between two boxes. The state X_n of the system is specified by the number of balls that are in the first box, so that the state space is $E = \{0, 1, 2, \dots, m\}$. At each iteration a number is chosen randomly (and uniformly) between 1 and m and the corresponding ball is switched from the box in which it is contained to the other.

Determine the transition matrix P of the Markov chain $(X_n)_{n\geq 0}$. Is it irreducible? Aperiodic?

♦ The stationary probability of this chain is computed in Exercise 4.36.

Exercise 4.16 (Simple random walk on $\mathbb{Z}$) A moving target performs a random dynamics on $\mathbb{Z}$ following a Markov chain with transition matrix

$$P(i,j) = \begin{cases} p & \text{if } j = i+1 \\ 1-p & \text{if } j = i-1 \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < p < 1$.

a) Is this chain irreducible? What is its period?

b) Let $(Z_n)_{n\geq 1}$ be a sequence of independent r.v.'s such that $\mathbf{P}(Z_n = 1) = p$, $\mathbf{P}(Z_n = -1) = 1 - p$ and let us set $X_0 = 0$, $X_n = Z_1 + \dots + Z_n$. Show that $(X_n)_{n\geq 0}$ is a Markov chain with transition matrix P and starting state 0.

c) What about the limit $\lim_{n\to\infty} \frac{1}{n} X_n$? If $p \neq \frac{1}{2}$, is the chain transient or recurrent?

d) Let us set $Y_i = \frac{1}{2}(Z_i + 1)$. Determine the law of Y_i and of $T_n = \frac{1}{2}(X_n + n)$. Compute $P^{(n)}(0,0) = \mathbf{P}(X_n = 0)$, in the case where n is even or odd.

e) Show that, if $p = \frac{1}{2}$, the chain is recurrent. Is it positive recurrent? (Recall Stirling's formula $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$).

Ⓜ **Exercise 4.17** Let $(Y_n)_{n\geq 1}$ be a sequence of independent identically distributed r.v.'s such that $\mathbf{P}(Y_1 = 1) = \mathbf{P}(Y_1 = -1) = \frac{1}{2}$. Let us define $S_0 = 0$, $\mathscr{F}_0 = \{\Omega, \emptyset\}$ and, for $n \geq 1$,

$$S_n = Y_1 + \dots + Y_n, \quad \mathscr{F}_n = \sigma(Y_1, \dots, Y_n).$$

We know that $(S_n)_{n\geq 0}$ is a Markov chain (see Exercises 1.8, 4.1 or 4.16). Let us note, for $x \in \mathbb{Z}$, $\tau_x = \inf\{n \geq 0; S_n = x\}$. We know that τ_x is a stopping time satisfying

$\mathbf{P}(\tau_x < +\infty) = 1$ (see Exercise 4.16). Finally, let us set, for $a, b \in \mathbb{N}^*$, $\tau = \tau_{-a} \wedge \tau_b$; τ is thus the exit time of $(S_n)_{n\geq 0}$ from $\{-a+1, \ldots, b-1\}$.

1) Show that $(S_n)_{n\geq 0}$ and $(S_n^2 - n)_{n\geq 0}$ are martingales.

2) Show that $\mathbf{E}(S_\tau) = 0$. What are the values of $\mathbf{P}(\tau_{-a} < \tau_b)$ and of $\mathbf{P}(\tau_{-a} > \tau_b)$?

3) Show that $\mathbf{E}(S_\tau^2) = \mathbf{E}(\tau)$ and compute $\mathbf{E}(\tau)$.

4a) Show that $Z_n = S_n^3 - 3nS_n$ is a martingale. What is the value of $\mathbf{E}(\tau S_\tau)$?

4b) Assume $b \neq a$. Are the r.v.'s τ and S_τ independent?

4c) Assume $b = a$; show that

$$\mathbf{P}(S_\tau = a, \tau = k) = \frac{1}{2}\mathbf{P}(\tau = k) \tag{4.48}$$

and that τ and S_τ are independent.

Exercise 4.18 Consider the canonical Markov chain on $E = \mathbb{N}$, with transition matrix described by Figure 4.1, with $p > 0$, $q > 0$, $r \geq 0$ and $p + q + r = 1$ (this is a particular case of a *birth and death chain*). We denote by $\tau = \inf\{n \geq 0;\ X_n = 0\}$ the passage time at 0.

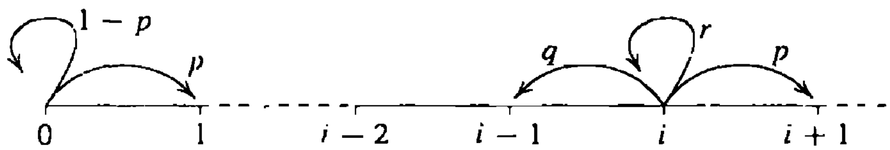

Figure 4.1

a) Show that the sequence $(f_i)_{i\geq 1}$, where $f_i = \mathbf{P}_i(\tau < +\infty)$, $i = 1, 2, \ldots$ satisfies a recurrence relation. What is the value of the probability $\mathbf{P}_i(\tau < +\infty)$, $i = 1, 2, \ldots$?

b) Which conditions on the three numbers p, q, r ensure that the chain is recurrent? (Hint: use the recurrence criterion of Exercise 4.11).

Exercise 4.19 A) Let X be a Markov chain associated to the transition matrix P on the countable state space E. Let $F \subset E$, τ the hitting time in F and, for a fixed $z \in [-1, +\infty[$, $\psi(i) = \mathbf{E}_i(z^\tau 1_{\{\tau<+\infty\}})$. If $\tau < +\infty$ a.s., $\psi(i)$ coincides with the value at z of the generating function of τ with respect to $\mathbf{P}_i$.

A) Prove that ψ is a solution of

$$\begin{cases} \psi(i) = zP\psi(i) & \text{if } i \in F^c \\ \psi(i) = 1 & \text{if } i \in F. \end{cases}$$

B) Let us consider a random walk on the vertices of the unit cube in $\mathbb{R}^3$: it is the Markov chain for which, at each step, the chain moves from a vertex to one of the three adjoining vertices with probability $\frac{1}{3}$ (see Figure 4.2). Let us fix one of the vertices, i_0, and let τ be the hitting time in $F = \{i_0\}$. Compute the probabilities $\mathbf{P}_{i_1}(\tau = 3)$, $\mathbf{P}_{i_1}(\tau = 15)$ and $\mathbf{P}_{i_1}(\tau \geq 15)$, where i_1 denotes the vertex that is opposed to i_0. What is the value of $\mathbf{E}_{i_1}(\tau)$?

C) Let us go back to the general case considered in (A). Let g be a positive or bounded solution of

$$\begin{cases} g(i) = zPg(i) & \text{if } i \in F^c \\ g(i) = 1 & \text{if } i \in F. \end{cases} \tag{4.49}$$

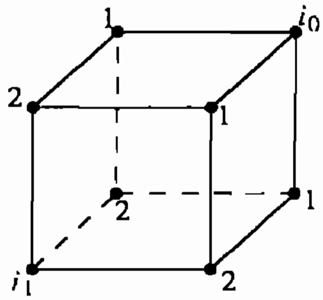

Figure 4.2

C1) Prove that, for every $i \in F^c$,

$$g(i) = z\mathbf{P}_i(\tau = 1) + z\mathbf{E}_i(1_{\{\tau>1\}}g(X_1)).$$

C2) Prove that, for every $i \in F^c$ and $n \geq 0$,

$$g(i) = \sum_{k=1}^{n} z^k \mathbf{P}_i(\tau = k) + z^n \mathbf{E}_i(1_{\{\tau>n\}}g(X_n)). \tag{4.50}$$

Deduce that if $|z| < 1$ ψ is the unique bounded solution of (4.49) whereas if $z \geq 0$ ψ is the smallest positive solution of (4.49).

Exercise 4.20 On the vertices of a triangle consider a Markov chain defined by the following rules: at every transition a displacement takes place to the clockwise adjoining vertex with probability $1-p$ and to the counterclockwise one with probability p, $0 < p < 1$ (see Figure 4.3).

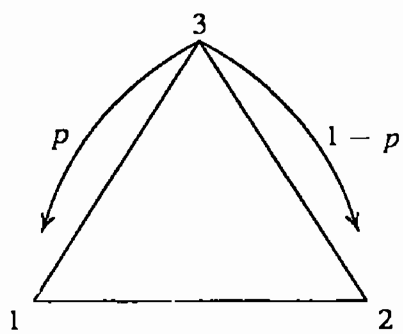

Figure 4.3

a) Write the transition matrix of this chain. Show that it is irreducible and aperiodic; compute its invariant probability.

b) Compute, for every starting law μ,

$$\lim_{n\to\infty} \mathbf{P}_\mu(X_n = 1, X_{n+1} = 2) \quad \text{and} \quad \lim_{n\to\infty} \mathbf{P}_\mu(X_n = 2, X_{n+1} = 1).$$

c) For which values of p is the invariant law reversible (•4.34)?

Exercise 4.21 (Random walk on a graph) Let $G = (E, A)$ a finite graph, where E denotes the set of the vertices and A the set of the edges. On E we consider the transition matrix $P = (p_{ij})_{i,j\in E}$ defined by

$$p_{ij} = \begin{cases} \frac{1}{k_i} & \text{if the vertex } j \text{ is connected to } i \text{ by an edge} \\ 0 & \text{otherwise,} \end{cases}$$

k_i being the number of vertices adjoining to vertex i (a vertex is adjoining to i if and only if it is connected to i by an edge).

a) Show that if the graph is connected, the transition matrix P is irreducible.

b) Show that, if $k = \sum_{i\in E} k_i$, then the probability on E defined by

$$\pi_i = \frac{k_i}{k}$$

is invariant.

Exercise 4.22 A mouse moves randomly on a chessboard with 16 squares, horizontally or vertically (diagonal displacements are not allowed). At every step, it moves from the actual square to one of the k adjoining admissible squares with probability $\frac{1}{k}$. We denote by D the set of the squares of the chessboard and by $(X_n)_{n\geq 0}$ the corresponding Markov chain on D.

Show that it is irreducible positive recurrent and compute its invariant probability.

Exercise 4.23 Let us consider the canonical Markov chain on $E = \{1, 2, 3, 4\}$ with transition matrix:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

1) Show that the chain is irreducible and positive recurrent.

2) What is its invariant probability?

3) What are the limits a.s. of $\frac{1}{n}\sum_{k=0}^{n-1} X_k$ and of $\frac{1}{n}\sum_{k=0}^{n-1} X_k^2$ as $n \to \infty$?

Exercise 4.24 Let us consider the canonical Markov chain on $E = \{1, 2, 3, 4, 5\}$ with transition matrix

$$P = \begin{pmatrix} \frac{1}{3} & 0 & \frac{1}{12} & \frac{1}{4} & \frac{1}{3} \\ 0 & \frac{1}{4} & 0 & 0 & \frac{3}{4} \\ \frac{1}{3} & \frac{1}{12} & \frac{1}{2} & 0 & \frac{1}{12} \\ 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} & \frac{1}{6} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{2}{3} \end{pmatrix}.$$

a) Classify the states. Show that there exists a unique closed class C formed by the recurrent states.

b) Let $\tau = \inf\{n \geq 0; X_n \in C\}$ be the passage time in C. Compute $\mathbf{E}_x(\tau)$ for every $x \in E$.

c) Compute, for $x \in E$ and $y \in C$, $\mathbf{P}_x(X_\tau = y)$ (the *hitting law in C*).

d) Determine, for every $x \in E$, the $\mathbf{P}_x$-a.s. limit of $\frac{1}{n}\sum_{k=0}^{n} X_k$ as $n \to \infty$.

Exercise 4.25 Let us denote by X_n the number of particles contained at time n in a given volume V; we assume that, during the time interval $[n, n+1[$, each of the X_n particles has a probability $p = 1 - q$, $0 < p < 1$ of leaving V and that, in the meantime, a random number of particles, with a Poisson law with parameter λ, enters into V. We assume that the various random phenomena considered above are independent of each other.

1a) Compute $\mathbf{E}(e^{itX_1} \mid X_0 = x)$.

1b) Assume that X_0 has a Poisson law with parameter θ, noted μ_θ. What is the characteristic function of X_1? Show that, for a suitable value of θ, μ_θ is an invariant probability.

2) Show that the transition matrix of the Markov chain $(X_n)_{n\geq 0}$ is given by

$$Q(x, y) = e^{-\lambda} \sum_{k=0}^{x \wedge y} \binom{x}{k} q^k (1-q)^{x-k} \frac{\lambda^{y-k}}{(y-k)!}. \tag{4.51}$$

Deduce that this chain is irreducible, hence positive recurrent.

3) What is the $\mathbf{P}_x$-a.s. limit of $\frac{1}{n}\sum_{k=0}^{n-1} X_k$ as $n \to \infty$? And of $\frac{1}{n}\sum_{k=0}^{n-1} X_k^2$?

Exercise 4.26 Let Q be the transition matrix on $E = \{0, 1\}$

$$Q = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

with $0 < \alpha \leq 1$ and $0 < \beta \leq 1$. Assume that $\min(\alpha, \beta) < 1$. Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in\{0,1\}})$ be the canonical Markov chain associated to Q. The aim of this exercise is to give an elementary proof of Theorems 4.25 and 4.29 for this simple chain.

a) Diagonalize Q and deduce the representation

$$Q^n = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta}\begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}. \tag{4.52}$$

Compute $\lim_{n\to\infty} Q^n$.

b) Compute the stationary distribution π of Q. Which is the value of

$$\mathrm{Cov}_\pi(X_n, X_{n+1}) = \mathbf{E}_\pi(X_n X_{n+1}) - \mathbf{E}_\pi(X_n)\mathbf{E}_\pi(X_{n+1})?$$

Are the r.v.'s $(X_n)_{n\geq 0}$ independent?

c) Let us set $S_n = X_1 + \ldots + X_n$. Compute $\mathbf{E}_\pi(S_n)$ and $\mathrm{Var}_\pi(X_n)$. Deduce that there exists a constant $C < +\infty$ such that

$$\mathrm{Var}_\pi(S_n) \leq C \cdot n$$

and then that

$$\frac{S_n}{n} \underset{n\to\infty}{\to} \frac{\alpha}{\alpha+\beta},$$

the convergence being in probability with respect to $\mathbf{P}_\pi$ and with respect to $\mathbf{P}_0$ and $\mathbf{P}_1$ as well.

Exercise 4.27 Let $(p(x))_{x\in\mathbb{N}}$ be a probability on $\mathbb{N}$ such that $p(x) > 0$ for every $x > 0$. Let us set $G(x) = \sum_{y>x} p(y)$. We consider the Markov chain $(X_n)_{n\geq 0}$ on $\mathbb{N}$ with transition matrix

$$Q(0, y) = p(y), y \geq 0$$

$$Q(x, y) = \begin{cases} \frac{1}{x} & \text{if } y < x \\ 0 & \text{otherwise} \end{cases} \quad \text{if } x \geq 1.$$

1) Show that $(X_n)_{n\geq 0}$ is irreducible and recurrent.

2a) Let μ be an invariant measure. Write the equations satisfied by μ and deduce that $\sum_{y>0}\mu(y)y^{-1} < +\infty$.

2b) Let us set $\varphi(x) = \sum_{y>x}\mu(y)y^{-1}$. Express φ and subsequently μ, as functions of G and $\mu(0)$.

3) Assume $p(0) = 0$ and $p(x) = \frac{1}{x} - \frac{1}{x+1}$, $x \geq 1$. Is this chain positive recurrent? What is the value of $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} 1_{\{X_k=0\}}$?

Exercise 4.28 Let us consider a birth and death chain, i.e., a Markov chain on $\mathbb{N}$ associated to the transition matrix Q defined by $Q(0,0) = r_0 \geq 0$, $Q(0,1) = p_0 > 0$, $p_0 + r_0 = 1$ and, for $i \geq 1$,

$$Q(i,j) = \begin{cases} p_i & \text{if } j = i+1 \\ r_i & \text{if } j = i \\ q_i & \text{if } j = i-1 \end{cases}$$

(see Figure 4.4) where $q_i, p_i > 0$, $r_i \geq 0$ and $p_i + r_i + q_i = 1$. Let, for $i \in \mathbb{N}$, τ_i be the passage time in i (•2.5) and define

$$\gamma_0 = 1, \qquad \gamma_i = \frac{q_1 \ldots q_i}{p_1 \ldots p_i}, \quad i \geq 1.$$

Let us fix finally $a, b \in \mathbb{N}$, $a + 1 < b$ and set $\tau = \tau_a \wedge \tau_b$.

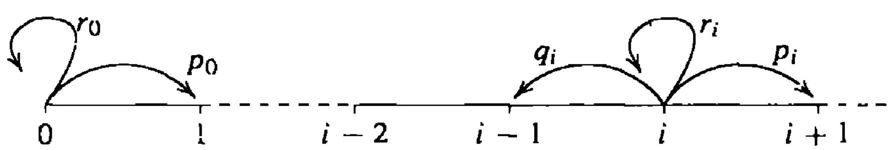

Figure 4.4

a) Investigate the class of functions u satisfying $Qu(i) = u(i)$ for $i \in \{a+1, \ldots, b-1\}$. Are they monotone?

b) Let $a \leq i \leq b$. Show that $\mathbf{P}_i(\tau < +\infty) = 1$; express $v(i) = \mathbf{P}_i(X_\tau = b)$ in function of the numbers γ_i.

c) Show that $\mathbf{P}_1(\tau_0 = +\infty) = \lim_{n\to\infty}\mathbf{P}_1(\tau_n < \tau_0)$. Deduce the value of $\mathbf{P}_1(\tau_0 = +\infty)$ and a condition ensuring the recurrence of the chain in function of the numbers γ_i.

d) Assume $p_i = p$ for every $i \geq 0$, $q_i = q$ for every $i \geq 1$. Investigate the recurrence of the chain according to the possible values of p, q.

e) Assume that $r_i = 0$ for $i \geq 1$ and

$$p_i = \frac{i+2}{2i+3} \qquad q_i = \frac{i+1}{2i+3}.$$

Is this chain recurrent? Same question if

$$p_i = \frac{i+3}{2i+3} \qquad q_i = \frac{i}{2i+3}. \tag{4.53}$$

Exercise 4.29 (Stationary measures of birth and death chains) Let us consider a birth and death chain with values $E = \mathbb{N}$ or $E = \{0, 1, \ldots, m\}$ (see the Figure 4.4 for

$E = \mathbb{N}$). In the first case, we assume $p_i > 0$ for $i \geq 0$ and $q_i > 0$ for $i > 0$; in the second case, assume $p_i > 0$ for $0 \leq i < m$ and $q_i > 0$ for $0 < i \leq m$. The aim of this exercise is to establish conditions for the existence of a stationary probability and to compute it.

a) Show that the chain is irreducible.

b) Show that every measure μ satisfying the detailed balance equation (4.44) has the form

$$\mu_i = \alpha\zeta_i, \qquad i \in E \tag{4.54}$$

where we set

$$\zeta_0 = 1, \qquad \zeta_i = \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i}, \quad i \geq 1. \tag{4.55}$$

Show that the measure defined by (4.54) is stationary.

c) Which conditions ensure the existence of a stationary probability? Express them in terms of the numbers ζ_i.

d) Assume $E = \mathbb{N}$, $p_i = p$ for every $i \geq 0$ and $q_i = q$ for every $i \geq 1$. Which assumptions on p and q ensure the existence of a stationary probability? In this case, which is the value, for $k \in E$, of $\mathbf{E}_k(\sigma_k)$ where, as usual, σ_k is the return time in k (•2.5)?

Exercise 4.30 (Bernoulli–Laplace diffusion) r white balls and r black balls ($r > 1$) are divided between two boxes with the constraint that every box contains r balls.

The state X_n of the system is specified by the number of white balls in the first box, so that the state space is $E = \{0, 1, \dots, r\}$. At every step one ball is chosen in each box and they are switched. All the draws are assumed to be independent and uniform.

a) Find the transition matrix P of the Markov chain $(X_n)_{n\geq 0}$. Is it irreducible? Aperiodic?

b) Determine the stationary probability of this chain (one can use Exercise 4.29).

Exercise 4.31 Let us consider d balls numbered from 1 to d ($d > 1$), divided between two urns A and B. A number i is randomly chosen between 1 and d and then one of the urns is chosen at random, into which the ball with the number i is replaced. All the draws are assumed to be independent and uniform. We note X_n the number of balls present in urn A after n draws.

a) Determine the transition matrix Q of the Markov chain $(X_n)_{n\geq 0}$. Show that this chain is irreducible and positive recurrent. Is it aperiodic?

b) Show that there exist two real constants a and b such that, for every $x \in E$, $\sum_{y\in E} yQ(x, y) = ax + b$. Deduce the values of $\mathbf{E}_x(X_1)$, $\mathbf{E}_x(X_n)$ and then of $\lim_{n\to\infty} \mathbf{E}_x(X_n)$.

c) Assume that X_0 has a binomial law with parameters d and $\frac{1}{2}$. Determine the law of X_1.

d) What is the invariant probability of this chain? What is the value of $\mathbf{E}_d(\sigma_d)$ (σ_d being the return time in d as in •2.5)? And of $\lim_{n\to\infty} Q^n(x, y)$?

Exercise 4.32 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on a countable set E, irreducible and positive recurrent. Let π be the invariant probability and let $a \in E$ be fixed. We note σ the return time (•2.5) in a.

a) Show that, for $f: E \to \mathbb{R}^+$,

$$\langle \pi, f \rangle = \sum_{b \in E} \pi(b) f(b) = \pi(a) \mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1} f(X_k)\Big).$$

b) Show that

$$\mathbf{E}_\pi(\sigma) = \frac{\pi(a)}{2} \mathbf{E}_a\big(\sigma(\sigma+1)\big),$$
$$\mathbf{E}_\pi(\sigma^2) = \frac{\pi(a)}{6} \mathbf{E}_a\big(\sigma(\sigma+1)(2\sigma+1)\big).$$

Exercise 4.33 (An extension of the maximum principle) Consider a canonical Markov chain $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n \geq 0}, (X_n)_{n \geq 0}, (\mathbf{P}_x)_{x \in E})$ on a countable set E, with transition matrix P. Let $f: E \to \mathbb{R}^+$; we note

$$Uf(x) = \sum_{n \geq 0} \mathbf{E}_x(f(X_n)) = \sum_{n \geq 0} P^n f(x)$$

and $A = \{x, f(x) > 0\}$ (the support of f). Let $\tau = \inf\{n \geq 0; X_n \in A\}$ be the passage time in A. Show that

$$Uf(x) = \mathbf{E}_x\big[1_{\{\tau < +\infty\}} Uf(X_\tau)\big] \tag{4.56}$$

and deduce that $\sup_{y \in E} Uf(y) = \sup_{y \in A} Uf(y)$.

Exercise 4.34 The goal of this exercise is to investigate whether a function of a Markov chain is still a Markov chain.

a) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n \geq 0}, (X_n)_{n \geq 0}, \mathbf{P})$ be a Markov chain on a countable set E, with transition matrix P and starting law μ. Let $\psi: E \to F$ be a *bijective* application. Show that, if $Y_n = \psi(X_n)$, then $Y = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n \geq 0}, (Y_n)_{n \geq 0}, \mathbf{P})$ is a Markov chain on F. Determine its transition matrix and starting law.

b) We construct now an example which shows that (a) does not hold necessarily if ψ is not bijective. Let us consider the random walk on the graph of Figure 4.5. This

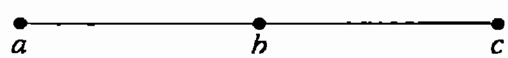

Figure 4.5

is the chain that at every step moves from the actual vertex to one of the adjoining ones, randomly chosen. Let ψ be the application, with values in $F = \{0, 1\}$, defined by $\psi(a) = 0$, $\psi(b) = \psi(c) = 1$. Show that

$$\begin{aligned} \mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1, Y_{n-1} = 0) &= \frac{1}{2}, \\ \mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1) &= \frac{1}{2} \frac{\mathbf{P}(X_n = b)}{\mathbf{P}(X_n = b) + \mathbf{P}(X_n = c)} \end{aligned} \tag{4.57}$$

and deduce that the process $(\Omega, \mathscr{F}, (Y_n)_{n \geq 0}, \mathbf{P})$ does not satisfy the Markov property (with respect to its natural filtration).

♦ The fact that a function Y of a Markov chain may not be a Markov chain itself is a fact of which one should take care and that leads to frequent mistakes. Conversely it is useful to find criteria in order to ensure that Y is a Markov chain, for a noninvertible ψ. This is the object of Exercise 4.35.

Exercise 4.35 (Dynkin's criterion) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ be a Markov chain on a countable set E, with transition matrix P. Let $\psi: E \to F$ be a surjective application such that, for every $j \in F$,

$$P(x, \psi^{-1}(j)) = P(y, \psi^{-1}(j)) \qquad \text{whenever } \psi(x) = \psi(y).$$

Define, for $i, j \in F$,

$$Q(i, j) = P(x, \psi^{-1}(j))$$

where x is any state in E such that $\psi(x) = i$.

a) Show that Q is a transition matrix and that, if $Y_n = \psi(X_n)$, then the process $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (Y_n)_{n\geq 0}, \mathbf{P})$ is a Markov chain with transition matrix Q.

b) Show that, if π is a stationary probability for P, the image law of π through ψ (noted $\pi \circ \psi^{-1}$) is a stationary probability for Q.

Exercise 4.36 (Random walk on the vertices of the hypercube)

a) Let $E = \{0, 1\}^m$ be the set of all binary vectors of length m. If $x \in E$, let $x^{(i)}$ be the vector whose coordinates coincide with those of x, but for the i-th one: $x_i^{(i)} = 1$ if $x_i = 0$, $x_i^{(i)} = 0$ if $x_i = 1$. On E let us consider the transition matrix P defined by

$$P(x, y) = \begin{cases} \frac{1}{m} & \text{if } y = x^{(i)}, i = 1, \dots, m \\ 0 & \text{otherwise.} \end{cases}$$

Is it irreducible? Aperiodic? What is its stationary probability?

♦ If we consider the states $x \in E$ as vectors in $\mathbb{R}^m$, it is not difficult to identify them as the vertices of the unit cube and to check that the Markov chain associated to P is the one which, at every step, moves from one vertex x of the cube to another vertex having an edge in common with x.

b) Consider the application $\psi: E \to F = \{0, 1, \dots, m\}$ defined by $\psi(x) = x_1 + \dots + x_m$. Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbf{P})$ be a Markov chain associated to P and set $Y_n = \psi(X_n)$. Show that Dynkin's condition of Exercise 4.35 is satisfied and that $Y = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (Y_n)_{n\geq 0}, \mathbf{P})$ is a Markov chain on F. What is its transition matrix? Compare it with the transition matrix of the Ehrenfest model of Exercise 4.15. Deduce the stationary law of the Ehrenfest chain.

c) Same questions if P is defined by

$$P(x, y) = \begin{cases} \frac{1}{m+1} & \text{if } y = x^{(i)}, i = 1, \dots, m \text{ or } y = x \\ 0 & \text{otherwise.} \end{cases} \tag{4.58}$$

Exercise 4.37 (The Metropolis algorithm) Let E be a finite set of states and P a transition matrix on E, *symmetric* and irreducible. Let π be a probability on E such that $\pi(x) > 0$ for every $x \in E$. We define on E a new transition matrix Q by setting

$$Q(x, y) = \begin{cases} P(x, y) & \text{if } \pi(y) \geq \pi(x) \\ P(x, y)\frac{\pi(y)}{\pi(x)} & \text{if } \pi(y) < \pi(x) \\ 1 - \sum_{z \neq x} Q(x, z) & \text{if } y = x \end{cases}$$

($Q(x, y)$ is defined for $y \neq x$ through the first two lines of the previous formula and then the value of $Q(x, x)$ is derived according to the third line).

a) Show that the probability π is reversible (•4.34) for Q and thus invariant.

b) Show that Q is irreducible (hint: use the criterion of Exercise 4.10 (c)).

c) Assume that π is not the uniform distribution (otherwise $Q = P$). Let us note M the set of the states $x \in E$ such that $\pi(x) = \max_{y \in E} \pi(y)$.

c1) Show that there exists $x_0 \in M$ such that $P(x_0, y) > 0$ for some $y \notin M$. Deduce that $Q(x_0, z) < P(x_0, z)$ for some $z \in E$, $z \neq x_0$. What can be said of $Q(x_0, x_0)$?

c2) Show that Q is aperiodic (even if P is not).

♦ This exercise shows how to construct a Markov chain having a stationary probability *fixed beforehand*. This is useful in some applications: sometimes one wants to simulate a r.v. with values in a set E, finite but very large and having a fixed law π. If one could afford a transition matrix Q irreducible and aperiodic, having π as a stationary law, one might simulate the chain X associated to Q. Theorem 4.29 ensures then that, for n large enough, X_n has, approximately, law π. This opens a new problem: how to determine this value of n that is large enough (speed of convergence ...)?

Exercise 4.38 (Perfect sampling) Let Q be an irreducible transition matrix on $E = \{0, 1, \dots, N\}$ with invariant probability π. Let us assume that, for every $i \in E$, $Q(i, 0) > 0$. Define, for every $i \in E$,

$$s_0(i) = 0,\ s_1(i) = Q(i, 0),\ \dots,\ s_k(i) = \sum_{j=0}^{k-1} Q(i, j),\ \dots,\ s_{N+1}(i) = 1$$

and an application $g : E \times [0, 1] \to E$ by

$$g(i, u) = j \text{ if } u \in [s_j(i), s_{j+1}(i)[.$$

Let $(U_n)_{n \in \mathbb{Z}}$ be a family of independent r.v.'s uniformly distributed on $[0, 1]$ and Z a r.v. independent of $(U_n)_{n \in \mathbb{Z}}$ with law π. Let us define, for every $n \in \mathbb{Z}$ and $i \in E$, $X_n^n(i) = i$, $X_{n+1}^n(i) = g(i, U_n)$ and, for $p \geq 2$,

$$X_{n+p}^n(i) = g(X_{n+p-1}^n(i), U_{n+p-1}) = X_{n+p}^{n+p-1}(X_{n+p-1}^n(i)).$$

1) Show that, for every $n < r < s$, it holds $X_s^n(i) = X_s^r(X_r^n(i))$. Deduce that, if $X_{n+p}^n(i) = X_{n+p}^n(j)$, then, for every $m \geq p$, $X_{n+m}^n(i) = X_{n+m}^n(j)$.

2) Let $X_n = X_n^0(Z)$. Show that $(X_n)_{n \geq 0}$ is a Markov chain with initial law π and transition matrix Q.

3) Let $A_n = \{\text{for every } i, j \in E, X_{n+1}^n(i) = X_{n+1}^n(j)\}$ and $\alpha = \inf_{i \in E} Q(i, 0) = \inf_{i \in E} s_1(i) > 0$. Show that A_n is $\sigma(U_n)$-measurable and that $\mathrm{P}(A_n) \geq \alpha$ for every $n \in \mathbb{Z}$.

Let us define $\nu = \inf\{n \geq 0\colon \text{ for every } i, j \in E.\ X_0^{-n}(i) = X_0^{-n}(j)\}$.

4) Show that $\mathrm{P}(\nu > n + 1) \leq (1 - \alpha)\mathrm{P}(\nu > n)$ and deduce that $\mathrm{P}(\nu < +\infty) = 1$.

Let $Y = X_0^{-\nu}(i)$ (going back to the definition of ν, this does not depend on i).

5) Show that, on $\{\nu \leq n\}$, it holds, for every $i \in E$, $X_0^{-n}(i) = Y$. Deduce that, for every $n \in \mathbb{N}$ and $j \in E$,

$$|\mathrm{P}(X_0^{-n}(Z) = j) - \mathrm{P}(Y = j)| \leq 2\mathrm{P}(\nu > n)$$

and that Y is a r.v. with law π.

Q is said to be monotone if the points of E can be numbered from 0 to N in such a way that:

$$i \le i' \quad \Rightarrow \quad Q(i,[k,N]) \le Q(i',[k,N]) \text{ for every } k \in E. \tag{4.59}$$

6) Show that, if Q is monotone, then, for every $n \in \mathbb{Z}$, $p \in \mathbb{N}$ and $i \in E$,

$$X^n_{n+p}(0) \le X^n_{n+p}(i) \le X^n_{n+p}(N). \tag{4.60}$$

Deduce that $\nu = \inf\{n \ge 0; X_0^{-n}(0) = X_0^{-n}(N)\}$.

7) Let π be a probability on E, different from the uniform distribution and such that $\pi(i) > 0$ for every $i \in E$. Let the states be numbered from 0 to N in such a way that π is decreasing on $\{0,\dots,N\}$. Let us consider the transition matrix

$$Q(i,j) = \begin{cases} \frac{1}{N+1} & \text{if } j < i \\ \frac{1}{N+1}\frac{\pi(j)}{\pi(i)} & \text{if } j > i \\ 1 - \sum_{h\ne i} Q(i,h) & \text{if } j = i. \end{cases}$$

We know (Exercise 4.37) that Q is irreducible with invariant probability π and it is clear that $Q(i,0) > 0$ for every $i \in E$. Show that Q is monotone.

♦ This method of simulation of a r.v. whose law is the invariant probability π is called *perfect sampling* and was introduced by Propp and Wilson in 1996. In practice, let $(u_{-n})_{n\in\mathbb{N}}$ be random numbers uniform on $[0,1]$. Compute first $g(i,u_{-1})$, $i \in E$. If all these numbers are equal to $i_1 \in E$, then set $Y = i_1$. Otherwise compute $g(g(i,u_{-2}),u_{-1})$. If all these numbers are equal to $i_2 \in E$, then set $Y = i_2$ and so on. This process comes to a stop necessarily in a finite time and Y is a sampling with law π. Of course this is complicated to perform if E is very large, unless the chain satisfies the monotonicity property (4.60) (which occurs in many interesting situations); if this is the case, then it suffices to apply the previous procedure to the states 0 and N only.

Actually this method can by applied to every Markov chain on a finite states space E, irreducible and aperiodic, since then there exists $p \ge 0$ such that $Q^p(i,j) > 0$ for every i, j. We shall see this in Problem 4.9 (it is point (A3), since E is assumed to be finite). Obviously Q^p has still π as an invariant probability.

Exercise 4.39 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge0}, (X_n)_{n\ge0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on a countable set E, with transition matrix Q. We assume that, for every $x \in E$, $Q(x,x) < 1$. Let $\tau = \inf\{n \ge 1; X_n \ne X_0\}$.

1) Show that, for every law $\mathbf{P}_x$, τ is a a.s. finite stopping time. Compute its law as well as the law of X_τ.

Let us define a sequence of integer r.v.'s by

$$\tau_0 = 0,\ \tau_1 = \tau, \dots, \tau_{n+1} = \tau_n + \tau\circ\theta_{\tau_n}, \dots .$$

2) Show that $\tau_n = \inf\{k; k > \tau_{n-1}, X_k \ne X_{\tau_{n-1}}\}$ and that $(\tau_n)_{n\ge0}$ is a sequence of stopping times that are a.s. finite.

3) Show that, if $Y_n = X_{\tau_n}$, then, for every $x \in E$,

$$Y = (\Omega, \mathscr{F}, (\mathscr{F}_{\tau_n})_{n\ge0}, (Y_n)_{n\ge0}, \mathbf{P}_x)$$

is a Markov chain. What is its transition matrix?

4) We assume that X is recurrent irreducible with invariant measure μ. Show that Y is recurrent irreducible and that $\tilde{\mu}(y) = (1 - Q(y, y))\mu(y)$, $y \in E$ is an invariant measure for Y (hint: note that the states visited by Y up to time n coincide with those visited by X up to time τ_n).

Exercise 4.40 Let P be a transition matrix on a countable set E. Let us fix $x \in E$ and assume that $U(x, x) = \sum_{n\geq 0} P^n(x, x) = +\infty$ and $P(x, x) = 0$.

Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_y)_{y\in E})$ be the canonical Markov chain with transition matrix P and $\rho_0 = 0$, $\rho_1, \ldots$ the subsequent times of passage in x defined, by recurrence, by $\rho_i = \inf\{k; k > \rho_{i-1}, X_k = x\}$.

a) Show that the r.v's. $(\rho_n)_{n\geq 0}$ are $\mathbf{P}_x$-a.s. finite stopping times.

b) Let us set $Z_n = X_{\rho_n+1}$, $n \in \mathbb{N}$. Using the strong Markov property, show that the r.v.'s $(Z_n)_{n\geq 0}$ are independent and determine their law.

Exercise 4.41 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on a countable set E, with transition matrix Q; we denote by $\mathbf{P}$ the law of the chain with starting distribution μ and assume that X is irreducible positive recurrent with invariant probability π. Let f be a *bounded* function on E such that $\langle \pi, f\rangle = 0$. Assume that there exists a *bounded* function g satisfying $(I - Q)g = f$. We note $S_n = \sum_{k=0}^n f(X_k)$. The aim of this exercise is to investigate the asymptotic behaviour of $\frac{1}{n}\,\mathbf{E}(S_n^2)$.

1) What is the limit of $\frac{1}{n}\,\mathbf{E}(S_n)$ as $n \to \infty$?

2) Show that the decomposition $S_n = M_n + Z_n$ holds, where

$$M_0 = 0,\quad M_n = \sum_{k=1}^n U_k,\quad U_k = g(X_k) - Qg(X_{k-1}),\quad Z_n = g(X_0) - Qg(X_n).$$

Show that $\mathbf{E}(U_k \mid \mathscr{F}_{k-1}) = 0$; deduce that $(M_n)_{n\geq 0}$ is a martingale and that

$$\mathbf{E}(M_n^2) = \mathbf{E}\Big(\sum_{k=1}^n U_k^2\Big). \tag{4.61}$$

3) Show that $\mathbf{E}(g(X_{k+1})Qg(X_k)) = \mathbf{E}((Qg)^2(X_k))$. Derive from this that

$$\frac{1}{n}\,\mathbf{E}\Big(\sum_{k=0}^{n-1} g(X_{k+1})Qg(X_k)\Big) \underset{n\to\infty}{\to} \langle (Qg)^2, \pi\rangle.$$

4) Show that

$$\frac{1}{n}\,\mathbf{E}(M_n^2) \underset{n\to\infty}{\to} \langle g^2 - (Qg)^2, \pi\rangle \overset{\text{def.}}{=} \Sigma^2.$$

5) Show that $\frac{1}{n}\,\mathbf{E}(S_n^2) \to_{n\to\infty} \Sigma^2$.

Exercise 4.42 (Product of Markov chains) Consider two canonical Markov chains

$$X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E}),$$
$$X' = (\Omega', \mathscr{F}', (\mathscr{F}'_n)_{n\geq 0}, (X'_n)_{n\geq 0}, (\mathbf{P}'_{x'})_{x'\in E'}),$$

on the countable sets E and E' with transition matrices Q and Q', respectively. Let us set

$$\bar{\Omega} = \Omega \times \Omega',\quad \bar{\mathscr{F}} = \mathscr{F} \otimes \mathscr{F}',\quad \bar{X}_n = (X_n, X'_n),\quad \bar{\mathbf{P}}_{(x,x')} = \mathbf{P}_x \otimes \mathbf{P}'_{x'}.$$

1) Show that $\bar{X} = X \otimes X' = (\bar{\Omega}, \bar{\mathscr{F}}, (\bar{X}_n)_{n\geq 0}, (\bar{\mathbf{P}}_{(x,x')})_{(x,x')\in\bar{E}})$ is a Markov chain on $\bar{E} = E \times E'$ and determine its transition matrix $\bar{Q}$.

2) Compute $\bar{Q}^2$ and then $\bar{Q}^n$.

3) Show that, if ν (resp. ν') is an invariant measure for X (resp. X'), then $\nu \otimes \nu'$ is an invariant measure for $\bar{X}$.

4) Show that, if $(a, a') \in \bar{E}$ is a recurrent state for $X \otimes X'$, a is recurrent for X and a' is recurrent for X'.

5) Let us choose $E = \mathbb{Z}$, $Q(x, y) = \frac{1}{2}$ if $|x - y| = 1$ and $Q(x, y) = 0$ otherwise. Compute $Q^n(0, 0)$; is the state 0 recurrent?

6) Prove, using (5), that a can be recurrent for X, a' recurrent for X' but (a, a') transient for the product chain.

Exercise 4.43 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on a countable set E with transition matrix Q. Assume that the potential matrix U of the chain satisfies

$$0 < U(x, y) < +\infty, \qquad \text{for every } x, y \in E.$$

Let F be a *finite* subset of E. We set

$$\rho = \sup\{n \geq 0;\ X_n \in F\} \text{ and } \rho = -\infty \text{ if } \{n \geq 0;\ X_n \in F\} = \emptyset,$$
$$v(x) = v_F(x) = 1_F(x)\mathbf{P}_x\Big(\bigcap_{n\geq 1}\{X_n \notin F\}\Big).$$

Thus ρ is the last passage time in F (this is not a stopping time!) and, for $x \in F$, $v(x)$ is the probability of never coming back to F when starting at x.

1) Show that, for every $x \in F$, $\mathbf{P}_x(0 \leq \rho < +\infty) = 1$.

2) Show that, for every $x \in E$ and $n \geq 0$, $\mathbf{P}_x(\rho = n) = Q^n v(x)$ and then that, for $x \in F$, $Uv(x) = 1$.

3) Let ν be a starting law with support contained in F, i.e., satisfying $\nu(x) = 0$ if $x \notin F$. Let $\partial \notin E$ and $E_\partial = E \cup \{\partial\}$. We define a sequence of r.v.'s with values in E_∂ through

$$Y_n = X_{\rho-n} \text{ if } 0 \leq n \leq \rho, \quad Y_n = \partial \text{ if } n > \rho.$$

3a) Letting $x_0, x_1, \dots, x_n \in E$, show that

$$\begin{aligned}\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n) &= \sum_{k=n}^{\infty} \mathbf{P}_\nu(X_{k-n} = x_n, \dots, X_k = x_0, \rho = k) = \\ &= \nu U(x_n) Q(x_n, x_{n-1}) \dots Q(x_1, x_0) v(x_0)\end{aligned}$$

and that

$$\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n, Y_{n+1} = \partial) = \nu(x_n) Q(x_n, x_{n-1}) \dots Q(x_1, x_0) v(x_0).$$

3b) Show that $((Y_n)_{n\geq 0}, \mathbf{P}_\nu)$ is a Markov chain with values in E_∂. Compute its transition matrix and its starting law.

Problems

Problem 4.1 (Riesz decomposition) Let us consider, on the countable set E, a canonical Markov chain $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ with transition matrix P; let U be its potential matrix.

A) Let f be a *finite* excessive function (•4.15) (recall that an excessive function is ≥ 0 by definition).

A1) Show that $(P^n f)_{n\geq 0}$ converges decreasing to a function $\tilde{f}$. Show that $\tilde{f}$ is harmonic (•4.15) and that $f = \tilde{f} + Uh$, where $h = f - Pf$.

A2) Show that, if it holds $f = g_0 + Ug_1$ with g_0 harmonic and $g_1 \geq 0$, then $g_0 = \tilde{f}$ and $g_1 = h$.

A3) Show that the limits

$$\lim_{n\to\infty} f(X_n), \qquad \lim_{n\to\infty} \tilde{f}(X_n) \tag{4.62}$$

exist $\mathbf{P}_x$-a.s. for every $x \in E$. Compute $\lim_{n\to\infty} \mathbf{E}_x(Uh(X_n))$ and deduce that the two limits in (4.62) coincide.

