

# Chapter 15

## Martingales

**Abstract** The definitions, simplest properties and first examples of martingales and sub/super-martingales are given in Sect. 15.1. Stopping (Markov) times are introduced in Sect. 15.2, which also contains Doob’s theorem on random change of time and Wald’s identity together with a number of its applications to boundary crossing problems and elsewhere. This is followed by Sect. 15.3 presenting fundamental martingale inequalities, including Doob’s inequality with a number of its consequences, and an inequality for the number of strip crossings. Section 15.4 begins with Doob’s martingale convergence theorem and also presents Lévy’s theorem and an application to branching processes. Section 15.5 derives several important inequalities for the moments of stochastic sequences.

### 15.1 Definitions, Simplest Properties, and Examples

In Chap. 13 we considered sequences of dependent random variables  $X_0, X_1, \dots$  forming Markov chains. Dependence was described there in terms of transition probabilities determining the distribution of  $X_{n+1}$  given  $X_n$ . That enabled us to investigate rather completely the properties of Markov chains.

In this chapter we consider another type of sequence of dependent random variables. Now dependence will be characterised only by the mean value of  $X_{n+1}$  given the whole “history”  $X_0, \dots, X_n$ . It turns out that one can also obtain rather general results for such sequences.

Let a probability space  $\langle \Omega, \mathfrak{F}, \mathbf{P} \rangle$  be given together with a sequence of random variables  $X_0, X_1, \dots$  defined on it and an increasing family (or flow) of  $\sigma$ -algebras  $\{\mathfrak{F}_n\}_{n \geq 0}$ :  $\mathfrak{F}_0 \subseteq \mathfrak{F}_1 \subseteq \dots \subseteq \mathfrak{F}_n \subseteq \dots \subseteq \mathfrak{F}$ .

**Definition 15.1.1** A sequence of pairs  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  is called a *stochastic sequence* if, for each  $n \geq 0$ ,  $X_n$  is  $\mathfrak{F}_n$ -measurable. A stochastic sequence is said to be a *martingale* (one also says that  $\{X_n\}$  is a *martingale with respect to the flow of  $\sigma$ -algebras  $\{\mathfrak{F}_n\}$* ) if, for every  $n \geq 0$ ,

(1)

$$\mathbf{E}|X_n| < \infty, \tag{15.1.1}$$

(2)  $X_n$  is measurable with respect to  $\mathfrak{F}_n$ ,

(3)

$$\mathbf{E}(X_{n+1} | \mathfrak{F}_n) = X_n. \quad (15.1.2)$$

A stochastic sequence  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  is called a *submartingale* (*supermartingale*) if conditions (1)–(3) hold with the sign “=” replaced in (15.1.2) with “ $\geq$ ” (“ $\leq$ ”, respectively).

We will say that a sequence  $\{X_n\}$  forms a *martingale* (*submartingale*, *supermartingale*) if, for  $\mathfrak{F}_n = \sigma(X_0, \dots, X_n)$ , the pairs  $\{X_n, \mathfrak{F}_n\}$  form a sequence with the same name. Submartingales and supermartingales are often called *semimartingales*.

It is evident that relation (15.1.2) persists if we replace  $X_{n+1}$  on its left-hand side with  $X_m$  for any  $m > n$ . Indeed, by virtue of the properties of conditional expectations,

$$\mathbf{E}(X_m | \mathfrak{F}_n) = \mathbf{E}[\mathbf{E}(X_m | \mathfrak{F}_{m-1}) | \mathfrak{F}_n] = \mathbf{E}(X_{m-1} | \mathfrak{F}_n) = \dots = X_n.$$

A similar assertion holds for semimartingales.

If  $\{X_n\}$  is a martingale, then  $\mathbf{E}(X_{n+1} | \sigma(X_0, \dots, X_n)) = X_n$ , and, by a property of conditional expectations,

$$\mathbf{E}(X_{n+1} | \sigma(X_n)) = \mathbf{E}[\mathbf{E}(X_{n+1} | \sigma(X_0, \dots, X_n)) | \sigma(X_n)] = \mathbf{E}(X_n | \sigma(X_n)) = X_n.$$

So, for martingales, as for Markov chains, we have

$$\mathbf{E}(X_{n+1} | \sigma(X_0, \dots, X_n)) = \mathbf{E}(X_{n+1} | \sigma(X_n)).$$

The similarity, however, is limited to this relation, because for a martingale, the equality does not hold for distributions, but the additional condition

$$\mathbf{E}(X_{n+1} | \sigma(X_n)) = X_n$$

is imposed.