We have thus proved that every finite excessive function can be decomposed uniquely into the sum of a harmonic function and of the potential of a positive function: it is the *Riesz decomposition*. In (B) we compute this decomposition for an excessive function of particular interest.

B) Let $F \subset E$, τ_F the passage time in F (•2.5), N_F the number of visits in F (see (4.22)) and set $\phi_F(x) = \mathbf{P}_x(\tau_F < +\infty)$.

B1) Show that $P^n\phi_F(x) = \mathbf{P}_x\big(\bigcup_{k\geq n}\{X_k \in F\}\big)$ and that ϕ_F is excessive.

B2) Let $\phi_F = \tilde{\phi}_F + Uh_F$ be the Riesz decomposition of ϕ_F. Show that

$$\begin{aligned}
&\tilde{\phi}_F(x) = \mathbf{P}_x(N_F = +\infty)\\
&\lim_{n\to\infty} \phi_F(X_n) = 1_{\{N_F=+\infty\}} \quad \mathbf{P}_x\text{-a.s.}\\
&h_F(x) = 1_F(x)\mathbf{P}_x\Big(\bigcap_{n\geq 1}\{X_n \notin F\}\Big)
\end{aligned}$$

(h_F thus vanishes outside F whereas, for $x \in F$, $h_F(x)$ is equal to the probability of never coming back to F when starting at x).

Problem 4.2 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on the countable set E with transition matrix P.

1) Let f be a positive function on E. For $B \subset E$, let us note I_B the matrix

$$I_B(x,x) = 1 \text{ if } x \in B, \quad I_B(x,x) = 0 \text{ if } x \in B^c, \quad I_B(x,y) = 0 \text{ if } x \neq y.$$

1a) Show that

$$I_B f(x) = 1_B(x) f(x), \qquad (I_B P) f(x) = \mathbf{E}_x(1_B(X_0) f(X_1)).$$

1b) Show that, for every $n \geq 1$,

$$(I_B P)^n f(x) = \mathbf{E}_x[1_B(X_0) 1_B(X_1) \dots 1_B(X_{n-1}) f(X_n)].$$

2) Let $A \subset E$ and $\tau = \inf\{n \geq 0; X_n \in A\}$ the corresponding passage time. Let us define for $f \geq 0$ (see also •4.17),

$$P_A f(x) = \mathbf{E}_x[1_{\{\tau<+\infty\}} f(X_\tau)], \quad U_A f(x) = 1_{A^c}(x)\mathbf{E}_x\Big[\sum_{n=0}^{\tau-1} f(X_n)\Big].$$

2a) Show that, for every $n \geq 0$,

$$\mathrm{E}_x[1_{\{\tau=n\}} f(X_n)] = (I_{A^c} P)^n I_A f(x)$$

and deduce that

$$P_A = \sum_{n \geq 0} (I_{A^c} P)^n I_A.$$

2b) Show that, for every $n \geq 0$,

$$\mathrm{E}_x[1_{\{\tau>n\}} f(X_n)] = (I_{A^c} P)^n I_{A^c} f(x)$$

and deduce that

$$U_A = \sum_{n \geq 0} (I_{A^c} P)^n I_{A^c} = \sum_{n \geq 0} I_{A^c} (P I_{A^c})^n.$$

3) Show that

$$P_A = I_A + I_{A^c} P P_A, \quad U_A = I_{A^c} + I_{A^c} P U_A.$$

4) Let g and h be positive functions on E. We set

$$u = P_A g + U_A h.$$

Remark that u only depends on the values of g on A and on those of h on A^c.

4a) Show that $u = I_A g + I_{A^c}(h + Pu)$ and derive that u satisfies

$$u(x) = \begin{cases} g(x) & \text{on } A \\ h(x) + Pu(x) & \text{on } A^c. \end{cases} \tag{4.63}$$

4b) Let v be another positive solution of (4.63). Show that, for every $n \geq 0$,

$$v \geq \sum_{k=0}^{n} (I_{A^c} P)^k (I_A g + I_{A^c} h) \tag{4.64}$$

and deduce that $v \geq u$.

5) Give a proof of Corollary 4.15.

Problem 4.3 Let $\mu = (p_k)_{k \geq 0}$ be a probability on $\mathbb{N}$ such that $p_0 + p_1 < 1$. In this problem we investigate the simplest model of a branching process, the Galton–Watson process. It models the following situation: at time 0 there is one particle that at time 1 undergoes a random subdivision in k particles, $k \geq 0$, with probability p_k. And so on for the resulting particles, each performing a random subdivision independent of the others. We denote by Z_n the number of particles at time n.

The point of interest of this model is to compute the extinction probability, i.e., the value of $\mathrm{P}(\tau_0 < +\infty)$, τ_0 being the passage time in 0 (•2.5).

We set $f(s) = \sum_{k \geq 0} p_k s^k$ and $m = \sum_{k \geq 0} p_k k$ ($m \leq \infty$) (respectively, the generating function and the mean of the probability μ).

A1) Let us model the process $(Z_n)_{n \geq 0}$ by a Markov chain. Show that its transition matrix is

$$P(i, j) = \mu^{*i}(j) \tag{4.65}$$

if $i \geq 1$ and $P(0, 0) = 1$ (μ^{*i} is the convolution of μ with itself i times).

A2) Show that, for $0 \le s \le 1$ and $i \ge 0$,

$$\sum_{j=0}^{\infty} P(i, j)s^j = f(s)^i. \tag{4.66}$$

Classify the states of this chain. Which states are recurrent?

A3) Let us define a sequence of functions $(f_n)_{n\ge 1}$ by $f_1 = f$, $f_{n+1}(t) = f(f_n(t))$. Show that

$$\sum_{j=0}^{\infty} P^n(i, j)s^j = f_n(s)^i.$$

A4) Show that, a.s., $\{Z_n = 0\} = \{\tau_0 \le n\}$ and deduce that

$$\mathbf{P}(\tau_0 < +\infty) = \lim_{n\to\infty} \mathbf{P}(Z_n = 0) = \lim_{n\to\infty} f_n(0).$$

B) In this part we compute $\mathbf{P}(\tau_0 < +\infty)$.

B1) Show that f is strictly increasing and strictly convex in $[0, 1]$.

B2) Show that, if $m \le 1$, $f(t) > t$ on $[0, 1[$ and that, if $m > 1$, the equation $f(t) = t$, $0 \le t < 1$, admits a unique root q.

B3) Moreover let q be the smallest root of $f(t) = t$ in $[0, 1]$. Then $q = 1$ if $m \le 1$ and $q < 1$ if $m > 1$, because of (B2).

Show that if $0 \le t < q$, then as $n \to \infty$ $f_n(t) \uparrow q$ and then that if $q < t < 1$, $f_n(t) \downarrow q$. Show that if $t = q$ or 1, then $f_n(t) = t$ for every $n \ge 1$. What do you deduce about $\mathbf{P}(\tau_0 < +\infty)$?

C) This problem can also be treated via martingale methods.

C1) Show that $\mathbf{E}(Z_n \mid Z_{n-1}) = mZ_{n-1}$ a.s.; deduce that $\mathbf{E}(Z_n) = m^n$ and that, if $m < 1$, $Z_n \to_{n\to\infty} 0$ a.s.

C2) Note that, for every $k \ge 1$, $\mu^{*k}(k) < 1$. Deduce that, for every $k \ge 1$, $\mathbf{P}(Z_{n+p} = k$ for every $p \ge 0) = 0$. Assume $m = 1$. Show that $(Z_n)_{n\ge 0}$ is a martingale converging to 0 a.s.

C3) Assume $m > 1$ and let q be the unique solution of $f(s) = s$ in $[0, 1[$. Show that $X_n = q^{Z_n}$ is a martingale. Show that $q = \mathbf{P}(\tau_0 < +\infty)$ (the extinction probability).

C4) Show that, if $\sigma^2 = \mathrm{Var}(Z_1)$,

$$\mathbf{E}(Z_n^2 \mid Z_{n-1}) = \sigma^2 Z_{n-1} + m^2 Z_{n-1}^2$$

and deduce that if $m = 1$ $\mathrm{Var}(Z_n) = n\sigma^2$ whereas, if $m \ne 1$,

$$\mathrm{Var}(Z_n) = \sigma^2 m^{n-1}\frac{m^n - 1}{m - 1}.$$

C5) Assume $m > 1$ and set $W_n = Z_n/m^n$. Show that $(W_n)_{n\ge 0}$ is a martingale bounded in L^2. Deduce that $W_n \to_{n\to\infty} W_\infty$ a.s. and in L^2, where W_∞ is a real r.v. whose characteristic function ϕ is a solution of the equation

$$f(\phi(t)) = \phi(mt).$$

Problem 4.4 A princess has n ($n \ge 5$) suitors numbered in decreasing order of appreciation $1, 2, \dots, n$ (suitor number 1 being thus the most deserving). She must choose one of them. The problem is that they are introduced one by one before her (in a random order) and that she can only compare a suitor with those who were introduced earlier. Moreover she must make her choice at once and cannot change her mind and call back a suitor whom she has already dismissed. She must then devise a strategy in order to maximize her chance of choosing the best one ... .

Let r, $0 \le r \le n-1$ be fixed. The princess decides to let go all the first r suitors and to choose the first one appearing after the r already dismissed, who appears to be better than the r first ones. She will choose the last one if no one will appear to be better than the r first. The problem is to determine which is the value of r that gives her the best chance of choosing the suitor number 1.

Let Ω be the set of permutations of $1, 2, \dots, n$, endowed with the uniform probability; $\sigma \in \Omega$ models a possible ordering: the candidates pass by in the order $\sigma(1), \sigma(2), \dots, \sigma(n)$.

For $1 \le k \le n$, let us introduce the r.v. Y_k that is the rank of $\sigma(k)$ in the set $\{\sigma(1), \dots, \sigma(k)\}$ sorted in increasing order (for the natural order of $\mathbb{N}$); thus Y_k is the rank, according to its merit, of the suitor appearing at time k among the first k.

1) Assume $n = 5$.

1a) Let $\sigma = (3, 5, 1, 4, 2)$. Compute $Y_k(\sigma)$, $k = 1, \dots, 5$.

1b) Let $(Y_k(\sigma), k = 1, \dots, 5) = (1, 1, 3, 4, 1)$. Determine σ.

2a) Show that, for $1 \le k \le n$,

$$\{Y_k = 1, Y_{k+1} > 1, \dots, Y_n > 1\} = \{\sigma(k) = 1\}.$$

2b) Show that $F : \sigma \to (Y_1(\sigma), Y_2(\sigma), \dots, Y_n(\sigma))$ is a bijection of Ω onto the set $\Pi = \{1\} \times \{1, 2\} \times \dots \times \{1, 2, \dots, n\}$.

2c) Deduce that the r.v.'s $Y_1, Y_2, \dots, Y_n$ are independent and that the law of Y_k is the uniform probability on $\{1, 2, \dots, k\}$.

Let us define, for $r = 1, \dots, n-1$, $\tau_r = \inf\{k > r; Y_k = 1\}$ with the understanding $\tau_r = n$ if $\{k > r; Y_k = 1\} = \emptyset$ and $A_r = \{\sigma(\tau_r) = 1\}$. Remark that A_r corresponds to the event "the suitor passing by at time τ_r is the best one".

3a) Show that

$$\mathbf{P}(A_r) = \frac{r}{n} \sum_{k=r}^{n-1} \frac{1}{k} \stackrel{\text{def}}{=} s_r.$$

3b) Prove that there exists a unique value r^* maximizing s_r (hint: look at $s_{r+1} - s_r$). We shall admit that $\sum_{k=r}^{n} \frac{1}{k} \neq 1$ for all integers r, n, $r \le n$, $n > 1$.

3c) Show that $r^* \sim \frac{n}{e}$ for $n \to +\infty$ (recall that $\sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + o(1)$, $\gamma = 0.577\dots$ being Euler constant). What is the value of $\lim_{n\to\infty} \mathbf{P}(A_r)$?

4) Let us define $X_1 = 1$ and

$$X_{j+1} = \inf\{k > X_j; Y_k = 1\}$$

($= n + 1$ if this set is empty). Show that $(X_j)_{j\ge 0}$ is a Markov chain and compute its transition matrix.

Problem 4.5 (Optimal stopping) An observer has admittance to the results of a random experiment given by a Markov chain $(X_n)_{n\ge 0}$ on a countable set E, with matrix transition Q. Given a bounded function $f : E \to \mathbb{R}^+$, he must decide a stopping time: if the chain is stopped at time ν, he gets a gain $f(X_\nu)$. The problem is to choose ν in order to maximize the mean gain $\mathbf{E}(f(X_\nu))$.

We note $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ the canonical chain associated to Q. We write $\mathscr{T}$ the family of the a.s. finite stopping times (i.e., such that $\mathbf{P}_x(\tau = +\infty) = 0$ for every $x \in E$). We define

$$V(x) = \sup_{\nu\in\mathscr{T}} \mathbf{E}_x(f(X_\nu)),$$

which is called the *value function* of the problem.

1) Let us define a sequence $(f_n)_{n\geq 0}$ of real functions on E by $f_0 = f$, $f_{n+1} = \max(f, Qf_n)$. Show that the sequence $(f_n)_{n\geq 0}$ is increasing and converges to a bounded function f^*, which satisfies the relation $f^* = \max(f, Qf^*)$. Prove that f^* is the smallest excessive function majorizing f.

2a) Show that, if $(\nu_x)_{x\in E}$ is a family of stopping times in $\mathscr{T}$, then

$$\nu = \sum_{x\in E} \nu_x 1_{\{X_0=x\}}$$

is still a stopping time belonging to $\mathscr{T}$. Show that, for every $\varepsilon > 0$, there exists $\nu \in \mathscr{T}$ such that, for every $x \in E$, $\mathbf{E}_x(f(X_\nu)) \geq V(x) - \varepsilon$.

2b) Deduce that V is excessive.

2c) Let u be an excessive function majorizing f. Show that, for every $\rho \in \mathscr{T}$, $\mathbf{E}_x(u(X_\rho)) \leq u(x)$ and thus that $u(x) \geq V(x)$. Deduce that $V = f^*$.

3) Let $\tau = \inf\{n;\ f(X_n) = f^*(X_n)\}$.

3a) Assume that E is finite. Remark that there exists $\delta > 0$ such that $f^*(x) > f(x) + \delta$ for every x such that $f^*(x) > f(x)$. Show that, for every $x \in E$, $\mathbf{P}_x(\tau < +\infty) = 1$.

We assume from now on that $\tau \in \mathscr{T}$ (which is certainly true if E is finite).

3b) Show that, if $n < \tau$, $f^*(X_n) = Qf^*(X_n)$.

3c) Deduce that $(f^*(X_{n\wedge\tau}))_{n\geq 0}$ is a martingale and conclude that

$$f^*(x) = V(x) = \mathbf{E}_x(f(X_\tau)).$$

Problem 4.6 (A criterion of recurrence/transience: the barrier function) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in\mathbb{N}})$ be a canonical Markov chain on $\mathbb{N}$, irreducible, with transition matrix Q and potential matrix U. We note, for $F \subset \mathbb{N}$, N_F the number of visits of X in F (see (4.22)).

A) Show that, for every $x \in \mathbb{N}$, $\mathbf{P}_x(\overline{\lim}_{n\to\infty} X_n = +\infty) = 1$.

B) Let ϕ be a finite positive function on $\mathbb{N}$ such that

$$Q\phi(x) \leq \phi(x) \quad \text{for every } x > p \tag{4.67}$$

and let τ be the passage time in $J_p = \{0, 1, \dots, p\}$.

Show that, for every $x > p$, $(\phi(X_{n\wedge\tau}))_{n\geq 0}$ is a $\mathbf{P}_x$-supermartingale and that it converges $\mathbf{P}_x$-a.s. to a $\mathbf{P}_x$-integrable r.v. Z.

C) Assume that there exists a positive function ϕ satisfying (4.67) and such that $\phi(x) \to_{x\to\infty} +\infty$.

C1) Show that $\mathbf{P}_x(\tau < +\infty) = 1$ for every $x \in \mathbb{N}$.

C2) Show that, for every $x \in \mathbb{N}$ and every n, $\mathbf{P}_x(\tau\circ\theta_n < +\infty) = 1$. Deduce that, for every $x \in \mathbb{N}$, $\mathbf{P}_x(N_{J_p} = +\infty) = 1$ and $U(x, J_p) = +\infty$.

C3) Show that there exists $z \leq p$ such that $U(x, z) = +\infty$. Deduce that X is recurrent.

D) We assume now that there exists a *strictly positive* function ϕ on $\mathbb{N}$, satisfying (4.67) and such that $\phi(x) \to_{x\to\infty} 0$. Show, by studying $\mathbf{E}_x[\phi(X_\tau)1_{\{\tau<+\infty\}}]$, that $\mathbf{P}_x(\tau < +\infty) \to 0$ as $x \to \infty$ and that X is transient.

E) Assume $Q(0, 1) = 1$ and that, for $x \geq 1$, Q is given by

$$Q(x, y) = \begin{cases} p_x & \text{if } y = x + 1 \\ q_x & \text{if } y = x - 1 \\ 0 & \text{otherwise} \end{cases}$$

where $p_x + q_x = 1$, $p_x > 0, q_x > 0$ (it is thus a birth and death chain).

E1) Show that X is irreducible.

E2) Assume that, for $x > p$, $p_x = \frac{1}{2} + \frac{\lambda}{x}(1 + s(x))$ with $s(x) \to 0$ as $x \to \infty$. Show that, for $\alpha \in \mathbb{R}^*$,

$$\mathbf{E}_x(X_1^\alpha) = x^\alpha\Big(1 + \frac{2\alpha}{x^2}\Big(\lambda - \frac{1}{4} + \frac{\alpha}{4}\Big) + \frac{\varepsilon_\alpha(x)}{x^2}\Big) \tag{4.68}$$

where $\varepsilon_\alpha(x) \to_{x\to\infty} 0$.

E3) Let us assume $\lambda < \frac{1}{4}$. Show that X is recurrent (hint: one may prove that there exists $\alpha > 0$ such that $\phi(x) = x^\alpha$ satisfies the assumptions of (C) ...).

E4) Let us assume $\lambda > \frac{1}{4}$. Show that X is transient.

Problem 4.7 In this problem we study a Markov chain that can be used as a model to study a queue, i.e., a system in which, at every instant, a random number of customers arrives to a facility and forms a queue waiting to be served.

Let $p = (p_n)_{n\in\mathbb{N}}$ a probability on $\mathbb{N}$ such that $p_0 > 0$, $p_0 + p_1 < 1$. Let $(X_n)_{n\geq 0}$ be the canonical Markov chain on $\mathbb{N}$ associated to the transition matrix Q, where $Q(0, k) = p_k$ and, for $n \geq 1$,

$$Q(n, k) = \begin{cases} p_{k-n+1} & \text{for } k \geq n - 1 \\ 0 & \text{if } k < n - 1. \end{cases}$$

a) Show that, for every $k > 0$, there exists $k' \geq k$ such that $0 \rightsquigarrow k'$. Show that the chain is irreducible.

b) Let us set $\varphi(x) = \sum_{k\geq 0} p_k x^k$ and $m = \sum_{k\geq 1} kp_k$.

b1) Show that the equation $\varphi(x) = x$ has a unique root $\alpha \in]0, 1[$ if and only if $m > 1$.

b2) Which are the values of $u \in]0, 1[$ such that the function $f(n) = u^n$ satisfies the relation $f(n) = Qf(n)$ for every $n > 0$?

c) Let us set $\tau_k = \inf\{n \geq 0: X_n = k\}$, the passage time in k.

c1) Show that, for every $k \geq 1, r \geq 0, n \geq 0$, $\mathbf{P}_{k+r}(\tau_k = n) = \mathbf{P}_r(\tau_0 = n)$ and deduce that $\mathbf{P}_{k+r}(\tau_k < +\infty) = \mathbf{P}_r(\tau_0 < +\infty)$ and $\mathbf{E}_{k+r}(\tau_k) = \mathbf{E}_r(\tau_0)$.

c2) Show that, $\mathbf{P}_{k+r}$-a.s., $\tau_0 = \tau_r + \tau_0 \circ \theta_{\tau_r}$ on $\{\tau_0 < +\infty\}$.

c3) Deduce that $\mathbf{P}_{k+r}(\tau_0 < +\infty) = \mathbf{P}_k(\tau_0 < +\infty)\mathbf{P}_r(\tau_0 < +\infty)$ and thus $\mathbf{P}_k(\tau_0 < +\infty) = \mathbf{P}_1(\tau_0 < +\infty)^k$.

c4) Let us set $f(k) = \mathbf{P}_k(\tau_0 < +\infty)$. Show that f satisfies an equation that allows, using (b), the computation of its value.

c5) Give conditions of recurrence and of transience for this chain.

d) Let us assume $\mathbf{P}_k(\tau_0 < +\infty) = 1$ for every $k \in \mathbb{N}$.

d1) Show that $\mathbf{E}_k(\tau_0) = k\mathbf{E}_1(\tau_0)$.

d2) Let us set $u(k) = \mathbf{E}_k(\tau_0)$. Show that u satisfies an equation that allows, using (d1), the computation of its value.

d3) Compute $\mathbf{E}_0(\sigma_0)$ where σ_0 is the return time in 0 (•2.5). Give conditions for the positive recurrence of this chain.

Problem 4.8 (Foster criterion) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical irreducible Markov chain on a countable set E with transition matrix P. In this problem we prove that if there exist $\alpha > 0$, a *finite* part F of E and a function $h: E \to \mathbb{R}^+$ such that

$$\begin{aligned} Ph(x) &< +\infty \text{ for every } x \in E, \\ Ph(x) &\leq h(x) - \alpha \text{ for every } x \in F^c, \end{aligned} \tag{4.69}$$

then the chain is positive recurrent. We set, for $F \subset E$,

$$\tau_F = \inf\{n \geq 0: X_n \in F\}, \qquad \sigma_F = \inf\{n \geq 1: X_n \in F\},$$

respectively, the hitting and return times in F. We shall write τ_x and σ_x if $F = \{x\}$.

A) Let us assume that there exists a *finite* part F of E such that $\mathbf{E}_x(\sigma_F) < +\infty$ for every $x \in F$. Let us set

$$\rho_0 = 0, \quad \rho_1 = \sigma_F, \dots, \rho_n = \rho_{n-1} + \sigma_F \circ \theta_{\rho_{n-1}}, \dots .$$

$(\rho_n)_{n\geq 1}$ is the sequence of the subsequent return times of the chain in F.

A1) Show that, for every $x \in F$ and $n \geq 0$, $\mathbf{P}_x(\rho_n < +\infty) = 1$.

Let us set $Y_n = X_{\rho_n}$.

A2) Show that $Y = (\Omega, \mathscr{F}, (\mathscr{F}_{\rho_n})_{n\geq 0}, (Y_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in F})$ is a Markov chain on F with transition matrix $\bar{P}(x, y) = \mathbf{P}_x(X_{\sigma_F} = y)$.

A3) Show that Y is irreducible positive recurrent.

For $x \in F$ define $\bar{\sigma}_x = \inf\{n \geq 1; Y_n = x\}$.

A4) Show that, for $x \in F$,

$$\sigma_x = \sum_{k=0}^{\infty} 1_{\{k<\bar{\sigma}_x\}} \sigma_F \circ \theta_{\rho_k}$$

and that $\{k < \bar{\sigma}_x\} \in \mathscr{F}_{\rho_k}$.

A5) Show that, for every $x \in F$,

$$\mathbf{E}_x(\sigma_x) \leq \max_{y\in F} \mathbf{E}_y(\sigma_F) \cdot \mathbf{E}_x(\bar{\sigma}_x)$$

and deduce that X is positive recurrent.

B) Assume that (4.69) is satisfied for a finite part F of E and set $Z_n = h(X_n) + n\alpha$.

B1) Show that, for every $x \in F^c$, $(Z_{n\wedge\tau_F})_{n\geq 0}$ is a $((\mathscr{F}_n)_{n\geq 0}, \mathbf{P}_x)$-supermartingale and deduce that $\mathbf{E}_x(\tau_F) \leq \alpha^{-1} h(x)$ for every $x \in E$.

B2) Show that $\mathbf{E}_x(\sigma_F) < +\infty$ for every $x \in F$ and deduce that X is positive recurrent.

C) Let $(Y_n)_{n\geq 1}, (Z_n)_{n\geq 1}$ be two sequences of integrable r.v.'s with values in $\mathbb{N}$, each formed by r.v.'s having the same law. Let us assume that the r.v.'s of the family $(Y_n, Z_n)_{n\geq 1}$ are independent. Assume that there exists $K \in \mathbb{N}^*$ such that $\mathbf{P}(Z_1 > K) = 0$ and $\mathbf{P}(Y_1 \geq K) > 0$, that $\mathbf{P}(Z_1 = 1) > 0$, $\mathbf{P}(Y_1 = 0) > 0$ and $\mathbf{E}(Y_1) <$

$\mathbf{E}(Z_1)$. We set $\mathscr{F}_n = \sigma(Y_i, Z_i, i \le n)$ and

$$X_0 = i \in \mathbb{N}, \quad X_{n+1} = (X_n - Z_{n+1})^+ + Y_{n+1}.$$

C1) Show that $(X_n)_{n\ge 0}$ is a Markov chain with respect to the filtration $(\mathscr{F}_n)_{n\ge 0}$ and that Foster condition (4.69) is satisfied for $F = \{0, \dots, K\}$ and $h(x) = x$.

C2) Show that $(X_n)_{n\ge 0}$ is irreducible and deduce that $(X_n)_{n\ge 0}$ is positive recurrent.

♦ Remark that the chain described in (C) can model a queue: Z_n represents the number of customers served during the n-th time interval and Y_n the number of customers arriving in the queue at the end of the same time interval. The reader can compare the result of (C) with Problem 4.7, where only the values 0 or 1 were admissible for Z_n.

Problem 4.9 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge 0}, (X_n)_{n\ge 0}, (\mathbf{P}_x)_{x\in E})$ be a canonical Markov chain on the countable set E with transition matrix Q; we assume X irreducible, positive recurrent, aperiodic and denote by π its invariant probability. We now give a proof of the convergence Theorem 4.29, i.e., for every $x, y \in E$, $Q^n(x, y) \to \pi(y)$ as $n \to \infty$. This will allow us to emphasize an important technique of the theory of processes: the *coupling*.

A) Let $z \in E$ and $I(z) = \{n \ge 1, Q^n(z, z) > 0\}$ be as in •4.31.

A1) Show that there exists $a_1, \dots, a_r \in I(z)$ and $\alpha_1, \dots, \alpha_r \in \mathbb{Z}$ such that $\sum_{k=1}^r \alpha_k a_k = 1$.

A2) Show that if $m, n \in \mathbb{N}$ and $a, b \in I(z)$, then $na + mb \in I(z)$. Let $n_1 = \sum_{k=1}^r |\alpha_k| a_k$. Show that $n_1 \in I(z)$ and $n_1 + 1 \in I(z)$. Deduce that $Q^n(z, z) > 0$ for every $n \ge n_1^2$.

A3) Show that, for every $x, y \in E$, there exists N (depending on x, y) such that $Q^n(x, y) > 0$ for every $n \ge N$.

B) If α and β are probabilities on E, by $\alpha \otimes \beta$ we denote the product probability on $\bar E = E \times E$, i.e., $\alpha \otimes \beta(\bar x) = \alpha \otimes \beta(x, x') = \alpha(x)\beta(x')$, for $x \in E, x' \in E'$ and $\bar x = (x, x')$. Let $X' = (\Omega', \mathscr{F}', (\mathscr{F}'_n)_{n\ge 0}, (X'_n)_{n\ge 0}, (\mathbf{P}'_x)_{x\in E})$ be another Markov chain associated to Q and independent of X. We have already seen (Exercise 4.42) that if we set

$$\bar\Omega = \Omega \times \Omega', \quad \bar{\mathscr{F}} = \mathscr{F} \otimes \mathscr{F}', \quad \bar X_n = (X_n, X'_n), \quad \bar{\mathbf{P}}_{(x,x')} = \mathbf{P}_x \otimes \mathbf{P}'_{x'}$$

then $\bar X = X \otimes X' = (\bar\Omega, \bar{\mathscr{F}}, (\bar X_n)_{n\ge 0}, (\bar{\mathbf{P}}_{(x,x')})_{(x,x')\in \bar E})$ is a Markov chain on $\bar E$ with transition matrix

$$\bar Q((x, x'), (y, y')) = Q(x, y) Q(x', y'). \tag{4.70}$$

Still in Exercise 4.42, we have seen that $\pi \otimes \pi$ is a stationary probability for $\bar Q$ and that

$$\bar Q^n[(x, x'), (y, y')] = Q^n(x, y) Q^n(x', y'). \tag{4.71}$$

Show that the chain $\bar X$ is irreducible aperiodic. Is it also positive recurrent?

C) Let us introduce the stopping time $\tau = \inf\{n \ge 0; X_n = X'_n\}$.

C1) Let $\bar{\mathbf{P}}_{\alpha\otimes\beta}$ be the law of the chain $\bar X$ with starting distribution $\alpha \otimes \beta$. Show that, for every α, β, $\bar{\mathbf{P}}_{\alpha\otimes\beta}(\tau < +\infty) = 1$.

C2) Show that, for every $n \ge 0$, $y \in E$ and α, β,

$$\bar{\mathbf{P}}_{\alpha\otimes\beta}(X_n = y, \tau \le n) = \bar{\mathbf{P}}_{\alpha\otimes\beta}(X'_n = y, \tau \le n).$$

C3) Prove the relations

$$\mathbf{P}_\alpha(X_n = y) \le \mathbf{P}_\beta(X'_n = y) + \bar{\mathbf{P}}_{\alpha\otimes\beta}(\tau > n),$$
$$\mathbf{P}_\beta(X'_n = y) \le \mathbf{P}_\alpha(X_n = y) + \bar{\mathbf{P}}_{\alpha\otimes\beta}(\tau > n)$$

and deduce that

$$\left|\mathbf{P}_\alpha(X_n = y) - \mathbf{P}_\beta(X'_n = y)\right| \underset{n\to\infty}{\to} 0.$$

C4) Deduce, by a suitable choice of α and β, that $Q^n(x, y) \to_{n\to\infty} \pi(y)$ for every $x \in E$.

Problem 4.10 Let p be a probability on $\mathbb{Z}$ such that $\sum_{x\in\mathbb{Z}} |x|\, p(x) < +\infty$. Let us assume that every number of $\mathbb{Z}$ can be written as a finite sum of numbers $z_1, \dots, z_k \in \mathbb{Z}$ such that $p(z_i) > 0$, $1 \le i \le k$. We set $\mu = \sum_{x\in\mathbb{Z}} x\, p(x)$ and $p_n = p^{*n}$ (the convolution of p with itself n times). Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\ge0}, (X_n)_{n\ge0}, (\mathbf{P}_x)_{x\in\mathbb{Z}})$ be the canonical chain with transition matrix $P(x, y) = p(y - x)$.

1) Show that, for every $x \in \mathbb{Z}$, the r.v.'s $Y_n = X_n - X_{n-1}$ are independent and have the same law under $\mathbf{P}_x$.

2) Express the potential matrix of X, U, with the help of the probabilities p_n and deduce that $U(x, y) = U(0, y - x)$.

3) Show that $U(0, x) \le U(0, 0)$ for every $x \in \mathbb{Z}$.

4) Show that the chain is irreducible.

5) Show, using (1), that, $\mathbf{P}_x$-a.s. for every $x \in \mathbb{Z}$, $(X_n)_{n\ge0}$ tends to $+\infty$ if $\mu > 0$ and to $-\infty$ if $\mu < 0$. Deduce that if $\mu \ne 0$ the chain is transient.

6) Show that, if the chain is transient, the function

$$\phi(k) = \frac{1}{k}\sum_{n=0}^{\infty} \mathbf{P}_0(|X_n| \le k)$$

is bounded.

7) Assume $\mu = 0$. Show that for every $\varepsilon > 0$ there exists an integer N_ε such that, for sufficiently large k,

$$\phi(k) \ge \frac{1}{2}\Big(\frac{1}{\varepsilon} - \frac{N_\varepsilon}{k}\Big).$$

Deduce from this that the chain is recurrent.

Problem 4.11 (Recurrence and transience of the simple symmetric random walk on $\mathbb{Z}^d$) On $\mathbb{Z}^d$ let us consider the probability μ defined by $\mu(x) = (2d)^{-1}$ for each of the $2d$ vectors $x \in \mathbb{Z}^d$ of the form

$$x = (0, \dots, 0, \pm1, 0, \dots, 0)$$

(i.e., the $2d$ vectors of $\mathbb{Z}^d$ nearest to 0 along the axes) and $\mu(z) = 0$ for all other $z \in \mathbb{Z}^d$.

A) Let $(Z_n)_{n\ge1}$ be a sequence of independent identically distributed r.v.'s with values in $\mathbb{Z}^d$ and law μ. We set $S_0 = 0$, $S_n = X_1 + \dots + X_n$. Show that $(S_n)_{n\ge0}$ is a Markov chain with transition matrix

$$P(i, j) = \mu(j - i) \tag{4.72}$$

and starting law concentrated at 0. The Markov chain associated to this transition matrix is called the *simple symmetric random walk.* The aim of this problem is to investigate the recurrence of this chain as d varies. We have already seen in Exercise 4.16 (the case $p = q = \frac{1}{2}$) that if $d = 1$ it is recurrent.

B) Let m be the counting measure on $\mathbb{Z}^d$ and $f: \mathbb{Z}^d \to \mathbb{R}$ a function in $L^1(m)$, i.e., such that $\sum_{x\in\mathbb{Z}^d} |f(x)| < +\infty$. Let us denote by $\hat{f}$ its Fourier transform, defined by

$$\hat{f}(\theta) = \sum_{x\in\mathbb{Z}^d} f(x)\, e^{i\langle\theta,x\rangle} \qquad \theta \in \mathbb{R}^d.$$

B1) Show that if $f, g \in L^1(m)$ and h is their convolution:

$$h(x) = \sum_{y\in\mathbb{Z}^d} f(y)g(x-y),$$

then $\hat{h}(\theta) = \hat{f}(\theta)\hat{g}(\theta)$, $\theta \in \mathbb{R}^d$.

B2) Let $Q_d =]-\pi, \pi]^d \subset \mathbb{R}^d$. Show that, if $f \in L^1(m)$, $\hat{f}$ is bounded; prove the inversion formula

$$f(x) = \frac{1}{(2\pi)^d} \int_{Q_d} \hat{f}(\theta)\, e^{-i\langle\theta,x\rangle}\, d\theta. \tag{4.73}$$

B3) Compute the Fourier transform $\hat{\mu}$ of the function $x \to \mu(x)$, μ being defined above. Show that $\hat{\mu}$ is a real function such that $\hat{\mu}(\theta) < 1$ for every $\theta \in Q_d$, $\theta \neq 0$ and that $1 - \mu(\theta) \sim |\theta|^2/(2d)$ as $\theta \to 0$.

C) Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathrm{P}_x)_{x\in\mathbb{Z}^d})$ be the canonical Markov chain with transition matrix P defined in (4.72).

C1) Let us define, for $|\lambda| < 1$,

$$U_\lambda(x, y) = \sum_{n=0}^{\infty} \lambda^n P^n(x, y).$$

Show that $\lim_{\lambda\uparrow 1} U_\lambda(x, y) = U(x, y)$ (U is the potential matrix of P).

C2) Let us write $u_\lambda(y) = U_\lambda(0, y)$. Show that $u_\lambda \in L^1(m)$ and that

$$\hat{u}_\lambda(\theta) = \frac{1}{1 - \lambda\hat{\mu}(\theta)}.$$

C3) Compute $U(0, 0)$ and show that X is recurrent if and only if $d \leq 2$.

Problem 4.12 Let P be a transition matrix on $E = \{1, \dots, d\}$. The aim of this problem is to investigate the relationships between certain properties of P (irreducibility, aperiodicity) and the spectrum of P. Let $M(d)$ be the vector space of the line vectors of $\mathbb{R}^d$. Every probability on E can be identified to an element of $M(d)$ whose components are positive and have sum equal to 1. If u is an endomorphism (resp. A a matrix), we note $\Pi_u(\lambda)$ (resp. $\Pi_A(\lambda)$) the characteristic polynomial of u (resp. A). Recall that $\Pi_P(\lambda) = \Pi_{{}^tP}(\lambda)$.

A) Show that all the eigenvalues of P have a modulus that is smaller than or equal to 1.

B1) Let π be an invariant probability. We consider the two vector sub-spaces of $M(d)$

$$M_0 = \{x \in M(d); x_1 + \dots + x_d = 0\}, \quad M_1 = \{t\pi; t \in \mathbb{R}\}.$$

Show that $M(d) = M_0 \oplus M_1$ and that M_0 and M_1 are invariant for the linear application $u: x \to xP$. Remark that the matrix associated to u in the canonical basis of $M(d)$ is ${}^tP$.

B2) Let us denote by u_0 and u_1 the restrictions of u to M_0 and M_1, respectively. Show that $\Pi_u(\lambda) = (1-\lambda)\Pi_{u_0}(\lambda)$ and that, if $x = x_0 + x_1$, $x_0 \in M_0$, $x_1 \in M_1$, then $u^n(x) = u_0^n(x_0) + x_1$.

B3) Assume P irreducible. Show that the eigenspace associated to the eigenvalue 1 has dimension 1.

C1) Let u be an endomorphism of $\mathbb{R}^d$. Show that all eigenvalues of u have modulus < 1 if and only if, for every $x \in \mathbb{R}^d$, $u^n(x) \to 0$ as $n \to \infty$ (one can use the fact that every $\bar{u} \in \mathscr{L}(\mathbb{C}^d, \mathbb{C}^d)$ can be written $\bar{u} = v + w$ where v is diagonalizable, w is nilpotent, $vw = wv$ and $\Pi_{\bar{u}}(\lambda) = \Pi_v(\lambda)$).

C2) Assume P irreducible. Show that the following are equivalent:

a) P is aperiodic.

b) For every starting law μ, $\mu P^n \to \pi$ for $n \to \infty$.

c) Every eigenvalue of P other than 1 has a modulus that is < 1.

Problem 4.13 (Random walk on the vertices of a polygon) Let us consider the Markov chain on the set $E = \{1, \dots, N\}$ represented by the vertices of a polygon, associated to the transition matrix described by the Figure 4.6, where the three numbers p, r, q are positive and their sum is equal to 1.

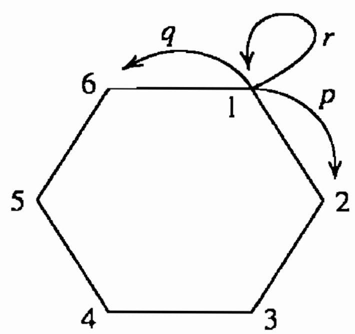

Figure 4.6

a) Write down the transition matrix P for $N = 6$.

b) Show that, if one at least among the numbers p, q is strictly positive, then the chain is irreducible. What is its stationary distribution? Show that if, moreover, $r > 0$, then the chain is also aperiodic.

c) We assume from now on $p > 0$, $q > 0$.

c1) Show that, if $r = 0$ and the number of vertices N is even, then the chain is periodic with period equal to 2.

c2) For $m = 0, 1, \dots, N-1$, let us consider the vector $v^{(m)}$ whose components are $v_h^{(m)} = e^{2i\pi mh/N}$, $h \in E$. Show that the vectors $v^{(m)}$, $m = 1, \dots, N-1$ are orthogonal in $\mathbb{C}^N$. Compute $Pv^{(m)}$. What is the spectrum of the matrix P? What can be said about the modulus of its eigenvalues?

c3) Show that, if the number of vertices N is odd, then the chain is aperiodic (regardless the value of r).

Problem 4.14 In this problem we use Dynkin's criterion (Exercise 4.35) and the product chains (Exercise 4.42) and investigate a problem of telecommunications.

A data transmission channel can receive messages from m independent sources. Each of these sources can be active (1) or inactive (0). If they are monitored at fixed time intervals, each of them switches from state 0 to state 1 with probability α and from 1 to 0 with probability β and remains in 0 and 1 with probabilities $1-\alpha$ and $1-\beta$, respectively. We assume that the behaviours of the sources are independent and that $0<\alpha, \beta<1$.

The behaviour of every source is modeled by a Markov chain with state space $\{0,1\}$ and transition matrix

$$P' = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}. \tag{4.74}$$

Let us note X' the canonical Markov chain on $\{0,1\}$ with transition matrix P'. Recall (Exercise 4.26) that the unique stationary distribution of this chain is

$$\mu_0 = \frac{\beta}{\alpha+\beta}, \qquad \mu_1 = \frac{\alpha}{\alpha+\beta} \tag{4.75}$$

(i.e., a Bernoulli law with parameter $\alpha(\alpha+\beta)^{-1}$).

a) Let us consider the set $E = \{0,1\}^m$ of all binary vectors with length m and, on E, the product chain $X = X' \otimes \ldots \otimes X'$ m times, as described in Exercise 4.42 (at least for the case $m=2$). Find a stationary distribution for this chain (use (3) of Exercise 4.42). Is it unique?

Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in E})$ be the product chain introduced in (a). It is clear that the binary vector X_n models the behaviour of the m independent sources at time n and that, in this model, the number of active sources is given by the sum Y_n of the coordinates of X_n.

b) Show that, as $n\to\infty$, $(Y_n)_{n\geq 0}$ converges in distribution and determine its limiting law. What is the mean number of active sources for n large?

c) Let us note $F=\{0,\ldots,m\}$ and $\psi\colon E\to F$ the application defined by $\psi(x) = x_1+\ldots+x_m$. We denote by P the transition matrix of the product chain introduced in (a). At every permutation σ of $\{1,\ldots,m\}$ let us associate the transformation $E\to E$ defined by $x\to x_\sigma = (x_{\sigma_1},\ldots,x_{\sigma_m})$.

c1) Show that a function g on E is constant on the sets $\psi^{-1}(i)$, $i\in F$ if and only if it is invariant with respect to every permutation σ, i.e., if and only if

$$g(x) = g(x_\sigma) \tag{4.76}$$

for every permutation σ. Show that, if P denotes the transition matrix of the product chain of (a), then $P(x,y) = P(x_\sigma, y_\sigma)$ for every permutation σ.

c2) Show that Dynkin's criterion of Exercise 4.35 is satisfied and that, if $Y_n = \psi(X_n)$, then $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (Y_n)_{n\geq 0}, \mathbf{P}_x)$ is a Markov chain on F with initial state $\psi(x)$. What is its transition matrix Q? Is it irreducible? Aperiodic? What is its stationary law?

Problem 4.15 (Markov chains observed at random times) We consider a Markov chain $(X_n)_{n\geq 0}$ on a countable set E, with starting law ν and transition matrix Q and let $(Z_n)_{n\geq 1}$ be a sequence of r.v.'s with values in $\mathbb{N}$, independent, having a common law μ such that $\mu(0) < 1$. We note μ^{*n} the convolution of μ with itself n times. Let us assume that the processes $(X_n)_{n\geq 0}$ and $(Z_n)_{n\geq 1}$ are independent. Let us set $T_0 = 0$ and $T_n = \sum_{i=1}^{n} Z_i$.

1a) Show that $(T_n)_{n\geq 0}$ is a Markov chain on $\mathbb{N}$; which is its transition matrix R? Show that its potential matrix V is of the form $V(m,n) = \gamma(n-m)$ where $\gamma(n) = 0$ if $n < 0$ and $\gamma(n) = 1_{\{0\}}(n) + \sum_{k\geq 1} \mu^{*k}(n)$ for $n \geq 1$. What is the value of $\gamma(0)$? Show that $\gamma(n) \leq \gamma(0)$ for every $n \geq 0$.