*Example 15.1.1* Let  $\xi_n, n \geq 0$  be independent. Then  $X_n = \xi_1 + \dots + \xi_n$  form a martingale (submartingale, supermartingale) if  $\mathbf{E}\xi_n = 0$  ( $\mathbf{E}\xi_n \geq 0$ ,  $\mathbf{E}\xi_n \leq 0$ ). It is obvious that  $X_n$  also form a Markov chain. The same is true of  $X_n = \prod_{k=0}^n \xi_k$  if  $\mathbf{E}\xi_n = 1$ .

*Example 15.1.2* Let  $\xi_n, n \geq 0$ , be independent. Then

$$X_n = \sum_{k=1}^n \xi_{k-1} \xi_k, \quad n \geq 1, \quad X_0 = \xi_0,$$

form a martingale if  $\mathbf{E}\xi_n = 0$ , because

$$\mathbf{E}(X_{n+1} | \sigma(X_0, \dots, X_n)) = X_n + \mathbf{E}(\xi_n \xi_{n+1} | \sigma(\xi_n)) = X_n.$$

Clearly,  $\{X_n\}$  is not a Markov chain here. An example of a sequence which is a Markov chain but not a martingale can be obtained, say, if we consider a random walk on a segment with reflection at the endpoints (see Example 13.1.1).

As well as  $\{0, 1, \dots\}$  we will use other sets of indices for  $X_n$ , for example,  $\{-\infty < n < \infty\}$  or  $\{n \leq -1\}$ , and also sets of integers including infinite values  $\pm\infty$ , say,  $\{0 \leq n \leq \infty\}$ . We will denote these sets by a common symbol  $\mathcal{N}$  and write martingales (semimartingales) as  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$ . By  $\mathfrak{F}_{-\infty}$  we will understand the  $\sigma$ -algebra  $\bigcap_{n \in \mathcal{N}} \mathfrak{F}_n$ , and by  $\mathfrak{F}_\infty$  the  $\sigma$ -algebra  $\sigma(\bigcup_{n \in \mathcal{N}} \mathfrak{F}_n)$  generated by  $\bigcup_{n \in \mathcal{N}} \mathfrak{F}_n$ , so that  $\mathfrak{F}_{-\infty} \subseteq \mathfrak{F}_n \subseteq \mathfrak{F}_\infty \subseteq \mathfrak{F}$  for any  $n \in \mathcal{N}$ .

**Definition 15.1.2** A stochastic sequence  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is called a *martingale* (*submartingale*, *supermartingale*), if the conditions of Definition 15.1.1 hold for any  $n \in \mathcal{N}$ .

If  $\{X_n, \mathfrak{F}; n \in \mathcal{N}\}$  is a martingale and the left boundary  $n_0$  of  $\mathcal{N}$  is finite (for example,  $\mathcal{N} = \{0, 1, \dots\}$ ), then the martingale  $\{X_n, \mathfrak{F}_n\}$  can be always extended “to the whole axis” by setting  $\mathfrak{F}_n := \mathfrak{F}_{n_0}$  and  $X_n := X_{n_0}$  for  $n < n_0$ . The same holds for the right boundary as well. Therefore if a martingale (semimartingale)  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is given, then without loss of generality we can always assume that one is actually given a martingale (semimartingale)  $\{X_n, \mathfrak{F}_n; -\infty \leq n \leq \infty\}$ .

*Example 15.1.3* Let  $\{\mathfrak{F}_n, -\infty \leq n \leq \infty\}$  be a given sequence of increasing  $\sigma$ -algebras, and  $\xi$  a random variable on  $(\Omega, \mathfrak{F}, \mathbf{P})$ ,  $\mathbf{E}|\xi| < \infty$ . Then  $\{X_n, \mathfrak{F}_n; -\infty \leq n \leq \infty\}$  with  $X_n = \mathbf{E}(\xi | \mathfrak{F}_n)$  forms a martingale.

Indeed, by the property of conditional expectations, for any  $m \leq \infty$  and  $m > n$ ,

$$\mathbf{E}(X_m | \mathfrak{F}_n) = \mathbf{E}[\mathbf{E}(\xi | \mathfrak{F}_m) | \mathfrak{F}_n] = \mathbf{E}(\xi | \mathfrak{F}_n) = X_n.$$

**Definition 15.1.3** The martingale of Example 15.1.3 is called a *martingale generated by the random variable  $\xi$*  (and the family  $\{\mathfrak{F}_n\}$ ).

**Definition 15.1.4** A set  $\mathcal{N}_+$  is called the *right closure* of  $\mathcal{N}$  if:

- (1)  $\mathcal{N}_+ = \mathcal{N}$  when the maximal element of  $\mathcal{N}$  is finite;
- (2)  $\mathcal{N}_+ = \mathcal{N} \cup \{\infty\}$  if  $\mathcal{N}$  is not bounded from the right.

If  $\mathcal{N} = \mathcal{N}_+$  then we say that  $\mathcal{N}$  is *right closed*. A martingale (semimartingale)  $\{X_n, \mathfrak{F}; n \in \mathcal{N}\}$  is said to be *right closed* if  $\mathcal{N}$  is right closed.

**Lemma 15.1.1** A martingale  $\{X_n, \mathfrak{F}; n \in \mathcal{N}\}$  is generated by a random variable if and only if it is right closed.

*The Proof* of the lemma is trivial. In one direction it follows from Example 15.1.3, and in the other from the equality

$$\mathbf{E}(X_N | \mathfrak{F}_n) = X_n, \quad N = \sup\{k; k \in \mathcal{N}\},$$

which implies that  $\{X_n, \mathfrak{F}\}$  is generated by  $X_N$ . The lemma is proved. □

Now we consider an interesting and more concrete example of a martingale generated by a random variable.

*Example 15.1.4* Let  $\xi_1, \xi_2, \dots$  be independent and identically distributed and assume  $\mathbf{E}|\xi_1| < \infty$ . Set

$$S_n = \xi_1 + \dots + \xi_n, \quad X_{-n} = S_n/n, \quad \mathfrak{F}_{-n} = \sigma(S_n, S_{n+1}, \dots) = \sigma(S_n, \xi_{n+1}, \dots).$$

Then  $\mathfrak{F}_{-n} \subset \mathfrak{F}_{-n+1}$  and, for any  $1 \leq k \leq n$ , by symmetry

$$\mathbf{E}(\xi_k | \mathfrak{F}_{-n}) = \mathbf{E}(\xi_1 | \mathfrak{F}_{-n}).$$

From this it follows that

$$S_n = \mathbf{E}(S_n | \mathfrak{F}_{-n}) = \sum_{k=1}^n \mathbf{E}(\xi_k | \mathfrak{F}_{-n}) = n\mathbf{E}(\xi_1 | \mathfrak{F}_{-n}), \quad \frac{S_n}{n} = \mathbf{E}(\xi_1 | \mathfrak{F}_{-n}).$$

This means that  $\{X_n, \mathfrak{F}_n; n \leq 1\}$  forms a martingale generated by  $\xi_1$ .

We will now obtain a series of auxiliary assertions giving the simplest properties of martingales and semimartingales. When considering semimartingales, we will confine ourselves to submartingales only, since the corresponding properties of supermartingales will follow immediately if one considers the sequence  $Y_n = -X_n$ , where  $\{X_n\}$  is a submartingale.

**Lemma 15.1.2**

(1) *The property that  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is a martingale is equivalent to invariability in  $m \geq n$  of the set functions (integrals)*

$$\mathbf{E}(X_m; A) = \mathbf{E}(X_n; A) \tag{15.1.3}$$

for any  $A \in \mathfrak{F}_n$ . In particular,  $\mathbf{E}X_m = \text{const.}$

(2) *The property that  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is a submartingale is equivalent to the monotone increase in  $m \geq n$  of the set functions*

$$\mathbf{E}(X_m; A) \geq \mathbf{E}(X_n; A) \tag{15.1.4}$$

for every  $A \in \mathfrak{F}_n$ . In particular,  $\mathbf{E}X_m \uparrow$ .