1b) For $0 \leq n_1 \leq \ldots \leq n_k$, express $\mathbf{P}(T_1 = n_1, \ldots, T_k = n_k)$ with the help of μ.

2) Let us set, for every $n \geq 0$, $Y_n = X_{T_n}$.

2a) Show that $S(x,y) = \sum_{n\geq 0} \mu(n) Q^n(x,y)$ is a transition matrix on E.

2b) Show that $\mathbf{P}(Y_0 = y_0, Y_1 = y_1) = \nu(y_0) S(y_0, y_1)$.

2c) For every $k \geq 0$, express $\mathbf{P}(Y_0 = y_0, \ldots, Y_k = y_k)$ with the help of ν and S. Deduce that $(Y_n)_{n\geq 0}$ is a Markov chain with transition matrix S. What is its starting distribution?

3a) Show that $S^2(x,y) = \sum_{n\geq 0} \mu^{*2}(n) Q^n(x,y)$.

3b) Show that the potential matrix W of $(Y_n)_{n\geq 0}$ is given by

$$W(x,y) = \sum_{n\geq 0} \gamma(n) Q^n(x,y).$$

3c) Deduce that every state transient for Q is also transient for S.

4) Let λ be an invariant measure for Q. Show that λ is also invariant for S; hence that, if Q is positive recurrent, the same holds true for S.

5) Let us assume that $E = \mathbb{Z}$, $Q(i, i+1) = Q(i, i-1) = \frac{1}{2}$, $i \in \mathbb{Z}$ and that the generating function of μ is $\phi(s) = 1 - (1-s)^{\rho}$, $0 < \rho < 1$.

5a) Compute $\mathbf{P}(X_n = 0 \mid X_0 = 0)$ and deduce that $(X_n)_{n\geq 0}$ is a recurrent irreducible chain.

5b) Show that $\sum_{n\geq 0} \gamma(n) s^n = (1-s)^{-\rho}$. Deduce the value of $\gamma(n)$.

5c) Assume $0 < \rho < \frac{1}{2}$. Show that all the states of the chain $(Y_n)_{n\geq 0}$ are transient. (Recall Stirling's formula: $\rho(\rho+1)\ldots(\rho+n) \sim \Gamma(\rho)^{-1}\sqrt{2\pi n}\, n^{n+\rho} e^{-n}$).

Problem 4.16 Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbf{P}_x)_{x\in\mathbb{Z}})$ be the canonical Markov chain on $\mathbb{Z}$ associated to the transition matrix Q given by

$$Q(x, x+1) = p,\ Q(x,x) = r,\ Q(x, x-1) = q$$

where $p + q + r = 1$, $p > 0$, $q > 0$, $r \geq 0$. Let us set $\rho = \frac{q}{p}$.

Let us fix $a, b \in \mathbb{Z}$, $a < b - 1$ and denote by $\tau = \inf\{n \geq 0; X_n \notin]a, b[\}$ the hitting time in $\{a+1, a+2, \ldots, b-1\}^c$. Let $\Phi_{a,b}$ be the set of the applications from $\{a, a+1, \ldots, b\}$ into $\mathbb{R}$. The aim of this problem is to show that the r.v. τ has moments of all orders and to give a computational approach.

A) In this section we prove that, if $g \in \Phi_{a,b}$, the system

$$\begin{cases} (I - Q)u(x) = g(x) & a < x < b \\ u(a) = \alpha,\ u(b) = \beta \end{cases} \tag{4.77}$$

has a unique solution for every $\alpha, \beta \in \mathbb{R}$.

A1) Prove that there exists a unique $\phi \in \Phi_{a,b}$ such that $\phi(a) = 0$, $\phi(a+1) = 1$, $(I - Q)\phi(x) = 0$ for every $a < x < b$, and that such a ϕ is strictly increasing.

A2) Prove that there exists a unique $\psi \in \Phi_{a,b}$ such that $\psi(a) = \psi(a+1) = 0$ and $(I - Q)\psi(x) = g(x)$ for every $a < x < b$.

A3) Show that, if $u \in \Phi_{a,b}$ is such that $u(a) = u(b) = 0$ and $(I - Q)u(x) = 0$ for every $a < x < b$, then $u \equiv 0$.

A4) Let $\alpha, \beta \in \mathbb{R}$ and $g \in \Phi_{a,b}$. Prove that there exists a unique function $u \in \Phi_{a,b}$ satisfying (4.77).

B1) Let f be a real bounded function on $\mathbb{Z}$ and define $M_0 = 0$ and $M_n = \sum_{k=0}^{n-1}[f(X_{k+1}) - Qf(X_k)]$. Show that $(M_n)_n$ is a martingale such that

$$f(X_n) - f(X_0) = M_n - \sum_{k=0}^{n-1}(I - Q)f(X_k).$$

B2) Let $u_1 \in \Phi_{a,b}$ be the solution of

$$\begin{cases} (I - Q)u_1(x) = 1, & \text{as } a < x < b \\ u_1(a) = u_1(b) = 0. \end{cases}$$

Show that, for every $x \in]a, b[$,

$$\mathbf{E}_x[u_1(X_{n\wedge\tau})] - u_1(x) = -\mathbf{E}_x(n \wedge \tau)$$

(apply (B1) to $f = u_1$). Deduce that $\mathbf{E}_x(\tau) < +\infty$ and that $\mathbf{E}_x(\tau) = u_1(x)$, $x \in]a, b[$.

B3) Let u be the solution of (4.77). Show that, for every $x \in]a, b[$,

$$u(x) = \alpha\mathbf{P}_x(X_\tau = a) + \beta\mathbf{P}_x(X_\tau = b) + \mathbf{E}_x\Big[\sum_{k=0}^{\tau-1} g(X_k)\Big].$$

C) (Computation of the moments). Let us assume for the time being that for every $s > 0$ and every x, $\mathbf{E}_x(\tau^s) < +\infty$. Let us set $u_s(x) = \mathbf{E}_x(\tau^s)$.

C1) Show that, for every $x \in]a, b[$, $\tau = 1 + \tau \circ \theta_1$ $\mathbf{P}_x$-a.s.

C2) Prove that u_2 is a solution of (4.77) for some α, β and g to be determined as functions of u_1.

C3) Prove that u_3 is a solution of (4.77) for some α, β and g to be determined as functions of u_1 and u_2.

D) (Existence of exponential moments).

D1) Let $m = b - a$. Prove that there exists $\gamma > 0$ such that, for every x, $\mathbf{P}_x(\tau \le m) \ge \gamma$. Deduce that, for every $k \in \mathbb{N}$, $\mathbf{P}_x(\tau > km) \le (1 - \gamma)^k$.

D2) Prove that there exist constants K and $\varepsilon > 0$ such that, for every $t > 0$, $\mathbf{P}_x(\tau > t) \le K\,e^{-\varepsilon t}$. Prove that, for every $\lambda < \varepsilon$, $\mathbf{E}_x(e^{\lambda\tau}) < +\infty$ and that, for every $s > 0$, $\mathbf{E}_x(\tau^s) < +\infty$.

Solutions

E4.1 a) Let us set $\mathscr{F}_n = \sigma(X_0, \dots, X_n)$ and let $B \subset E$. Since Y_{n+1} is independent of $\mathscr{F}_n$ and X_n is $\mathscr{F}_n$-measurable, by Lemma 1.2 one has a.s.

$$\mathbf{P}(X_{n+1} \in B \mid \mathscr{F}_n) = \mathbf{E}(1_{\{X_{n+1} \in B\}} \mid \mathscr{F}_n) = \mathbf{E}(1_{\{\Phi(X_n, Y_{n+1}) \in B\}} \mid \mathscr{F}_n) = \Psi_B(X_n)$$

where

$$\Psi_B(i) = \mathbf{E}(1_{\{\Phi(i,Y_{n+1})\in B\}}) = \mathbf{P}(\Phi(i, Y_{n+1}) \in B).$$

This shows simultaneously that $(X_n)_{n\geq 0}$ is a Markov chain and, the r.v.'s Y_n being identically distributed, that its transition matrix is given by

$$P(i, j) = \mathbf{P}(\Phi(i, Y_1) = j).$$

b) By (a), $(X_n)_n$ is a Markov chain with transition matrix

$$Q(i, j) = \mathbf{P}(g(i, U_1) = j) = \mathbf{P}(U_1 \in [s_j^{(i)}, s_{j+1}^{(i)}[) = s_{j+1}^{(i)} - s_j^{(i)} = P(i, j)$$

and initial law equal to the law of $\psi(U_0)$. Similarly

$$\mathbf{P}(\psi(U_0) = j) = \mathbf{P}(U_0 \in [t_j, t_{j+1}[) = t_{j+1} - t_j = \mu(j).$$

E4.2 Assume first that Y is a Markov chain and let us try to identify its transition matrix. By definition, it is sufficient to compute $\mathbf{P}_x(Y_1 = y) = \mathbf{P}_x(X_{1\wedge\tau} = y)$ for $(x, y) \in E^2$. Two possible situations can arise: if $x \in F$, $\tau = 0$ $\mathbf{P}_x$-a.s. and $\mathbf{P}_x(Y_1 = y) = 1_{\{x=y\}}$. Conversely, if $x \notin F$, $\tau \geq 1$ $\mathbf{P}_x$-a.s. and $\mathbf{P}_x(Y_1 = y) = \mathbf{P}_x(X_1 = y) = Q(x, y)$. The transition matrix of $(Y_n)_{n\geq 0}$ must therefore be

$$\tilde{Q}(x, y) = 1_{\{x\in F\}} 1_{\{x=y\}} + 1_{\{x\notin F\}} Q(x, y).$$

Let us show now that Y is indeed a Markov chain with transition matrix $\tilde{Q}$. One has

$$\mathbf{P}^{\mathscr{F}_n}(Y_{n+1} = y) = \mathbf{P}^{\mathscr{F}_n}(Y_{n+1} = y, Y_n \in F) + \mathbf{P}^{\mathscr{F}_n}(Y_{n+1} = y, Y_n \notin F) \quad \text{a.s.}$$

The following equalities of events hold: $\{Y_n \in F, Y_{n+1} = y\} = \{Y_n \in F, Y_n = y\}$, $\{Y_n \in F^c\} \subset \{X_n = Y_n\}$ and $\{Y_n \in F^c, Y_{n+1} = y\} = \{Y_n \in F^c, X_{n+1} = y\}$. Therefore,

$$\begin{aligned}\mathbf{P}^{\mathscr{F}_n}(Y_{n+1} = y) &= 1_F(Y_n)1_y(Y_n) + 1_{F^c}(Y_n)\mathbf{P}^{\mathscr{F}_n}(X_{n+1} = y) = \\ &= 1_F(Y_n)1_y(Y_n) + 1_{F^c}(Y_n)Q(X_n, y) = \tilde{Q}(Y_n, y).\end{aligned}$$

E4.3 a) Let us remark that $\{\tau_F \geq n+1\} = \{\tau_F \leq n\}^c \in \mathscr{F}_n$ and that the r.v.

$$1_{\{\tau_F<n+1\}} f(X_{\tau_F}) = \sum_{k=0}^{n} 1_{\{\tau_F=k\}} f(X_k)$$

is $\mathscr{F}_n$-measurable. Then a.s.

$$\begin{aligned}&\mathbf{E}_x(f(X_{(n+1)\wedge\tau_F}) \mid \mathscr{F}_n) = \\ &= \mathbf{E}_x(1_{\{\tau_F<n+1\}} f(X_{\tau_F}) \mid \mathscr{F}_n) + 1_{\{\tau_F\geq n+1\}}\mathbf{E}_x(f(X_{n+1}) \mid \mathscr{F}_n) = \\ &= 1_{\{\tau_F<n+1\}} f(X_{\tau_F}) + 1_{\tau_F\geq n+1\}} Pf(X_n) \leq \\ &\leq 1_{\{\tau_F<n+1\}} f(X_{\tau_F}) + 1_{\{\tau_F\geq n+1\}} f(X_n) = f(X_{n\wedge\tau_F}).\end{aligned}$$

If f satisfies to $f = Pf$ on F, it is sufficient to repeat the previous calculation (the last inequality becoming an equality).

b) If w is a positive solution of (4.28), then $(w(X_{n\wedge\tau_F}))_{n\geq 0}$ is a positive martingale and

$$\begin{aligned}w(x) = \mathbf{E}_x(w(X_0)) = \mathbf{E}_x(w(X_{n\wedge\tau_F})) &\geq \mathbf{E}_x(1_{\{\tau_F\leq n\}} w(X_{n\wedge\tau_F})) = \\ &= \mathbf{E}_x(1_{\{\tau_F\leq n\}} g(X_{\tau_F})).\end{aligned}$$

As this inequality holds for every $n \geq 0$, taking the limit,

$$w(x) \geq \mathbf{E}_x(1_{\{\tau_F<+\infty\}} g(X_{\tau_F})) = P_F g(x).$$

E4.4 1) Obviously $b(x) = \mathbf{E}_x(X_1) = \sum_{y\in\mathbb{Z}} y\,Q(x,y)$ and

$$a(x) = \mathbf{E}_x[(X_1 - b(x))^2] = \sum_{y\in\mathbb{Z}} (y - b(x))^2 Q(x,y).$$

2) By the Markov property,

$$\mathbf{E}_x(X_{n+1}) = \mathbf{E}_x(X_1 \circ \theta_n) = \mathbf{E}_x[\mathbf{E}_{X_n}(X_1)] = \mathbf{E}_x[b(X_n)]$$

as well as

$$\mathbf{E}_x(X_{n+1}^2) = \mathbf{E}_x(X_1^2 \circ \theta_n) = \mathbf{E}_x(\mathbf{E}_{X_n}(X_1^2)) = \mathbf{E}_x(\mathrm{Var}_{X_n}(X_1) + \mathbf{E}_{X_n}(X_1)^2) = \\ = \mathbf{E}_x[a(X_n) + b(X_n)^2].$$

Therefore

$$\mathrm{Var}_x(X_{n+1}) = \mathbf{E}_x(X_{n+1}^2) - \mathbf{E}_x(X_{n+1})^2 = \mathbf{E}_x(X_{n+1}^2) - \{\mathbf{E}_x[b(X_n)]\}^2 = \\ = \mathbf{E}_x(a(X_n) + b(X_n)^2) - \big[\mathbf{E}_x(b(X_n)^2) - \mathrm{Var}_x(b(X_n))\big] = \\ = \mathrm{Var}_x(b(X_n)) + \mathbf{E}_x(a(X_n)).$$

E4.5 Let $f \in \mathscr{H}$. We look for a characterization of the operator A. The process $Z_n = f(X_n) - \sum_{k=0}^{n-1} Af(X_k)$ is a martingale if and only if, for every $n \geq 0$, $\mathbf{E}(Z_{n+1} - Z_n \mid \mathscr{F}_n) = 0$. Now, a.s.,

$$\mathbf{E}(Z_{n+1} - Z_n \mid \mathscr{F}_n) = \mathbf{E}(f(X_{n+1}) \mid \mathscr{F}_n) - f(X_n) - Af(X_n) = \\ = Pf(X_n) - f(X_n) - Af(X_n).$$

This is equal to 0 if we set $Af(x) = (P - I)f(x)$.

E4.6 1) One has, everything being positive,

$$\mathbf{E}_x\Big[\sum_{n\geq 0} f(X_n) \sum_{p\geq n} f(X_p)\Big] = \sum_{n\geq 0}\sum_{p\geq 0} \mathbf{E}_x[f(X_n) f(X_{n+p})] = \\ = \sum_{n\geq 0}\sum_{p\geq 0} \mathbf{E}_x[f(X_n) f(X_p)\circ\theta_n] = \sum_{n\geq 0}\sum_{p\geq 0} \mathbf{E}_x[f(X_n)\mathbf{E}_{X_n}(f(X_p))] = \\ = \sum_{n\geq 0} \mathbf{E}_x\Big[f(X_n)\mathbf{E}_{X_n}\Big(\sum_{p\geq 0} f(X_p)\Big)\Big] = \sum_{n\geq 0} \mathbf{E}_x[f(X_n) Uf(X_n)] = \\ = \mathbf{E}_x\Big[\sum_{n\geq 0} f(X_n) Uf(X_n)\Big] = U(f\cdot Uf)(x).$$

2) It holds

$$\mathbf{E}_x\Big[\Big(\sum_{n\geq 0} f(X_n)\Big)^2\Big] = \mathbf{E}_x\Big[\sum_{n\geq 0} f(X_n)^2 + 2\sum_{n\geq 0}\sum_{p>n} f(X_n) f(X_p)\Big] = \\ = 2\mathbf{E}_x\Big[\sum_{n\geq 0}\sum_{p\geq n} f(X_n) f(X_p)\Big] - \mathbf{E}_x\Big[\sum_{n\geq 0} f(X_n)^2\Big] = 2U(f\cdot Uf)(x) - U(f^2)(x).$$

3) One has

$$\mathbf{E}_x(N_y^2) = \mathbf{E}_x\Big[\Big(\sum_{n\geq 0} 1_{\{y\}}(X_n)\Big)^2\Big] = 2U(f\cdot Uf)(x) - U(f^2)(x)$$

for $f = 1_{\{y\}}$. But, with this choice of f, $Uf^2(x) = U1_{\{y\}}(x) = U(x,y)$ and

$$U(f\cdot Uf)(x) = \sum_{z\in E} U(x,z)(fUf)(z) = \sum_{z\in E} U(x,z)1_{\{y\}}(z)U(y,z) =$$
$$= U(x,y)U(y,y).$$

Therefore $\mathbf{E}_x(N_y^2) = U(x,y)(2U(y,y)-1)$.

E4.7 a) The state n is obviously recurrent. One easily checks that every state x, $0 \leq x \leq n-1$, leads to every other state, thanks to the transitivity of the relation $\rightsquigarrow$. The criterion •4.21 then allows to state that x, $0 \leq x \leq n-1$, is transient, as $x \rightsquigarrow n$ whereas $n \not\rightsquigarrow x$.

The function $v(x) = \mathbf{E}_x(\tau)$ is the smallest positive solution of $v = Pv + 1$ on $\{0, 1, \ldots, n-1\}$, $v(n) = 0$ (Corollary 4.15). Writing $q = 1-p$, one is thus led to the system

$$\begin{aligned}
v(n-1) &= 1 + qv(0)\\
v(n-2) &= 1 + qv(0) + pv(n-1)\\
&\ldots\\
v(x) &= 1 + qv(0) + pv(x+1)\\
&\ldots\\
v(0) &= 1 + qv(0) + pv(1).
\end{aligned}$$

Replacing the value of $v(n-1)$ given by the first equation into the second one, one gets

$$v(n-2) = 1 + p + q(1+p)v(0).$$

Replacing again this value into the third equation,

$$v(n-3) = 1 + p + p^2 + q(1+p+p^2)v(0)$$

and, by induction,

$$v(x) = 1 + p + \ldots + p^{n-x-1} + q(1 + p + \ldots + p^{n-x-1})v(0). \tag{4.78}$$

For $x = 0$,

$$v(0) = 1 + p + \ldots + p^{n-1} + q(1 + p + \ldots + p^{n-1})v(0) =$$
$$= \frac{1-p^n}{1-p} + q\frac{1-p^n}{1-p}v(0) = \frac{1-p^n}{1-p} + (1-p^n)v(0)$$

so that

$$v(0) = \frac{1-p^n}{p^n(1-p)}. \tag{4.79}$$

One replaces this value in (4.78) and obtains

$$\mathbf{E}_x(\tau) = v(x) = \frac{1-p^{n-x}}{p^n(1-p)}.$$

b) One can model the successive tosses by a Markov chain on the set $\{0, 1, \dots, n\}$ where i corresponds to the fact that, in the $i+1$ last tosses, the result has been one heads followed by i successive tails. If $0 \le i \le n-1$, at each toss the Markov chain makes a transition from i to $i+1$ with probability p (if the toss gives tails again) and to 0 with probability $1-p$ (if, on the other hand, it gives heads). We assume moreover that the state n is absorbing. The event "one gets n consecutive tails" corresponds to the event $\{\tau < +\infty\}$ for this chain, where τ is the hitting time at n.

Observe that the chain is the same as in (a). Therefore it is immediate that $\mathbf{P}(\tau < +\infty) = 1$, whatever the initial distribution, as every state x other than n is transient and thus the set $\{0, 1, \dots, n-1\}$ is only visited a finite number of times.

The mean time to obtain n consecutive tails is therefore given by (4.79). If $p = \frac{1}{2}$ one obtains $\mathbf{E}_0(\tau) = v(0) = 2^{n+1} - 2$. For $n = 6$, for example, $2^7 - 2 = 126$ tosses are necessary in mean.

E4.8 a) It is immediate that $P(0,0) = 1$, $P(0,j) = 0$, $j \ge 1$ and, similarly, $P(N,N) = 1$. $P(N,j) = 0$, $j \le N-1$. This shows that the two states 0 and N are absorbing. The chain is therefore not irreducible and 0 and N are recurrent.

Let $j \in E$ be a state other than 0 and 1. Then $j \rightsquigarrow 0$ as $P(j,0) > 0$ and $0 \not\rightsquigarrow j$, as $P(0,0) = 1$. This implies (•4.21) that j is transient. In conclusion: 0 and N are recurrent and $1, \dots, N-1$ are transient.

b) The expectation of a binomial distribution is the product of its parameters and the Markov property easily gives

$$\mathbf{E}_i(X_{n+1} \mid \mathscr{F}_n) = \sum_{j \in E} P(X_n, j) j = N \cdot \frac{X_n}{N} = X_n$$

so that X is a martingale. As it is bounded, it converges a.s. and in L^1. In order to identify X_∞, it is sufficient to observe that, as the number of visits to a transient state is finite, X_∞ necessarily takes its values in $\{0, N\}$. To obtain its law, it is therefore sufficient to compute $\mathbf{P}_i(X_\infty = 0)$ and $\mathbf{P}_i(X_\infty = N)$. But, the martingale being regular,

$$i = \mathbf{E}_i(X_0) = \mathbf{E}_i(X_\infty) = N \cdot \mathbf{P}_i(X_\infty = N)$$

so that $\mathbf{P}_i(X_\infty = N) = \frac{i}{N}$ and $\mathbf{P}_i(X_\infty = 0) = 1 - \frac{i}{N}$.

E4.9 Observe that, E being finite and Q irreducible, the chain is recurrent irreducible and, for every $k \in E$, $\mathbf{P}_k(\tau < +\infty) = 1$.

1) For $i = 1, \dots, N-2$, $Qh(i) = (Cy)_i = y_i = h(i)$. Thanks to Exercise 4.3, $(h(X_{n \wedge \tau}))_{n \ge 0}$ is, for every $k \in E$, a $\mathbf{P}_k$-martingale, bounded and therefore regular. It converges to $h(X_\tau) = 0$, $\mathbf{P}_k$-a.s. Therefore, for every $k \in E \setminus F$, it holds $y_k = \mathbf{E}_k(h(X_0)) = \mathbf{E}_k(h(X_\tau)) = 0$ (Proposition 3.13). We have proved that $I - C$ is injective and therefore invertible.

2) The function $u(i) = \mathbf{P}_i(X_\tau = N-1)$ is a solution (Corollary 4.14) of

$$u(i) = Qu(i), \quad i \in E \setminus F, \qquad u(N-1) = 1, \quad u(N) = 0,$$

which can be written, with the proposed notation,

$$\bar{p}_{N-1} = C\bar{p}_{N-1} + De_1.$$

Therefore $\bar{p}_{N-1} = (I - C)^{-1} De_1$ and also $\bar{p}_N = (I - C)^{-1} De_2$.

♦ If $\mathbf{1} \in \mathbb{R}^{N-2}$ is the vector with components $1, \dots, 1$, the relation $1 = Q1$ implies $\mathbf{1} = C\mathbf{1} + D(e_1 + e_2)$, i.e., $\bar{p}_{N-1} + \bar{p}_N = 1$, which was obvious as $\mathbf{P}_k(\tau < +\infty) = 1$.

3) Recall that $\mathbf{P}_\mu(G) = \sum_{i \in E} \mu(i) \mathbf{P}_i(G)$ for every $G \in \mathscr{F}$. If $G = \{X_\tau = N-1\}$, then obviously $\mathbf{P}_{N-1}(G) = 1$ and $\mathbf{P}_N(G) = 0$ whereas the probabilities $\mathbf{P}_i(G)$, $i = 0, \dots, N-2$ are computed in (2). Thus, if μ is the uniform law,

$$p_{N-1} = N^{-1}(\langle \mathbf{1}, (I - C)^{-1} De_1 \rangle + 1).$$

Similarly $p_N = N^{-1}(\langle \mathbf{1}, (I - C)^{-1} De_2 \rangle + 1)$.

E4.10 a) Using criterion •4.21 and Proposition 4.19, one gets easily

• $1 \rightsquigarrow 7$ and $1 \rightsquigarrow 9$, $7 \rightsquigarrow 1$ and $7 \rightsquigarrow 9$, $9 \rightsquigarrow 7$. The states $\{1, 7, 9\}$ communicate; however they do not lead to any other state and therefore they form a closed class.

• $2 \leftrightsquigarrow 4$ and these two states do not lead to any other state; as above $\{2, 4\}$ is an irreducible closed class.

• $3 \rightsquigarrow 5$, $5 \rightsquigarrow 2$, but $2 \not\rightsquigarrow 3$ and $2 \not\rightsquigarrow 5$. Using criterion •4.21, 3 and 5 are transient.

• $6 \rightsquigarrow 1$ but $1 \not\rightsquigarrow 6$: 6 is also transient.

• $8 \rightsquigarrow 3$, $3 \rightsquigarrow 5$, $5 \rightsquigarrow 2$; therefore $8 \rightsquigarrow 2$ but $2 \not\rightsquigarrow 8$; 8 is transient.

• 10 only leads to itself: it is absorbing.

The set E may therefore be decomposed into three irreducible closed classes

$$\{1, 7, 9\} \quad \{2, 4\} \quad \{10\}$$

and the class $\{3, 5, 6, 8\}$ of the transient states.

b) By definition, $x \rightsquigarrow y$ if and only if there exists n such that $P^n(x, y) > 0$. As $P^n = P \dots P$ (n times) (•4.4) one has

$$P^n(x, y) = \sum_{h_1, \dots, h_{n-1} \in E} P(x, h_1) P(h_1, h_2) \dots P(h_{n-1}, y)$$

and therefore $P^n(x, y) > 0$ if and only if at least one term in the previous sum is > 0, i.e., if and only if (4.45) is satisfied.

If $x \rightsquigarrow y$, let m be the smallest nonnegative integer such that $P^m(x, y) > 0$. If $h_1, \dots, h_{m-1} \in E$ are such that

$$P(x, h_1) > 0, P(h_1, h_2) > 0, \dots, P(h_{m-1}, y) > 0,$$

the states $x = h_0, h_1, \dots, h_{m-1}, y = h_m$ are all different. Indeed if it was $h_i = h_j$ for $i < j$, then $x, h_1, \dots, h_{i-1}, h_j, \dots, y$ would be a path with positive probability and length $k < m$. This would imply $P^k(x, y) > 0$ and contradict the minimality assumption on m.

c) It has been seen in (b) that $x \rightsquigarrow y$ for the transition matrix P if and only if (4.45) is satisfied. But if (4.46) is satisfied, (4.45) remains true if one replaces P by Q and this implies that $x \rightsquigarrow y$ for Q. One may repeat the same argument if the weaker condition (4.47) is satisfied as, thanks to (b), one may suppose the states $x, h_1, \dots, h_{m-1}, y$ to be all different.

d) If P is irreducible, the last argument in (c) implies the irreducibility of Q. The two transition matrices

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

satisfy the condition (4.47), but P is aperiodic, thanks to the first criterion of •4.33, whereas it is easy to see that Q is not. If the stronger condition (4.46) is satisfied, one easily sees that, for $a \in E$,

$$P^n(a,a) > 0 \qquad \Rightarrow \qquad Q^n(a,a) > 0$$

and

$$\{n;\ P^n(a,a) > 0\} \subset \{n;\ Q^n(a,a) > 0\}.$$

Therefore, P being assumed to be aperiodic, the greatest common divisor of the leftmost set above is 1, which implies that the g.c.d. of $\{n;\ Q^n(a,a) > 0\}$ is also 1 and Q is also aperiodic.

E4.11 If $x, y \in E$, then x leads to x_0 (thanks to (b)), x_0 leads to y (thanks to (a)) and therefore x leads to y. It is clear that the chain is irreducible. It remains to prove that x_0 is recurrent, i.e.,

$$\mathbf{P}_{x_0}(\sigma_{x_0} < +\infty) = 1.$$

Now $\sigma_{x_0} = 1 + \tau_{x_0} \circ \theta_1$ and, by the Markov property (Theorem 4.5),

$$\mathbf{P}_{x_0}(\sigma_{x_0} < +\infty) = \mathbf{E}_{x_0}[1_{\{\tau_{x_0} < +\infty\}} \circ \theta_1] = \mathbf{E}_{x_0}[\underbrace{\mathbf{P}_{X_1}(\tau_{x_0} < +\infty)}_{=1}] = 1.$$

E4.12 a) If $x \in A$, $\tau_A = 0$ $\mathbf{P}_x$-a.s. For $x \in A^c$,

$$\begin{aligned}\mathbf{P}_x(\tau_A > (k+1)n) &\le \mathbf{P}_x(\tau_A > kn,\ X_{(k+1)n} \in A^c) = \\ &= \mathbf{E}_x[1_{\{\tau_A > kn\}} \underbrace{\mathbf{P}_{X_{kn}}(X_n \in A^c)}_{\le 1-\alpha}] \le (1-\alpha)\mathbf{P}_x(\tau_A > kn),\end{aligned}$$

therefore, by induction, $\mathbf{P}_x(\tau_A > kn) \le (1-\alpha)^k$. One has $\mathbf{P}_x(\tau_A > r) \le \mathbf{P}_x(\tau_A > kn)$ for $r \ge kn$ so that

$$\mathbf{E}_x(\tau_A) = \sum_{r=0}^{\infty} \mathbf{P}_x(\tau_A > r) \le \sum_{k=0}^{\infty} \sum_{r=kn}^{(k+1)n-1} \mathbf{P}_x(\tau_A > kn) \le n \sum_{k=0}^{\infty} (1-\alpha)^k = \frac{n}{\alpha}.$$

b) Let $u > 0$ and $k = [\frac{u}{n}]$ ($[x]$ denotes the largest integer which is smaller than x); then $\mathbf{P}_x(\tau_A > u) \le \mathbf{P}_x(\tau_A > kn) \le (1-\alpha)^k \le (1-\alpha)^{\frac{u}{n}-1}$. Remark that

$$\begin{aligned}\mathbf{E}_x(e^{\rho\tau_A}) &= 1 + \mathbf{E}_x \int_0^{\tau_A} \rho\, e^{\rho u}\, du = 1 + \mathbf{E}_x \int_0^{+\infty} \rho\, e^{\rho u}\, 1_{\{u < \tau_A\}}\, du = \\ &= 1 + \rho \int_0^{+\infty} e^{\rho u}\, \mathbf{P}_x(\tau_A > u)\, du.\end{aligned}$$

One can then deduce that, if $\rho + \frac{1}{n}\log(1-\alpha) = -\delta < 0$,

$$\mathbf{E}_x(e^{\rho\tau_A}) \le 1 + \rho(1-\alpha)^{-1} \int_0^{+\infty} e^{-\delta u}\, du < +\infty.$$

E4.13 One begins by observing that no state $x > 0$ leads to 0 and therefore $F = \{1, 2, \dots\}$ is a closed class. As $0 \rightsquigarrow 1$ whereas $1 \not\rightsquigarrow 0$, 0 is transient thanks to the transience criterion of •4.21.

On the other hand all the states in $\mathbb{N}^*$ communicate. Indeed every state leads to 1, that leads to 2 and 3; 2 leads to 4, which leads to 6, ..., and therefore 2 leads to every even state. By the same argument 3 leads to every odd state. By the transitivity of the relation $\rightsquigarrow$, all the states of $\mathbb{N}^*$ communicate.

Let σ_x be the return time to x (•2.2). Then $\mathbf{P}_1(\sigma_1 = 1) = 0$ and, for every $n \geq 2$,

$$\begin{aligned}\mathbf{P}_1(\sigma_1 = n) &= \mathbf{P}_1(X_1 = 2, X_2 = 4, \dots, X_{n-1} = 2(n-1), X_n = 1) + \\ &\quad + \mathbf{P}_1(X_1 = 3, X_2 = 5, \dots, X_{n-1} = 2n-1, X_n = 1) = \\ &= \beta q_2 q_4 \dots q_{2(n-2)} p_{2(n-1)} + (1-\beta) q_3 q_5 \dots q_{2n-3} p_{2n-1}.\end{aligned}$$

One gets, for every $n \geq 2$,

$$\begin{aligned}\mathbf{P}_1(\sigma_1 \leq n) &= \beta \sum_{k=2}^{n} q_2 q_4 \dots q_{2k-4} p_{2k-2} + (1-\beta) \sum_{k=2}^{n} q_3 q_5 \dots q_{2k-3} p_{2k-1} = \\ &= \beta \sum_{k=2}^{n} q_2 q_4 \dots q_{2k-4}(1 - q_{2k-2}) + (1-\beta) \sum_{k=2}^{n} q_3 q_5 \dots q_{2k-3}(1 - q_{2k-1}) = \\ &= 1 - \beta q_2 q_4 \dots q_{2n-2} - (1-\beta) q_3 q_5 \dots q_{2n-1}\end{aligned}$$

and therefore

$$\mathbf{P}_1(\sigma_1 < +\infty) = 1 - \beta \prod_{k=1}^{\infty} q_{2k} - (1-\beta) \prod_{k=1}^{\infty} q_{2k+1}.$$

State 1 is therefore recurrent if and only if these two infinite products are equal to 0, i.e., if $\sum_{k=1}^{\infty} p_{2k} = \sum_{k=1}^{\infty} p_{2k+1} = +\infty$.

E4.14 1) 0 is absorbing (and therefore recurrent). If $k \neq 0$, $P(k, 0) > 0$ and so $k \rightsquigarrow 0$; as $0 \not\rightsquigarrow k$, k is transient, thanks to the criterion of •4.21.

2) Denote by f the identity function. One has to prove that $Pf = f$. Suppose first $k \neq 0$. Then

$$Pf(k) = \sum_{n \geq 0} P(k, n) f(n) = \sum_{n \geq 0} P(k, n) n.$$

Therefore $Pf(k)$ is the mean of the probability $P(k, \cdot)$. This is a Poisson distribution with parameter k and $Pf(k) = k$. If $k = 0$, the relation is straightforward.

3) As f is P-harmonic, one knows that $f(X_n) = X_n$ is a positive martingale (Proposition 4.10). Therefore, $(X_n)_{n \geq 0}$ converges $\mathbf{P}_k$-a.s., for every $k \in \mathbb{N}$, to a r.v. W. This r.v. may take the values 0, $+\infty$ or $k \neq 0$. This last possibility cannot occur since a transient state is visited by the chain only a finite number of times a.s. and cannot be a cluster point. Finally we know (Theorem 3.8) that $W < +\infty$ $\mathbf{P}_k$-a.s.

E4.15 Suppose $X_n = i$, with $0 \leq i \leq m$. At step $n+1$ then $X_{n+1} = i - 1$ if the chosen ball is among the i balls in the first box (probability equal to $\frac{i}{m}$). One has $X_{n+1} = i + 1$ if the chosen ball is among those in the second box (probability $1 - \frac{i}{m}$).

The transition matrix is therefore

$$P(i,j)=\begin{cases}\frac{i}{m} & \text{if } j=i-1\\ \frac{m-i}{m} & \text{if } j=i+1\\ 0 & \text{otherwise.}\end{cases}$$

Obviously, all the states communicate and the chain is irreducible. To prove this rigorously, it is sufficient to observe that a state $i\in E$, $i<m$ leads to $i+1$, as $P(i,i+1)>0$ and, for the same reason, $i+1\rightsquigarrow i+2$, $i+2\rightsquigarrow i+3\dots$. Thanks to the transitivity of the relation $\rightsquigarrow$ (•4.21) i also leads to $i+2, i+3,\dots, r$. By the same argument i leads to $i-1, i-2,\dots,0$.

Concerning the aperiodicity, one remarks that at each transition the chain takes one step exactly. Therefore if $i\in E$ is even, $\mathbf{P}_i(X_1 \text{ is even})=0$ and, more generally, $\mathbf{P}_i(X_n \text{ is even})=0$ for every n odd. As it is immediate that $P^2(i,i)>0$, the set $I(i)$ of •4.31 is composed of all even numbers, which implies that the period of i, whatever is i, is equal to 2. The chain is not aperiodic.

E4.16 a) It is immediate that $i\rightsquigarrow i+1$ and $i\rightsquigarrow i-1$. Therefore all the states communicate, thanks to the transitivity of the relation $\rightsquigarrow$.

One easily finds out that if $P^n(i,i)>0$, then n is even. As moreover, $P^2(i,i)\geq P(i,i+1)P(i+1,i)=p(1-p)>0$, the period of the chain is 2.

b) As Z_{n+1} is independent of $(X_0,X_1,\dots,X_n)$, if $\mathbf{P}(X_n=i, X_{n-1}=a_{n-1},\dots, X_1=a_1)>0$,

$$\begin{aligned}
&\mathbf{P}(X_{n+1}=j\mid X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)=\\
&\frac{\mathbf{P}(X_{n+1}=j, X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)}{\mathbf{P}(X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)}=\\
&=\frac{\mathbf{P}(Z_{n+1}=j-i, X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)}{\mathbf{P}(X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)}=\\
&=\frac{\mathbf{P}(Z_{n+1}=j-i)\mathbf{P}(X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)}{\mathbf{P}(X_n=i, X_{n-1}=a_{n-1},\dots,X_1=a_1)}=\\
&=\mathbf{P}(Z_{n+1}=j-i)=P(i,j).
\end{aligned}$$

c) Thanks to the law of large numbers, $(\frac{1}{n}X_n)_{n\geq 0}$ converges a.s. to $\mathbf{E}(Z_1)=2p-1$. If $p\neq\frac{1}{2}$, this value is not equal to 0 and $(X_n)_{n\geq 0}$ converges a.s. to $+\infty$ or $-\infty$ according to the sign of $2p-1$. Therefore the chain cannot be recurrent: indeed, if it were, $(X_n)_{n\geq 0}$ would visit 0 (or any other state) an infinite number of times a.s., which is incompatible with $X_n\to_{n\to\infty}+\infty$ (or $-\infty$) a.s. The chain is therefore transient.

d) $Y_i=\frac{1}{2}(Z_i+1)$ takes the values 0 with probability $1-p$ and 1 with probability p. It is therefore a Bernoulli r.v. with parameter p.

Now $T_n=\frac{1}{2}(X_n+n)=\sum_{i=1}^n Y_i$. T_n is the sum of n independent Bernoulli r.v.'s with parameter p: it is therefore a binomial r.v. $B(n,p)$. Therefore, $P^n(0,0)=\mathbf{P}(X_n=0)=\mathbf{P}(T_n=\frac{n}{2})$. Therefore $P^n(0,0)=0$ if n is odd and $P^n(0,0)=\binom{n}{n/2}p^{n/2}(1-p)^{n/2}$ if n is even.

e) Assume $p=\frac{1}{2}$. One knows that a criterion of recurrence is given by the divergence of the series $\sum_{n\geq 1} P^n(0,0)$ (Theorem 4.16). By Stirling's formula

$$(2k)! \sim \sqrt{4\pi k}\left(\frac{2k}{e}\right)^{2k}, \qquad k! \sim \sqrt{2\pi k}\left(\frac{k}{e}\right)^{k}$$

so that

$$\binom{2k}{k} = \frac{(2k)!}{(k!)^2} \sim \frac{\sqrt{4\pi k}}{2\pi k}\frac{4^k k^{2k}}{e^{2k}}\frac{e^{2k}}{k^{2k}} \sim \frac{1}{\sqrt{\pi k}}4^k.$$

Therefore $P^{2k}(0,0) = \binom{2k}{k}\frac{1}{4^k} \sim (\pi k)^{-1/2}$ as $k \to \infty$, which guarantees the divergence of the series $\sum_{n\geq 1} P^n(0,0)$. The state 0 is therefore recurrent, together with all the other states, the chain being irreducible.

Finally one observes that the measure μ putting a mass 1 at all points, i.e., $\mu(x) = 1$ for every $x \in \mathbb{Z}$ (also called sometimes the *counting measure* of $\mathbb{Z}$), is stationary for P. As its total mass is infinite, the chain is null recurrent (•4.27).

E4.17 1) As Y_{n+1} is independent of $\mathscr{F}_n$,

$$\mathbf{E}^{\mathscr{F}_n}(Y_{n+1}) = \mathbf{E}(Y_{n+1}) = 0 \text{ a.s.} \quad \text{and} \quad \mathbf{E}^{\mathscr{F}_n}(Y_{n+1}^2) = 1 \text{ a.s.}$$

Therefore $\mathbf{E}^{\mathscr{F}_n}(S_{n+1} - S_n) = \mathbf{E}^{\mathscr{F}_n}(Y_{n+1}) = 0$ a.s., i.e., $(S_n)_{n\geq 0}$ is a martingale. In the same way, setting $Z_n = S_n^2 - n$, one has a.s.

$$\mathbf{E}^{\mathscr{F}_n}(Z_{n+1} - Z_n) = \mathbf{E}^{\mathscr{F}_n}(S_{n+1}^2 - S_n^2) - 1 =$$
$$= \mathbf{E}^{\mathscr{F}_n}(Y_{n+1}(2S_n + Y_{n+1})) - 1 = 2S_n\mathbf{E}^{\mathscr{F}_n}(Y_{n+1}) + \mathbf{E}^{\mathscr{F}_n}(Y_{n+1}^2) - 1 = 0$$

and $(S_n^2 - n)_{n\geq 0}$ is a martingale.

2) Let us set $c = \max(a,b)$, $p_a = \mathbf{P}(\tau_{-a} < \tau_b)$ and $p_b = \mathbf{P}(\tau_{-a} > \tau_b)$. By the optional sampling Theorem 3.3, $\mathbf{E}(S_{n\wedge\tau}) = \mathbf{E}(S_0) = 0$. But $S_{n\wedge\tau} \to S_\tau$ a.s. as $n \to \infty$ (since $\mathbf{P}(\tau < +\infty) = 1$) and $|S_{n\wedge\tau}| \leq c$. One can therefore apply Lebesgue's theorem and get $\mathbf{E}(S_\tau) = 0$. On one hand $p_a + p_b = 1$; on the other hand $0 = \mathbf{E}(S_\tau) = -ap_a + bp_b$, from which

$$p_a = \frac{b}{a+b}, \qquad p_b = \frac{a}{a+b}.$$

3) By the optional sampling theorem, $\mathbf{E}(S_{n\wedge\tau}^2) = \mathbf{E}(n\wedge\tau)$. On one hand $n\wedge\tau \uparrow \tau$, hence $\mathbf{E}(n\wedge\tau) \uparrow \mathbf{E}(\tau)$; on the other hand, τ being finite a.s., $S_{n\wedge\tau}^2 \to_{n\to\infty} S_\tau^2$ a.s. Since $0 \leq S_{n\wedge\tau}^2 \leq c^2$, $\mathbf{E}(S_{n\wedge\tau}^2) \to_{n\to\infty} \mathbf{E}(S_\tau^2)$. This implies that $\mathbf{E}(S_\tau^2) = \mathbf{E}(\tau)$. Therefore

$$\mathbf{E}(\tau) = \mathbf{E}(S_\tau^2) = a^2 p_a + b^2 p_b = ab.$$

4a) Since $\mathbf{E}(Y_{n+1}) = \mathbf{E}(Y_{n+1}^3) = 0$,

$$\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n) = \mathbf{E}[(S_n + Y_{n+1})^3 - 3(n+1)(S_n + Y_{n+1}) \mid \mathscr{F}_n] =$$
$$= \mathbf{E}[S_n^3 + 3Y_{n+1}S_n^2 + 3Y_{n+1}^2 S_n + Y_{n+1}^3 - 3(n+1)(S_n + Y_{n+1}) \mid \mathscr{F}_n] =$$
$$= S_n^3 + 3\mathbf{E}(Y_{n+1}^2)S_n - 3(n+1)S_n = S_n^3 - 3nS_n = Z_n.$$

It is immediate that $|Z_{n\wedge\tau}| \leq K(1+\tau)$, K being equal to the greatest among the values $3a, 3b, a^3, b^3$. Since it has been proved in (3) that τ is integrable, one can apply Lebesgue's theorem and get

$$\mathbf{E}(S_\tau^3 - 3\tau S_\tau) = \mathbf{E}(Z_\tau) = \lim_{n\to\infty}\mathbf{E}(Z_{n\wedge\tau}) = \mathbf{E}(Z_0) = 0.$$

Thus

$$\mathbf{E}(\tau S_\tau) = \frac{1}{3}\,\mathbf{E}(S_\tau^3) = \frac{1}{3}\left(\frac{b^3 a}{a+b} - \frac{a^3 b}{a+b}\right) = \frac{1}{3}\,ab(b-a).$$

4b) Intuition suggests that these two r.v.'s are not independent: if $b > a$ then larger values of τ should be expected on $S_\tau = b$ than on $S_\tau = -a$. Nonindependence is easy to check if the two r.v.'s are correlated: this is what happens here since $\mathbf{E}(S_\tau) = 0$, so that $\mathbf{E}(\tau S_\tau) \neq \mathbf{E}(\tau)\mathbf{E}(S_\tau)$.

4c) Relation (4.48) is quite intuitive, because of symmetry reasons. Looking for a rigorous proof, let us set $\tilde{Y}_n = -Y_n$, $\tilde{S}_n = -S_n$. $(\tilde{S}_n)_{n\geq 0}$ is a Markov chain having the same transition matrix as $(S_n)_{n\geq 0}$. These two sequences have therefore the same law and, if we denote by $\tilde{\tau}$ the exit time of $(\tilde{S}_n)_{n\geq 0}$ from $]-a, a[$, the two r.v.'s (S_τ, τ) and $(\tilde{S}_{\tilde{\tau}}, \tilde{\tau})$ have the same law. On the other hand it is clear that $\tau = \tilde{\tau}$, since $S_n = \pm a$ if and only if $\tilde{S}_n = \mp a$ and

$$\mathbf{P}(S_\tau = a, \tau = k) = \mathbf{P}(\tilde{S}_\tau = a, \tau = k) = \mathbf{P}(S_\tau = -a, \tau = k).$$

Since $\mathbf{P}(S_\tau = a, \tau = k) + \mathbf{P}(S_\tau = -a, \tau = k) = \mathbf{P}(\tau = k)$, we get (4.48). This entails

$$\mathbf{P}(S_\tau = a, \tau = k) = \mathbf{P}(S_\tau = a)\mathbf{P}(\tau = k).$$

The relation holds for every k and also with $-a$ instead of a; therefore the two r.v.'s S_τ and τ are independent.

E4.18 a) We know (•4.17) that the function $f_i = \mathbf{P}_i(\tau_0 < +\infty)$ is the smallest nonnegative function on $\mathbb{N}$ that is equal to 1 at 0 and such that $Pf = f$ on $\mathbb{N}\setminus\{0\}$. The last condition can be written

$$f_i = qf_{i-1} + rf_i + pf_{i+1}$$

for $i = 1, 2, \dots$, i.e., as $r = 1 - p - q$,

$$f_{i+1} - (1 + \tfrac{q}{p})f_i + \tfrac{q}{p}f_{i-1} = 0. \tag{4.80}$$

This equation can be solved with the usual methods for the sequences defined by recurrence: one looks for the solutions v_1, v_2 of the quadratic equation

$$v^2 - (1 + \tfrac{q}{p})v + \tfrac{q}{p} = 0.$$

If the two roots are different, then every sequence satisfying equation (4.80) is necessarily of the form

$$f_i = c_1 v_1^i + c_2 v_2^i.$$

Here one immediately finds that $v_1 = 1$, $v_2 = \frac{q}{p}$. Thus, if $p \neq q$, every solution of (4.80) is of the form

$$f_i = c_1 + c_2\Big(\frac{q}{p}\Big)^i.$$

Two cases are possible.

1. $p < q$. Then $c_2 \geq 0$, since otherwise f_i would be strictly negative for large values of i. The smallest positive solution is therefore of the form $f_i = c_1$; the condition $f_0 = 1$ implies that the smallest positive solution of (4.80) taking the value 1 at 0 is $f_i = 1$ for every $i = 1, 2, \dots$. Therefore, if $p < q$, whatever $i \geq 1$, the chain starting at i visits 0 with probability 1.

2. $p > q$. In this case $c_1 \geq 0$, as $f_i \to_{i\to\infty} c_1$. The smallest nonnegative solution is therefore of the form $f_i = c_2(\frac{q}{p})^i$ and the condition $f_0 = 1$ implies $c_2 = 1$. Therefore, starting at $i \geq 1$, the chain visits the state 0 with probability $(\frac{q}{p})^i$ and never visits 0 with probability $1 - (\frac{q}{p})^i$.

If $p = q$, the sequences satisfying (4.80) are of the form

$$f_i = c_1 + c_2 i.$$

Since $f_0 = 1$, $c_1 = 1$ and the smallest positive solution is obtained for $c_2 = 0$. The hitting probability of 0 is therefore $f_i = 1$ for every $i = 1, 2, \dots$, as in the case $p < q$.

b) With the assumptions made on the parameters p, q, it is clear that all the states communicate. They are therefore all simultaneously recurrent or transient. If $p > q$, $\mathbf{P}_i(\tau_0 < +\infty) = \mathbf{P}_i(\sigma_0 < +\infty) < 1$ for every $i > 0$. If the chain were recurrent, one would have $\mathbf{P}_i(\sigma_0 < +\infty) = 1$ (•4.21). Therefore the chain is transient. Conversely, if $p \leq q$, by the criterion of Exercise 4.11 (with $x_0 = 0$), the chain is recurrent.

E4.19 A) Let $i \in F^c$. Then $\tau = 1 + \tau \circ \theta$ $\mathbf{P}_i$-a.s. (this relation holds even if $\tau = +\infty$). Thus, by the Markov property (4.18),

$$\begin{aligned}\psi(i) &= \mathbf{E}_i(z^{1+\tau\circ\theta} 1_{\{\tau\circ\theta < +\infty\}}) = z\mathbf{E}_i((z^\tau 1_{\{\tau < +\infty\}}) \circ \theta) = \\ &= z\mathbf{E}_i(\mathbf{E}_{X_1}(z^\tau 1_{\{\tau < +\infty\}})) = z\mathbf{E}_i(\psi(X_1)) = zP\psi(i).\end{aligned}$$

B) Let us compute the generating function of τ. For a fixed z, thanks to (A), $\psi(i) = \mathbf{E}_i(z^\tau)$ is a solution of (4.49). This is a linear system with 7 unknown variables, but it is obvious, by symmetry, that the value of ψ at the three states marked with 1 in Figure 4.2 is the same and the same happens for the three states marked with 2. If, z being fixed, we denote x_1 and x_2, respectively, the values of ψ on these two classes and x_3 the value of ψ on the vertex opposed to i_0, we are reduced to the system

$$\begin{aligned}x_1 &= \tfrac{z}{3} + \tfrac{2z}{3}x_2 \\ x_2 &= \tfrac{2z}{3}x_1 + \tfrac{z}{3}x_3 \\ x_3 &= zx_2.\end{aligned}$$

This can be easily solved, the solution being unique; we find $x_1 = \frac{z}{3} + \frac{2}{3}z^3(9-7z^2)^{-1}$, $x_2 = z^2(9-7z^2)^{-1}$ and

$$x_3 = \frac{2z^3}{9 - 7z^2}.$$

Recalling that, for $|u| < 1$, $(1-u)^{-1} = \sum_{n=0}^\infty u^n$,

$$\frac{2z^3}{9 - 7z^2} = \frac{2}{9}\sum_{n=0}^\infty \Big(\frac{7}{9}\Big)^n z^{3+2n},$$

this permits computing the required probabilities. The clever reader might have recalled that the generating function of a r.v. X, which is geometric with parameter p (i.e., $\mathbf{P}(X = k) = p(1-p)^k$, $k = 0, 1, \dots$), is given by

$$\frac{p}{1 - z(1-p)}.$$

so that the generating function of $2X$ is

$$\frac{p}{1 - z^2(1-p)}.$$

Since one can write

$$\frac{2z^3}{9 - 7z^2} = z^3 \frac{\frac{2}{9}}{1 - \frac{7}{9}z^2},$$

we find that, starting at i_1, τ has the same law as $3 + 2X$, where X is geometric with parameter $p = \frac{2}{9}$. Thus

$$\begin{aligned}
\mathbf{P}(\tau = 3) &= \mathbf{P}(X = 0) = \tfrac{2}{9} = 0.222\\
\mathbf{P}(\tau = 15) &= \mathbf{P}(X = 6) = \tfrac{2}{9}(\tfrac{7}{9})^6 = 0.049\\
\mathbf{P}(\tau \geq 15) &= \mathbf{P}(X \geq 6) = (\tfrac{7}{9})^6 = 0.221.
\end{aligned}$$

The expectation of τ is easily obtained, either by computing the derivative of the generating function at $z = 1$, or recalling the value of the mean for the geometric laws:

$$\mathbf{E}_{i_1}(\tau) = 3 + 2\mathbf{E}(X) = 3 + \frac{\frac{7}{9}}{\frac{2}{9}} = \frac{13}{2}.$$

C1) If g is a positive or bounded solution of (4.49) then, if $i \in F^c$ and since $g \equiv 1$ on F,

$$\begin{aligned}
g(i) = zPg(i) = z\mathbf{E}_i(g(X_1)) &= z\mathbf{E}_i(1_{\{X_1 \in F\}} + 1_{\{X_1 \in F^c\}} g(X_1)) =\\
&= z\mathbf{P}_i(X_1 \in F) + z\mathbf{E}_i(1_{\{X_1 \in F^c\}} g(X_1)) =\\
&= z\mathbf{P}_i(\tau = 1) + z\mathbf{E}_i(1_{\{\tau > 1\}} g(X_1)).
\end{aligned} \tag{4.81}$$

C2) Thanks to (C1), (4.50) holds for $n = 1$. Let us assume it to hold at the step n. Then $X_n \in F^c$ on $\{\tau > n\}$, so that

$$g(X_n) = z\mathbf{P}_{X_n}(\tau = 1) + z\mathbf{E}_{X_n}(1_{\{\tau > 1\}} g(X_1))$$

whence

$$g(i) = \sum_{k=1}^{n} z^k \mathbf{P}_i(\tau = k) + z^{n+1}\mathbf{E}_i\big[1_{\{\tau > n\}}\big(\mathbf{P}_{X_n}(\tau = 1) + \mathbf{E}_{X_n}(1_{\{\tau > 1\}} g(X_1))\big)\big].$$

Since the r.v. $1_{\{\tau > n\}}$ is $\mathscr{F}_n$-measurable, the Markov property (4.18) gives

$$\begin{aligned}
\mathbf{E}_i\big(1_{\{\tau > n\}}\mathbf{P}_{X_n}(\tau = 1)\big) &= \mathbf{E}_i\big(1_{\{\tau > n\}} 1_{\{\tau = 1\}} \circ \theta_n\big) =\\
&= \mathbf{E}_i(1_{\{\tau = n+1\}}) = \mathbf{P}_i(\tau = n + 1)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}_i\big[1_{\{\tau > n\}}\mathbf{E}_{X_n}\big(1_{\{\tau > 1\}} g(X_1)\big)\big] &= \mathbf{E}_i\big(1_{\{\tau > n\}}(1_{\{\tau > 1\}} g(X_1)) \circ \theta_n\big) =\\
&= \mathbf{E}_i\big[1_{\{\tau > n+1\}} g(X_{n+1})\big],
\end{aligned}$$

which completes the proof of (4.50).

If $|z| < 1$ and g is a bounded solution of (4.49), $z^n\mathbf{E}_i(1_{\{\tau > n\}} g(X_n))$ tends to 0 as $n \to \infty$ and

$$\sum_{k=1}^{n} z^k \mathbf{P}_i(\tau = k) = \mathbf{E}_i(z^\tau 1_{\{\tau \leq n\}}) \underset{n\to\infty}{\longrightarrow} \mathbf{E}_i(z^\tau 1_{\{\tau < +\infty\}}) = \psi(i).$$

Thus $g(i) = \psi(i)$.

If $z \geq 0$ and g is a positive solution of (4.49), then for every $i \in F^c$ and every $n \geq 1$,

$$g(i) \geq \sum_{k=1}^{n} z^k \mathbf{P}(\tau = k)$$

so that

$$g(i) \geq \sum_{k=1}^{\infty} z^k \mathbf{P}_i(\tau = k) = \mathbf{E}_i(z^\tau 1_{\{\tau < +\infty\}}) = \psi(i).$$

E4.20 a) It is immediate that

$$P = \begin{pmatrix} 0 & p & 1-p \\ 1-p & 0 & p \\ p & 1-p & 0. \end{pmatrix}.$$

It is immediate that all the states communicate, hence the irreducibility of the chain. In order to obtain the aperiodicity, one cannot apply the first of the two criteria of •4.33, as all the elements on the diagonal of P are equal to 0. One therefore uses the second one; indeed it is immediate that P^2 has all its elements > 0 (see also •4.33). As for the stationary probability, one immediately sees that the transition matrix is bistochastic (•4.34 (i)) and therefore the uniform distribution $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is invariant.

b) One has $\mathbf{P}_\mu(X_n = 1, X_{n+1} = 2) = \sum_{x \in E} \mu(x) P^n(x, 1) P(1, 2)$, where μ denotes the distribution of X_0. Thanks to Theorem 4.29,

$$\lim_{n \to \infty} P^n(x, 1) = \pi(x) = \frac{1}{3}$$

for every initial state x. It is therefore immediate that

$$\lim_{n \to \infty} \mathbf{P}_\mu(X_n = 1, X_{n+1} = 2) = \frac{p}{3}.$$

By the same argument

$$\lim_{n \to \infty} \mathbf{P}_\mu(X_n = 2, X_{n+1} = 1) = \frac{1-p}{3}.$$

c) The probability π is reversible if and only if $\pi(x)P(x, y) = \pi(y)P(y, x)$ for every $x, y \in E$. As π is uniform, this amounts to requiring that P is symmetric, i.e., $p = 1 - p = \frac{1}{2}$ (this is also the value for which the two limits computed in (b) are equal).

E4.21 a) If the graph is connected, for any two given vertices i and j, there exists a finite sequence $h_1, \ldots, h_m$ of vertices such that i is connected to h_1 by an edge, h_1 is connected to h_2 by an edge, ..., h_m is connected to j by an edge. This implies that $p_{i,h_1} > 0$, $p_{h_1,h_2} > 0, \ldots, p_{h_m,j} > 0$. Therefore

$$p_{ij}^{(m+1)} = \sum_{k_1, \ldots, k_m \in E} p_{i,k_1} p_{k_1,k_2} \cdots p_{k_m,j} \geq p_{i,h_1} p_{h_1,h_2} \cdots p_{h_m,j} > 0.$$

b) The simplest way is to check that π is reversible (see •4.34 (ii)). If the states i

and j are connected by an edge,

$$\pi_i p_{ij} = \frac{k_i}{k}\frac{1}{k_i} = \frac{1}{k}.$$

The quantity on the right-hand side being symmetric in i, j, it is also equal to $\pi_j p_{ji}$. If, conversely, the states i and j are not connected by an edge, then the two quantities $\pi_i p_{ij}$ and $\pi_j p_{ji}$ are both equal to 0; hence, π is reversible and therefore stationary.

E4.22 Observe that, by the definition of the transition matrix, the probability to pass from one cell to each of its neighbours is positive. One can therefore build a path of positive probability leading from one cell to any other one; hence, the chain is irreducible.

It is easy to see that the cells of the board are of three different kinds.

- Those on a corner, which have two neighbours (the transition probability to any of these is therefore equal to $\frac{1}{2}$).
- The edge cells that are not on a corner and have three neighbours (transition probability equal to $\frac{1}{3}$).
- The central cells that have four neighbours (transition probability equal to $\frac{1}{4}$).

For symmetry reasons, the cells of the same kind are given the same weight by the stationary probability: denote these weights, respectively, by π_2, π_3 and π_4 (the index corresponding to the number of neighbours). Therefore

$$\begin{aligned}\pi_2 &= \tfrac{1}{3}\pi_3 + \tfrac{1}{3}\pi_3\\ \pi_3 &= \tfrac{1}{2}\pi_2 + \tfrac{1}{3}\pi_3 + \tfrac{1}{4}\pi_4\\ \pi_4 &= \tfrac{2}{3}\pi_3 + \tfrac{2}{4}\pi_4\end{aligned}$$

with the normalization condition

$$4\pi_2 + 8\pi_3 + 4\pi_4 = 1. \tag{4.82}$$

The previous system can be rewritten as

$$\begin{aligned}\pi_2 &= \tfrac{2}{3}\pi_3\\ \pi_4 &= \tfrac{4}{3}\pi_3.\end{aligned}$$

Solving the system of the two previous equations and using (4.82), one finds

$$\pi_3 = \frac{1}{16}, \quad \pi_2 = \frac{1}{24}, \quad \pi_4 = \frac{1}{12}.$$

With a different approach, the clever reader may have noticed that this chain is a random walk on the graph of Figure 4.7 (every vertex corresponding to the centers of the cells of the board). One can therefore use Exercise 4.21 (b), which gives directly

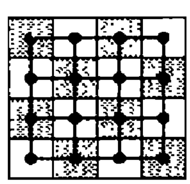

Figure 4.7

the irreducibility (the graph is connected) as well as the expression of the stationary probability:

$$\pi_j = \frac{k_j}{k}$$

where k_j is the number of edges issuing from vertex j and k is the sum of the numbers k_j. Therefore $k = 4 \cdot 2 + 8 \cdot 3 + 4 \cdot 4 = 48$.

E4.23 1) As all the points communicate, X is irreducible. As E is finite, X is positive recurrent.

2) $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$ is stationary if and only if $\mu = \mu Q$, i.e.,

$$\begin{aligned} \mu_1 &= \tfrac{1}{2}\mu_2 + \tfrac{1}{2}\mu_3 \\ \mu_2 &= \mu_1 + \tfrac{1}{2}\mu_3 \\ \mu_3 &= \tfrac{1}{4}\mu_2 + \mu_4 \\ \mu_4 &= \tfrac{1}{4}\mu_2. \end{aligned}$$

One finds $\mu_1 = 3\mu_4$, $\mu_2 = 4\mu_4$, $\mu_3 = 2\mu_4$ and $\mu = (\frac{3}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10})$.

3) Applying the ergodic Theorem 4.25 to the functions $f(x) = x$ and $g(x) = x^2$, one has, $\mathbf{P}_x$-a.s. for every x,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} X_k = \langle \mu, f\rangle = \frac{3}{10}\cdot 1 + \frac{4}{10}\cdot 2 + \frac{2}{10}\cdot 3 + \frac{1}{10}\cdot 4 = \frac{21}{10}$$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} X_k^2 = \langle \mu, g\rangle = \frac{3}{10}\cdot 1 + \frac{4}{10}\cdot 2^2 + \frac{2}{10}\cdot 3^2 + \frac{1}{10}\cdot 4^2 = \frac{53}{10}.$$

E4.24 a) It is immediate that the set $C = \{2, 5\}$ is a closed class, as $P(x_0, C) = 1$ for every $x_0 \in C$ (Proposition 4.19). Observe that $p_{2,5} = \frac{3}{4} > 0$ and $p_{5,2} = \frac{1}{3} > 0$: the chain restricted to C is therefore irreducible. The states 2 and 5 are necessarily recurrent.

Observe next that $P(1, 5) = \frac{1}{3} > 0$. As $1 \rightsquigarrow 5$, but $5 \not\rightsquigarrow 1$, the state 1 is transient thanks to the criterion of •4.21. The same argument proves that 3 and 4 are also transient.

b) If $x \in C$, $\tau \equiv 0$ and $\mathbf{E}_x(\tau) = 0$. More precisely, one knows (Corollary 4.15) that $v(x) = \mathbf{E}_x(\tau)$ is a solution of

$$v(x) = \begin{cases} 0 & \text{if } x \in C \\ 1 + Pv(x) & \text{if } x \notin C. \end{cases}$$

This system becomes here

$$\begin{aligned} v_1 &= 1 + \tfrac{1}{3}v_1 + \tfrac{1}{12}v_3 + \tfrac{1}{4}v_4 \\ v_3 &= 1 + \tfrac{1}{3}v_1 + \tfrac{1}{2}v_3 \\ v_4 &= 1 + \tfrac{1}{4}v_3 + \tfrac{1}{4}v_4 \end{aligned}$$

and its unique solution is $v_1 = 3$, $v_3 = 4$, $v_4 = \frac{8}{3}$.

c) One knows (Corollary 4.14) that the function $u(x) = \mathbf{P}_x(\tau < +\infty, X_\tau = y)$

solves the system

$$u(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \in C \setminus \{y\} \\ Pu(x) & \text{if } x \notin C. \end{cases}$$

For $y = 2$, this system can be written

$$\begin{aligned} u_1 &= && \tfrac{1}{3}u_1 + && \tfrac{1}{12}u_3 + && \tfrac{1}{4}u_4 \\ u_3 &= \tfrac{1}{12} + && \tfrac{1}{3}u_1 + && \tfrac{1}{2}u_3 && \\ u_4 &= \tfrac{1}{3} + && && \tfrac{1}{4}u_3 + && \tfrac{1}{4}u_4 \end{aligned}$$

whence the solution $u_1 = \frac{1}{4}$, $u_3 = \frac{1}{3}$, $u_4 = \frac{5}{9}$. In order to find the value of $w(x) = \mathbf{P}_x(\tau < +\infty, X_\tau = 5)$, one may write and solve the corresponding system; but, with the class C consisting of two states, it is simpler to observe that $\mathbf{P}_x(\tau < +\infty, X_\tau = 5) = 1 - \mathbf{P}_x(\tau < +\infty, X_\tau = 2)$; hence, $w_1 = \frac{3}{4}$, $w_3 = \frac{2}{3}$, $w_4 = \frac{4}{9}$.

d) Let $x \in C$. As the chain restricted to C is irreducible positive recurrent, the ergodic Theorem 4.25 applied to the function $f(x) = x$ gives $\mathbf{P}_x$-a.s.

$$\frac{1}{n}\sum_{k=1}^{n} X_k \underset{n\to\infty}{\to} \langle \pi, f\rangle \tag{4.83}$$

where π denotes the stationary probability of the restricted chain. If $x \notin C$, after a time τ the chain enters C and stays there forever; τ being a.s. finite, this suggests that (4.83) remains true even if $x \notin C$. More rigorously, by the strong Markov property,

$$\mathbf{P}_x\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} X_k = \langle\pi, f\rangle\Big) = \mathbf{P}_x\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{k=\tau+1}^{n+\tau} X_k = \langle\pi, f\rangle\Big) =$$

$$= \mathbf{P}_x\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} X_k \circ \theta_\tau = \langle\pi, f\rangle\Big) =$$

$$= \mathbf{E}_x\Big(\underbrace{\mathbf{P}_{X_\tau}\Big(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} X_k = \langle\pi, f\rangle\Big)}_{=1}\Big) = 1.$$

We still have to find the stationary probability π of the chain restricted to C. It must satisfy $\pi_2 + \pi_5 = 1$ and

$$\begin{cases} \pi_2 = \tfrac{1}{4}\pi_2 + \tfrac{1}{3}\pi_5 \\ \pi_5 = \tfrac{3}{4}\pi_2 + \tfrac{2}{3}\pi_5, \end{cases}$$

i.e., $\pi_2 = \frac{4}{13}$ and $\pi_5 = \frac{9}{13}$. Therefore

$$\langle\pi, f\rangle = 2\cdot\frac{4}{13} + 5\cdot\frac{9}{13} = \frac{53}{13}.$$

E4.25 Denote by Y_n the number of particles remaining in V, and Z_n the number of particles entering V during the time interval $[n, n+1[$. Then $X_{n+1} = Y_n + Z_n$.

1a) Y_0 and Z_0 are independent; Z_0 follows a Poisson distribution with parameter λ whereas the distribution of Y_0 given $X_0 = x$ is binomial $B(x, q)$. As the characteristic function of a sum of independent r.v.'s is the product of the characteristic functions and recalling the expression of the characteristic functions of the binomial and Poisson r.v.'s,

$$\mathbf{E}(e^{itX_1} \mid X_0 = x) = (p + q\,e^{-it})^x\, e^{\lambda(e^{it}-1)}.$$

1b)

$$\begin{aligned}\mathbf{E}(e^{itX_1}) = \mathbf{E}(\mathbf{E}(e^{itX_1} \mid X_0)) &= \mathbf{E}(e^{\lambda(e^{it}-1)}(p + q\,e^{-it})^{X_0}) = \\ &= e^{\lambda(e^{it}-1)}\,\mathbf{E}((p + q\,e^{-it})^{X_0});\end{aligned}$$

hence, thanks to the fact that $\mathbf{E}(s^{X_0}) = e^{\theta(s-1)}$,

$$\mathbf{E}(e^{itX_1}) = e^{\lambda(e^{it}-1)}\,e^{\theta((p+q\,e^{it})-1)} = e^{(\lambda+\theta q)(e^{it}-1)}.$$

X_1 has therefore a Poisson distribution with parameter $\lambda + \theta q$. μ_θ is stationary if and only if $\theta = \lambda + \theta q$, i.e., $\theta = \frac{\lambda}{p}$.

2) With the notations introduced in the solution of (1a),

$$\begin{aligned}\mathbf{P}(X_{n+1} = y \mid X_n = x) &= \mathbf{P}(Z_n + Y_n = y \mid X_n = x) = \\ &= \sum_{k\in\mathbb{N}} \mathbf{P}(Y_n = k, Z_n = y - k \mid X_n = x) = \\ &= \sum_{k\in\mathbb{N}} \mathbf{P}(Y_n = k \mid X_n = x)\mathbf{P}(Z_n = y - k) = \\ &= e^{-\lambda} \sum_{0\le k\le x, k\le y} \binom{x}{k} q^k (1-q)^{x-k} \frac{\lambda^{y-k}}{(y-k)!}.\end{aligned}$$

Observe that $Q(x, y) > 0$ for all $x, y \in \mathbb{N}$ and therefore the chain is irreducible. As it admits a stationary probability, it is positive recurrent (•4.23).

3) The ergodic Theorem 4.25, applied to the function $f(x) = x$, gives

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \ \underset{n\to\infty}{\overset{a.s.}{\to}}\ \langle \mu_{\lambda/p}, f\rangle = \frac{\lambda}{p},$$

as the mean of a Poisson distribution coincides with the parameter, which is here $\frac{\lambda}{p}$. By the same argument, the sequence $\sum_{k=0}^{n-1} X_k^2$ converges to the second moment of this Poisson distribution, i.e., $\frac{\lambda}{p}(1 + \frac{\lambda}{p})$.

E4.26 a) The calculation of the eigenvalues of Q can be performed by finding the roots of the characteristic polynomial, which leads to a second degree equation. It is more clever to remember that the sum of the eigenvalues is equal to the trace ($= 2 - \alpha - \beta$) and that 1 is always the eigenvalue of a stochastic matrix. The second eigenvalue is therefore $1 - \alpha - \beta$.

As $\alpha > 0$ and $\beta > 0$, $1 - \alpha - \beta < 1$: the two eigenvalues 1 and $1 - \alpha - \beta$ are different and the matrix is therefore diagonalizable.

It is known that $e_1 = \binom{1}{1}$ is an eigenvector associated to the eigenvalue 1. The equation $Q\binom{x}{y} = (1 - \alpha - \beta)\binom{x}{y}$ can be written $\beta x + \alpha y = 0$. One may therefore take $e_2 = \binom{-\alpha}{\beta}$ as the eigenvector associated to $1-\alpha-\beta$. Denoting by

$$R = \begin{pmatrix} 1 & -\alpha \\ 1 & \beta \end{pmatrix}$$

the matrix associated to the change of basis, $Q = R\Delta R^{-1}$ where Δ is the diagonal matrix with the values 1 and $1-\alpha-\beta$ on the diagonal. Therefore $Q^n = R\Delta^n R^{-1}$ and, as

$$R^{-1} = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ -1 & 1 \end{pmatrix},$$

one gets

$$\begin{aligned} Q^n &= \frac{1}{\alpha+\beta}\begin{pmatrix} 1 & -\alpha \\ 1 & \beta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{pmatrix}\begin{pmatrix} \beta & \alpha \\ -1 & 1 \end{pmatrix} = \\ &= \frac{1}{\alpha+\beta}\begin{pmatrix} \beta+\alpha(1-\alpha-\beta)^n & \alpha-\alpha(1-\alpha-\beta)^n \\ \beta-\beta(1-\alpha-\beta)^n & \alpha+\beta(1-\alpha-\beta)^n \end{pmatrix} = \\ &= \frac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta}\begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}. \end{aligned}$$

Since $\min(\alpha,\beta) < 1$, then $-1 < 1-\alpha-\beta < 1$, from which

$$\lim_{n\to\infty} Q^n = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}.$$

b) The stationarity equation $\pi Q = \pi$ becomes here

$$\begin{aligned} \pi_0(1-\alpha) + \pi_1\beta &= \pi_0 \\ \pi_0\alpha + \pi_1(1-\beta) &= \pi_1. \end{aligned}$$

These two equations are obviously proportional (1 is an eigenvalue) and we must add the condition $\pi_0 + \pi_1 = 1$. One easily finds

$$\pi_0 = \frac{\beta}{\alpha+\beta}, \qquad \pi_1 = \frac{\alpha}{\alpha+\beta}.$$

As the r.v.'s X_i take the values 0 or 1 and, under $\mathbf{P}_\pi$, have a common distribution,

$$\begin{aligned} \mathbf{E}_\pi(X_n X_{n+1}) &= \mathbf{P}_\pi(X_n = 1, X_{n+1} = 1) = \mathbf{P}_\pi(X_n = 1)Q(1,1) = \\ &= \pi_1 Q(1,1) = \frac{\alpha}{\alpha+\beta}(1-\beta). \end{aligned}$$

Moreover,

$$\mathbf{E}_\pi(X_n) = \mathbf{E}_\pi(X_{n+1}) = \frac{\alpha}{\alpha+\beta};$$

hence,

$$\mathrm{Cov}_\pi(X_n, X_{n+1}) = \frac{\alpha\beta}{(\alpha+\beta)^2}(1-\alpha-\beta).$$

The r.v.'s X_n and X_{n+1} are not independent if $1-\alpha-\beta \neq 0$, since in this case their covariance is $\neq 0$. If $1-\alpha-\beta = 0$ one easily checks that X_n and X_{n+1} are independent (actually that $X_0, X_1, \dots, X_n$ are independent). Note that this condition implies that the rows of the transition matrix are equal.

c) Under the probability $\mathbf{P}_\pi$, all the r.v.'s X_k follow the distribution π. Therefore

$$\mathbf{E}_\pi(S_n) = \sum_{k=1}^{n} \mathbf{E}_\pi(X_k) = n\mathbf{E}_\pi(X_1) = \frac{n\alpha}{\alpha+\beta}.$$

As for the variance,

$$\operatorname{Var}_\pi(S_n) = \sum_{k=1}^{n} \operatorname{Var}_\pi(X_k) + 2 \sum_{1\le i<j\le n} \operatorname{Cov}_\pi(X_i, X_j). \tag{4.84}$$

Let us compute $\operatorname{Cov}_\pi(X_i, X_j)$, $1 \le i < j \le n$. One can repeat the argument of (b):

$$\begin{aligned}\mathbf{E}_\pi(X_i X_j) &= \mathbf{P}_\pi(X_i = 1, X_j = 1) = \mathbf{P}_\pi(X_i = 1) Q^{j-i}(1,1) = \\ &= \frac{\alpha}{\alpha+\beta}\Big(\frac{\alpha}{\alpha+\beta} + \frac{\beta(1-\alpha-\beta)^{j-i}}{\alpha+\beta}\Big)\end{aligned} \tag{4.85}$$

(one can use (4.52) in order to compute $Q^{j-i}(1,1)$). Therefore

$$\operatorname{Cov}_\pi(X_i, X_j) = \frac{\alpha(\alpha+\beta(1-\alpha-\beta)^{j-i})}{(\alpha+\beta)^2} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta(1-\alpha-\beta)^{j-i}}{(\alpha+\beta)^2},$$

and

$$\begin{aligned}\sum_{1\le i<j\le n} \operatorname{Cov}_\pi(X_i, X_j) &= \frac{\alpha\beta}{(\alpha+\beta)^2} \sum_{i=1}^{n} \sum_{j=i+1}^{n} (1-\alpha-\beta)^{j-i} = \\ &= \frac{\alpha\beta}{(\alpha+\beta)^2} \sum_{i=1}^{n} \sum_{k=1}^{n-i} (1-\alpha-\beta)^k = \\ &= \frac{\alpha\beta}{(\alpha+\beta)^2} \sum_{i=1}^{n} \frac{(1-\alpha-\beta) - (1-\alpha-\beta)^{n-i+1}}{\alpha+\beta} = \\ &= \frac{\alpha\beta}{(\alpha+\beta)^3}\Big[n(1-\alpha-\beta) + \frac{1-\alpha-\beta}{\alpha+\beta}((1-\alpha-\beta)^n - 1)\Big].\end{aligned}$$

These computations and (4.84) give

$$\operatorname{Var}_\pi(S_n) = n \operatorname{Var}_\pi(X_1) + \frac{2\alpha\beta(1-\alpha-\beta)}{(\alpha+\beta)^3}\Big[n + \frac{1}{\alpha+\beta}((1-\alpha-\beta)^n - 1)\Big].$$

Since $|1-\alpha-\beta| < 1$, we deduce the existence of a constant C such that $\operatorname{Var}_\pi(S_n) \le Cn$. If $\varepsilon > 0$, thanks to the Bienaymé–Chebyshev inequality,

$$\mathbf{P}_\pi\Big(\Big|\frac{S_n}{n} - \frac{\alpha}{\alpha+\beta}\Big| > \varepsilon\Big) \le \frac{\operatorname{Var}_\pi(S_n)}{n^2\varepsilon^2} \le \frac{C}{n\varepsilon^2}$$

from which follows the requested convergence in probability, with respect to $\mathbf{P}_\pi$. Let us denote

$$A_n = \Big\{\Big|\frac{S_n}{n} - \frac{\alpha}{\alpha+\beta}\Big| > \varepsilon\Big\}.$$

It is known (see (4.10)) that $\mathbf{P}_\pi(A_n) = \pi(0)\mathbf{P}_0(A_n) + \pi(1)\mathbf{P}_1(A_n)$. As $\pi(0) > 0$, $\pi(1) > 0$ and $\lim_{n\to\infty} \mathbf{P}_\pi(A_n) = 0$, then $\lim_{n\to\infty} \mathbf{P}_0(A_n) = \lim_{n\to\infty} \mathbf{P}_1(A_n) =$

0, which implies the convergence in probability of $(\frac{1}{n}S_n)_{n\geq 0}$ to $\alpha(\alpha+\beta)^{-1}$, with respect to $\mathbf{P}_0$ and $\mathbf{P}_1$.

E4.27 1) Remark that $Q(0, y) = p(y) > 0$ and therefore 0 leads to y for every $y > 0$. Moreover, if $x > 0$, $Q(x, 0) = \frac{1}{x} > 0$ and therefore $x \rightsquigarrow 0$. One concludes that $x \rightsquigarrow y$ for every $x, y \in \mathbb{N}$, thanks to the transitivity of the relation $\rightsquigarrow$. The chain is therefore irreducible.

Let $x > 0$ and denote by τ_0 the hitting time at 0; then $x > X_1 > X_2 > \ldots > X_{\tau_0 - 1}$ $\mathbf{P}_x$-a.s. and therefore, since $X_n \in \mathbb{N}$, one has $\mathbf{P}_x(\tau_0 \leq x) = 1$ and $\mathbf{P}_x(\tau_0 < \infty) = 1$. By the Markov property,

$$\mathbf{P}_0(\sigma_0 < +\infty) = \mathbf{P}_0(\tau_0 \circ \theta_1 < \infty) = \mathbf{E}_0(\underbrace{\mathbf{P}_{X_1}(\tau_0 < +\infty)}_{=1}) = 1.$$

The state 0 is recurrent, as every other state, the chain being irreducible.

2a) The stationarity equation can be written here, for every $x \geq 0$,

$$\mu(x) = \sum_{y\in\mathbb{N}} \mu(y)Q(y, x) = \mu(0)Q(0, x) + \sum_{y>x} \mu(y)Q(y, x) =$$
$$= p(x)\mu(0) + \sum_{y>x} \frac{\mu(y)}{y}.$$

If we replace in (4.86) $\mu(x)$ by its expression in terms of (4.87), one obtains, for each $x \geq 1$,

$$x\varphi(x-1) = p(x)\mu(0) + (x+1)\varphi(x)$$

and, adding up for x from 1 to y,

$$\varphi(0) = \mu(0)\big(1 - G(y) - p(0)\big) + (y+1)\varphi(y),$$

but $\varphi(0) = \mu(0) - \mu(0)p(0)$, which gives

$$(y+1)\varphi(y) = \mu(0)G(y).$$

Plugging this relation into (4.86), one has, as $p(y) = G(y-1) - G(y)$,

$$\mu(y) = \mu(0)\Big(p(y) + \frac{G(y)}{y+1}\Big) = \mu(0)y\Big(\frac{G(y-1)}{y} - \frac{G(y)}{y+1}\Big).$$

3) In this case $G(y) = (y+1)^{-1}$ and therefore, for every $y \geq 1$,

$$\mu(y) = \mu(0)\Big(\frac{1}{(y+1)^2} + \frac{1}{y(y+1)}\Big) = \mu(0)\Big(\frac{1}{(y+1)^2} + \frac{1}{y} - \frac{1}{y+1}\Big).$$

Thus

$$\sum_{y\ge 0}\mu(y)=\mu(0)+\mu(0)\Big(\sum_{y\ge 1}\frac{1}{(y+1)^2}+1\Big)=\mu(0)\Big(1+\frac{\pi^2}{6}\Big)<\infty$$

and the measure μ is finite: the chain is positive recurrent. The stationary probability is given by $\mu(0)=(1+\frac{\pi^2}{6})^{-1}$ and

$$\mu(y)=\Big(1+\frac{\pi^2}{6}\Big)^{-1}\Big(\frac{1}{(y+1)^2}+\frac{1}{y}-\frac{1}{y+1}\Big)$$

for $y\ge 1$. Therefore the ergodic Theorem 4.25 implies

$$\frac{1}{n}\sum_{k=0}^{n-1}1_{\{X_k=0\}}\ \underset{n\to\infty}{\overset{\text{a.s.}}{\to}}\ \langle\mu,1_{\{x=0\}}\rangle=\Big(1+\frac{\pi^2}{6}\Big)^{-1}$$

E4.28 a) The equation $u(i)=Qu(i)$ may be written in this case

$$u(i)=p_iu(i+1)+r_iu(i)+q_iu(i-1)$$

so that (recall that $r_i=1-p_i-q_i$)

$$u(i+1)-u(i)=\frac{q_i}{p_i}(u(i)-u(i-1))$$

and therefore, for $i=a+1,\dots,b$,

$$u(i+1)-u(i)=\frac{q_{a+1}\cdots q_i}{p_{a+1}\cdots p_i}(u(a+1)-u(a))=\frac{\gamma_i}{\gamma_a}(u(a+1)-u(a))$$

and

$$u(i)=u(a)+(u(a+1)-u(a))\frac{1}{\gamma_a}\sum_{k=a}^{i-1}\gamma_k. \tag{4.88}$$

This is obviously a monotone function, increasing or decreasing according to the sign of $u(a+1)-u(a)$.

b) The function $u(i)=\mathbf{P}_i(\tau<+\infty)$ satisfies $u(a)=u(b)=1$ and $u(i)=Qu(i)$ for $a+1\le i\le b-1$ (it is Corollary 4.13); by (a) such a function is monotone and therefore equal to 1 for every $a\le i\le b$.

The function $v(i)=\mathbf{P}_i(X_\tau=b)$ is the smallest positive solution of $Qv=v$ on $\{a+1,\dots,b-1\}$ with the condition $v(a)=0$, $v(b)=1$ (Corollary 4.14); thus, if in (4.88) one imposes the condition $u(a)=0$, $u(b)=1$,

$$u(a+1)-u(a)=\frac{\gamma_a}{\sum_{k=a}^{b-1}\gamma_k}$$

and, replacing into (4.88),

$$v(i)=\frac{\sum_{k=a}^{i-1}\gamma_k}{\sum_{k=a}^{b-1}\gamma_k}.$$

c) The chain is obviously irreducible (one uses, as always, the transitivity of the relation $\rightsquigarrow$). As the chain may take at most one step to the right at each transition, $\tau_{n+1}\ge\tau_n+1$ $\mathbf{P}_1$-a.s.; therefore, $\mathbf{P}_1$-a.s., the sequence $\{\tau_n<\tau_0\}$ is decreasing and

$\bigcap_{n\geq 1}\{\tau_n < \tau_0\} \subset \{\tau_0 = +\infty\}$. But $\{\tau_0 = +\infty\} \subset \{\tau_n < \tau_0\}$ for every $n \geq 1$. Thanks to (b), $\mathbf{P}_1(\tau_0 \wedge \tau_n < +\infty) = 1$ and

$$\mathbf{P}_1(\tau_0 = +\infty) = \lim_{n\to\infty} \mathbf{P}_1(\tau_n < \tau_0) = \frac{1}{\sum_{k=0}^{\infty} \gamma_k}\cdot \tag{4.89}$$

In order to check whether the chain is recurrent, let us investigate the value of $\mathbf{P}_0(\sigma_0 < +\infty)$, where, as usual, σ_0 is the return time to 0 (•2.5). As $\sigma_0 = 1 + \tau_0 \circ \theta_1$, thanks to the Markov property,

$$\begin{aligned}\mathbf{P}_0(\sigma_0 < +\infty) &= \mathbf{P}_0(\tau_0 \circ \theta_1 < +\infty) = \mathbf{E}_0(1_{\{\tau_0<+\infty\}} \circ \theta_1) = \\ &= \mathbf{E}_0(\mathbf{E}_{X_1}(1_{\{\tau_0<+\infty\}})) = r_0 + p_0\mathbf{P}_1(\tau_0 < +\infty).\end{aligned}$$

Therefore, by (4.89), the chain is recurrent if and only if $\sum_{k=0}^{\infty} \gamma_k = +\infty$ (we have in fact repeated the proof of the recurrence criterion of Exercise 4.11).

d) It holds $\gamma_i = (\frac{q}{p})^i$. Therefore the series with general term γ_i is summable if and only if $q < p$ and in this case the chain is transient. It is recurrent if and only if $p \leq q$.

e) For $i \geq 1$ one easily obtains (all other terms cancel)

$$\gamma_i = \frac{q_1 \dots q_i}{p_1 \dots p_i} = \frac{(1+1)\dots(i+1)}{(1+2)\dots(i+2)} = \frac{2}{i+2},$$

which is the general term of divergent series; this chain is therefore recurrent. For the chain given by (4.53), one finds similarly

$$\gamma_i = \frac{6}{(i+1)(i+2)(i+3)}\cdot$$

It is the general term of a convergent series and the chain is transient.