*The Proof* follows immediately from the definitions. If (15.1.3) holds then, by the definition of conditional expectation,  $X_n = \mathbf{E}(X_m | \mathfrak{F}_n)$ , and vice versa. Now let (15.1.4) hold. Put  $Y_n = \mathbf{E}(X_m | \mathfrak{F}_n)$ . Then (15.1.4) implies that  $\mathbf{E}(Y_n; A) \geq \mathbf{E}(X_n; A)$  and  $\mathbf{E}(Y_n - X_n; A) \geq 0$  for any  $A \in \mathfrak{F}_n$ . From this it follows that  $Y_n = \mathbf{E}(X_m | \mathfrak{F}_n) \geq X_n$  with probability 1. The converse assertion can be obtained as easily as the direct one. The lemma is proved. □

**Lemma 15.1.3** *Let  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  be a martingale,  $g(x)$  be a convex function, and  $\mathbf{E}|g(X_n)| < \infty$ . Then  $\{g(X_n), \mathfrak{F}_n; n \in \mathcal{N}\}$  is a submartingale.*

*If, in addition,  $g(x)$  is nondecreasing, then the assertion of the theorem remains true when  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is a submartingale.*

The *Proof* of both assertions follows immediately from Jensen’s inequality

$$\mathbf{E}(g(X_{n+1})|\mathfrak{F}_n) \geq g(\mathbf{E}(X_{n+1}|\mathfrak{F}_n)) \geq g(\mathbf{E}(X_n|\mathfrak{F}_n)). \quad \square$$

Clearly, the function  $g(x) = |x|^p$  for  $p \geq 1$  satisfies the conditions of the first part of the lemma, and the function  $g(x) = e^{\lambda x}$  for  $\lambda > 0$  meets the conditions of the second part of the lemma.

**Lemma 15.1.4** *Let  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  be a right closed submartingale. Then, for  $X_n(a) = \max\{X_n, a\}$  and any  $a$ ,  $\{X_n(a), \mathfrak{F}_n; n \in \mathcal{N}\}$  is a uniformly integrable submartingale.*

*If  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is a right closed martingale, then it is uniformly integrable.*

*Proof* Let  $N := \sup\{k : k \in \mathcal{N}\}$ . Then, by Lemma 15.1.3,  $\{X_n(a), \mathfrak{F}_n; n \in \mathcal{N}\}$  is a submartingale. Hence, for any  $c > 0$ ,

$$c\mathbf{P}(X_n(a) > c) \leq \mathbf{E}(X_n(a); X_n(a) > c) \leq \mathbf{E}(X_N(a); X_n(a) > c) \leq \mathbf{E}X_N^+(a)$$

(here  $X^+ = \max(0, X)$ ) and so

$$\mathbf{P}(X_n(a) > c) \leq \frac{1}{c}\mathbf{E}(X_N^+(a)) \rightarrow 0,$$

uniformly in  $n$  as  $c \rightarrow \infty$ . Therefore we get the required uniform integrability:

$$\sup_n \mathbf{E}(X_n(a); X_n(a) > c) \leq \sup_n \mathbf{E}(X_N(a); X_n(a) > c) \rightarrow 0,$$

since  $\sup_n \mathbf{P}(X_n(a) > c) \rightarrow 0$  as  $c \rightarrow \infty$  (see Lemma A3.2.3 in Appendix 3; by truncating at the level  $a$  we avoided estimating the “negative tails”).

If  $\{X_n, \mathfrak{F}_n; n \in \mathcal{N}\}$  is a martingale, then its uniform integrability will follow from the first assertion of the lemma applied to the submartingale  $\{|X_n|, \mathfrak{F}_n; n \in \mathcal{N}\}$ . The lemma is proved.  $\square$

The nature of martingales can be clarified to some extent by the following example.