E4.29 a) The assumptions made on the numbers p_i, q_i imply that every state leads both to its left and right neighbours. By the transitivity of the relation $\rightsquigarrow$ (see •4.21) all the states communicate.

b) For $i = 0$, $j = 1$, the detailed balance equation (4.44) becomes

$$\mu_0 p_0 = \mu_1 q_1.$$

Therefore, if one sets $\mu_0 = \alpha$, $\mu_1 = \alpha p_0 q_1^{-1} = \alpha\zeta_1$. By the same argument the detailed balance equation for $i = 1$, $j = 2$ is $\mu_1 p_1 = \mu_2 q_2$ so that

$$\mu_2 = \mu_1 \frac{p_1}{q_2} = \alpha \frac{p_0 p_1}{q_1 q_2} = \alpha\zeta_2$$

and one easily obtains the relation (4.54) by induction.

By construction, the measure μ defined by (4.54) satisfies the detailed balance equation if i and j are neighbours. But if $|i - j| > 1$, then $P(i, j) = P(j, i) = 0$ and the two members of (4.44) both vanish; the equation is obviously satisfied. μ satisfies the detailed balance equation and is therefore stationary.

c) If $E = \{0, 1, \dots, m\}$, it is sufficient to set

$$\alpha = (1 + \zeta_1 + \dots + \zeta_m)^{-1}.$$

With this choice, (4.54) defines a stationary distribution, unique as the chain is irreducible. On the other hand, as $E = \mathbb{N}$, let us consider

$$\alpha = \Big(\sum_{i=0}^{\infty} \zeta_i\Big)^{-1}. \tag{4.90}$$

If the series in (4.90) converges, one obtains again a stationary distribution. Conversely, if it diverges, there is no stationary probability. Indeed if one existed, the chain would be irreducible positive recurrent (•4.23) and every stationary measure would be finite (Theorem 4.21). Here the condition $\sum_{i=0}^{\infty} \zeta_i = +\infty$ implies that the stationary measure (4.54) is infinite for every value of $\alpha > 0$.

d) With the given values of p_i, q_i, it holds $\zeta_i = \left(\frac{p}{q}\right)^i$. Therefore the series in (4.90) is geometric with parameter $\frac{p}{q}$ and is convergent if and only if $p < q$. Its sum is, in this case, $(1 - \frac{p}{q})^{-1} = q(q - p)^{-1}$ and the stationary probability is

$$\pi_i = \frac{q-p}{q}\Big(\frac{p}{q}\Big)^i$$

and $\mathbf{E}_k(\sigma_k) = q^{k+1}p^{-k}(q - p)^{-1}$ (see Theorem 4.23).

♦ Condition $p < q$ in (d) has an intuitive meaning: p and q denote the probabilities of taking one step to the right or to the left, respectively. Thus the stationarity condition found in (c) (summability of the series $\sum_{i=1}^{\infty} \zeta_i$) asserts that the tendency to go to the right is compensated by an attraction towards 0. One also observes that the numbers r_i have no influence (but they are connected to p_i, q_i by the relation $r_i = 1 - p_i - q_i$) and even that two different birth and death chains, but such that the quotients p_{i-1}/q_i are the same, have the same stationary distribution.
♦ It has already been seen in Exercise 4.18 (b) that, if $p = q$, then the chain is recurrent. This exercise shows moreover that it is null recurrent.

E4.30 a) Suppose first $X_n = r$; all the white balls are therefore in the first box. It is therefore a black ball that is taken from the second box and a white one that is taken from the first one; thus $X_{n+1} = r - 1$ with probability 1. If $X_n = 0$, similarly, $X_{n+1} = 1$ with probability 1.

On the other hand if $X_n = i$ with $0 < i < r$, three cases are possible:

• One takes two white balls or two black ones; this occurs with probability $2(r - i)ir^{-2}$. In this case $X_{n+1} = i$.

• One takes a white ball in the first box and a black one in the second; this occurs with probability i^2r^{-2}. In this case $X_{n+1} = i - 1$.

• One takes a black ball in the first box and a white one in the second; this occurs with probability $(r - i)^2r^{-2}$. In this case $X_{n+1} = i + 1$.

The transition matrix of the chain $(X_n)_{n\ge 0}$, whose state space is $\{0, 1, \dots, r\}$, is given by $P(0, 1) = P(r, r - 1) = 1$ and, for $1 \le i \le r - 1$,

$$P(i,j) = \begin{cases} \frac{i^2}{r^2} & \text{if } j = i - 1 \\ \frac{2(r-i)i}{r^2} & \text{if } j = i \\ \frac{(r-i)^2}{r^2} & \text{if } j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

Obviously all states communicate and the chain is irreducible (one still uses the transitivity of the relation $\rightsquigarrow$). As moreover $P(i, i) > 0$ for at least one $i \in E$, the chain is aperiodic, thanks to the first criterion of •4.33.

b) One observes that this is a birth and death chain on $E = \{0, 1, \ldots, r\}$, with

$$p_i = \frac{(r-i)^2}{r^2} \qquad q_i = \frac{i^2}{r^2}.$$

With the notations of Exercise 4.29, one has, for $1 \le i \le r$,

$$\zeta_i = \frac{p_0 \ldots p_{i-1}}{q_1 \ldots q_i} = \frac{r^2(r-1)^2 \ldots (r-i+1)^2}{1 \cdot 2^2 \ldots i^2} = \binom{r}{i}^2.$$

Therefore the stationary probability of the chain is

$$\pi_i = \alpha \binom{r}{i}^2$$

where $\alpha^{-1} = \binom{r}{0}^2 + \binom{r}{1}^2 + \ldots + \binom{r}{r}^2$. One may easily show that the last sum is in fact equal to $\binom{2r}{r}$ (see Feller's book, volume I, II.12).

E4.31 a) Let us suppose $X_n = x$; four cases can arise.

• The ball of the $(n+1)$-th choice belongs to urn A (probability $\frac{x}{d}$) and it is urn A that is chosen (probability $\frac{1}{2}$). The resulting probability is therefore $\frac{x}{2d}$ and $X_{n+1} = x$.

• The ball of the $(n+1)$-th choice belongs to urn A (probability $\frac{x}{d}$) and it is urn B that is chosen (probability $\frac{1}{2}$). The resulting probability is $\frac{x}{2d}$ and $X_{n+1} = x - 1$.

• The ball in the $(n+1)$-th choice belongs to urn B (probability $1 - \frac{x}{d}$) and it is urn A that is chosen (probability $\frac{1}{2}$). The resulting probability is $\frac{1}{2}(1 - \frac{x}{d})$ and $X_{n+1} = x + 1$.

• The ball of the $(n+1)$-th choice belongs to urn B (probability $1 - \frac{x}{d}$) and it is urn B that is chosen (probability $\frac{1}{2}$). The resulting probability is $\frac{1}{2}(1 - \frac{x}{d})$ and $X_{n+1} = x$.

The transition matrix of the chain is thus

$$Q(x, y) = \begin{cases} \frac{x}{2d} & \text{if } y = x - 1 \\ \frac{1}{2}(1 - \frac{x}{d}) & \text{if } y = x + 1 \\ \frac{1}{2} & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

for $1 \le x \le d - 1$ and $Q(0, 0) = Q(0, 1) = Q(d, d) = Q(d, d - 1) = \frac{1}{2}$.

This chain is obviously irreducible as, starting from every state, both of its neighbours can be visited with positive probability. It is therefore positive recurrent, the state space being finite. It is also aperiodic, thanks to the first criterion of •4.33, as there is at least one strictly positive element on the diagonal of the transition matrix.

b) Suppose $1 \le x \le d - 1$. Then

$$\sum_{y \in E} y Q(x, y) = (x-1)Q(x, x-1) + (x+1)Q(x, x+1) + xQ(x, x) =$$

$$= \frac{x}{2d}(x-1) + \frac{1}{2}\Big(1 - \frac{x}{d}\Big)(x+1) + \frac{x}{2} = \Big(1 - \frac{1}{d}\Big)x + \frac{1}{2}.$$

For $x = d$ or $x = 0$, one has

$$\sum_{y\in E} yQ(d, y) = d\,Q(d, d) + (d-1)Q(d, d-1) = d - \frac{1}{2}$$

$$\sum_{y\in E} yQ(0, y) = Q(0, 1) = \frac{1}{2}.$$

The requested relation is therefore proved with $a = 1 - \frac{1}{d}$, $b = \frac{1}{2}$. If we keep this notation, $\mathbf{E}_x(X_1) = \sum_{y\in E} yQ(x, y) = ax + b$ and

$$\mathbf{E}_x(X_2) = \mathbf{E}_x(X_1 \circ \theta_1) = \mathbf{E}_x(\mathbf{E}_{X_1}(X_1)) = a^2x + b(1 + a).$$

Assume that $\mathbf{E}_x(X_n) = a^n x + b(1 + a + \ldots + a^{n-1})$, then

$$\mathbf{E}_x(X_{n+1}) = \mathbf{E}_x(X_n \circ \theta_1) = \mathbf{E}_x(\mathbf{E}_{X_1}(X_n)) =$$
$$= a^n\mathbf{E}_x(X_1) + b(1 + a + \ldots + a^{n-1}) = a^{n+1}x + b(1 + a + \ldots + a^n).$$

Therefore, by induction

$$\mathbf{E}_x(X_n) = a^n x + b\frac{1 - a^n}{1 - a}$$

and

$$\lim_{n\to\infty} \mathbf{E}_x(X_n) = \frac{b}{1-a} = \frac{d}{2}.$$

c) Let us denote by ν the distribution $B(d, \frac{1}{2})$. One knows, by (4.10), that, for $0 \le y \le d$,

$$\mathbf{P}_\nu(X_1 = y) = \sum_{x\in E} \nu(x)Q(x, y).$$

Therefore

$$\mathbf{P}_\nu(X_1 = y) =$$
$$= 2^{-d}\left(\binom{d}{y-1}Q(y-1, y) + \binom{d}{y}Q(y, y) + \binom{d}{y+1}Q(y+1, y)\right) =$$
$$= 2^{-(d+1)}\left(\binom{d}{y-1}\frac{d-y+1}{d} + \binom{d}{y} + \binom{d}{y+1}\frac{y+1}{d}\right) =$$
$$= 2^{-(d+1)}\left(\binom{d-1}{y-1} + \binom{d}{y} + \binom{d-1}{y}\right) = 2^{-d}\binom{d}{y}$$

with the understanding $\binom{0}{-1} = \binom{d}{d+1} = 0$. Therefore X_1 has the same distribution as X_0.

d) By the previous question, the unique stationary distribution π of the chain is the binomial distribution $B(d, \frac{1}{2})$. The chain is aperiodic, as seen in (a). Therefore

$$\lim_{n\to\infty} Q^n(x, y) = \pi(y) = 2^{-d}\binom{d}{y}$$

and, by (4.41), $\mathbf{E}_d(\sigma_d) = \pi(d)^{-1} = 2^d$.

♦ It is possible to derive the stationary distribution of this chain from other exercises.

Note for instance that this model is very close to the Ehrenfest chain of Exercise 4.15, whose stationary distribution is computed in Exercise 4.36. Moreover, in both cases, these are birth and death chains on $E = \{0, \dots, d\}$ such that, setting $p_i = Q(i, i+1)$, $q_i = Q(i, i-1)$, the quotients p_{i-1}/q_i, $i \ge 1$ coincide. This implies that the two chains have the same stationary distribution (see the observation at the end of Exercise 4.29).

E4.32 a) The chain being irreducible and recurrent, for a fixed $a \in E$ the unique invariant measure μ such that $\mu(a) = 1$ is given by

$$\mu(b) = \mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1} 1_{\{X_k=b\}}\Big)$$

(see •4.26). This measure has necessarily finite total mass as the chain is positive recurrent and the unique invariant probability is given by $\pi(x) = \mu(x)\big(\sum_{b\in E}\mu(b)\big)^{-1}$. Let $f: E \to \mathbb{R}^+$; then

$$\pi(f) = \sum_{b\in E}\pi(b)f(b) = \pi(a)\sum_{b\in E}\frac{\pi(b)}{\pi(a)}f(b) = \pi(a)\sum_{b\in E}\mu(b)f(b) =$$

$$= \pi(a)\sum_{b\in E}\mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1}1_{\{X_k=b\}}\Big)f(b) = \pi(a)\mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1}f(X_k)\Big).$$

b) It holds (see (4.10)) $\mathbf{E}_\pi(\sigma) = \sum_{b\in E}\pi(b)\mathbf{E}_b(\sigma) = \langle\pi, f\rangle$, with $f(b) = \mathbf{E}_b(\sigma)$. One gets from (a)

$$\mathbf{E}_\pi(\sigma) = \pi(a)\mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1}\mathbf{E}_{X_k}(\sigma)\Big) = \pi(a)\sum_{k=0}^{\infty}\mathbf{E}_a\big(1_{\{k<\sigma\}}\mathbf{E}_{X_k}(\sigma)\big).$$

Therefore, by the Markov property,

$$\mathbf{E}_\pi(\sigma) = \pi(a)\sum_{k=0}^{\infty}\mathbf{E}_a(1_{\{k<\sigma\}}\sigma\circ\theta_k).$$

But on $\{k < \sigma\}$ $\sigma = k + \sigma\circ\theta_k$ and therefore

$$\mathbf{E}_\pi(\sigma) = \pi(a)\mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1}(\sigma-k)\Big) = \pi(a)\mathbf{E}_a\Big(\sum_{j=1}^{\sigma}j\Big) = \pi(a)\mathbf{E}_a\Big(\frac{\sigma(\sigma+1)}{2}\Big).$$

Let us now perform the same calculations for $f(b) = \mathbf{E}_b(\sigma^2)$, $b \in E$. Repeating the same argument,

$$\mathbf{E}_\pi(\sigma^2) = \pi(a)\mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1}\mathbf{E}_{X_k}(\sigma^2)\Big) = \pi(a)\sum_{k=0}^{\infty}\mathbf{E}_a(1_{\{k<\sigma\}}\sigma^2\circ\theta_k) =$$

$$= \pi(a)\mathbf{E}_a\Big(\sum_{k=0}^{\sigma-1}(\sigma-k)^2\Big) = \pi(a)\mathbf{E}_a\Big(\sum_{j=1}^{\sigma}j^2\Big) = \pi(a)\mathbf{E}_a\Big(\frac{\sigma(\sigma+1)(2\sigma+1)}{6}\Big).$$

E4.33 Remark that $\tau(\omega) = \infty$ means that, for every n, $f(X_n(\omega)) = 0$. One has, for

$x \in E \setminus A$.

$$Uf(x) = \mathbf{E}_x\Big(\sum_{n\geq 0} f(X_n)\Big) = \mathbf{E}_x\Big(1_{\{\tau<+\infty\}}\sum_{n\geq\tau} f(X_n)\Big) =$$
$$= \mathbf{E}_x\Big(1_{\{\tau<+\infty\}}\sum_{n\geq 0} f(X_n)\circ\theta_\tau\Big).$$

By the strong Markov property (4.19),

$$Uf(x) = \mathbf{E}_x\Big[1_{\{\tau<+\infty\}}\mathbf{E}_{X_\tau}\Big(\sum_{n\geq 0} f(X_n)\Big)\Big] = \mathbf{E}_x\big[1_{\{\tau<+\infty\}}Uf(X_\tau)\big].$$

Obviously

$$\sup_{y\in E} Uf(y) \geq \sup_{y\in A} Uf(y). \tag{4.91}$$

Let $x \in E \setminus A$, i.e., such that $f(x) = 0$: let us show that $Uf(x) \leq \sup_{z\in A} Uf(z)$. If $\tau < \infty$, $X_\tau \in A$ so that $Uf(X_\tau) \leq \sup_{z\in A} Uf(z)$. Therefore, by (4.56), $Uf(x) \leq \sup_{z\in A} Uf(z)$. This being true for every $x \in E \setminus A$, one deduces with (4.91) the requested result.

E4.34 a) If $A \subset F$, thanks to the Markov property (Definition 4.2 (ii)) for the chain X,

$$\mathbf{P}(Y_{n+1} \in A \mid \mathscr{F}_n) = \mathbf{P}(X_{n+1} \in \psi^{-1}(A) \mid \mathscr{F}_n) = P(X_n, \psi^{-1}(A)).$$

As ψ is supposed invertible, the relation

$$Q(y,z) = P(\psi^{-1}(y), \psi^{-1}(z)), \qquad y, z \in F$$

defines a transition matrix on F and

$$\mathbf{P}(Y_{n+1} \in A \mid \mathscr{F}_n) = P(X_n, \psi^{-1}(A)) = P(\psi^{-1}(Y_n), \psi^{-1}(A)) = Q(Y_n, A),$$

which shows that Y is a Markov chain with transition matrix Q. Of course the initial distribution of Y is the image by ψ of μ, the initial distribution of X.

b) As $P(c,a) = 0$ and $\mathbf{P}(X_{n+1} = a, X_n = c) = \mathbf{P}(X_n = c)P(c,a) = 0$,

$$\{Y_{n+1} = 0, Y_n = 1\} = \{X_{n+1} = a, X_n = b\} \quad \text{a.s.}$$

Therefore, again by the Markov property,

$$\mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1, Y_{n-1} = 0) =$$
$$= \mathbf{P}(X_{n+1} = a \mid X_n = b, X_{n-1} = a) = \mathbf{P}(X_{n+1} = a \mid X_n = b) = \frac{1}{2}.$$

Moreover $\{Y_n = 1\} = \{X_n \in \{b,c\}\}$ and

$$\mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1) = \mathbf{P}(X_{n+1} = a \mid X_n \in \{b,c\}) =$$
$$= \frac{\mathbf{P}(X_{n+1} = a, X_n \in \{b,c\})}{\mathbf{P}(X_n \in \{b,c\})} =$$
$$= \frac{\mathbf{P}(X_n = b)P(b,a) + \mathbf{P}(X_n = c)P(c,a)}{\mathbf{P}(X_n \in \{b,c\})}.$$

Recalling that $P(b,a) = \frac{1}{2}$, $P(c,a) = 0$,

$$\mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1) = \frac{1}{2}\frac{\mathbf{P}(X_n = b)}{\mathbf{P}(X_n = b) + \mathbf{P}(X_n = c)}.$$

The chain X being irreducible, there exists n such that $\mathbf{P}(X_n = c) > 0$, i.e., such that

$$\mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1, Y_{n-1} = 0) \neq \mathbf{P}(Y_{n+1} = 0 \mid Y_n = 1).$$

The Markov property is not satisfied by the process Y.

E4.35 a) The elements of the matrix Q are obviously ≥ 0. If $x \in E$ is such that $\psi(x) = i$,

$$\sum_{j\in F} Q(i,j) = \sum_{j\in F}\sum_{y,\psi(y)=j} P(x,y) = \sum_{y\in E} P(x,y) = 1.$$

Thus Q is a transition matrix. Now

$$\begin{aligned}\mathbf{P}(Y_{n+1} = j \mid \mathscr{F}_n) = \mathbf{P}(\psi(X_{n+1}) = j \mid \mathscr{F}_n) = \mathbf{P}(X_{n+1} \in \psi^{-1}(j) \mid \mathscr{F}_n) = \\ = P(X_n, \psi^{-1}(j)) = Q(X_n, j).\end{aligned}$$

Therefore $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geq 0}, (Y_n)_{n\geq 0}, \mathbf{P})$ is a Markov chain with transition matrix Q.

b) If $\tilde{\pi} = \pi \circ \psi^{-1}$, then

$$\tilde{\pi}(i) = \pi(\psi^{-1}(i)) = \sum_{x,\psi(x)=i} \pi(x)$$

and

$$\tilde{\pi}(i)Q(i,j) = \sum_{x,\psi(x)=i}\sum_{y,\psi(y)=j} \pi(x)P(x,y)$$

(one uses the fact that $x \to \sum_{y,\psi(y)=j} P(x,y)$, as a function of x, is constant on $\{x, \psi(x) = i\}$). We are now ready to check the stationarity equation:

$$\begin{aligned}\sum_{i\in F}\tilde{\pi}(i)Q(i,j) &= \sum_{i\in F}\sum_{x,\psi(x)=i}\sum_{y,\psi(y)=j} \pi(x)P(x,y) = \\ &= \sum_{y,\psi(y)=j}\sum_{x\in E}\pi(x)P(x,y) = \sum_{y,\psi(y)=j}\pi(y) = \tilde{\pi}(j).\end{aligned}$$

E4.36 a) It is easily shown that all the states communicate: for every $x, y \in E$ there exist states $z_1, \dots, z_r$ such that $P(x, z_1) > 0$, $P(z_1, z_2) > 0, \dots, P(z_r, y) > 0$ and therefore $P^{r+1}(x,y) > 0$. One may also observe that this chain is as a random walk on the points of the hypercube, and a random walk on a connected graph is irreducible as seen in Exercise 4.21 (a).

One next notes that the transition matrix P is symmetric, as $x^{(i)} = y$ implies $y^{(i)} = x$. In particular it is bistochastic (see •4.34) and the uniform probability on E is stationary (and it is the only one, as the chain is irreducible). The stationary probability is therefore

$$\pi(x) = \frac{1}{|E|} = 2^{-m}, \qquad x \in E.$$

This chain is not aperiodic as, if x is such that $x_1 + \dots + x_m$ is even and if $y \in E$

is such that $P(x, y) > 0$, then necessarily $y_1 + \ldots + y_m$ must be odd. Its period is therefore ≥ 2. One observes moreover that, for $1 \leq i \leq m$,

$$P^2(x, x) \geq P(x, x^{(i)})P(x^{(i)}, x) = \frac{1}{m^2} > 0.$$

This implies that the set $I(x)$ of •4.31 contains the number 2 and therefore the period is ≤ 2. One concludes that the chain has period 2.

b) Suppose $\psi(x) = i$; then $P(x, y) > 0$ implies $\psi(y) = i - 1$ or $\psi(y) = i + 1$. More precisely there are i coordinates of x that may be changed from 1 into 0 and $m - i$ from 0 into 1. Therefore, if $\psi(x) = i$,

$$P(x, \psi^{-1}(j)) = \begin{cases} \frac{i}{m} & \text{if } j = i - 1 \\ \frac{m-i}{m} & \text{if } j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

In any case, $x \to P(x, \psi^{-1}(j))$ is a function depending only on $\psi(x) = i$. Dynkin's criterion applies and Y is a Markov chain with transition matrix

$$Q(i, j) = \begin{cases} \frac{i}{m} & \text{if } j = i - 1 \\ \frac{m-i}{m} & \text{if } j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

A comparison with Exercise 4.15 shows that Y is an Ehrenfest model. In order to compute the stationary probability, one can use Exercise 4.35 (b): it is the image $\tilde{\pi}$ by ψ of the stationary probability π of the random walk on the hypercube. Now, if $i \in F$, the set $\psi^{-1}(i)$ has cardinality $\binom{m}{i}$ (it is the number of vertices with i coordinates equal to 1). Therefore

$$\tilde{\pi}_i = \binom{m}{i} 2^{-m}. \tag{4.92}$$

c) One easily sees that the transition matrix defined in (4.58) is still irreducible and bistochastic. The uniform probability is therefore still stationary.

On the other hand, thanks to the aperiodicity criterion of •4.33, as now $P(x, x) = (m + 1)^{-1} > 0$, the chain is aperiodic and

$$P^{(n)}(x, y) \underset{n\to\infty}{\to} 2^{-m}$$

for every $x, y \in E$. By a similar computation as in (b) one gets, for $x \in E$ such that $\psi(x) = i$,

$$P(x, \psi^{-1}(j)) = \begin{cases} \frac{i}{m+1} & \text{if } j = i - 1 \\ \frac{1}{m+1} & \text{if } j = i \\ \frac{m-i}{m+1} & \text{if } j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

This is again a quantity depending only on $\psi(x) = i$. Therefore Y is still a Markov chain, associated to the transition matrix $Q(i, j) = P(x, \psi^{-1}(j))$ described above.

The chain is aperiodic, since $Q(i, i) > 0$ for every $i \in E$ (it would have been sufficient for this relation to hold for just one state i). Its stationary distribution remains the one defined in (4.92), as it is again the image of the uniform distribution by ψ.

♦ Dynkin's criterion has been useful in this exercise to determine the stationary probability of the Ehrenfest model; other methods are available to this aim: for instance one could note that it is a birth and death chain and apply Exercise 4.29. In Problem 4.14 we use Dynkin's criterion to study a model for which it is not so simple to find the stationary probability otherwise.

E4.37 a) Let $x, y \in E$, $x \neq y$. One may assume $\pi(x) \geq \pi(y)$. Then

$$Q(x, y) = P(x, y)\frac{\pi(y)}{\pi(x)}, \qquad Q(y, x) = P(y, x)$$

and, P being symmetric,

$$\pi(x)Q(x, y) = \pi(x)P(x, y)\frac{\pi(y)}{\pi(x)} = \pi(y)P(y, x) = \pi(y)Q(y, x).$$

π is therefore $\mathbf{Q}$-reversible (the relation is obvious if $x = y$).

b) It is clear, by the definition, that

$$P(x, y) > 0 \quad \Rightarrow \quad Q(x, y) > 0$$

and the irreducibility of Q is implied by the irreducibility of P, by (c) of Exercise 4.10.

c1) If $P(x, y) = 0$ for every $x \in M$, $y \notin M$, then M would be a closed class (Proposition 4.19). But, as one supposes that $x \to \pi(x)$ is not constant, $M \neq E$ and this is not possible since P is supposed to be irreducible. Therefore there exist $x_0 \in M$ and $y \notin M$ such that $P(x_0, y) > 0$. Since $\pi(x_0) > \pi(y)$,

$$Q(x_0, y) = P(x_0, y)\frac{\pi(y)}{\pi(x_0)} < P(x_0, y).$$

This implies $Q(x_0, x_0) > 0$ as, thanks to the definition of Q,

$$Q(x_0, x_0) = 1 - \sum_{z \neq x_0} Q(x_0, z) > 1 - \sum_{z \neq x_0} P(x_0, z) = P(x_0, x_0) \geq 0. \qquad (4.93)$$

where the strict inequality comes from the fact that, on one hand $Q(x, y) \leq P(x, y)$ for every $x \neq y$ and that, on the other hand, there is at least one state z, $z \neq x_0$ such that $Q(x, y) < P(x, y)$.

c2) An irreducible transition matrix with an element > 0 on the diagonal is aperiodic (•4.33).

E4.38 1) The first relation is proved easily by recurrence. If $X^n_{n+p}(i) = X^n_{n+p}(j)$, then, for $m \geq p$, $X^n_{n+m}(i) = X^{n+p}_{n+m}(X^n_{n+p}(i)) = X^{n+p}_{n+m}(X^n_{n+p}(j)) = X^n_{n+m}(j)$.

2) On one hand $X_0 = X^0_0(Z) = Z \sim \pi$. On the other, if $\mathscr{F}_n = \sigma(Z, U_0, \dots, U_{n-1})$, since $X^0_n(Z)$ is $\mathscr{F}_n$-measurable hence independent of U_n and thanks to Lemma 1.2,

$$\mathbf{P}(X_{n+1} = j \mid \mathscr{F}_n) = \mathbf{P}(g(X^0_n(Z), U_n) = j \mid \mathscr{F}_n) = \phi(X^0_n(Z)) \quad \text{a.s.}$$

where $\phi(i) = \mathbf{P}(g(i, U_n) = j) = Q(i, j)$, i.e., $\mathbf{P}(X_{n+1} = j \mid \mathscr{F}_n) = Q(X_n, j)$ a.s.

3) The first point is obvious. One gets $A_n \supset \{\text{for every } i \in E, X^n_{n+1}(i) = 0\} = \{\text{for every } i \in E, g(i, U_n) = 0\} \supset \{U_n \leq \alpha\}$ whence $\mathbf{P}(A_n) \geq \alpha$.

4) Let us remark first that $\{\nu > n\} \in \sigma(U_{-1}, U_{-2}, \dots, U_{-n})$ is independent of

U_{-n-1} and thus of A_{-n-1}. It holds $\{\nu \le n+1\} \supset \{\nu \le n\} \cup A_{-n-1}$ and thus $\{\nu > n+1\} \subset \{\nu > n\} \cap A^c_{-n-1}$. Hence

$$\mathbf{P}(\nu > n+1) \le \mathbf{P}(\nu > n)\mathbf{P}(A^c_{-n-1}) \le (1-\alpha)\mathbf{P}(\nu > n).$$

One gets $\mathbf{P}(\nu > n) \le (1-\alpha)^n$ and $\mathbf{P}(\nu < +\infty) = 1$.

5) On $\{\nu = m\}$ there exists $j \in E$ such that, for every $i \in E$, $X_0^{-m}(i) = j = Y$ whence, if $n \ge m$, $X_0^{-n}(Z) = X_0^{-m}(X_{-m}^{-n}(Z)) = j = Y$, and the first statement is proved. Then

$$\begin{aligned}&\left|\mathbf{P}(X_0^{-n}(Z) = j) - \mathbf{P}(Y = j)\right| \le \\ &\le \left|\mathbf{P}(X_0^{-n}(Z) = j, \nu \le n) - \mathbf{P}(Y = j, \nu \le n)\right| + 2\mathbf{P}(\nu > n) = 2\mathbf{P}(\nu > n).\end{aligned}$$

By (2), $X_n^0(Z) \sim \pi$ and, since $X_0^{-n}(Z)$ and $X_n^0(Z)$ have the same law, $X_0^{-n}(Z) \sim \pi$. Thus $|\pi(j) - \mathbf{P}(Y = j)| \le 2\mathbf{P}(\nu > n)$ and, as $\mathbf{P}(\nu > n) \to_{n\to\infty} 0$, $\pi(j) = \mathbf{P}(Y = j)$.

6) Since $s_k(i) = Q(i, [0, k-1]) = 1 - Q(i, [k, N])$, the relation $i \le i'$ implies $s_k(i) \ge s_k(i')$ for every $k \in E$ and $g(i, u) \le g(i', u)$ for every $u \in [0, 1]$. Thus, if $i \le i'$, it holds $X^n_{n+1}(i) \le X^n_{n+1}(i')$ and (4.60) is easily obtained. The last point is obvious.

7) For every k, $k \ne i$ one can write

$$Q(i, k) = \Big(\frac{\pi(k)}{\pi(i)} \wedge 1\Big)\frac{1}{N+1}.$$

Thus, if $i < i'$, $Q(i, k) \le Q(i', k)$ for every k different from i and from i'. This proves immediately the monotonicity relation (4.59) if $i' < k$. The latter is also immediate if $k \le i$, since then $Q(i, h) = Q(i', h) = \frac{1}{N+1}$ for every $h < k$, hence,

$$Q(i, [k, N]) = 1 - Q(i, [0, k[) = 1 - \frac{k}{N+1} = Q(i', [k, N]).$$

The case $i < k \le i'$ requires a bit more of attention. It holds $Q(i', [0, k[) = \frac{k}{N+1}$ whereas

$$\begin{aligned}Q(i, [0, k[) &= Q(i, [0, i[) + Q(i,]i, k[) + Q(i, i) = \\ &= \frac{i}{N+1} + \frac{1}{N+1}\sum_{i<j<k}\frac{\pi(j)}{\pi(i)} + 1 - \frac{i}{N+1} - \frac{1}{N+1}\sum_{j>i}\frac{\pi(j)}{\pi(i)} = \\ &= 1 - \frac{1}{N+1}\sum_{j\ge k}\frac{\pi(j)}{\pi(i)} \ge 1 - \frac{1}{N+1}(N+1-k) = \frac{k}{N+1}\end{aligned}$$

which concludes easily.

E4.39 1) τ is a stopping time since

$$\{\tau = n\} = \{X_1 = X_0, \dots, X_{n-1} = X_0, X_n \ne X_0\} \in \mathscr{F}_n.$$

One has next

$$\mathbf{P}_x(\tau \ge n+1) = \mathbf{P}_x(X_1 = x, \dots, X_n = x) = Q(x, x)^n.$$

Therefore $\mathbf{P}_x(\tau = \infty) = \lim_{n\to\infty} Q(x, x)^n = 0$, as it is assumed that $Q(x, x) < 1$.

Moreover, for $n \geq 1$, by subtraction

$$\mathbf{P}_x(\tau = n) = Q(x, x)^{n-1}(1 - Q(x, x)).$$

Moreover $\mathbf{P}_x(X_\tau = x) = 0$ and, for $y \neq x$,

$$\mathbf{P}_x(X_\tau = y) = \sum_{n \geq 1} \mathbf{P}_x(\tau = n, X_n = y) =$$

$$= \sum_{n \geq 1} \mathbf{P}_x(X_1 = x, X_2 = x, \dots, X_{n-1} = x, X_n = y) =$$

$$= \sum_{n \geq 1} Q(x, x)^{n-1} Q(x, y) = \frac{Q(x, y)}{1 - Q(x, x)}.$$

2) The fact that τ_n is a stopping time and its expression follow by Exercise 2.2. Let $n \geq 0$. Suppose that τ_n is $\mathbf{P}_x$-a.s. finite for every $x \in E$. Therefore, by the strong Markov property,

$$\mathbf{P}_x(\tau_{n+1} < +\infty) = \mathbf{P}_x(\tau_n < +\infty, \tau \circ \theta_{\tau_n} < +\infty) =$$
$$= \mathbf{E}_x(1_{\{\tau_n < +\infty\}} \underbrace{\mathbf{P}_{X_{\tau_n}}(\tau < +\infty)}_{=1}) = \mathbf{P}_x(\tau_n < +\infty) = 1 \quad \text{a.s.}$$

3) Let $n \geq 0$. If $y \neq x$, by the strong Markov property and using the relation $X_{\tau_{n+1}} = X_\tau \circ \theta_{\tau_n}$ (see again Exercise 2.2),

$$\mathbf{P}_x(Y_{n+1} = y \mid \mathscr{F}_{\tau_n}) = \mathbf{P}_x(X_{\tau_{n+1}} = y \mid \mathscr{F}_{\tau_n}) =$$
$$= \mathbf{P}_x(X_\tau \circ \theta_{\tau_n} = y \mid \mathscr{F}_{\tau_n}) = \mathbf{P}_{X_{\tau_n}}(X_\tau = y) = \frac{Q(X_{\tau_n}, y)}{1 - Q(X_{\tau_n}, X_{\tau_n})}.$$

This shows that $(Y_n)_{n \geq 0}$ is a Markov chain with transition matrix

$$\tilde{Q}(x, y) = \frac{Q(x, y)}{1 - Q(x, x)}, \; y \neq x, \qquad \tilde{Q}(x, x) = 0. \tag{4.94}$$

4) One observes that, if $X_{\tau_{n-1}} = x$, then $X_k = x$ for every $\tau_{n-1} \leq k < \tau_n$ and, between the times τ_{n-1} and $\tau_n - 1$, X takes the only value Y_{n-1}.

The states visited by $(Y_k)_{k \geq 0}$ coincide with those visited by $(X_k)_{k \geq 0}$. Therefore, if X visits x with probability 1, the same holds true for Y: thus if X is irreducible recurrent, this is also the case for Y.

One has

$$\tilde{\mu}\tilde{Q}(x) = \sum_{y \in E} \tilde{\mu}(y)\tilde{Q}(y, x) = \sum_{y, y \neq x} \tilde{\mu}(y)\tilde{Q}(y, x) = \sum_{y, y \neq x} \mu(y) Q(y, x) =$$
$$= \mu(x) - \mu(x) Q(x, x) = \mu(x)(1 - Q(x, x)) = \tilde{\mu}(x)$$

and $\tilde{\mu}$ is stationary.

E4.40 a) Going back to Exercise 2.2, one sees that it is also possible to write $\rho_{n+1} = \rho_n + \rho_1 \circ \theta_{\rho_n}$ and that the considered r.v.'s are stopping times.

The hypothesis $U(x, x) = \sum_{n \geq 0} P^n(x, x) = +\infty$ implies that x is recurrent. Therefore, the chain starting at x visits an infinite number of times the state x, and the stopping times $\rho_1, \rho_2, \dots$ are finite $\mathbf{P}_x$-a.s.

b) Let $y_0, y_1, \dots, y_n$ be elements of E. One has

$$\begin{aligned}\mathbf{P}_x(Z_0 = y_0, Z_1 = y_1, \dots, Z_n = y_n) = \\ = \mathbf{P}_x(X_1 = y_0, X_{\rho_1+1} = y_1, \dots, X_{\rho_n+1} = y_n) = \\ = \mathbf{E}_x(1_{\{X_1=y_0, X_{\rho_1+1}=y_1,\dots,X_{\rho_n+1}=y_n\}}).\end{aligned}$$

As

$$1_{\{X_1=y_0, X_{\rho_1+1}=y_1,\dots,X_{\rho_n+1}=y_n\}} = 1_{\{X_1=y_0, X_{\rho_1+1}=y_1,\dots\}} 1_{\{X_1=y_n\}} \circ \theta_{\rho_n},$$

one can use the strong Markov property: (4.19) gives

$$\begin{aligned}\mathbf{E}_x(1_{\{X_1=y_0, X_{\rho_1+1}=y_1,\dots,X_{\rho_n+1}=y_n\}}) = \\ = \mathbf{E}_x(1_{\{X_1=y_0, X_{\rho_1+1}=y_1,\dots,X_{\rho_{n-1}+1}=y_{n-1}\}} \mathbf{P}_{X_{\rho_n}}(X_1 = y_n)) = \\ = \mathbf{E}_x(1_{\{X_1=y_0, X_{\rho_1+1}=y_1,\dots,X_{\rho_{n-1}+1}=y_{n-1}\}}) \underbrace{\mathbf{P}_x(X_1 = y_n)}_{=P(x,y_n)},\end{aligned}$$

as $X_{\rho_n} = x$ $\mathbf{P}_x$-a.s. By induction,

$$\mathbf{P}_x(Z_0 = y_0, Z_1 = y_1, \dots, Z_n = y_n) = \prod_{i=1}^{n} P(x, y_i).$$

This simultaneously shows that the r.v.'s $(Z_n)_{n\geq 0}$ are independent and that they have the same distribution on E, given by $P(x, \cdot)$.

Alternatively, just remark that

$$\begin{aligned}\mathbf{P}_x(Z_n = y \mid \mathscr{F}_{\rho_n}) = \mathbf{P}_x(X_{\rho_n+1} = y \mid \mathscr{F}_{\rho_n}) = \\ = \mathbf{P}_x(X_1 \circ \theta_{\rho_n} = y \mid \mathscr{F}_{\rho_n}) = \mathbf{P}_{X_{\rho_n}}(X_1 = y) = \mathbf{P}_x(X_1 = y) = P(x, y).\end{aligned}$$

As $(Z_1, \dots, Z_{n-1})$ is $\mathscr{F}_{\rho_n}$-measurable, this implies that the r.v.'s $(Z_n)_{n\geq 1}$ are independent and have distribution $P(x, \cdot)$.

E4.41 1) One has $\frac{1}{n}\mathbf{E}(S_n) = \mathbf{E}(\frac{1}{n}\sum_{k=0}^{n} f(X_k))$. The chain being positive recurrent, the ergodic Theorem 4.25 asserts that $\frac{1}{n}\sum_{k=0}^{n} f(X_k) \to_{n\to\infty} \langle \pi, f\rangle = 0$ a.s. Observe that, in absolute value, $\frac{1}{n}S_n$ is dominated by $2\|f\|_\infty$ and thus, by Lebesgue's theorem, $\frac{1}{n}\mathbf{E}(S_n) \to_{n\to\infty} 0$.

2) As $f = (I - Q)g$, one immediately has

$$\begin{aligned}S_n = \sum_{k=0}^{n} f(X_k) = \sum_{k=0}^{n} g(X_k) - Qg(X_k) = \\ = g(X_0) + \sum_{k=1}^{n}[g(X_k) - Qg(X_{k-1})] - Qg(X_n) = M_n + Z_n.\end{aligned}$$

By the Markov property

$$\mathbf{E}(U_k \mid \mathscr{F}_{k-1}) = \mathbf{E}(g(X_k) \mid \mathscr{F}_{k-1}) - Qg(X_{k-1}) = 0.$$

This relation implies that $(M_n)_{n\geq 0}$ is a martingale.

Relation (4.61), which is easily checked, is a typical property of martingales (the quadratic mean of a square integrable martingale is equal to the sum of the quadratic means of its increments); see •3.8.

3) Easily the Markov property gives

$$\begin{aligned}\mathbf{E}(g(X_{k+1})Qg(X_k)) &= \mathbf{E}(\mathbf{E}(g(X_{k+1}) \mid \mathscr{F}_k)Qg(X_k)) = \\ &= \mathbf{E}(Qg(X_k)\cdot Qg(X_k)) = \mathbf{E}((Qg)^2(X_k)).\end{aligned}$$

Therefore

$$\frac{1}{n}\mathbf{E}\Big(\sum_{k=0}^{n-1} g(X_{k+1})Qg(X_k)\Big) = \frac{1}{n}\sum_{k=0}^{n-1}\mathbf{E}((Qg)^2(X_k)) = \frac{1}{n}\mathbf{E}\Big(\sum_{k=0}^{n-1}(Qg)^2(X_k)\Big).$$

Remark that $Qg = g - f$; Qg is thus a bounded function. By the same method as in (1), applied to $f = (Qg)^2$,

$$\frac{1}{n}\mathbf{E}\Big(\sum_{k=0}^{n-1} g(X_{k+1})Qg(X_k)\Big) \underset{n\to\infty}{\to} \langle (Qg)^2, \pi\rangle.$$

4) Thanks to (3)

$$\begin{aligned}\mathbf{E}(U_{k+1}^2) &= \mathbf{E}(g(X_{k+1})^2) + \mathbf{E}(Qg(X_k)^2) - 2\mathbf{E}(g(X_{k+1})Qg(X_k)) = \\ &= \mathbf{E}(g(X_{k+1})^2) - \mathbf{E}(Qg(X_k)^2);\end{aligned}$$

therefore

$$\frac{1}{n}\mathbf{E}(M_n^2) = \frac{1}{n}\mathbf{E}\Big(\sum_{k=0}^{n-1} g(X_{k+1})^2\Big) - \frac{1}{n}\mathbf{E}\Big(\sum_{k=0}^{n-1} Qg(X_k)^2\Big).$$

Hence, by the same argument as before,

$$\frac{1}{n}\mathbf{E}(M_n^2) \underset{n\to\infty}{\to} \langle g^2 - (Qg)^2, \pi\rangle.$$

5) $S_n^2 = M_n^2 + Z_n^2 + 2M_nZ_n$. Observe that $|Z_n|$ is a r.v. bounded by a constant K (as g and Qg are bounded). Therefore, using the Cauchy–Schwarz inequality,

$$\mathbf{E}(Z_n^2 + 2M_nZ_n) \le K^2 + 2K\,\mathbf{E}(M_n^2)^{1/2}.$$

Now, as $\mathbf{E}(M_n^2)^{1/2}n^{-1/2} \to_{n\to\infty} \Sigma$, $\mathbf{E}(M_n^2)^{1/2}n^{-1} \to_{n\to\infty} 0$. Therefore

$$\lim_{n\to\infty}\frac{1}{n}\mathbf{E}(S_n^2) = \lim_{n\to\infty}\frac{1}{n}\mathbf{E}(M_n^2) = \Sigma^2.$$

E4.42 1) Let $\bar{z}, \bar{x}_0, \ldots, \bar{x}_{n-1}, \bar{x}, \bar{y} \in \bar{E} = E \times E'$ be such that $\bar{\mathbf{P}}_{\bar{z}}(\bar{X}_n = \bar{x}, \ldots, \bar{X}_0 = \bar{x}_0) > 0$. One must show that, for a transition matrix $\bar{Q}$ to be determined,

$$\bar{\mathbf{P}}_{\bar{z}}(\bar{X}_{n+1} = \bar{y} \mid \bar{X}_n = \bar{x}, \ldots, \bar{X}_0 = \bar{x}_0) = \bar{Q}(\bar{x}, \bar{y}).$$

The left-hand side is the quotient of

$$\begin{aligned}&\bar{\mathbf{P}}_{\bar{z}}(\bar{X}_0 = \bar{x}_0, \ldots, \bar{X}_n = \bar{x}, \bar{X}_{n+1} = \bar{y}) = \\ &= \mathbf{P}_z(X_0 = x_0, \ldots, X_n = x, X_{n+1} = y)\times \\ &\qquad \times \mathbf{P}'_{z'}(X'_0 = x'_0, \ldots, X'_n = x', X'_{n+1} = y')\end{aligned}$$

and

$$\begin{aligned}&\bar{\mathbf{P}}_{\bar{z}}(\bar{X}_0 = \bar{x}_0, \ldots, \bar{X}_n = \bar{x}) = \\ &= \mathbf{P}_z(X_0 = x_0, \ldots, X_n = x)\mathbf{P}'_{z'}(X'_0 = x'_0, \ldots, X'_n = x').\end{aligned}$$

Thanks to the Markov property for the processes X, X', this quotient is equal to

$$\bar{Q}(\bar{x}, \bar{y}) \stackrel{\text{def}}{=} Q(x, y)Q'(x', y').$$

2) One has

$$\bar{Q}^2(\bar{x}, \bar{y}) = \sum_{\bar{z}\in\bar{E}} \bar{Q}(\bar{x}, \bar{z})\bar{Q}(\bar{z}, \bar{y}) =$$

$$= \sum_{z\in E, z'\in E'} Q(x, z)Q'(x', z')Q(z, y)Q'(z', y') = Q^2(x, y)Q'^2(x', y')$$

and, by induction, $\bar{Q}^n(\bar{x}, \bar{y}) = Q^n(x, y)Q'^n(x', y')$.

3) If $\bar{\nu} = \nu \otimes \nu'$,

$$\bar{\nu}\bar{Q}(\bar{y}) = \sum_{\bar{x}\in\bar{E}} \bar{\nu}(\bar{x})\bar{Q}(\bar{x}, \bar{y}) = \sum_{x\in E, x'\in E'} \nu(x)\nu'(x')Q(x, y)Q'(x', y') =$$

$$= \nu Q(y)\nu'Q'(y') = \nu(y)\nu'(y') = \bar{\nu}(\bar{y}).$$

4) As $\bar{a} = (a, a')$ is recurrent, $\sum_{n=1}^{\infty} \bar{Q}^n(\bar{a}, \bar{a}) = +\infty$ but

$$\sum_{n=1}^{\infty} \bar{Q}^n(\bar{a}, \bar{a}) = \sum_{n=1}^{\infty} Q^n(a, a)Q'^n(a', a') \leq$$

$$\leq \min\Big(\sum_{n=1}^{\infty} Q^n(a, a), \sum_{n=1}^{\infty} Q'^n(a', a')\Big),$$

as $Q^n(a, a) \leq 1$ and $Q'^n(a', a') \leq 1$; therefore

$$\sum_{n=1}^{\infty} Q^n(a, a) = \sum_{n=1}^{\infty} Q'^n(a', a') = +\infty$$

and a and a' are recurrent for X and X', respectively.

5) It holds $Q^{2n+1}(0, 0) = 0$, $Q^{2n}(0, 0) = \binom{2n}{n}2^{-2n} \sim (2\pi n)^{-1/2}$ thanks to Stirling's formula; thus $\sum_{n=1}^{\infty} Q^n(0, 0) = +\infty$ and 0 is recurrent (see Exercise 4.16 for details).

6) One chooses X as in (5) and let X' and X'' be two copies of X. It results by (5) that 0 is recurrent for X; taking into account (2) and (5),

$$\bar{Q}^{2n}((0, 0), (0, 0)) = Q^{2n}(0, 0)^2 \sim (2\pi n)^{-1},$$

which is the term of a divergent series, implying that $(0, 0)$ is recurrent for $X' \otimes X''$.

Consider now the chain $\bar{X} = X \otimes (X' \otimes X'')$ and denote by $\bar{Q}$ its transition matrix transition. By (2),

$$\bar{Q}^{2n}((0, (0, 0)), (0, (0, 0))) = Q^{2n}(0, 0)^3 \sim (2\pi n)^{-3/2}$$

and $\bar{Q}^{2n+1}((0, (0, 0)), (0, (0, 0))) = 0$. As $\bar{Q}^n((0, (0, 0)), (0, (0, 0)))$ is the term of a convergent series, the state $(0, (0, 0))$ is transient for $\bar{X}$ whereas 0 and $(0, 0)$ are recurrent for X and $X' \otimes X''$, respectively.

E4.43 1) If $x \in F$, $\rho \geq 0$ $\mathbf{P}_x$-a.s. As F is finite, if N_F is the number of visits in F

(•4.13),

$$\mathbf{E}_x(N_F) = \mathbf{E}_x\Big(\sum_{y\in F} N_y\Big) = \sum_{y\in F} U(x,y) < +\infty.$$

Therefore $N_F < +\infty$ and $\rho < +\infty$, $\mathbf{P}_x$-a.s.

2) Thanks to the Markov property,

$$\mathbf{P}_x(\rho = n) = \mathbf{P}_x(X_n \in F,\, X_{n+k} \notin F \text{ for every } k > 0) =$$
$$= \mathbf{E}_x[1_F(X_n) 1_{\{X_k\notin F \text{ for every } k>0\}} \circ \theta_n] =$$
$$= \mathbf{E}_x[1_F(X_n)\mathbf{P}_{X_n}(X_k \notin F \text{ for each } k > 0)] = \mathbf{E}_x(v(X_n)) = Q^n v(x).$$

Therefore, if $x \in F$,

$$Uv(x) = \sum_{k=0}^{\infty} Q^n v(x) = \sum_{k=0}^{\infty} \mathbf{P}_x(\rho = n) = \mathbf{P}_x(0 \le \rho < +\infty) = 1.$$

3a) Remark that, by the definition of Y_n, $\{Y_0 = x_0, \dots, Y_n = x_n\} \subset \{\rho \ge n\}$. Then

$$\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n) = \mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n, \rho \ge n) =$$
$$= \sum_{k=n}^{\infty} \mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n, \rho = k) =$$
$$= \sum_{k=n}^{\infty} \mathbf{P}_\nu(X_{k-n} = x_n, \dots, X_k = x_0, \rho = k).$$

But $1_{\{\rho=k\}} = 1_{\{\rho=0\}} \circ \theta_k$ and for $k \ge n$, taking (2) into account,

$$\mathbf{P}_\nu(X_{k-n} = x_n, \dots, X_k = x_0, \rho = k) = \mathbf{E}_\nu(1_{\{X_{k-n}=x_n,\dots,X_k=x_0\}} 1_{\{\rho=0\}} \circ \theta_k) =$$
$$= \mathbf{E}_\nu(1_{\{X_{k-n}=x_n,\dots,X_k=x_0\}} \mathbf{P}_{X_k}(\rho = 0)) =$$
$$= \mathbf{P}_\nu(X_{k-n} = x_n, \dots, X_k = x_0) v(x_0) =$$
$$= \nu Q^{k-n}(x_n) Q(x_n, x_{n-1}) \dots Q(x_1, x_0) v(x_0),$$

whence the requested formula by adding up in k from n to ∞. Similarly

$$\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n, Y_{n+1} = \partial) = \mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n, \rho = n) =$$
$$= \mathbf{P}_\nu(X_0 = x_n, \dots, X_n = x_0, \rho = n) =$$
$$= \nu(x_n) Q(x_n, x_{n-1}) \dots Q(x_1, x_0) v(x_0).$$

3b) For every $x_0, \dots, x_{n+1} \in E_\partial$ such that $\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n) > 0$,

$$p_n = \mathbf{P}_\nu(Y_{n+1} = x_{n+1} \mid Y_0 = x_0, \dots, Y_n = x_n) =$$
$$= \frac{\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_{n+1} = x_{n+1})}{\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n)}.$$

If $x_n = \partial$, by the definition of Y_n, $p_n = 1$ if $x_{n+1} = \partial$, $p_n = 0$ if $x_{n+1} \ne \partial$. Suppose $x_n \in E$, then $\mathbf{P}_\nu(Y_0 = x_0, \dots, Y_n = x_n) > 0$ implies that $x_0, \dots, x_{n-1} \in E$ and, applying (3a),

(i) if $x_{n+1} \in E$

$$p_n = \frac{\nu U(x_{n+1}) Q(x_{n+1}, x_n)}{\nu U(x_n)}.$$

(ii) If $x_{n+1} = \partial$,

$$p_n = \frac{\nu(x_n)}{\nu U(x_n)}.$$

It has therefore been proved that $((Y_n)_{n\geq 0}, \mathbf{P}_\nu)$ is a Markov chain with transition matrix $\bar{Q}$ given, for $x, y \in E$, by

$$\bar{Q}(x, y) = \frac{\nu U(y) Q(y, x)}{\nu U(x)}, \quad \bar{Q}(x, \partial) = \frac{\nu(x)}{\nu U(x)}, \quad \bar{Q}(\partial, \partial) = 1.$$

Finally, applying (3a), $\mathbf{P}_\nu(Y_0 = x_0) = \nu U(x_0) v(x_0)$; the support of this distribution is obviously contained in F.

P4.1 A1) By induction $P^n f \geq P^{n+1} f$ for every $n \geq 1$; therefore the limit $\bar{f} = \lim_{n\to\infty} P^n f(x)$ exists and

$$P\bar{f} = P\Big(\lim_{n\to\infty} P^n f\Big) = \lim_{n\to\infty} P^{n+1} f = \bar{f}$$

where the exchange between the operators P and lim is a consequence of Lebesgue's theorem, using the upper bounds $0 \leq P^n f \leq f$ and $Pf \leq f < +\infty$. Finally

$$Uh = U(f - Pf) = \lim_{n\to\infty} \sum_{k=0}^{n} P^k (f - Pf) = \lim_{n\to\infty} (f - P^{n+1} f) = f - \bar{f}.$$

A2) As $U = I + PU$ (•4.13),

$$g_1 + Pf = g_1 + Pg_0 + PUg_1 = g_0 + Ug_1 = f.$$

Hence $g_1 = f - Pf = h$ and $g_0 = \bar{f}$.

A3) Recalling •4.15, the two sequences $(f(X_n))_{n\geq 0}$, $(\bar{f}(X_n))_{n\geq 0}$ are positive supermartingales (the second one is even a martingale) and are therefore convergent to a finite limit (Theorem 3.8). Then

$$\mathbf{E}_x(Uh(X_n)) = P^n Uh(x) = \sum_{k\geq n} P^k h(x) \underset{n\to\infty}{\to} 0 \tag{4.95}$$

as, $Uh(x)$ being finite, the series $\sum_{k=0}^{\infty} P^k h(x)$ is convergent. The positive supermartingale $(Uh(X_n))_{n\geq 0}$ converges to a finite limit $\mathbf{P}_x$-a.s. for every $x \in E$; since (4.95) implies that this sequence converges to 0 in L^1, $Uh(X_n) \to_{n\to\infty} 0$ $\mathbf{P}_x$-a.s. for every x. As $Uh(X_n) = f(X_n) - \bar{f}(X_n)$, the two limits in (4.62) coincide.

B1) One has $\{\tau_F < +\infty\} = \bigcup_{n\geq 0}\{X_n \in F\}$; hence, thanks to the Markov property,

$$\begin{aligned} P^n \phi_F(x) &= \mathbf{E}_x(\phi_F(X_n)) = \mathbf{E}_x(\mathbf{P}_{X_n}(\tau_F < +\infty)) = \\ &= \mathbf{P}_x(\tau_F \circ \theta_n < +\infty) = \mathbf{P}_x\Big(\bigcup_{k\geq n}\{X_k \in F\}\Big). \end{aligned}$$

One deduces that

$$P\phi_F(x) = \mathbf{P}_x\Big(\bigcup_{k\geq 1}\{X_k \in F\}\Big) \leq \mathbf{P}_x(\tau_F < +\infty) = \phi_F(x),$$

which shows that ϕ_F is superharmonic.

B2) It holds

$$\bar{\phi}_F(x) = \lim_{n\to\infty} \downarrow P^n\phi_F(x) = \lim_{n\to\infty} \downarrow \mathbf{P}_x\Big(\bigcup_{k\geq n}\{X_k \in F\}\Big) = \\ = \mathbf{P}_x\Big(\bigcap_{n\geq 0}\bigcup_{k\geq n}\{X_k \in F\}\Big) = \mathbf{P}_x(N_F = +\infty),$$

which proves the first relation. By (A3), $\lim_{n\to\infty}\phi_F(X_n) = \lim_{n\to\infty}\bar{\phi}_F(X_n)$ and, using the Markov property,

$$\bar{\phi}_F(X_n) = \mathbf{P}_{X_n}(N_F = +\infty) = \mathbf{P}_x(N_F \circ \theta_n = +\infty \mid \mathscr{F}_n) = \\ = \mathbf{P}_x(N_F = +\infty \mid \mathscr{F}_n),$$

as the events $\{N_F \circ \theta_n = +\infty\}$ and $\{N_F = +\infty\}$ coincide a.s. Thanks to Theorem 3.10,

$$\bar{\phi}_F(X_n) \underset{n\to\infty}{\to} 1_{\{N_F=+\infty\}}$$

$\mathbf{P}_x$-a.s. for every $x \in E$, since $\{N_F = +\infty\}$ is $\mathscr{F}_\infty$-measurable. As for the third relation,

$$h_F(x) = \phi_F(x) - P\phi_F(x) = \mathbf{P}_x\Big(\bigcup_{k\geq 0}\{X_k \in F\} \setminus \bigcup_{k\geq 1}\{X_k \in F\}\Big) = \\ = \mathbf{P}_x\Big(\{X_0 \in F\} \cap \Big(\bigcap_{k\geq 1}\{X_k \notin F\}\Big)\Big) = 1_F(x)\mathbf{P}_x\Big(\bigcap_{n\geq 1}\{X_n \notin F\}\Big).$$

P4.2 1a) It holds $I_B f(x) = \sum_y I_B(x,y)f(y) = 1_B(x)f(x)$ and

$$(I_B P)f(x) = I_B(Pf)(x) = \\ = 1_B(x)Pf(x) = 1_B(x)\mathbf{E}_x(f(X_1)) = \mathbf{E}_x(1_B(X_0)f(X_1)).$$

1b) The formula is true for $n = 1$. Let us assume it true for n. Then

$$(I_B P)^{n+1}f(x) = I_B P(I_B P)^n f(x) = \mathbf{E}_x[1_B(X_0)(I_B P)^n f(X_1)] = \\ = \mathbf{E}_x[1_B(X_0)\mathbf{E}_{X_1}(1_B(X_0)1_B(X_1)\dots 1_B(X_{n-1})f(X_n))] = \\ = \mathbf{E}_x[1_B(X_0)\{1_B(X_0)1_B(X_1)\dots 1_B(X_{n-1})f(X_n)\} \circ \theta_1] = \\ = \mathbf{E}_x[1_B(X_0)1_B(X_1)1_B(X_2)\dots 1_B(X_n)f(X_{n+1})].$$

2a) Using (1b),

$$\mathbf{E}_x[1_{\{\tau=n\}}f(X_n)] = \mathbf{E}_x[1_{\{X_0\in A^c, X_1\in A^c,\dots,X_{n-1}\in A^c, X_n\in A\}}f(X_n)] = \\ = \mathbf{E}_x[1_{A^c}(X_0)1_{A^c}(X_1)\dots 1_{A^c}(X_{n-1})1_A(X_n)f(X_n)] = \\ = (I_{A^c}P)^n(1_A f)(x) = (I_{A^c}P)^n I_A f(x).$$

Thus

$$P_A f(x) = \mathbf{E}_x[1_{\{\tau<+\infty\}}f(X_\tau)] = \sum_{n\geq 0}\mathbf{E}_x[1_{\{\tau=n\}}f(X_n)] = \\ = \sum_{n\geq 0}(I_{A^c}P)^n I_A f(x).$$

2b) Similarly, for $x \in A^c$,

$$\begin{aligned}\mathbf{E}_x[1_{\{\tau>n\}}f(X_n)] &= \mathbf{E}_x[1_{\{X_0\in A^c, X_1\in A^c,\dots,X_{n-1}\in A^c, X_n\in A^c\}}f(X_n)] = \\ &= \mathbf{E}_x[1_{A^c}(X_0)1_{A^c}(X_1)\dots 1_{A^c}(X_{n-1})1_{A^c}(X_n)f(X_n)] \\ &= (I_{A^c}P)^n(1_{A^c}f)(x) = (I_{A^c}P)^n I_{A^c}f(x).\end{aligned}$$

The formula is obvious for $x \in A$ whereas, for $x \in A^c$,

$$U_A f(x) = \sum_{n\geq 0}\mathbf{E}_x[1_{\{\tau>n\}}f(X_n)] = \sum_{n\geq 0}(I_{A^c}P)^n I_{A^c}f(x).$$

Moreover,

$$\sum_{n\geq 0}(I_{A^c}P)^n I_{A^c} = I_{A^c} + I_{A^c}PI_{A^c} + I_{A^c}PI_{A^c}PI_{A^c} + \dots = \sum_{n\geq 0} I_{A^c}(PI_{A^c})^n I_{A^c}.$$

3) It holds, by (2a),

$$\begin{aligned}P_A &= I_A + I_{A^c}PI_A + I_{A^c}PI_{A^c}PI_A + \dots = \\ &= I_A + I_{A^c}P[I_A + I_{A^c}PI_A + I_{A^c}PI_{A^c}PI_A + \dots] = I_A + I_{A^c}PP_A\end{aligned}$$

and in the same way

$$\begin{aligned}U_A &= I_{A^c} + I_{A^c}PI_{A^c} + I_{A^c}PI_{A^c}PI_{A^c} + \dots = \\ &= I_{A^c} + I_{A^c}P[I_{A^c} + I_{A^c}PI_{A^c} + \dots] = I_{A^c} + I_{A^c}PU_A.\end{aligned}$$

4a) Using the relations obtained in (3),

$$\begin{aligned}u &= I_A g + I_{A^c}PP_A g + I_{A^c}h + I_{A^c}PU_A h = \\ &= I_A g + I_{A^c}(h + P(P_A g + U_A h)) = I_A g + I_{A^c}(h + Pu),\end{aligned}$$

i.e., $u(x) = 1_A(x)g(x) + 1_{A^c}(x)(h(x) + Pu(x))$; this shows that u is a solution of (4.63).

4b) If v is a positive solution of (4.63), then $v = I_A g + I_{A^c}(h + Pv)$. As $v \geq 0$, $Pv \geq 0$ and $v \geq I_A g + I_{A^c}h$; hence (4.64) is true for $n = 0$. Assume that (4.64) is true at rank n. Then,

$$\begin{aligned}v &= I_A g + I_{A^c}h + I_{A^c}Pv \geq \\ &\geq I_A g + I_{A^c}h + I_{A^c}P\Big[\sum_{k=0}^{n}(I_{A^c}P)^k(I_A g + I_{A^c}h)\Big] \geq \\ &\geq I_A g + I_{A^c}h + \sum_{k=1}^{n+1}(I_{A^c}P)^k(I_A g + I_{A^c}h) = \\ &= \sum_{k=0}^{n+1}(I_{A^c}P)^k(I_A g + I_{A^c}h).\end{aligned}$$

Therefore

$$v \geq \sum_{k\geq 0}(I_{A^c}P)^k(I_A g + I_{A^c}h) = P_A g + U_A h = u.$$

5) Clearly if $h = 1$ and $x \in A^c$ then $\sum_{n=0}^{\tau-1}h(X_n) = \tau$ and $\mathbf{E}_x(\tau) = U_A 1(x)$.

By (4a) and (4b), $v(x) = \mathbf{E}_x(\tau)$ is the smallest positive solution of (4.29) (here $g = 0, h = 1$).

P4.3 A1) If $Z_n = i$, the number of particles at time $n+1$ is the sum of numbers of the particles generated by each of the i particles present at time n. As they are independent, the conditional distribution of Z_{n+1} given $Z_n = i$ is the convolution of μ with itself i times.

A2) (4.66) is obvious if $i = 0$, as $P(0,0) = 1$ and $P(0, j) = 0$ for every $j \geq 1$. Hence

$$\sum_{j=0}^{\infty} P(0, j)s^j = 1 = f(s)^0.$$

For $i \geq 1$

$$\sum_{j=0}^{\infty} P(i, j)s^j = \sum_{j=0}^{\infty} \mu^{*i}(j)s^j$$

and one knows that, if $s \to f(s)$ is the generating function of μ, then $s \to f(s)^i$ is the generating function of μ^{*i}. Hence the formula.

Clearly 0 is absorbing. If $p_0 > 0$, $P(i, 0) = f(0)^i = p_0^i > 0$. Therefore all the states lead to 0. As $0 \not\to i$ for every $i \geq 1$, all the states $i \geq 1$ are transient, thanks to the criterion of •4.21. Conversely, if $p_0 = 0$, it is easily seen that the chain starting at i does not come back to i a.s., which still implies the transience of every state $i \geq 1$.

A3) The relation is true for $n = 1$ by (A2). Let us assume it to hold for n, $n \geq 1$. By definition, $P^{n+1}(i, j) = \sum_{k \geq 0} P^n(i,k)P(k, j)$. Then,

$$\begin{aligned}\sum_{j=0}^{\infty} P^{n+1}(i, j)s^j &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} P(i,k)P^n(k, j)s^j = \\ &= \sum_{k=0}^{\infty} P(i,k)\sum_{j=0}^{\infty} P^n(k, j)s^j.\end{aligned} \tag{4.96}$$

But, by the induction hypothesis, $\sum_{j \geq 0} P^n(k, j)s^j = f_n(s)^k$, so that (4.96) can be written,

$$\sum_{j=0}^{\infty} P^{n+1}(i, j)s^j = \sum_{k=0}^{\infty} P(i,k)f_n(s)^k = [f(f_n(s))]^i = f_{n+1}(s)^i.$$

A4) Just remark that

$$\{\tau_0 \leq n\} = \bigcup_{k=0}^{n}\{Z_k = 0\}$$

and that, a.s., $\{Z_k = 0\} \subset \{Z_n = 0\}$, $0 \leq k \leq n$, since $\{0\}$ is a closed class. Hence

$$\mathbf{P}(\tau_0 < +\infty) = \lim_{n\to\infty} \mathbf{P}(\tau_0 \leq n) = \lim_{n\to\infty} \mathbf{P}(Z_n = 0).$$

Now the distribution of Z_n is given by $P^n(1, \cdot)$ (at the beginning the presence of a single particle is assumed). Hence $\mathbf{P}(Z_n = 0) = P^n(1,0) = f_n(0)$.

B1) The function f is, on $[0, 1]$, the sum of convex increasing functions. If $p_0 < 1$, at least one of these functions is strictly increasing and if $p_0 + p_1 < 1$ at least one of these functions is strictly convex.

B2) It holds $q_n = f(q_{n-1})$. Thus $q_1 = f(0) = p_0 > 0$, $q_2 = f(q_1) \geq f(0) = q_1$ and, by recurrence, the sequence $(q_n)_{n\geq 1}$ is increasing. Also, f being increasing on $[0, 1]$ with $f(1) = 1$, necessarily $q_n = f_n(0) \leq 1$ for every $n \geq 1$. Therefore $(q_n)_{n\geq 1}$ converges to a limit q and, f being continuous, $q = f(q)$. Moreover, if $r \geq 0$ is such that $f(r) = r$, then $r = f(r) \geq f(0) = q_1$ and, by recurrence, $r \geq q_n$ for every $n \geq 1$ and $r \geq q$. q is thus the smallest fixed point of f on $[0, 1]$.

B3) Let $\phi(t) = f(t) - 1$. Then $\phi(0) = p_0 > 0$ and $\phi(1) = 0$. Moreover $\phi' = f' - 1$ is a strictly increasing function on $[0, 1]$, $\phi'(0) = p_1 - 1 < 0$ and $\phi'(1) = m - 1$.

Assume $m \leq 1$. Then $\phi'(1) \leq 0$. Therefore, for $0 \leq t < 1$, $\phi'(t) < 0$ and, as $\phi(1) = 0$, $\phi(t) > 0$, i.e., $f(t) > t$. Thus $q = 1$.

Assume $m > 1$. Then $\phi'(0) < 0$ and $\phi'(1) > 0$. Therefore there exists $\alpha \in]0, 1[$ such that ϕ is strictly decreasing on $]0, \alpha[$ and strictly increasing on $]\alpha, 1[$. As $\phi(0) > 0$, $\phi(1) = 0$, it must be $\phi(\alpha) < 0$; thus ϕ has a unique root in $]0, \alpha[$ and $q < 1$.

Taking into account A4), $\mathbf{P}(\tau_0 < +\infty) = 1$ if $m \leq 1$ and $\mathbf{P}(\tau_0 < +\infty) < 1$ if $m > 1$.

♦ It is also possible to argue using the strict convexity of f and recalling that m is the slope of f at 1 (see Figure 4.8).

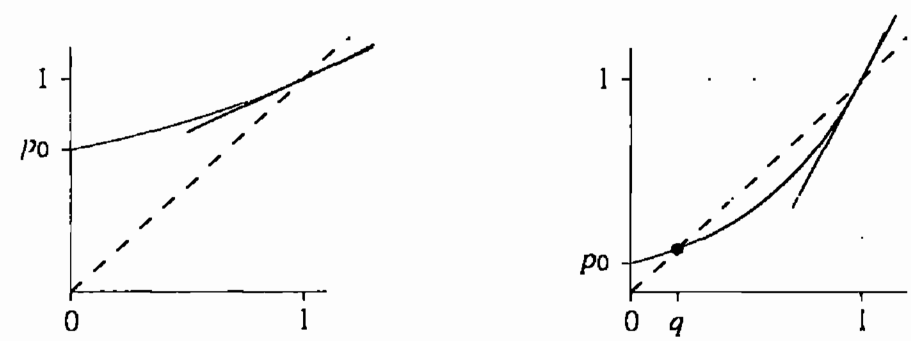

Figure 4.8 The two typical situations: $m \leq 1$ (left) and $m > 1$ (right).

C1) It holds

$$\mathbf{E}(Z_n \mid Z_{n-1} = i) = \sum_{j=1}^{\infty} j\mathbf{P}(Z_n = j \mid Z_{n-1} = i) = \sum_{j=1}^{\infty} j\mu^{*i}(j) = im.$$

Therefore $\mathbf{E}(Z_n \mid Z_{n-1}) = mZ_{n-1}$, which implies $\mathbf{E}(Z_n) = m\mathbf{E}(Z_{n-1})$ and, by recurrence, $\mathbf{E}(Z_n) = m^n$. One gets also that $(Z_n/m^n)_{n\geq 0}$ is a martingale that is positive and converges a.s. to a r.v. Z, which is a.s. finite (Theorem 3.8). Since we assume $m < 1$, $\lim_{n\to\infty} Z_n = 0$ a.s. One might also remark that $\mathbf{P}(Z_n \geq 1) \leq \mathbf{E}(Z_n) = m^n$ so that $\sum_{n\geq 0} \mathbf{P}(Z_n \geq 1) < +\infty$. The Borel–Cantelli lemma gives immediately $Z_n \to_{n\to\infty} 0$, without using the convergence theorems of martingales.

C2) By the hypotheses, there exists $i > 1$ such that $p_i > 0$; therefore $\mu^{*k}(ki) \geq p_i^k > 0$ and $\mu^{*k}(k) \leq 1 - \mu^{*k}(ki) < 1$. One gets, for every $r \geq 0$,

$$\mathbf{P}(Z_{n+p} = k \text{ for every } p \geq 0) \leq \mathbf{P}(Z_n = \ldots = Z_{n+r} = k) = P^n(1, k)\mu^{*k}(k)^r,$$

which gives $\mathbf{P}(Z_{n+p} = k \text{ for every } p \geq 0) = 0$.

Let $\mathscr{F}_n = \sigma(Z_0, \ldots, Z_n)$. By (C1) and the Markov property, $\mathbf{E}(Z_n \mid \mathscr{F}_{n-1}) = \mathbf{E}(Z_n \mid Z_{n-1}) = Z_{n-1}$ a.s., so that $(Z_n)_{n\geq 0}$ is a positive martingale. It converges a.s. to a r.v. Z_∞ and, since $Z_0 = 1$, $Z_\infty < +\infty$ a.s. (Theorem 3.8). If $k \geq 1$, $\mathbf{P}(Z_\infty = k) = 0$ since every state, 0 excepted, is transient and is visited a finite number of times at most a.s. (Theorem 4.16); thus it cannot be a limit point of $(Z_n)_{n\geq 0}$ a.s. Thus $\mathbf{P}(Z_\infty = 0) = 1$.

C3) Remark first that, if $q = 0$, $p_0 = 0$ and $\mathbf{P}(\tau_0 < +\infty) = 0$. Suppose $0 < q < 1$ (hence $p_0 > 0$). The uniqueness of the solution of $f(s) = s$ in $[0, 1[$ has been shown in (B). $(Z_n)_{n\geq 0}$ being a Markov chain,

$$\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = \mathbf{E}(q^{Z_{n+1}} \mid \mathscr{F}_n) = \mathbf{E}(q^{Z_{n+1}} \mid Z_n).$$

Moreover, thanks to (A2),

$$\mathbf{E}(q^{Z_{n+1}} \mid Z_n = i) = \sum_{j=0}^{\infty} P(i, j)q^j = \sum_{j=0}^{\infty} \mu^{*i}(j)q^j = f(q)^i = q^i,$$

so that $\mathbf{E}(q^{Z_{n+1}} \mid Z_n) = q^{Z_n}$ and $\mathbf{E}(X_{n+1} \mid \mathscr{F}_n) = X_n$.

The sequence $(X_n)_{n\geq 0}$ is a bounded martingale and converges a.s. to an integrable r.v. From the a.s. convergence of $(q^{Z_n})_{n\geq 0}$, one gets, taking logarithms, the a.s. convergence of $(Z_n)_{n\geq 0}$ to an element of $\bar{\mathbb{N}}$. This limit cannot be an integer > 0. The most simple way to see it is to observe that all the $i \geq 1$ are transient and therefore are visited by the chain at most a finite number of times (Theorem 4.16 and Proposition 4.17). $(Z_n)_{n\geq 0}$ converges therefore a.s. to a r.v. Z_∞ taking the values 0 or $+\infty$. In particular, as $\{Z_\infty = 0\} = \{\tau_0 < +\infty\}$, $X_n \to 1_{\{\tau_0 < +\infty\}}$ a.s. (as $q^\infty = 0$!) and, applying Lebesgue's theorem (as $X_n \in [0, 1]$),

$$\mathbf{P}(\tau_0 < +\infty) = \lim_{n\to\infty} \mathbf{E}(X_n) = \mathbf{E}(X_0) = q.$$

We therefore recover the result proved in (B3).

C4) In the same way, recalling that the second order moment is equal to the sum of the variance and of the square of the mean,

$$\mathbf{E}(Z_n^2 \mid Z_{n-1} = i) = \sum_{j=1}^{\infty} j^2\mu^{*i}(j) = i\sigma^2 + i^2m^2. \tag{4.97}$$

Consequently,

$$\mathbf{E}(Z_n^2 \mid Z_{n-1}) = \sigma^2 Z_{n-1} + m^2 Z_{n-1}^2.$$

One deduces

$$\mathbf{E}(Z_n^2) = \sigma^2\mathbf{E}(Z_{n-1}) + m^2\mathbf{E}(Z_{n-1}^2) = \sigma^2 m^{n-1} + m^2\mathbf{E}(Z_{n-1}^2)$$

and, again by induction,

$$\mathbf{E}(Z_n^2) = \sigma^2 m^{n-1} + m^2\mathbf{E}(Z_{n-1}^2) =$$

$$= \sigma^2 m^{n-1} + m^2(\sigma^2 m^{n-2} + m^2\mathbf{E}(Z_{n-2}^2)) = \sigma^2 \sum_{k=n-1}^{2n-2} m^k + m^{2n} \underbrace{\mathbf{E}(Z_0^2)}_{=1}.$$

Therefore

$$\mathrm{Var}(Z_n) = \mathbf{E}(Z_n^2) - m^{2n} = \sigma^2 \sum_{k=n-1}^{2n-2} m^k.$$

Considering separately the two cases $m \neq 1$ and $m = 1$, one can conclude.

C5) First it holds

$$\mathbf{E}(W_n^2) = \frac{1}{m^{2n}}\,\mathbf{E}(Z_n^2) = \frac{\sigma^2}{m^{n+1}}\frac{m^n - 1}{m - 1} + 1 \leq \frac{\sigma^2}{m-1} + 1$$

and the sequence $(W_n)_{n\geq 0}$ is bounded in L^2. Moreover,

$$\mathbf{E}(W_{n+1} \mid \mathscr{F}_n) = \frac{1}{m^{n+1}}\,\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n)$$

and, thanks to the Markov property, $\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n) = \mathbf{E}(Z_{n+1} \mid Z_n) = m\,Z_n$ whence

$$\mathbf{E}(W_{n+1} \mid \mathscr{F}_n) = \frac{Z_n}{m^n} = W_n,$$

which proves that $(W_n)_{n\geq 0}$ is a martingale; being bounded in L^2 it converges a.s. and in L^2 to a r.v. W_∞. Let us set $\phi_n(t) = \mathbf{E}(\mathrm{e}^{itZ_n/m^n})$. Then

$$\phi_n(t) = \sum_{j=0}^{\infty} \mathbf{P}(Z_n = j)(\mathrm{e}^{it/m^n})^j = f_n(\mathrm{e}^{it/m^n}) \tag{4.98}$$

(f_n is defined in (A3)). By (4.98),

$$\phi_n(mt) = f_n(\mathrm{e}^{it/m^{n-1}}) = f[f_{n-1}(\mathrm{e}^{it/m^{n-1}})] = f(\phi_{n-1}(t)). \tag{4.99}$$

As $Z_n/m^n \to W_\infty$ a.s., there is also convergence in distribution and $\phi_n(t)$ converges, for every $t \in \mathbb{R}$, to $\phi(t)$, ϕ being the characteristic function of W_∞; f being continuous, one may take the limit in (4.99) and obtain $\phi(mt) = f(\phi(t))$.

♦ This model was introduced in the nineteenth century by Galton and Watson in order to explain the extinction of the noble families in United Kingdom. The central role of the value m and some surprising consequences (extinction with probability 1 in the critical case $m = 1$, positive probability of extinction even for large values of m...) have led to numerous studies and developments.

P4.4 1a) $Y_1 = 1, Y_2 = 2, Y_3 = 1, Y_4 = 3, Y_5 = 2$.

1b) It holds $\sigma(5) < \sigma(2) < \sigma(1) < \sigma(3) < \sigma(4)$; hence $\sigma = (3, 2, 4, 5, 1)$.

2a) $\{Y_k = 1\}$ is the event $\sigma(k) < \sigma(i)$ for $i = 1, 2, \dots, k-1$; $\{Y_k = 1, Y_{k+1} > 1\}$ is the event $\sigma(k) < \sigma(i)$ for $i = 1, 2, \dots, k-1, k+1$; ... $\{Y_k = 1, Y_{k+1} > 1, \dots, Y_n > 1\}$ is the event $\sigma(k) < \sigma(i)$ for $i \neq 1$, i.e., $\sigma(k) = 1$.

2b) The knowledge of (Y_1, Y_2) determines $\sigma(1), \sigma(2)$; the knowledge of (Y_1, Y_2, Y_3) determines $\sigma(1), \sigma(2), \sigma(3)\dots$. The knowledge of $(Y_1, \dots, Y_n)$ determines $\sigma(1), \sigma(2), \dots, \sigma(n)$, i.e., determines σ. Thus, for every $(k_1, \dots, k_n) \in \Pi$, there exists exactly one $\sigma \in \Omega$ such that $F(\sigma) = (k_1, \dots, k_n)$.

2c) One deduces that, for every $(k_1, \dots, k_n) \in \Pi$,

$$\mathbf{P}(Y_1 = k_1, Y_2 = k_2, \dots, Y_n = k_n) = \mathbf{P}(F^{-1}(k_1, \dots, k_n)) = \frac{1}{n!}\cdot$$

One gets $\mathbf{P}(Y_i = k_i)$ by summing over $k_1, k_2, \ldots, k_{i-1}, k_{i+1}, \ldots, k_n$ whence, for $k_i = 1, \ldots, i$,

$$\mathbf{P}(Y_i = k_i) = 1 \cdot 2 \ldots (i-1) \cdot (i+1) \ldots n \cdot \frac{1}{n!} = \frac{1}{i}.$$

This proves that the law of Y_i is uniform on $\{1, 2, \ldots, i\}$ and that, for every $k_1, \ldots, k_n$,

$$\mathbf{P}(Y_1 = k_1, Y_2 = k_2, \ldots, Y_n = k_n) = \prod_{i=1}^{n} \mathbf{P}(Y_i = k_i)$$

whence the independence of $Y_1, \ldots, Y_n$.

3a) By (2a),

$$\mathbf{P}(A_r) = \sum_{k=r+1}^{n} \mathbf{P}(\tau_r = k, \sigma(k) = 1) =$$

$$= \sum_{k=r+1}^{n} \mathbf{P}(\tau_r = k, Y_k = 1, Y_{k+1} > 1, \ldots, Y_n > 1) =$$

$$\sum_{k=r+1}^{n} \mathbf{P}(Y_{r+1} > 1, \ldots, Y_{k-1} > 1, Y_k = 1, Y_{k+1} > 1, \ldots, Y_n > 1) =$$

$$= \sum_{k=r+1}^{n} \frac{r}{r+1} \frac{r+1}{r+2} \cdots \frac{k-2}{k-1} \frac{1}{k} \frac{k}{k+1} \cdots \frac{n-1}{n} =$$

$$= \frac{r}{n} \sum_{k=r+1}^{n} \frac{1}{k-1} = \frac{r}{n} \sum_{k=r}^{n-1} \frac{1}{k}.$$

3b) It holds $\mathbf{P}(A_r) = s_r = s_1 + \sum_{k=1}^{r-1} \delta_k$ with

$$\delta_k \doteq s_{k+1} - s_k = \frac{1}{n}\Big(\sum_{i=k+1}^{n-1} \frac{1}{i} - 1 \Big). \tag{4.100}$$

Let us note that the application $k \to \delta_k$ is strictly decreasing and that $\delta_1 > 0$, $\delta_{n-2} < 0$, $\delta_k \neq 0$. Thus there exists a unique r^* such that $\delta_{r^*} < 0$ and $\delta_{r^*-1} > 0$. This is the unique value maximizing s_r.

3c) In this question n varies. Let us denote by r_n^* the value obtained in (3b). Thus, by (4.100),

$$m_n^* \stackrel{\text{def}}{=} \sum_{k=r_n^*+1}^{n-1} \frac{1}{k} < 1 < \sum_{k=r_n^*+1}^{n-1} \frac{1}{k} + \frac{1}{r_n^*}. \tag{4.101}$$

This implies that $r_n^* \to +\infty$. Otherwise there would be a sequence $n_i \to \infty$ such that $r_{n_i}^* \leq r_0 < +\infty$ and this would entail $\sum_{k=r_0}^{+\infty} \frac{1}{k} < +\infty$. Thus (4.101) gives

$$\log \frac{n-1}{r_n^*} + o(1) \leq 1 \leq \log \frac{n-1}{r_n^*} + o(1)$$

which implies that, as $n \to \infty$, $\log \frac{n}{r_n^*} \to 1$ and $\frac{n}{r_n^*} \to e$. Still by (4.101) and since

$r_n^* \to_{n\to\infty} +\infty$, $m_n^* \to_{n\to\infty} 1$ and

$$s_{r_n^*} = \mathbf{P}(A_{r_n^*}) = \frac{r_n^*}{n}\left(m_n^* + \frac{1}{r_n^*}\right) \to \frac{1}{e}.$$

♦ Proof of the result admitted in (3b). Let $\sigma_{r,n} = \sum_{i=r}^{n} \frac{1}{i}$ and k the largest integer such that $2^k \le n$. Let us remark that every integer $m \le n$ and different from 2^k can be written $m = 2^s q$ with $s < k$ and q odd. If $r > 2^k$, then $\sigma_{r,n} < 2^k \cdot \frac{1}{2^k} = 1$. If $r \le 2^k$, then

$$2^k \sigma_{r,n} = \sum_j \frac{2^{k-r_j}}{q_j} + 1 = \frac{2M}{Q} + 1, \qquad M \in \mathbb{N},\ Q \text{ odd},$$

whence $Q2^k\sigma_{r,n} = 2M + Q$ and $\sigma_{r,n} \notin \mathbb{N}$.

4) The event $\{X_1 = i_1, \dots, X_j = i_j\}$ depends on $Y_1, \dots, Y_{i_j}$ only whereas, for a fixed i_j, $\{X_{j+1} = i_{j+1}\} = \{\inf\{k > i_j;\ Y_k = 1\} = i_{j+1}\}$ only depends on Y_k, $k > i_j$. Thus these two events are independent and

$$\begin{aligned} &\mathbf{P}(X_1 = i_1, \dots, X_j = i_j, X_{j+1} = i_{j+1}) = \\ &= \mathbf{P}(X_1 = i_1, \dots, X_j = i_j)\mathbf{P}\big(\inf\{k > i_j;\ Y_k = 1\} = i_{j+1}\big). \end{aligned}$$

One can deduce

$$\begin{aligned} &\mathbf{P}(X_{j+1} = i_{j+1} \mid X_1 = i_1, \dots, X_j = i_j) = \\ &= \mathbf{P}\big(\inf\{k > i_j;\ Y_k = 1\} = i_{j+1}\big) \stackrel{\text{def}}{=} Q(i_j, i_{j+1}) \end{aligned}$$

(if $\mathbf{P}(X_1 = i_1, \dots, X_j = i_j) > 0$). This proves the Markov property. We only need to determine the transition matrix Q.

1st case: $i_j < i_{j+1} < n + 1$. In this case

$$\begin{aligned} \{X_{j+1} = i_{j+1}\} &= \{\inf\{k > i_j;\ Y_k = 1\} = i_{j+1}\} = \\ &= \{Y_{i_j+1} > 1, \dots Y_{i_{j+1}-1} > 1, Y_{i_{j+1}} = 1\}. \end{aligned}$$

Thus, by (b),

$$\begin{aligned} Q(i_j, i_{j+1}) &= \left(1 - \frac{1}{i_j + 1}\right)\cdots\left(1 - \frac{1}{i_{j+1} - 1}\right)\frac{1}{i_{j+1}} = \\ &= \frac{i_j}{i_j + 1} \cdot \frac{i_j + 1}{i_j + 2} \cdots \frac{i_{j+1} - 2}{i_{j+1} - 1}\frac{1}{i_{j+1}} = \frac{i_j}{i_{j+1}(i_{j+1} - 1)}. \end{aligned}$$

2nd case: $i_j < n + 1$, $i_{j+1} = n + 1$. It holds

$$\{X_{j+1} = n + 1\} = \{Y_{i_j+1} > 1, Y_{i_j+2} > 1, \dots, Y_n > 1\},$$

whence

$$\begin{aligned} Q(i_j, n + 1) &= \left(1 - \frac{1}{i_j + 1}\right)\cdots\left(1 - \frac{1}{n}\right) = \\ &= \frac{i_j}{i_j + 1} \cdot \frac{i_j + 1}{i_j + 2} \cdots \frac{n - 1}{n} = \frac{i_j}{n}. \end{aligned}$$

3rd case: $i_j = n+1, i_{j+1} = n+1$. Here $Q(n+1, n+1) = 1$.

Putting things together, $(X_j)_{j\geq 0}$ is a Markov chain with values in $\{1, 2\dots, n+1\}$ and with transition matrix

$$\begin{cases} Q(k,l) = 0 & \text{if } 1 \leq l \leq k \leq n \\ Q(k,l) = \frac{k}{l(l-1)} & \text{if } 1 \leq k < l \leq n \\ Q(k, n+1) = \frac{k}{n} & \text{if } 1 \leq k \leq n \\ Q(n+1, n+1) = 1. \end{cases}$$

♦ One can note that, considering the stopping time $\tilde{\tau}_r = \inf\{j > 1;\, X_j > r\}$ of the filtration $\tilde{\mathscr{F}}_j = \sigma(X_1, \dots, X_j)$, $\tau_r = X_{\tilde{\tau}_r}$. The optimal stopping problem thus can be viewed as related to the Markov chain $(X_n)_{n\geq 0}$. Thus it belongs to the general setting of the optimal stopping problems of a Markov chain, taking as payoff the quantity $P(\sigma(X_\tau) = 1)$ (see Problem 4.5).

We just proved that the stopping time τ_{r^*} maximizes the probability that at time τ_r the suitor passing by is the best one, among all the stopping times of the type τ_r. The unafraid reader will be able to prove, using Problem 4.5, the complete and much stronger result, stating that τ_{r^*} is the optimal stopping time among all the stopping times of the filtration $\mathscr{F}_k = \sigma(Y_1, \dots, Y_k)$.

P4.5 1) It is immediate that $f_1 \geq f_0$. Supposing that $f_n \geq f_{n-1}$, one can write $f_{n+1} = \max(f, Qf_n) \geq \max(f, Qf_{n-1}) = f_n$; this proves by induction that the sequence $(f_n)_{n\geq 0}$ is increasing; it is moreover majorized by $\sup_E f$. Thus it converges to a bounded function f^*.

Moreover, by the Beppo Levi theorem $(Qf_n)_{n\geq 0}$ increases to Qf^* and therefore, taking the limit in the relation $f_{n+1} = \max(f, Qf_n)$, one gets $f^* = \max(f, Qf^*)$. In particular $f^* \geq Qf^*$ and f^* is then excessive.

Finally, let h be an excessive function dominating f. Then $h \geq f_0 = f$. Moreover, if $h \geq f_n$, $h \geq Qh \geq Qf_n$ and, consequently, $h \geq \max(f, Qf_n) = f_{n+1}$. One therefore deduces by induction that $h \geq f_n$ for every n and thus $h \geq f^*$: f^* is the smallest excessive function majorizing f.

2a) Checking that ν is a stopping time is obvious. Moreover $\nu \in \mathscr{T}$, as $P_x(\nu = +\infty) = \mathbf{P}_x(\nu_x = +\infty) = 0$. Let $x \in E$. By the definition of V, there exists $\nu_x \in \mathscr{T}$ such that $\mathbf{E}_x(f(X_{\nu_x})) \geq V(x) - \varepsilon$. Therefore, for every $x \in E$,

$$\mathbf{E}_x(f(X_\nu)) = \mathbf{E}_x(f(X_{\nu_x})) \geq V(x) - \varepsilon.$$

2b) Let $\nu' = 1 + \nu \circ \theta_1$, where ν is as in (2a). Then $X_{\nu'} = X_\nu \circ \theta_1$ and, for every $x \in E$,

$$\begin{aligned} V(x) &\geq \mathbf{E}_x(f(X_{\nu'})) = \mathbf{E}_x(f(X_\nu \circ \theta_1)) = \mathbf{E}_x(\mathbf{E}_{X_1}(f(X_\nu))) \geq \\ &\geq \mathbf{E}_x(V(X_1) - \varepsilon) = QV(x) - \varepsilon. \end{aligned}$$

Consequently, for every $x \in E$, $V(x) \geq QV(x)$ and V is excessive.

2c) As u is excessive, $(u(X_n))_{n\geq 0}$ is a positive supermartingale and, thanks to the optional sampling theorem and to Fatou's lemma,

$$\begin{aligned} \mathbf{E}_x(f(X_\rho)) \leq \mathbf{E}_x(u(X_\rho)) &= \mathbf{E}_x\Big(\lim_{n\to\infty} u(X_{\rho\wedge n})\Big) \leq \lim_{n\to\infty} \mathbf{E}_x(u(X_{\rho\wedge n})) \leq \\ &\leq \mathbf{E}_x(u(X_0)) = u(x). \end{aligned}$$

This relation, applied to the stopping time ν introduced in (2a), gives that $u(x) \geq \mathbf{E}_x(f(X_\nu)) \geq V(x) - \varepsilon$ and, ε being arbitrary, $u \geq V$. Therefore $V \leq u$ and $V \leq f^*$ (as one may take $u = f^*$). On the other hand, V is an excessive function dominating f (taking the null constant stopping time $\rho \equiv 0$ gives $V(x) \geq f(x)$) and thus $V \geq f^*$. Hence $V = f^*$.

3a) If x is such that $f(x) = f^*(x)$, then obviously $\mathbf{P}_x(\tau = 0) = 1$. Suppose that there exists $x \in E$ such that $f^*(x) > f(x)$. Let $\delta > 0$ be such that $f^*(x) \geq f(x) + \delta$ for every x such that $f^*(x) > f(x)$. Let $\nu \in \mathscr{T}$. Then

$$\begin{aligned}\mathbf{E}_x(f(X_\nu)) = \mathbf{E}_x(f^*(X_\nu)1_{\{f^*(X_\nu)=f(X_\nu)\}}) + \mathbf{E}_x(f(X_\nu)1_{\{f^*(X_\nu)>f(X_\nu)\}}) \leq \\ \leq \mathbf{E}_x(f^*(X_\nu)) - \delta\mathbf{P}_x(f^*(X_\nu) > f(X_\nu)) \leq \mathbf{E}_x(f^*(X_\nu)) - \delta\mathbf{P}_x(\tau = \infty) \leq \\ \leq f^*(x) - \delta\mathbf{P}_x(\tau = \infty),\end{aligned}$$

the last inequality resulting from question (2c). As this is true for every $\nu \in \mathscr{T}$, taking the supremum over $\nu \in \mathscr{T}$ and since $V = f^*$, one has therefore $f^*(x) \leq f^*(x) - \delta\mathbf{P}_x(\tau = \infty)$; hence $\mathbf{P}_x(\tau = \infty) = 0$.

3b) If $n < \tau$ then, by definition, $f(X_n) < f^*(X_n)$. Hence $f^*(X_n) = \max(f(X_n), Qf^*(X_n)) = Qf^*(X_n)$.

3c) One has

$$\begin{aligned}&\mathbf{E}(f^*(X_{(n+1)\wedge\tau}) \mid \mathscr{F}_n) = \\ &= \mathbf{E}(f^*(X_{(n+1)\wedge\tau})1_{\{\tau\leq n\}} + f^*(X_{(n+1)\wedge\tau})1_{\{\tau>n\}} \mid \mathscr{F}_n) = \\ &= 1_{\{\tau\leq n\}}f^*(X_\tau) + 1_{\{\tau>n\}}\mathbf{E}(f^*(X_{n+1}) \mid \mathscr{F}_n) = \\ &= 1_{\{\tau\leq n\}}f^*(X_\tau) + 1_{\{\tau>n\}}Qf^*(X_n) = \\ &= 1_{\{\tau\leq n\}}f^*(X_\tau) + 1_{\{\tau>n\}}f^*(X_n) = f^*(X_{n\wedge\tau})\end{aligned}$$

(we repeat here the proof of Exercise 4.3). As f^* is bounded, the martingale is regular and (Proposition 3.13)

$$f^*(x) = \mathbf{E}_x(f^*(X_\tau)) = \mathbf{E}_x(f(X_\tau)),$$

i.e., τ is an optimal stopping time.

P4.6 A) If X is recurrent, it is obvious, as the chain visits every state infinitely many times $\mathbf{P}_x$-a.s. for every x. If X is transient, then $U(x, y) < +\infty$ for every x and y. Then, if $J_p = \{0, \dots, p\}$, $\mathbf{E}_x(N_{J_p}) = \sum_{y\in J_p} U(x, y) < +\infty$, hence, $\mathbf{P}_x(N_{J_p} < +\infty) = 1$. This means that, $\mathbf{P}_x$-a.s., there exists $n_0 = n_0(\omega)$ such that $X_n(\omega) > p$ for every $n \geq n_0$. This being true for every p, $\lim_{n\to\infty} X_n = +\infty$, $\mathbf{P}_x$-a.s.

B) The fact that $(\phi(X_{n\wedge\tau}))_{n\geq 0}$ is a positive supermartingale is the content of Exercise 4.3, which the reader is invited to peruse. As a positive supermartingale is a.s. convergent, there exists $Z \geq 0$ such that $\phi(X_{n\wedge\tau}) \to_{n\to\infty} Z \geq 0$. By Fatou's lemma, Z is integrable as

$$\mathbf{E}_x(Z) = \mathbf{E}_x\big(\lim_{n\to\infty} \phi(X_{n\wedge\tau})\big) \leq \varliminf_{n\to\infty} \mathbf{E}_x(\phi(X_{n\wedge\tau})) \leq \mathbf{E}_x(\phi(X_0)) = \phi(x).$$

C1) For $x \leq p$, $\mathbf{P}_x(\tau = 0) = 1$. For $x > p$, on $\{\tau = +\infty\}$ it holds $\phi(X_n) \to_{n\to\infty} Z < +\infty$ $\mathbf{P}_x$-a.s., but, as $\phi(x) \to_{x\to+\infty} +\infty$, by (A) $\overline{\lim}_{n\to\infty} \phi(X_n) = +\infty$ $\mathbf{P}_x$-a.s.; hence, necessarily, $\mathbf{P}_x(\tau = +\infty) = 0$.

C2) One has

$$\mathbf{P}_x(\tau \circ \theta_n < +\infty) = \mathbf{E}_x(1_{\{\tau<+\infty\}} \circ \theta_n) = \mathbf{E}_x[\mathbf{P}_{X_n}(\tau < +\infty)] = 1.$$

It is easy to check that $\{\tau \circ \theta_n < +\infty\} \downarrow \{N_{J_p} = +\infty\}$; hence, $\mathbf{P}_x(N_{J_p} = +\infty) = 1$ and $U(x, J_p) = \mathbf{E}_x(N_{J_p}) = +\infty$.

C3) One has $U(x, J_p) = \sum_{y=0}^{p} U(x, y) = +\infty$. Thus, there exists $z \le p$ such that $U(x, z) = +\infty$, but (maximum principle, •4.20) $U(x, z) \le U(z, z)$ hence $U(z, z) = +\infty$; z is recurrent and so is the chain, being irreducible.

D) One has

$$\mathbf{E}_x(\phi(X_\tau)1_{\{\tau<+\infty\}}) = \mathbf{E}_x\Big(\lim_{n\to\infty} \phi(X_{n\wedge\tau})1_{\{\tau<+\infty\}}\Big) \le$$
$$\le \varliminf_{n\to\infty} \mathbf{E}_x(\phi(X_{n\wedge\tau})1_{\{\tau<+\infty\}}) \le \phi(x).$$

But $\mathbf{E}_x(\phi(X_\tau)1_{\{\tau<+\infty\}}) \ge \gamma \mathbf{P}_x(\tau < +\infty)$ where $\gamma = \inf_{y\le p} \phi(y) > 0$ and therefore

$$\mathbf{P}_x(\tau < +\infty) \le \gamma^{-1}\phi(x) \underset{x\to\infty}{\to} 0.$$

The chain X is transient; indeed if it were recurrent, one would have $\mathbf{P}_x(\tau < +\infty) = 1$ for every $x \in E$.

E1) It is easy to see that all the states communicate.

E2) Developing to the second order the function $s \to (1+s)^\alpha$,

$$\mathbf{E}_x(X_1^\alpha) = (x+1)^\alpha p_x + (x-1)^\alpha q_x = x^\alpha[(1+x^{-1})^\alpha p_x + (1-x^{-1})^\alpha q_x] =$$
$$= x^\alpha\Big[\Big(1 + \frac{\alpha}{x} + \frac{1}{2x^2}\alpha(\alpha-1) + \frac{\varepsilon_1(x)}{x^2}\Big)\Big(\frac{1}{2} + \frac{\lambda}{x} + \frac{\lambda s(x)}{x}\Big) +$$
$$+\Big(1 - \frac{\alpha}{x} + \frac{1}{2x^2}\alpha(\alpha-1) + \frac{\varepsilon_2(x)}{x^2}\Big)\Big(\frac{1}{2} - \frac{\lambda}{x} - \frac{\lambda s(x)}{x}\Big)\Big] =$$
$$= x^\alpha\Big[1 + \frac{2\alpha}{x^2}\Big(\lambda - \frac{1}{4} + \frac{\alpha}{4}\Big) + \frac{\varepsilon_\alpha(x)}{x^2}\Big]$$

where $\varepsilon_\alpha(x) \to_{x\to+\infty} 0$.

E3) If $\lambda < \frac{1}{4}$, one can find $\alpha > 0$ such that $\lambda - \frac{1}{4} + \frac{\alpha}{4} < 0$ and $\mathbf{E}_x(X_1^\alpha) \le x^\alpha$ for x large enough. The function $\phi(x) = x^\alpha$ satisfies the hypotheses in (C).

E4) If $\lambda > \frac{1}{4}$, one may find $\alpha < 0$ such that $\lambda - \frac{1}{4} + \frac{\alpha}{4} > 0$ and $\mathbf{E}_x(X_1^\alpha) \le x^\alpha$ for x large enough. One may apply (D).

P4.7 a) Thanks to the hypotheses on $p = (p_n)_{n\in\mathbb{N}}$, there exists $k_0 > 1$ such that $p_{k_0} > 0$. Since $Q(0, k_0) = p_{k_0} > 0$ and $Q(h, h + k_0 - 1) = p_{k_0} > 0$, $0 \rightsquigarrow k_0 \rightsquigarrow k_0 + (k_0 - 1) \rightsquigarrow \dots \rightsquigarrow k_0 + m(k_0 - 1)$. Since $k_0 - 1 \ge 1$, for every $k > 0$ there exists m such that $k' = k_0 + m(k_0 - 1) \ge k$.

If $k > 0$, $Q(k, k-1) = p_0 > 0$; hence, k leads to all the states i with $i \le k$. Let $i, k \in \mathbb{N}$: let us show that $i \rightsquigarrow k$. Indeed we have just seen that $i \rightsquigarrow 0$; next $0 \rightsquigarrow k'$ for some $k' \ge k$. At last $k' \rightsquigarrow k$, as k' leads to all the states i with $i \le k'$. By the transitivity of the relation $\rightsquigarrow$, one gets finally $i \rightsquigarrow k$. Therefore all the states communicate and the chain is irreducible.

b1) As there exists at least one index $i, i \ge 2$, such that $p_i > 0$, it holds $\varphi'' > 0$ on $]0, 1[$ and φ is strictly convex. For the same reason, φ is increasing. Define $g(t) =$

$\varphi(t) - t$. If $m \leq 1$, one has, for $t \in [0, 1[$,

$$g'(t) = \varphi'(t) - 1 = \sum_{k \geq 1} k p_k t^{k-1} - 1 < \sum_{k \geq 1} k p_k - 1 = m - 1 \leq 0.$$

Hence g is decreasing and, as $g(1) = \varphi(1) - 1 = 0$, $\varphi(t) > t$ for every $t \in [0, 1[$ and there are no solutions of the equation $\varphi(t) = t$ in $]0, 1[$.

Conversely if $m > 1$, then $g(1) = 0$ and $g'(1) = m - 1 > 0$. Therefore g is strictly negative for some $\xi \in [0, 1[$. As $g(0) = p_0 > 0$, by the intermediate values theorem, there exists $0 < \alpha \leq \xi < 1$ such that $g(\alpha) = 0$, i.e., $\varphi(\alpha) = \alpha$. This value α is unique, as $\varphi'' = g'' > 0$ and so g is strictly convex.

b2) For $n \geq 1$,

$$Qf(n) = \sum_{k \geq n-1} p_{k-n+1} f(k) = \sum_{k \geq 0} p_k f(n - 1 + k) = \sum_{k \geq 0} p_k u^{n-1+k} = u^{n-1}\varphi(u).$$

Therefore $f(n) = Qf(n)$ for every $n \geq 1$ is equivalent to $u = \varphi(u)$. If $m > 1$ and α is the unique root of $\varphi(u) = u$ in $]0, 1[$, the function $f(n) = \alpha^n$ is harmonic and it is the only one of the form $f(n) = u^n$ with $0 < u < 1$. If $m \leq 1$, there is no harmonic function of this form.

c1) Let us remark that $Q(i, j) = Q(i + k, j + k)$ for every $i \geq 1$, $j \geq 0$, $k \geq 0$, and that, $\mathbf{P}_k$-a.s., the chain $(X_n)_{n \geq 0}$ cannot pass from $i > r$ to $j < r$ without passing through r, as it can go to the left at most one step every time. More precisely,

$$\begin{aligned}
\mathbf{P}_{k+r}(\tau_k = n) &= \sum_{x_1, \dots, x_{n-1} > k} \mathbf{P}_{k+r}(X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = k) = \\
&= \sum_{x_1, \dots, x_{n-1} > k} Q(r + k, x_1) \dots Q(x_{n-1}, k) = \\
&= \sum_{y_1, \dots, y_{n-1} > 0} Q(r + k, y_1 + k) \dots Q(y_{n-1} + k, k) = \\
&= \sum_{y_1, \dots, y_{n-1} > 0} Q(r, y_1) \dots Q(y_{n-1}, 0) = \mathbf{P}_r(\tau_0 = n);
\end{aligned}$$

hence,

$$\mathbf{P}_{k+r}(\tau_k < +\infty) = \sum_{n \geq 0} \mathbf{P}_{k+r}(\tau_k = n) = \sum_{n \geq 0} \mathbf{P}_r(\tau_0 = n) = \mathbf{P}_r(\tau_0 < +\infty)$$

and

$$\mathbf{E}_{k+r}(\tau_k) = \sum_{n \geq 0} n\mathbf{P}_{k+r}(\tau_k = n) = \sum_{n \geq 0} n\mathbf{P}_r(\tau_0 = n) = \mathbf{E}_r(\tau_0).$$

c2) If $r = 0$ there is nothing to prove. If $r \geq 1$ one observes that, in order to reach 0 from $k + r$, the chain passes through r, hence, $\tau_r < \tau_0$ $\mathbf{P}_{k+r}$-a.s. and $\tau_r + \tau_0 \circ \theta_{\tau_r}$ is the hitting time at 0 after visiting r.

c3) Thanks to the strong Markov property (Theorem 4.6) and using (c1),

$$\begin{aligned}
\mathbf{P}_{k+r}(\tau_0 < +\infty) &= \mathbf{P}_{k+r}(\tau_r < +\infty, \tau_0 \circ \theta_{\tau_r} < +\infty) = \\
&= \mathbf{P}_{k+r}(\tau_r < +\infty)\mathbf{P}_r(\tau_0 < +\infty) = \mathbf{P}_k(\tau_0 < +\infty)\mathbf{P}_r(\tau_0 < +\infty).
\end{aligned}$$

The relation $\mathbf{P}_k(\tau_0 < +\infty) = \mathbf{P}_1(\tau_0 < +\infty)^k$ now follows easily.

c4) It is known (Corollary 4.13) that f is the smallest positive solution of $f(0) = 0$ and $f(n) = Qf(n), n > 0$. By (c3), $f(n)$ is of the form u^n with $u = \mathbf{P}_1(\tau_0 < +\infty)$. By (b2), if $m > 1$, $u = \alpha < 1$ and $\mathbf{P}_k(\tau_0 < +\infty) = \alpha^k$; conversely, if $m \leq 1$, $\mathbf{P}_k(\tau_0 < +\infty) = 1$.

c5) Let σ_0 be the return time to 0: as $\sigma_0 = 1 + \tau_0 \circ \theta_1$, thanks to the Markov property (Theorem 4.5),

$$\mathbf{P}_0(\sigma_0 < +\infty) = \mathbf{E}_0[1_{\{\tau_0<+\infty\}} \circ \theta_1] = \mathbf{E}_0[\mathbf{P}_{X_1}(\tau_0 < +\infty)].$$

If $m \leq 1$, by the argument above, $\mathbf{P}_x(\tau_0 < +\infty) = 1$ for every $x \in \mathbb{N}$ and $\mathbf{P}_0(\sigma_0 < +\infty) = 1$; the chain is thus recurrent. Conversely if $m > 1$, $\mathbf{P}_1(\tau_0 < +\infty) = \alpha < 1$ and the chain, which is irreducible, is transient.

d) We suppose now $m \leq 1$.

d1) Using (c1) and (c2).

$$\begin{aligned}\mathbf{E}_{k+r}(\tau_0) &= \mathbf{E}_{k+r}(\tau_r) + \mathbf{E}_{k+r}(\tau_0 \circ \theta_{\tau_r}) = \\ &= \mathbf{E}_k(\tau_0) + \mathbf{E}_{k+r}(\mathbf{E}_{X_{\tau_r}}(\tau_0)) = \mathbf{E}_k(\tau_0) + \mathbf{E}_r(\tau_0);\end{aligned}$$

hence, $\mathbf{E}_k(\tau_0) = k\mathbf{E}_1(\tau_0)$.

d2) It is known (Corollary 4.15) that u satisfies $u(0) = 0$ and $u(n) = Qu(n) + 1$ for every $n > 0$. Since, by (d1), $u(n) = nv$ with $v = \mathbf{E}_1(\tau_0)$,

$$\begin{aligned}v = u(1) &= Qu(1) + 1 = \sum_{k\geq 0} Q(1,k)u(k) + 1 = \\ &= \sum_{k\geq 0} kQ(1,k)v + 1 = \sum_{k\geq 0} kp_k v + 1 = mv + 1.\end{aligned}$$

If $m = 1$, there is no finite solution and $\mathbf{E}_k(\tau_0) = +\infty$. Conversely if $m < 1$, one finds $v = (1-m)^{-1}$ and $\mathbf{E}_k(\tau_0) = k(1-m)^{-1}$.

d3) As $\sigma_0 = 1 + \tau_0 \circ \theta_1$, just apply the Markov property and obtain

$$\mathbf{E}_0(\sigma_0) = 1 + \mathbf{E}_0(\tau_0 \circ \theta_1) = 1 + \mathbf{E}_0(\mathbf{E}_{X_1}(\tau_0)) = 1 + \sum_{k\geq 1} p_k \mathbf{E}_k(\tau_0).$$

Therefore, if $m = 1$, $\mathbf{E}_0(\sigma_0) = +\infty$ whence, if $m < 1$,

$$\mathbf{E}_0(\sigma_0) = 1 + \frac{1}{1-m}\sum_{k\geq 1} kp_k = 1 + \frac{m}{1-m} = \frac{1}{1-m}.$$

Therefore the chain is null recurrent for $m = 1$ and positive recurrent for $m < 1$ (Theorem 4.23).

P4.8 A1) By hypothesis, for $x \in F$, $\mathbf{E}_x(\sigma_F) < +\infty$ and $\mathbf{P}_x(\sigma_F < +\infty) = 1$. Let us suppose that, for every $x \in F$, $\mathbf{P}_x(\rho_{n-1} < +\infty) = 1$; then, by the strong Markov property,

$$\begin{aligned}\mathbf{P}_x(\rho_n < +\infty) &= \mathbf{P}_x(\rho_{n-1} < +\infty, \sigma_F \circ \theta_{\rho_{n-1}} < +\infty) = \\ &= \mathbf{E}_x[1_{\{\rho_{n-1}<+\infty\}}\mathbf{P}_{X_{\rho_{n-1}}}(\sigma_F < +\infty)] = 1\end{aligned}$$

as, on $\{\rho_{n-1} < +\infty\}$, $X_{\rho_{n-1}} \in F$ and $\mathbf{P}_y(\sigma_F < +\infty) = 1$ for every $y \in F$.

A2) Observe that $X_{\rho_{n+1}} = X_{\sigma_F} \circ \theta_{\rho_n}$ (Exercise 2.2): hence,

$$\mathbf{P}_x(Y_{n+1} = y \mid \mathscr{F}_{\rho_n}) = \mathbf{P}_x(X_{\rho_{n+1}} = y \mid \mathscr{F}_{\rho_n}) = \mathbf{P}_x(X_{\sigma_F} \circ \theta_{\rho_n} = y \mid \mathscr{F}_{\rho_n}) = \\ = \mathbf{P}_{X_{\rho_n}}(X_{\sigma_F} = y) = \bar{P}(X_{\rho_n}, y) = \bar{P}(Y_n, y).$$

A3) The chain Y takes its values in the finite set F and it is sufficient to prove that it is irreducible. Let $x, y \in F$, $x \neq y$. The chain X being irreducible, there exists $n \geq 1$ such that $P^n(x, y) > 0$ and therefore there exist $x_1, \dots, x_{n-1} \in E$ such that $\mathbf{P}_x(X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = y) > 0$. Let $x_{j_1}, \dots, x_{j_m}$ be the sequence of the states $x_1, \dots, x_{n-1}$ that belong to F; then

$$\bar{P}^{m+1}(x, y) \geq \mathbf{P}_x(Y_1 = x_{j_1}, \dots, Y_m = x_{j_m}, Y_{m+1} = y) \geq \\ \geq \mathbf{P}_x(X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = y) > 0,$$

which proves the irreducibility of Y.

A4) Let $x \in F$. To say that $\bar{\sigma}_x = p$ means that the chain X comes back to x for the first time at its p-th return in F. Therefore $\{\bar{\sigma}_x = p\} = \{\sigma_x = \rho_p\}$. Hence, on $\{\bar{\sigma}_x = p\}$,

$$\sigma_x = \rho_p = \sum_{k=0}^{p-1}(\rho_{k+1} - \rho_k) = \sum_{k=0}^{p-1} \sigma_F \circ \theta_{\rho_k} = \sum_{k=0}^{\infty} \sigma_F \circ \theta_{\rho_k} 1_{\{k<p\}},$$

which gives the formula we were looking for. Note also that

$$\{\bar{\sigma}_x \leq k\} = \bigcup_{p=0}^{k} \{\sigma_x = \rho_p\}$$

and that $\{\sigma_x = \rho_p\} \in \mathscr{F}_{\rho_p} \subset \mathscr{F}_{\rho_k}$ as $0 \leq p \leq k$. Therefore $\{\bar{\sigma}_x \leq k\} \in \mathscr{F}_{\rho_k}$ and $\{\bar{\sigma}_x > k\}$ as well.

A5) Using (A4) and the strong Markov property,

$$\mathbf{E}_x(\sigma_x) = \sum_{k=0}^{\infty} \mathbf{E}_x[\sigma_F \circ \theta_{\rho_k} 1_{\{k<\bar{\sigma}_x\}}] = \sum_{k=0}^{\infty} \mathbf{E}_x[1_{\{k<\bar{\sigma}_x\}} \mathbf{E}_{X_{\rho_k}}(\sigma_F)] \leq \\ \leq \max_{y \in F} \mathbf{E}_y(\sigma_F) \mathbf{E}_x\Big[\sum_{k=0}^{\infty} 1_{\{k<\bar{\sigma}_x\}}\Big] = \max_{y \in F} \mathbf{E}_y(\sigma_F) \cdot \mathbf{E}_x(\bar{\sigma}_x).$$

The chain Y is irreducible positive recurrent; hence, $\mathbf{E}_x(\bar{\sigma}_x) < +\infty$ and, by the previous relation, $\mathbf{E}_x(\sigma_x) < +\infty$ and $\mathbf{P}_x(\sigma_x < +\infty) = 1$. This implies that x is recurrent and, thanks to the irreducibility of X and Theorem 4.23, that the chain X is positive recurrent.

B1) One has

$$\mathbf{E}_x(Z_{(n+1)\wedge\tau_F} \mid \mathscr{F}_n) = \mathbf{E}_x[1_{\{n<\tau_F\}} Z_{n+1} + 1_{\{n\geq\tau_F\}} Z_{\tau_F} \mid \mathscr{F}_n]$$

and since the r.v. $1_{\{n\geq\tau_F\}} Z_{\tau_F} = 1_{\{n\geq\tau_F\}} Z_{n\wedge\tau_F}$ is $\mathscr{F}_n$-measurable and by (4.69),

$$\mathbf{E}_x(Z_{(n+1)\wedge\tau_F} \mid \mathscr{F}_n) = 1_{\{n<\tau_F\}} \mathbf{E}_x(Z_{n+1} \mid \mathscr{F}_n) + 1_{\{n\geq\tau_F\}} Z_{\tau_F} = \\ = 1_{\{n<\tau_F\}}(Ph(X_n) + (n+1)\alpha) + 1_{\{n\geq\tau_F\}} Z_{\tau_F} \leq \\ \leq 1_{\{n<\tau_F\}}(h(X_n) + n\alpha) + 1_{\{n\geq\tau_F\}} Z_{\tau_F} = Z_{n\wedge\tau_F}.$$

If $x \in F^c$, $\mathbf{E}_x(Z_{n\wedge\tau_F}) \le \mathbf{E}_x(Z_0) = h(x)$. As $h \ge 0$, $\mathbf{E}_x(Z_{n\wedge\tau_F}) \ge \alpha\mathbf{E}_x(n \wedge \tau_F)$; hence, $\mathbf{E}_x(n \wedge \tau_F) \le \alpha^{-1}h(x)$ and we get the requested result by monotone limit. If $x \in F$, the relation is obvious.

B2) One has $\sigma_F = 1 + \tau_F \circ \theta_1$; hence,

$$\mathbf{E}_x(\sigma_F) = 1 + \mathbf{E}_x(\tau_F \circ \theta_1) = 1 + \mathbf{E}_x(\mathbf{E}_{X_1}(\tau_F)) \le 1 + \alpha^{-1}\mathbf{E}_x(h(X_1)) = \\ = 1 + \alpha^{-1}Ph(x) < +\infty.$$

The fact that X is recurrent is therefore a consequence of (A).

C1) Thanks to Lemma 1.2, for every function $f: \mathbb{N} \to \mathbb{R}$,

$$\mathbf{E}(f(X_{n+1}) \mid \mathscr{F}_n) = \mathbf{E}(f(Y_{n+1} + (X_n - Z_{n+1})^+) \mid \mathscr{F}_n) = \psi_f(X_n)$$

where $\psi_f(x) = \mathbf{E}(f(Y_1 + (x - Z_1)^+))$ (see also Exercise 4.1). Therefore, with respect to $(\mathscr{F}_n)_{n\ge0}$, $(X_n)_{n\ge0}$ is a Markov chain with transition matrix

$$P(x, y) = \mathbf{P}(Y_1 + (x - Z_1)^+ = y).$$

If $x > K$, then $(x - Z_1)^+ = x - Z_1$ a.s. since $Z_1 \le K$ a.s.; therefore, if $h(x) = x$, one has for $x > K$

$$Ph(x) = \mathbf{E}(Y_1 + (x - Z_1)) = x - \underbrace{\mathbf{E}(Z_1 - Y_1)}_{=\alpha>0} = h(x) - \alpha.$$

Moreover the condition $Ph(y) < +\infty$ for every $y \in \mathbb{N}$ is obvious.

C2) If $x > 0$, $P(x, x - 1) = \mathbf{P}(Y_1 + (x - Z_1)^+ = x - 1) \ge \mathbf{P}(Z_1 = 1, Y_1 = 0) > 0$. Therefore every state leads to its left neighbour. Next if $k_1 > K$ is such that $\mathbf{P}(Y_1 = k_1) > 0$, then $P(x, x + k_1 - 1) \ge \mathbf{P}(Y = k_1, Z_1 = 1) > 0$.Therefore $x \rightsquigarrow x + k_1 - 1 \rightsquigarrow x + k_1 - 2 \rightsquigarrow \dots \rightsquigarrow x + 1$. Every state leads to its right and left neighbours and the chain is therefore irreducible. Applying the previous results, we deduce that it is recurrent.

P4.9 A1) By hypothesis, the g.c.d. of $I(z)$ is equal to 1. Thus there exist numbers $a_1, \dots, a_r \in I(z)$ such that g.c.d.$(a_1, \dots, a_r) = 1$. It is therefore sufficient to apply the Bézout lemma.

A2) Observe that, if $a, b \in I(z)$ and, if $n, m \in \mathbb{N}$, $na + mb \in I(z)$ since

$$Q^{na+mb}(z, z) \ge Q^{na}(z, z)Q^{mb}(z, z) \ge Q^a(z, z)^n Q^b(z, z)^m > 0.$$

Therefore $n_1 \in I(z)$ and, since the numbers $|\alpha_k| + \alpha_k$ are positive, $n_1 + 1 = \sum_{k=1}^r (|\alpha_k| + \alpha_k)a_k \in I(z)$. If $n \ge n_1^2$, then $n - n_1^2 = mn_1 + s$ with $s < n_1$, hence $n = (n_1 + 1)s + n_1(n_1 - s + m) \in I(z)$.

A3) Let $x, y \in E$. Let us choose n_2 and n_3 such that $Q^{n_2}(x, z) > 0$ and $Q^{n_3}(z, y) > 0$. For $n \ge n_1^2$,

$$Q^{n_2+n+n_3}(x, y) \ge Q^{n_2}(x, z)Q^n(z, z)Q^{n_3}(z, y) > 0.$$

It is now sufficient to choose $N = n_1^2 + n_2 + n_3$.

B) Let $x, y, x', y' \in E$. By (A), there exists $\bar{N}$ such that, for every $n \ge \bar{N}$, $Q^n(x, y) > 0$, $Q^n(x', y') > 0$ and therefore also $\bar{Q}^n((x, x'), (y, y')) > 0$. This implies that all the states communicate and the chain $\bar{X}$ is irreducible. As the set $I(d)$ of •4.31 contains two successive integers, its g.c.d. is equal to 1 and $\bar{X}$ is also aperiodic. As $\bar{X}$ possesses a stationary probability, it is positive recurrent.

C1) Let $z \in E$ and denote $\tau_{(z,z)}$ the hitting time of (z, z) for the product chain $\bar{X}$. One has $\tau \le \tau_{(z,z)}$ and $\bar{\mathbf{P}}_{\alpha\otimes\beta}(\tau_{(z,z)} < +\infty) = 1$, as the chain $\bar{X}$ is recurrent.

C2) Let us point out that $X_\tau = X'_\tau$ and, for every $z \in E$,

$$\bar{\mathbf{P}}_{(z,z)}(X_k = y) = Q^k(z, y) = \bar{\mathbf{P}}_{(z,z)}(X'_k = y).$$

Therefore, for $m \le n$,

$$\begin{aligned}\bar{\mathbf{P}}_{\alpha\otimes\beta}(X_n = y, \tau = m) &= \bar{\mathbf{E}}_{\alpha\otimes\beta}(1_{\{\tau=m\}} X_{n-m} \circ \theta_m = y) = \\ &= \bar{\mathbf{E}}_{\alpha\otimes\beta}(1_{\{\tau=m\}} \bar{\mathbf{P}}_{(X_\tau, X'_\tau)}(X_{n-m} = y)) = \\ &= \bar{\mathbf{E}}_{\alpha\otimes\beta}(1_{\{\tau=m\}} \bar{\mathbf{P}}_{(X_\tau, X'_\tau)}(X'_{n-m} = y)) = \bar{\mathbf{P}}_{\alpha\otimes\beta}(X'_n = y, \tau = m)\end{aligned}$$

and it is sufficient to sum for m from 0 to n.

C3) Indeed

$$\begin{aligned}\mathbf{P}_\alpha(X_n = y) &= \bar{\mathbf{P}}_{\alpha\otimes\beta}(X_n = y, \tau \le n) + \bar{\mathbf{P}}_{\alpha\otimes\beta}(X_n = y, \tau > n) = \\ &= \bar{\mathbf{P}}_{\alpha\otimes\beta}(X'_n = y, \tau \le n) + \bar{\mathbf{P}}_{\alpha\otimes\beta}(X_n = y, \tau > n) \le \\ &\le \mathbf{P}_\beta(X'_n = y) + \bar{\mathbf{P}}_{\alpha\otimes\beta}(\tau > n).\end{aligned}$$

The second formula is proved in the same way. One deduces that

$$\left|\mathbf{P}_\alpha(X_n = y) - \mathbf{P}_\beta(X'_n = y)\right| \le \mathbf{P}_{\alpha\otimes\beta}(\tau > n) \underset{n\to\infty}{\to} 0$$

as $\mathbf{P}_{\alpha\otimes\beta}(\tau < +\infty) = 1$.

C4) Let us choose $\alpha = \delta_x$ and $\beta = \pi$. Then

$$\begin{aligned}\mathbf{P}_\alpha(X_n = y) &= \mathbf{P}_x(X_n = y) = Q^n(x, y) \\ \mathbf{P}_\beta(X'_n = y) &= \mathbf{P}_\pi(X'_n = y) = \pi(y)\end{aligned}$$

and, by (C3), $|Q^n(x, y) - \pi(y)| \to_{n\to\infty} 0$.

P4.10 1) One has

$$\begin{aligned}\mathbf{P}_x(Y_1 = y_1, \ldots, Y_n = y_n) &= \mathbf{P}_x(X_1 - X_0 = y_1, \ldots, X_n - X_{n-1} = y_n) = \\ &= \mathbf{P}_x(X_1 = x + y_1, X_2 = x + y_1 + y_2, \ldots, X_n = x + y_1 + \ldots + y_n) = \\ &= P(x, x+y_1)P(x+y_1, x+y_1+y_2)\ldots P(x+y_1+\ldots+y_{n-1}, x+y_1+\ldots+y_n) = \\ &= p(y_1)\ldots p(y_n),\end{aligned}$$

which simultaneously proves that the r.v.'s $Y_1, \ldots, Y_n$ are independent and have the same distribution, p, under $\mathbf{P}_x$ for every $x \in \mathbb{Z}$.

2) For every $n \ge 1$, $X_n = x + Y_1 + \ldots + Y_n$ $\mathbf{P}_x$-a.s. and

$$P^n(x, y) = \mathbf{P}_x(X_n = y) = \mathbf{P}_x(Y_1 + \ldots + Y_n = y - x) = p_n(y - x)$$

whence

$$U(x, y) = 1_{\{x=y\}} + \sum_{n\ge1} p_n(y - x)$$

and thus $U(x, y) = U(0, y - x)$.

3) By the maximum principle (•4.20), $U(0, x) \le U(x, x) = U(0, 0)$.

4) Let $x, y \in \mathbb{Z}$. By hypothesis, there exist $z_1, \ldots, z_k \in \mathbb{Z}$ such that $y - x =$

$z_1 + \ldots + z_k$ and $p(z_i) > 0$, $1 \le i \le k$. Therefore

$$\begin{aligned} P^k(x, y) &= \mathbf{P}_x(X_k = y) = \mathbf{P}_x(Y_1 + \ldots + Y_k = y) \ge \\ &\ge \mathbf{P}_x(Y_1 = z_1, \ldots, Y_k = z_k) = p(z_1) \ldots p(z_k) > 0. \end{aligned}$$

Thus $x \rightsquigarrow y$; x and y being arbitrary, the chain is irreducible.

5) Let us suppose $\mu > 0$. As $X_n = X_0 + Y_1 + \ldots + Y_n$, by the strong law of large numbers, $\mathbf{P}_x$-a.s. $\frac{1}{n} X_n \to_{n\to\infty} \mathbf{E}_x(Y_1) = \mu$ and $X_n \to_{n\to\infty} +\infty$. Therefore each state can be visited only a finite number of times $\mathbf{P}_x$-a.s. and the chain is transient. The case $\mu < 0$ can be treated in the same way.

6) One has

$$\mathbf{P}_0(|X_n| \le k) = \sum_{y=-k}^{k} \mathbf{P}_0(X_n = y)$$

and therefore, recalling (3),

$$\begin{aligned} \phi(k) &= \frac{1}{k} \sum_{y=-k}^{k} \sum_{n\ge 0} \mathbf{P}_0(X_n = y) = \frac{1}{k} \sum_{y=-k}^{k} U(0, y) \le \\ &\le \frac{2k+1}{k} U(0,0) \le 3U(0,0) < +\infty. \end{aligned}$$

7) Let $\varepsilon > 0$. As $\mathbf{P}_0(\frac{1}{n}|X_n| \ge \varepsilon) \to_{n\to\infty} 0$ by the law of large numbers, there exists N_ε such that, if $n \ge N_\varepsilon$, $\mathbf{P}_0(|X_n| \le \varepsilon n) \ge \frac{1}{2}$. Thus, for k large enough,

$$\begin{aligned} \phi(k) &\ge \frac{1}{k} \sum_{n=N_\varepsilon}^{[k/\varepsilon]} \mathbf{P}_0(|X_n| \le \varepsilon \tfrac{k}{\varepsilon}) \ge \frac{1}{k} \sum_{n=N_\varepsilon}^{[k/\varepsilon]} \mathbf{P}_0(|X_n| \le \varepsilon n) \ge \\ &\ge \frac{1}{2k}\Big(\frac{k}{\varepsilon} - N_\varepsilon\Big) = \frac{1}{2}\Big(\frac{1}{\varepsilon} - \frac{N_\varepsilon}{k}\Big) \end{aligned}$$

($[x]$ denotes the integer part of x). According to (6), in order to prove that the chain is recurrent, it is sufficient to show that ϕ is not bounded. But this comes from the fact that, for every $\varepsilon > 0$, $\underline{\lim}_{k\to\infty} \phi(k) \ge \frac{1}{2\varepsilon}$.

P4.11 A) The answer to this question may be found in the solution of other exercises: Exercise 4.16 in particular (where it is assumed $d = 1$). See also Exercise 1.8, where another method is used, and Exercise 4.1 for a more general framework.

B1) As

$$\sum_{x, y\in\mathbb{Z}^d} |e^{i\langle\theta, x\rangle} f(y) g(x-y)| = \sum_{y\in\mathbb{Z}^d} |f(y)| \sum_{x\in\mathbb{Z}^d} |g(x)| < +\infty,$$

one may apply Fubini's theorem (with respect to the product of two counting measures) and obtain

$$\sum_{x\in\mathbb{Z}^d} h(x) e^{i\langle\theta,x\rangle} = \sum_{x\in\mathbb{Z}^d} e^{i\langle\theta,x\rangle} \sum_{y\in\mathbb{Z}^d} f(y) g(x-y) =$$

$$= \sum_{y\in\mathbb{Z}^d} f(y) \sum_{x\in\mathbb{Z}^d} g(x-y)\,\mathrm{e}^{i\langle\theta,x\rangle} = \sum_{y\in\mathbb{Z}^d} f(y) \underbrace{\sum_{x\in\mathbb{Z}^d} g(z)\,\mathrm{e}^{i\langle\theta,z\rangle}}_{=\hat g(\theta)}\,\mathrm{e}^{i\langle\theta,y\rangle} =$$

$$= \hat f(\theta)\hat g(\theta).$$

B2) One has

$$|\hat f(\theta)| \le \sum_{x\in\mathbb{Z}^d} |f(x)\,\mathrm{e}^{i\langle\theta,x\rangle}| = \sum_{x\in\mathbb{Z}^d} |f(x)|$$

so that $\hat f$ is bounded. Thanks to Fubini's Theorem, with respect to the product of the counting measure and of the Lebesgue measure on Q_d (the integrability of the function is easily checked as in (B1)),

$$\int_{Q_d} \hat f(\theta)\,\mathrm{e}^{-i\langle\theta,x\rangle}\,d\theta = \int_{Q_d} \mathrm{e}^{-i\langle\theta,x\rangle} \sum_{y\in\mathbb{Z}^d} f(y)\,\mathrm{e}^{i\langle\theta,y\rangle}\,d\theta =$$

$$= \sum_{y\in\mathbb{Z}^d} f(y) \int_{Q_d} \mathrm{e}^{-i\langle\theta,y-x\rangle}\,d\theta.$$

But one immediately checks that

$$\int_{Q_d} \mathrm{e}^{-i\langle\theta,y-x\rangle}\,d\theta = \begin{cases} 0 & \text{if } y \ne x \\ (2\pi)^d & \text{if } y = x; \end{cases}$$

hence the requested result.

B3) With the notation $\theta = (\theta_1, \dots, \theta_d)$,

$$\hat\mu(\theta) = \sum_{x\in\mathbb{Z}^d} \mu(x)\,\mathrm{e}^{i\langle\theta,x\rangle} = \sum_{k=1}^{d} \frac{1}{2d}(\mathrm{e}^{i\theta_k} + \mathrm{e}^{-i\theta_k}) = \frac{1}{d}\sum_{k=1}^{d}\cos\theta_k.$$

If $\theta \in Q_d$ and $\theta \ne 0$, then $\cos\theta_k < 1$ for one k at least and so $\hat\mu(\theta) < 1$. Moreover, as $\theta \to 0$,

$$1 - \hat\mu(\theta) = \frac{1}{d}\sum_{k=1}^{d}(1-\cos\theta_k) \sim \frac{1}{d}\sum_{k=1}^{d}\frac{\theta_k^2}{2} = \frac{|\theta|^2}{2d}.$$

C1) As we deal with positive series and the functions $n \to \lambda^n P^n(x,y)$ are increasing in λ, applying the monotone convergence theorem to the counting measure on $\mathbb{N}$,

$$\lim_{\lambda\uparrow 1} U_\lambda(x,y) = \lim_{\lambda\uparrow 1}\sum_{n=0}^{\infty}\lambda^n P^n(x,y) = \sum_{n=0}^{\infty} P^n(x,y) = U(x,y).$$

C2) By (A), $P^n(0,y) = \mathrm{P}(X_1 + \dots + X_n = y) = \mu^{*n}(y)$. Hence $u_\lambda(y) = \sum_{n=0}^{\infty}\lambda^n\mu^{*n}(y)$ and, by Fubini's theorem and (B1),

$$\hat u_\lambda(\theta) = \sum_{n=0}^{\infty}\lambda^n\hat\mu^n(\theta) = \frac{1}{1-\lambda\hat\mu(\theta)}.$$

C3) Thanks to the inversion formula (4.73) and to (C1),

$$U(0,0) = \lim_{\lambda\uparrow 1} U_\lambda(0,0) = \lim_{\lambda\uparrow 1} u_\lambda(0) =$$

$$= \lim_{\lambda \uparrow 1} \frac{1}{(2\pi)^d} \int_{Q_d} \hat{u}_\lambda d\theta = \lim_{\lambda \uparrow 1} \frac{1}{(2\pi)^d} \int_{Q_d} \frac{1}{1 - \lambda \hat{\mu}(\theta)} d\theta.$$

If we were allowed to take the limit under the integral sign we would get

$$U(0,0) = \frac{1}{(2\pi)^d} \int_{Q_d} \frac{1}{1 - \hat{\mu}(\theta)} d\theta.$$

The function in the integral is bounded except in 0 where it diverges as $2d|\theta|^{-2}$. Thus it is integrable in 0 if and only if $d \geq 3$. One would therefore have

$$U(0,0) \begin{cases} = +\infty & \text{if } d = 1,2 \\ < +\infty & \text{if } d \geq 3. \end{cases} \tag{4.102}$$

It remains to prove that we can take the limit in this way, i.e., that

$$\lim_{\lambda \uparrow 1} \int_{Q_d} \frac{1}{1 - \lambda \hat{\mu}(\theta)} d\theta = \int_{Q_d} \frac{1}{1 - \hat{\mu}(\theta)} d\theta.$$

A possible way to check this is to divide the cube Q_d into two parts: a small neighbourhood $\mathscr{U}$ of 0 and its complementary set $Q_d \setminus \mathscr{U}$. One chooses $\mathscr{U}$ small enough so that $\hat{\mu}(\theta) \geq 0$ for $\theta \in \mathscr{U}$. With this choice, the functions $\theta \to (1 - \lambda \hat{\mu}(\theta))^{-1}$ are increasing in λ on $\mathscr{U}$. Therefore by the monotone convergence theorem

$$\lim_{\lambda \uparrow 1} \int_{\mathscr{U}} \frac{1}{1 - \lambda \hat{\mu}(\theta)} d\theta = \int_{\mathscr{U}} \frac{1}{1 - \hat{\mu}(\theta)} d\theta.$$

Conversely, on $Q_d \setminus \mathscr{U}$, the upper bound

$$0 \leq \frac{1}{1 - \lambda \hat{\mu}(\theta)} \leq \max(1, (1 - \hat{\mu}(\theta))^{-1})$$

holds (one dominates by 1 if $\hat{\mu}(\theta) \leq 0$ and by $(1 - \hat{\mu}(\theta))^{-1}$ if $\hat{\mu}(\theta) \geq 0$). Lebesgue's theorem therefore applies, as the function $\theta \to (1 - \hat{\mu}(\theta))^{-1}$ is bounded outside of $\mathscr{U}$ (by the explicit expression of (B3)).

As the chain is irreducible (it is easy to see that all the states communicate), it is sufficient to study the recurrence of state 0. Now (4.102) implies that 0 is recurrent for $d \leq 2$ and transient for $d \geq 3$.

♦ This problem deals with a particular instance of Markov chains called *random walks*. These are the chains that appear when one considers a sequence $(X_n)_{n \geq 1}$ of r.v.'s taking values in a discrete group; define $S_n = X_1 + \ldots + X_n$, where "+" denotes the operation of the group. The arguments of (A) may be repeated without modification in order to obtain that $(S_n)_{n \geq 0}$ is a Markov chain with a transition matrix given by (4.72). In the framework of random walks the Fourier techniques developed in (C) are a classical tool.

♦ The proof of the transience of the simple random walks for $d \geq 3$, as seen in (C), may be easily extended to random walks on $\mathbb{Z}^d$ (as they were described in the previous comment) with more general distributions μ. Indeed in (C) one only uses the fact that $1 - \hat{\mu}(\theta)$ behaves near 0 as a quadratic form in θ and that the function $\theta \to |\theta|^{-2}$ is integrable at the origin as soon as $d \geq 3$. These properties remain true for every probability measure μ on $\mathbb{Z}^d$ that is symmetric ($\mu(x) = \mu(-x)$) and has an invertible covariance matrix.

4.12 A) Let λ be an eigenvalue of P and $z \neq 0$ an associated eigenvector. One may ppose $\sup_{1\le i\le d} |z_i| = 1 = |z_k|$. Then,

$$|\lambda| = |\lambda z_k| = |(Pz)_k| = \Big|\sum_{i=1}^{d} p_{ki} z_i\Big| \le \sum_{i=1}^{d} p_{ki}|z_i| \le \sum_{i=1}^{d} p_{ki} = 1.$$

B1) The subspaces M_0 and M_1 have obviously dimensions $d-1$ and 1, respectively. order to prove that $M(d) = M_0 \oplus M_1$, it suffices to prove that $M_0 \cap M_1 = \{0\}$. But $y = (y_1, \dots, y_d) \in M_0 \cap M_1$, y is of the form $y = t\pi$, hence, $y_1 + \dots + y_d = t$. is sum can vanish only if $t = 0$ so that $y = 0$, which implies $M_0 \cap M_1 = \{0\}$. M_1 invariant by the definition of a stationary distribution. If now $x \in M_0$,

$$\sum_{i=1}^{d}(xP)_i = \sum_{i=1}^{d}\sum_{j=1}^{d} x_j p_{ji} = \sum_{j=1}^{d} x_j \sum_{i=1}^{d} p_{ji} = \sum_{j=1}^{d} x_j = 0;$$

ence, $xP \in M_0$ and M_0 is also invariant.

B2) The representative matrix of the transformation $x \to xP$, in a basis of $M(d)$ hosen in such a way that its $d-1$ first elements form a basis of M_0 and with the -th equal to π, is of the form

$$\begin{pmatrix} & & & 0 \\ & B & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

he characteristic polynomial $\Pi(\lambda)$ of P is equal to the characteristic polynomial f B multiplied by $1-\lambda$. But B is the matrix associated to u_0 in the chosen basis; ence, $\Pi_u(\lambda) = (1-\lambda)\Pi_{u_0}(\lambda)$. Obviously $u(x) = u(x_0 + x_1) = u_0(x_0) + x_1$ and, y induction, $u^n(x) = u_0^n(x_0) + x_1$.

B3) Note first that, if P is irreducible, there exists a unique stationary probability moreover having all components strictly positive (Theorem 4.21). Note next that, f $x_0 \in M_0$, $x_0 \neq 0$, $\mu = \alpha x_0 + \pi$ is a probability measure as soon as $0 \le \alpha$ is small nough (so that all the coordinates of μ are positive) and that obviously $\mu \neq \pi$ if $\alpha \neq 0$.

Let us suppose that the eigenspace associated to 1 has dimension ≥ 2; this would mply that 1 is a multiple root of $\Pi_u(\lambda)$ and, by (B2), that 1 is an eigenvector of u_0. et $x_0 \in M_0$, $x_0 \neq 0$, be such that $u_0(x_0) = x_0$, i.e., $x_0 P = x_0$. For $0 < \alpha \le \alpha_0$, he probability $\mu = \alpha x_0 + \pi$ satisfies $\mu P = \alpha x_0 P + \pi P = \alpha x_0 + \pi = \mu$, i.e., it is nvariant and distinct from π, which is impossible.

C1) One defines an endomorphism $\tilde u$ of $\mathbb{C}^d$ by setting $\tilde u(x) = u(x_1) + iu(x_2)$ for $x = x_1 + ix_2$, $x \in \mathbb{C}^d$, $x_1, x_2 \in \mathbb{R}^d$. It holds $\Pi_u(\lambda) = \Pi_{\tilde u}(\lambda)$.

(i) Let λ be an eigenvalue of u such that $|\lambda| \ge 1$. There exists $x \in \mathbb{C}^d$, $x = x_1 + ix_2 \neq 0$ such that $\tilde u(x) = \lambda x$; thus $\tilde u^n(x) = \lambda^n x$ does not tend to 0. But, $\tilde u^n(x) = u^n(x_1) + iu^n(x_2)$ and so either $u^n(x_1) \not\to 0$ or $u^n(x_2) \not\to 0$ as $n \to \infty$. Therefore, if $u^n(x) \to_{n\to\infty} 0$ for every $x \in \mathbb{R}^d$, then $|\lambda| < 1$ for every eigenvalue of u.

(ii) Suppose that all the eigenvalues of u have a modulus that is < 1. Let us write $\tilde u = v + w$, with v and w as in the text of this problem. Observe first that, for every

$x \in \mathbb{C}^d$ and every polynomial p, $p(n)v^n(x) \to_{n\to\infty} 0$ (it is sufficient to work in a basis where v is diagonal). Let $x \in \mathbb{R}^d$; then, if $w^p = 0$,

$$u^{n+p}(x) = \bar{u}^{n+p}(x) = \sum_{k=0}^{p-1} \binom{n+p}{k} v^n(v^{p-k} \circ w^k(x)) \to 0.$$

C2) a)$\Rightarrow$b). By Theorem 4.29, $P^n(i,j) \to_{n\to\infty} \pi(j)$ but $0 \le P^n(i,j) \le 1$; by Lebesgue's theorem $\mu P^n(j) = \sum_i \mu(i)P^n(i,j) \to_{n\to\infty} \sum_i \mu(i)\pi(j) = \pi(j)$.

b)$\Rightarrow$c). In order to prove (c) it is sufficient, by (B2), to show that all the eigenvalues of u_0 have a modulus that is < 1, or also, because of (C1), that, for every $x_0 \in M_0$, $u_0^n(x_0) = x_0 P^n \to_{n\to\infty} 0$. Consider therefore $x_0 \in M_0$ and the probability measure $\mu = \alpha x_0 + \pi$, $0 < \alpha \le \alpha_0$. One has $x_0 = \alpha^{-1}(\mu - \pi)$ and, by (b),

$$x_0 P^n = \alpha^{-1} = (\mu P^n - \pi) \underset{n\to\infty}{\to} 0.$$

c)$\Rightarrow$a). As P is irreducible, (c) implies that all the eigenvalues of u_0 have a modulus that is < 1 and that, for every $x_0 \in M_0$, $u^n(x_0) = x_0 P^n \to_{n\to\infty} 0$. Let μ be an arbitrary probability measure; then $\mu - \pi \in M_0$ and

$$\mu P^n - \pi = \mu P^n - \pi P^n = (\mu - \pi)P^n \underset{n\to\infty}{\to} 0,$$

i.e., $\mu P^n \to \pi$ as $n \to \infty$. Suppose that P has period $a > 1$. By Theorem 4.30 for every $i \in E$ there exists a subsequence $(n_k)_{k\ge 0}$ such that $P^{n_k}(i,j) \to_{k\to\infty} 0$, i.e., $\delta_i P^{n_k} \to_{k\to\infty} 0$, which contradicts the previous assertion.

♦ It can be proved that if P has period d, the d-th roots of 1 are eigenvalues of P; all other eigenvalues have a modulus that is < 1.

♦ From this problem, one may gather information on the speed of convergence in Theorem 4.29: under the hypotheses of this theorem there exist $b \in]0, 1[$ and $n_0 > 0$ such that, for every $n \ge n_0$,

$$\|\mu P^n - \pi\| \le b^n.$$

It is indeed sufficient to choose b so that $r < b < 1$, r being the largest of the moduli of the eigenvalues different from 1.

P4.13 a) Obviously for $N = 6$

$$P = \begin{pmatrix} r & p & 0 & 0 & 0 & q \\ q & r & p & 0 & 0 & 0 \\ 0 & q & r & p & 0 & 0 \\ 0 & 0 & q & r & p & 0 \\ 0 & 0 & 0 & q & r & p \\ p & 0 & 0 & 0 & q & r \end{pmatrix}.$$

b) If $p > 0$, every state x leads to its contiguous clockwise state y. The same thing is true for y and, by the transitivity of the relation $\rightsquigarrow$, all the states communicate and the chain is irreducible. The same argument applies if $q > 0$, arguing counterclockwise. If moreover $r > 0$, the chain is aperiodic thanks to the criterion of ▪4.33.

c1) If N is even, one observes, as in the figure, that all the states contiguous to a state indexed by an even number are odd and conversely. Therefore, as the chain takes exactly one step at every transition, if one denotes by A the set consisting of ıe odd states and if x is even, $P(x, A) = 1$. It is easily shown by induction that, for odd and x even, $P^n(x, A) = 1$. Therefore $P^n(x, x) = 0$ for every odd n whereas $P^2(x, x) = 2pq > 0$. The set $I(x)$ in •4.31 contains the number 2 and no odd umber: its g.c.d. is therefore equal to 2.

c2) If $m \neq m'$, as $e^{2i\pi(m-m')} = 1$,

$$\langle v^{(m)}, v^{(m')}\rangle = \sum_{h=1}^{N} v_h^{(m)} \overline{v_h^{(m')}} = \sum_{h=1}^{N} e^{2i\pi(m-m')h/N} =$$
$$= e^{2i\pi(m-m')/N} \frac{1 - e^{2i\pi(m-m')}}{1 - e^{2i\pi(m-m')/N}} = 0;$$

ence, the vectors $v^{(m)}$, $m = 0, \ldots, N-1$ are orthogonal. One easily has

$$(Pv^{(m)})_j = p\, e^{2i\pi m(j+1)/N} + r\, e^{2i\pi mj/N} + q\, e^{2i\pi m(j-1)/N} =$$
$$= e^{2i\pi mj/N}(p\, e^{2i\pi m/N} + r + q\, e^{-2i\pi m/N}) = (p\, e^{2i\pi m/N} + r + q\, e^{-2i\pi m/N})v_j^{(m)};$$

ence, the numbers

$$\lambda_m = p\, e^{2i\pi m/N} + r + q\, e^{-2i\pi m/N} \qquad m = 0, \ldots, N-1 \tag{4.103}$$

re eigenvalues of P. As the corresponding N eigenvectors form a basis of $\mathbb{C}^N$ (they re linearly independent by (c2)), the numbers λ_m, $0 \le m \le N-1$ (which may not be listinct), are all the eigenvalues of P. Observe that if $p = q = \frac{1}{2}$, these eigenvalues re

$$\lambda_m = \cos(2\pi \tfrac{m}{N}).$$

Suppose $r = 0$. One observes that, for every $m = 0, \ldots, N-1$, λ_m is a convex combination of the complex numbers $e^{2i\pi m/N}$ and $e^{-2i\pi m/N}$, which have modulus 1. As the ball is strictly convex, one may have $|\lambda_m| = 1$ only if these two numbers are equal. As one is the inverse of the other, this may occur only if $e^{2i\pi m/N} = \pm 1$, i.e., f either $m = 0$ or $m = \frac{1}{2}N$. The first condition implies that $\lambda_m = 1$; the second one s satisfied if N is even. In conclusion, if N is odd, $|\lambda_m| < 1$ for $1 \le m \le N-1$. Conversely if N is even $|\lambda_m| < 1$ for $1 \le m \le N-1$, unless $m = \frac{1}{2}N$ (and in this case $\lambda_{N/2} = -1$).

If $r > 0$, then λ_m is a convex combination of $e^{2i\pi m/N}$, $e^{-2i\pi m/N}$ and 1. These hree numbers having again modulus equal to 1, this is possible only for $m = 0$. One leduces that 1 is the only eigenvalue with modulus 1.

c3) Two methods: one may first observe that

$$P^N(1,1) \ge P(1,2)P(2,3)\ldots P(N,1) \ge p^N > 0.$$

Moreover, $P^2(1,1) \ge P(1,2)P(2,1) = pq > 0$. Hence the set of periods contains 2 and N, which have 1 as g.c.d. The g.c.d. of the set $I(1)$ is therefore 1 and the chain is aperiodic.

Second method: it has been found in (c2) that, if N is odd, all the eigenvalues of P different from 1 have a modulus < 1. This immediately implies (Problem 4.12) that the chain is aperiodic.

P4.14 a) Thanks to Exercise 4.42 (3) the stationary distribution of the product chain

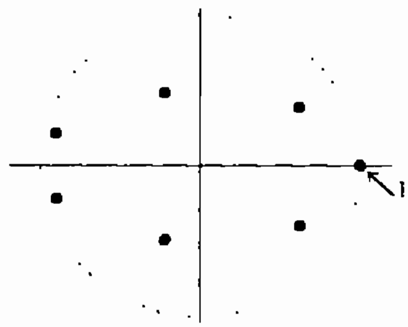

Figure 4.9 Example: the eigenvalues of P for $N = 7$, $p = \frac{1}{4}$, $q = \frac{3}{4}$; the modulus of the two leftmost eigenvalues is 0.927.

is the product of the stationary distributions on each factor. Hence the stationary distribution π is

$$\pi(x) = \mu(x_1)\dots\mu(x_n)$$

where $x = (x_1, \dots, x_m)$, $x_i \in \{0, 1\}$ and where μ is the stationary distribution on each factor, given by (4.75). Hence π is the product of m Bernoulli distributions with parameter $\alpha(\alpha+\beta)^{-1}$:

$$\pi(x) = \Big(\frac{\alpha}{\alpha+\beta}\Big)^{x_1+\dots+x_m}\Big(\frac{\beta}{\alpha+\beta}\Big)^{m-x_1-\dots-x_m}.$$

This stationary distribution is unique since the chain is irreducible. Indeed, if we denote by P its transition matrix, then

$$P(x, y) = P'(x_1, y_1)\dots P'(x_m, y_m) > 0.$$

and hence all the states communicate.

b) The distribution of Y_n is the image of the distribution of X_n by the application ϕ given by the sum of the coordinates. As X is aperiodic and irreducible. $(X_n)_{n\geq 0}$ converges in law to the stationary distribution π. $(Y_n)_{n\geq 0}$ converges therefore in law to the image of the latter by the application ϕ, i.e., to the binomial law $B(m, \alpha^{-1}(\alpha+\beta))$. Since $(Y_n)_n$ is bounded, the mean of Y_n converges to the mean of the limit distribution, i.e., $m\alpha^{-1}(\alpha+\beta)$.

c1) The set $\psi^{-1}(i)$ is formed by the binary vectors having exactly i coordinates equal to 1; thus if $x \in \psi^{-1}(i)$ then also $x_\sigma \in \psi^{-1}(i)$ for every permutation σ. Conversely if $x, y \in \psi^{-1}(i)$, there exists a permutation σ such that $y = x_\sigma$. Indeed, if one denotes by J_x (J_y, respectively) the set of indices such that $x_i = 1$ (resp. $y_i = 1$), then J_x and J_y have the same cardinality (equal to i) and there exists a permutation that transforms J_x in J_y. Hence g is constant on the sets $\psi^{-1}(i)$ if and only if (4.76) holds. Therefore

$$P(x_\sigma, y_\sigma) = \prod_{i=1}^{m} P'(x_{\sigma_i}, y_{\sigma_i}) = \prod_{i=1}^{m} P'(x_i, y_i) = P(x, y) \tag{4.104}$$

as the two products in the previous formula contain the same factors.

c2) In order to apply Dynkin's criterion we need to show that $P(x, \psi^{-1}(j)) =$

$P(y, \psi^{-1}(j))$ if $\psi(x) = \psi(y)$. But, if $\psi(x) = \psi(y)$, there exists a permutation σ such that $y = x_\sigma$. Thanks to (4.104),

$$P(x, \psi^{-1}(j)) = \sum_{z \in \psi^{-1}(j)} P(x, z) = \sum_{z \in \psi^{-1}(j)} P(x_\sigma, z_\sigma) =$$
$$= \sum_{z \in \psi^{-1}(j)} P(y, z) = P(y, \psi^{-1}(j))$$

where one uses the fact that $z \in \psi^{-1}(j)$ is equivalent to $z_\sigma \in \psi^{-1}(j)$. Let us now determine the transition matrix Q of the chain Y: if $x \in E$ is such that $\psi(x) = i$,

$$Q(i, j) = \sum_{y \in \psi^{-1}(j)} P(x, y).$$

But every $y \in \psi^{-1}(j)$ may be reached from x by keeping equal to 1 h among the i coordinates of x which are equal to 1 and by changing from 0 to 1, $j - h$ of the $m - i$ coordinates of x which are 0. The probability for this to occur, for h fixed, is

$$\binom{i}{h}(1-\beta)^h \beta^{i-h} \binom{m-i}{j-h} \alpha^{j-h}(1-\alpha)^{m-j-i+h} \tag{4.105}$$

as there are $\binom{i}{h}$ ways to choose h coordinates among i and $\binom{m-i}{j-h}$ to choose $j-h$ among $m - i$. To complete the computation of $Q(i, j)$ one must sum over all the possible values of h, which enforces the constraints $0 \le h \le i$ and $0 \le j - h \le m - i$. Therefore

$$Q(i, j) = \sum_{h=0 \vee (i+j-m)}^{i \wedge j} \binom{i}{h}(1-\beta)^h \beta^{i-h} \binom{m-i}{j-h} \alpha^{j-h}(1-\alpha)^{m-i-j+h}.$$

Y is irreducible and aperiodic as $Q(i, j) > 0$ for every $i, j \in F$. The chain converges therefore to its unique stationary probability, which, by (b), is binomial $B(m, \alpha^{-1}(\alpha + \beta))$.

♦ One can determine directly the transition matrix Q of the chain Y without constructing the chain X. But the direct computation of the stationary distribution of Q would certainly have been more elaborate.

P4.15 (a) One has

$$\mathbf{P}(T_{n+1} = t_{n+1} \mid T_1 = t_1, \dots, T_n = t_n) =$$
$$= \mathbf{P}(T_n + Z_{n+1} = t_{n+1} \mid T_1 = t_1, \dots, T_n = t_n) =$$
$$= \mathbf{P}(Z_{n+1} = t_{n+1} - t_n) = \mu(t_{n+1} - t_n);$$

hence, $(T_n)_{n \ge 0}$ is a Markov chain with transition matrix $R(x, y) = \mu(y - x)$ (see Exercise 1.8, of which this is a particular case). In particular $R(x, y) = 0$ if $y < x$. Then

$$R^n(x, y) = \mathbf{P}(T_{n+k} = y \mid T_k = x) = \mathbf{P}(Z_1 + \dots + Z_n = y - x) =$$
$$= \mu^{*n}(y - x).$$

Hence $R^n(x, y) = 0$ if $y < x$ whereas, if $y \geq x$,

$$V(x, y) = 1_{\{0\}}(y - x) + \sum_{n \geq 1} \mu^{*n}(y - x) = \gamma(y - x).$$

As here $\mu^{*n}(0) = \mu(0)^n$,

$$\gamma(0) = 1 + \sum_{n \geq 1} \mu(0)^n = \frac{1}{1 - \mu(0)}$$

We know that $V(x, y) \leq V(y, y)$ by the maximum principle (•4.20); hence $\gamma(x) \leq \gamma(0)$.

1b) One may write

$$\mathbf{P}(T_1 = n_1, \ldots, T_k = n_k) = R(0, n_1)R(n_1, n_2) \ldots R(n_{k-1}, n_k) = \\ = \mu(n_1)\mu(n_2 - n_1) \ldots \mu(n_k - n_{k-1}).$$

2a) It is obvious that $S(x, y) \geq 0$; moreover

$$\sum_{y \in E} S(x, y) = \sum_{y \in E} \sum_{n \geq 0} \mu(n) Q^n(x, y) = \\ = \sum_{n \geq 0} \mu(n) \sum_{y \in E} Q^n(x, y) = \sum_{n \geq 0} \mu(n) = 1.$$

2b)

$$\mathbf{P}(Y_0 = y_0, Y_1 = y_1) = \sum_{n \geq 0} \mathbf{P}(Y_0 = y_0, Y_1 = y_1, T_1 = n) = \\ = \sum_{n \geq 0} \mathbf{P}(X_0 = y_0, X_n = y_1)\mathbf{P}(T_1 = n) = \\ = \sum_{n \geq 0} \nu(y_0) Q^n(y_0, y_1)\mu(n) = \nu(y_0) S(y_0, y_1).$$

2c) Let us determine the joint distribution of $(Y_0, \ldots, Y_k)$. One has

$$\mathbf{P}(Y_0 = y_0, \ldots, Y_k = y_k) = \\ = \sum_{n_1 \leq \ldots \leq n_k} \mathbf{P}(X_0 = y_0, \ldots, X_{n_k} = y_k, T_1 = n_1, \ldots, T_k = n_k) = \\ = \sum_{n_1 \leq \ldots \leq n_k} \mathbf{P}(X_0 = y_0, \ldots, X_{n_k} = y_k)\mathbf{P}(T_1 = n_1, \ldots, T_k = n_k) = \\ = \sum_{n_1 \leq \ldots \leq n_k} \nu(y_0) Q^{n_1}(y_0, y_1) \ldots Q^{n_k - n_{k-1}}(y_{k-1}, y_k)\mu(n_1) \ldots \mu(n_k - n_{k-1}) = \\ = \sum_{r_1 \geq 0, \ldots, r_k \geq 0} \nu(y_0) Q^{r_1}(y_0, y_1) \ldots Q^{r_k}(y_{k-1}, y_k)\mu(r_1) \ldots \mu(r_k) = \\ = \nu(y_0) S(y_0, y_1) \ldots S(y_{k-1}, y_k).$$

This implies that $(Y_n)_{n \geq 0}$ is a Markov chain with transition matrix S and initial distribution ν (Theorem 4.3).

3a) One has

$$S^2(x,y) = \sum_{z\in E} S(x,z)S(z,y) = \sum_{z\in E}\sum_{n_1,n_2\geq 0} \mu(n_1)Q^{n_1}(x,z)\mu(n_2)Q^{n_2}(z,y) =$$

$$= \sum_{n_1,n_2\geq 0} \mu(n_1)\mu(n_2)Q^{n_1+n_2}(x,y) = \sum_{n\geq 0} Q^n(x,y)\sum_{m=0}^{n}\mu(n-m)\mu(m) =$$

$$= \sum_{n\geq 0}\mu^{*2}(n)Q^n(x,y).$$

3b) It can be proved by induction that $S^k(x,y) = \sum_{n\geq 0}\mu^{*k}(n)Q^n(x,y)$; hence

$$W(x,y) = 1_{\{x\}}(y) + \sum_{k\geq 1}\sum_{n\geq 0}\mu^{*k}(n)Q^n(x,y) = \sum_{n\geq 0}\gamma(n)Q^n(x,y).$$

3c) As $\gamma(n)\leq\gamma(0)$, $W(x,x)\leq\gamma(0)U(x,x)$, where U is the potential of Q. If x is transient for Q, $U(x,x) < +\infty$ and therefore $W(x,x) < +\infty$ and x is transient for S.

4) If $\lambda Q = \lambda$, $\lambda Q^n = \lambda$ and $\lambda S = \sum_{n\geq 0}\mu(n)\lambda Q^n = \lambda$.

5a) (See for more details Exercise 4.16) $\mathbf{P}(X_{2n+1} = 0 \mid X_0 = 0) = 0$, as X_{2n+1} cannot take odd values given $X_0 = 0$ a.s. On the other hand, by Stirling's formula

$$\mathbf{P}(X_{2n} = 0 \mid X_0 = 0) = 2^{-2n}\binom{2n}{n} \sim \frac{1}{(2\pi n)^{1/2}};$$

hence, $U(0,0) = \sum_{n\geq 0}\mathbf{P}(X_n = 0 \mid X_0 = 0) = +\infty$ and 0 is recurrent. As all the states communicate, the chain is recurrent irreducible.

5b) It has been seen that $\gamma(n) = V(0,n) = \sum_{k\geq 0} R^k(0,n) = \sum_{k\geq 0}\mathbf{P}(T_k = n)$; hence

$$\sum_{n\geq 0}\gamma(n)s^n = \sum_{n\geq 0}\sum_{k\geq 0}\mathbf{P}(T_k = n)s^n = \sum_{k\geq 0}\phi(s)^k = (1-s)^{-\rho}.$$

Recalling the development in power series of the function $s \to (1-s)^{-\rho}$, one gets $\gamma(n) = \frac{1}{n!}\rho(\rho+1)\dots(\rho+n-1)$.

5c) One has

$$W(0,0) = \sum_{n\geq 0}\gamma(n)Q^n(0,0) = \sum_{n\geq 0}\gamma(2n)Q^{2n}(0,0) =$$

$$= \sum_{n\geq 0}\frac{1}{(2n)!}\rho(\rho+1)\dots(\rho+2n-1)2^{-2n}\binom{2n}{n} = \sum_{n\geq 0} u_n$$

with $u_n \sim 2^{\rho-1}\pi^{-1/2}\Gamma(\rho)^{-1}n^{\rho-3/2}$. If $0<\rho<\frac{1}{2}$ then the series with general term u_n is convergent and therefore 0 is transient for $(Y_n)_{n\geq 0}$. As $W(x,y) > 0$ for every $x, y \in \mathbb{Z}$, the chain is irreducible and all the states are transient. Q is recurrent whereas S is transient.

P4.16 A) Let us remark first that, for every x, $(I-Q)u(x) = g(x)$ is equivalent to

$$u(x+1) - u(x) = \rho[u(x) - u(x-1)] - p^{-1}g(x). \tag{4.106}$$

A1) One can compute $\phi(x)$ recursively using (4.106) with $g \equiv 0$, obtaining $\phi(x+1) > \phi(x)$.

A2) Again $\psi(x), a \le x \le b$ is computed recursively, using (4.106)

A3) Let us assume $u(a+1) \neq 0$ and let $v = u(a+1)^{-1} u$. Then $v(a) = 0$, $v(a+1) = 1$ and $(I-Q)v(x) = 0$, $a < x < b$; thus $v = \phi$, which is impossible since $v(b) = 0$ whereas $\phi(b) > \phi(0) = 0$. Thus $u(a+1) = 0$, but (4.106) implies $u \equiv 0$.

A4) Let $w = \alpha + \gamma\phi + \psi$, with γ to be determined. Obviously $(I-Q)w(x) = g(x)$, $a < x < b$, $w(a) = \alpha$ and $w(b) = \alpha + \gamma\phi(b) + \psi(b) = \beta$, if we choose $\gamma = \phi(b)^{-1}(\beta - \alpha - \psi(b))$, which is allowed since $\phi(b) > 0$. The uniqueness comes from (A3).

B1) Since f is bounded, M_n is integrable and

$$\mathbf{E}(M_{n+1} - M_n \mid \mathscr{F}_n) = \mathbf{E}(f(X_{n+1}) \mid \mathscr{F}_n) - Qf(X_n) = 0,$$

so that $(M_n)_{n\ge 0}$ is a martingale. Moreover

$$f(X_n) - f(X_0) = \sum_{k=0}^{n-1}[f(X_{k+1}) - f(X_k)] =$$

$$= \sum_{k=0}^{n-1}[f(X_{k+1}) - Qf(X_k)] + \sum_{k=0}^{n-1}[Qf(X_k) - f(X_k)].$$

B2) Let $x \in]a,b[$ and denote by $\tilde{u}_1$ the extension of u_1 to $\mathbb{Z}$ obtained by setting $\tilde{u}_1(x) = 0$ for $x \notin [a,b]$. By (B1)

$$M_n = \tilde{u}_1(X_n) - \tilde{u}_1(X_0) + \sum_{k=0}^{n-1}(I-Q)\tilde{u}_1(X_k)$$

is a martingale and $\mathbf{E}_x(M_1) = 0$. By the stopping theorem, $\mathbf{E}_x(M_{\tau\wedge n}) = 0$. But, if $x \in]a,b[$ and $k < \tau$, $X_k \in]a,b[$ and $(I-Q)u_1(X_k) = 1$ $\mathbf{P}_x$ a.s. Thus

$$\mathbf{E}_x(u_1(X_{\tau\wedge n})) - u_1(x) + \mathbf{E}_x(\tau\wedge n) = 0.$$

We deduce that $\mathbf{E}_x(\tau\wedge n) \le 2\|u_1\|_\infty$. But $\tau\wedge n \uparrow \tau$ and $\mathbf{E}_x(\tau\wedge n) \uparrow \mathbf{E}_x(\tau) \le 2\|u_1\|_\infty < +\infty$ as $n \to \infty$. The r.v. τ is therefore integrable and $\tau < +\infty$ a.s., which implies $u_1(X_{\tau\wedge n}) \to u_1(X_\tau) = 0$ as $n \to \infty$. Since u_1 is bounded, by Lebesgue's theorem, $\mathbf{E}_x(u_1(X_{\tau\wedge n})) \to_{n\to\infty} 0$ and $\mathbf{E}_x(\tau) = u_1(x)$.

B3) As above one gets, for $x \in]a,b[$,

$$0 = \mathbf{E}_x(M_{\tau\wedge n}) = \mathbf{E}_x(u(X_{\tau\wedge n})) - u(x) + \mathbf{E}_x\Big[\sum_{k=0}^{\tau\wedge n-1}(I-Q)u(X_k)\Big] =$$

$$= \mathbf{E}_x(u(X_{\tau\wedge n})) - u(x) + \mathbf{E}_x\Big[\sum_{k=0}^{\tau\wedge n-1} g(X_k)\Big].$$

Taking the limit as $n \to \infty$.

$$u(x) = \mathbf{E}_x(u(X_\tau)) + \mathbf{E}_x\Big[\sum_{k=0}^{\tau-1} g(X_k)\Big].$$

The passage to the limit for the first term was justified above. As for the second one, let us remark that $|\sum_{k=0}^{\tau\wedge n-1} g(X_k)| \le \tau \|g\|_\infty$ and that τ is $\mathbf{P}_x$-integrable. Finally, as

$$\mathbf{E}_x(u(X_\tau)) = \alpha \mathbf{P}_x(X_\tau = a) + \beta \mathbf{P}_x(X_\tau = b),$$

the relation is proved.

C1) For $x \in]a,b[$ $\tau \ge 1$ $\mathbf{P}_x$-a.s. and

$$\tau = 1 + \inf\{n \ge 0;\ X_{n+1} \notin]a,b[\} = 1 + \inf\{n \ge 0;\ X_n \circ \theta_1 \notin]a,b[\} = \\ = 1 + \tau \circ \theta_1.$$

C2) It holds $u_2(a) = u_2(b) = 0$. If $x \in]a,b[$,

$$u_2(x) = \mathbf{E}_x(\tau^2) = \mathbf{E}_x((1 + \tau\circ\theta_1)^2) = \\ = 1 + 2\mathbf{E}_x(\tau\circ\theta_1) + \mathbf{E}_x(\tau^2\circ\theta_1) = 1 + 2\mathbf{E}_x[\mathbf{E}_{X_1}(\tau)] + \mathbf{E}_x[\mathbf{E}_{X_1}(\tau^2)] = \\ = 1 + 2\mathbf{E}_x(u_1(X_1)) + \mathbf{E}_x(u_2(X_1)) = 1 + 2Qu_1(x) + Qu_2(x)$$

so that $(I - Q)u_2(x) = 1 + 2Qu_1(x)$.

C3) It holds $u_3(a) = u_3(b) = 0$. For $x \in]a,b[$,

$$u_3(x) = \mathbf{E}_x(\tau^3) = \mathbf{E}_x((1 + \tau\circ\theta_1)^3) = \\ = 1 + 3\mathbf{E}_x(\tau\circ\theta_1) + 3\mathbf{E}_x(\tau^2\circ\theta_1) + \mathbf{E}_x(\tau^3\circ\theta_1) = \\ = \dots = 1 + 3Qu_1(x) + 3Qu_2(x) + Qu_3(x)$$

so that $(I - Q)u_3(x) = 1 + 3Qu_1(x) + 3Qu_2(x)$.

D1) If $x \notin]a,b[$, $\mathbf{P}_x(\tau = 0) = 1$. If $x \in]a,b[$,

$$\mathbf{P}_x(\tau \le m) \ge \mathbf{P}_x(X_1 = x+1, X_2 = x+2, \dots, X_{b-x} = b) \ge \\ \ge p^{b-x} \ge p^m \overset{\text{def}}{=} \gamma > 0.$$

Therefore

$$\mathbf{P}_x(\tau > km) = \mathbf{P}_x(\tau > (k-1)m, \tau\circ\theta_{(k-1)m} > m) = \\ = \mathbf{E}_x[1_{\{\tau > (k-1)m\}} \mathbf{P}_{X_{(k-1)m}}(\tau > m)] \le \\ \le (1-\gamma)\mathbf{P}_x(\tau > (k-1)m) \le \dots \le (1-\gamma)^k.$$

D2) If $t > m$, then $t = km + r$ with $r < m$ and

$$\mathbf{P}_x(\tau > t) \le \mathbf{P}_x(\tau > km) \le (1-\gamma)^k \le (1-\gamma)^{\frac{t-m}{m}} = K\,\mathrm{e}^{-\varepsilon t},$$

where $\varepsilon = -\frac{1}{m}\log(1-\gamma) > 0$ and $K = (1-\gamma)^{-1}$. By (D1), for $\lambda < \varepsilon$,

$$\mathbf{E}_x(\mathrm{e}^{\lambda\tau}) = 1 + \lambda\mathbf{E}_x\Big(\int_0^\tau \mathrm{e}^{\lambda t}\,dt\Big) = 1 + \lambda\int_0^\infty \mathbf{E}_x(1_{]0,\tau[}(t))\,\mathrm{e}^{\lambda t}\,dt = \\ = 1 + \lambda\int_0^{+\infty} \mathrm{e}^{\lambda t}\,\mathbf{P}_x(\tau > t)\,dt \le 1 + K\lambda\int_0^{+\infty} \mathrm{e}^{-(\varepsilon-\lambda)t}\,dt < +\infty.$$

Finally, let $s > 0$ and $\lambda < \varepsilon$. It holds, for every $x \geq 0$, $x^s \leq c_s e^{\lambda x}$, so that $\mathbf{E}_x(\tau^s) \leq c_s \mathbf{E}_x(e^{\lambda\tau}) < +\infty$, for every $s \geq 0$.

References

Barbe, Ph. and Ledoux, M., *Probabilité*, Belin, Paris 1998.

Billingsley, P., *Probability and Measure*, John Wiley & Sons, New York. 1979.

Bouleau, N., *Processus Stochastiques et Applications*, Hermann, Paris, 1988.

Brémaud, P., *Markov Chains. Gibbs Fields, Monte Carlo Simulation, and Queues*. Texts in Applied Mathematics, 31. Springer-Verlag, New York, 1999.

Breiman, L., *Probability*, Addison-Wesley, Reading MA. 1968.

Cottrel, M., Duhamel. Ch., Genon-Catalot, V. and Meyre, T., *Exercices de Probabilités*, Cassini, Paris, 1999.

Dacunha-Castelle, D. and Duflo, M., *Probability and Statistics. Vol. II*. Springer-Verlag, New York, 1986.

Dacunha-Castelle, D., Duflo, M., and Genon-Catalot, V., *Exercices de Probabilités et Statistiques, Tome 2: Problèmes à Temps Mobile*, Masson, Paris, 1984.

Dacunha-Castelle, D., Revuz, D. and Schreiber, M., *Recueil de Problèmes de Calcul des Probabilités*, Masson, Paris, 1970.

Dellacherie, C. and Meyer, P-A., *Probabilities and Potential. B. Theory of Martingales*, North-Holland Mathematics Studies, 72. North-Holland Publishing Co., Amsterdam, 1982.

Duflo, M., *Random Iterative Models*, Applications of Mathematics, 34 Springer-Verlag, Berlin, 1997.

Durrett, R., *Probability: Theory and Examples*, Wadsworth and Brooks, Pacific Grove CA, 1991.

Feller, W., *An Introduction to Probability Theory and its Applications, Vol.I*. John Wiley & Sons. New York, 1968.

Grimmett, G. R. and Stirzaker, D. R., *Probability and Random Processes. Solved Problems*. Second edition. The Clarendon Press, Oxford University Press, New York, 1991.

Hoel, P. G., Port, S. C. and Stone, C. J., *Introduction to Stochastic Processes*, Houghton Mifflin, Boston, 1972.

Neveu, J., *Discrete-Parameter Martingales*, North-Holland Mathematical Library, Vol. 10, North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975.

Norris, J. R., *Markov Chains*, Cambridge University Press, Cambridge, 1998.

Revuz, D., *Probabilités*, Hermann, Paris, 1997.

Shiryaev, A. N., *Probability*, Springer-Verlag, New York, 1984.

Williams, D., *Probability with Martingales*, Cambridge Math.Textbooks, Cambridge, 1991.

Index