*Example 15.1.5* Let  $\xi_1, \xi_2, \dots$  be an arbitrary sequence of random variables,  $\mathbf{E}|\xi_k| < \infty$ ,  $\mathfrak{F}_n = \sigma(\xi_1, \dots, \xi_n)$  for  $n \geq 1$ ,  $\mathfrak{F}_0 = (\emptyset, \Omega)$  (the trivial  $\sigma$ -algebra),

$$S_n = \sum_{k=1}^n \xi_k, \quad Z_n = \sum_{k=1}^n \mathbf{E}(\xi_k|\mathfrak{F}_{k-1}), \quad X_n = S_n - Z_n.$$

Then  $\{X_n, \mathfrak{F}_n; n \geq 1\}$  is a martingale. This is a consequence of the fact that

$$\mathbf{E}(S_{n+1} - Z_{n+1}|\mathfrak{F}_n) = \mathbf{E}(X_n + \xi_{n+1} - \mathbf{E}(\xi_{n+1}|\mathfrak{F}_n)|\mathfrak{F}_n) = X_n.$$

In other words, for an arbitrary sequence  $\{\xi_n\}$ , the sequence  $S_n$  can be “compensated” by a so-called “predictable” (in the sense that its value is determined by  $S_1, \dots, S_{n-1}$ ) sequence  $Z_n$  so that  $S_n - Z_n$  will be a martingale.

## 15.2 The Martingale Property and Random Change of Time. Wald's Identity

Throughout this section we assume that  $\mathcal{N} = \{n \geq 0\}$ . Recall the definition of a stopping time.

**Definition 15.2.1** A random variable  $\nu$  will be called a *stopping time* or a *Markov time* (with respect to an increasing family of  $\sigma$ -algebras  $\{\mathfrak{F}_n; n \geq 0\}$ ) if, for any  $n \geq 0$ ,  $\{\nu \leq n\} \in \mathfrak{F}_n$ .

It is obvious that a constant  $\nu \equiv m$  is a stopping time. If  $\nu$  is a stopping time, then, for any fixed  $m$ ,  $\nu(m) = \min(\nu, m)$ , is also a stopping time, since for  $n \geq m$  we have

$$\nu(m) \leq m \leq n, \quad \{\nu(m) \leq n\} = \Omega \in \mathfrak{F}_n,$$

and if  $n < m$  then

$$\{\nu(m) \leq n\} = \{\nu \leq n\} \in \mathfrak{F}_n.$$

If  $\nu$  is a stopping time, then

$$\{\nu = n\} = \{\nu \leq n\} - \{\nu \leq n - 1\} \in \mathfrak{F}_n, \quad \{\nu \geq n\} = \Omega - \{\nu \leq n - 1\} \in \mathfrak{F}_{n-1}.$$

Conversely, if  $\{\nu = n\} \in \mathfrak{F}_n$ , then  $\{\nu \leq n\} \in \mathfrak{F}_n$  and therefore  $\nu$  is a stopping time.

Let a martingale  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  be given. A typical example of a stopping time is the time  $\nu$  at which  $X_n$  first hits a given measurable set  $B$ :

$$\nu = \inf\{n \geq 0 : X_n \in B\}$$

( $\nu = \infty$  if all  $X_n \notin B$ ). Indeed,

$$\{\nu = n\} = \{X_0 \notin B, \dots, X_{n-1} \notin B, X_n \in B\} \in \mathfrak{F}_n.$$

If  $\nu$  is a proper stopping time ( $\mathbf{P}(\nu < \infty) = 1$ ), then  $X_\nu$  is a random variable, since

$$X_\nu = \sum_{n=0}^{\infty} X_n I_{\{\nu \geq n\}}.$$

By  $\mathfrak{F}_\nu$  we will denote the  $\sigma$ -algebra of sets  $A \in \mathfrak{F}$  such that  $A \cap \{\nu = n\} \in \mathfrak{F}_n$ ,  $n = 0, 1, \dots$ . This  $\sigma$ -algebra can be thought of as being generated by the events  $\{\nu \leq n\} \cap B_n$ ,  $n = 0, 1, \dots$ , where  $B_n \in \mathfrak{F}_n$ . Clearly,  $\nu$  and  $X_\nu$  are  $\mathfrak{F}_\nu$ -measurable. If  $\nu_1$  and  $\nu_2$  are two stopping times, then  $\{\nu_2 \geq \nu_1\} \in \mathfrak{F}_{\nu_1}$  and  $\{\nu_2 \geq \nu_1\} \in \mathfrak{F}_{\nu_2}$ , since  $\{\nu_2 \geq \nu_1\} = \bigcup_n [\{\nu_2 = n\} \cap \{\nu_1 \leq n\}]$ .

We already know that if  $\{X_n, \mathfrak{F}_n\}$  is a martingale then  $\mathbf{E}X_n$  is constant for all  $n$ . Will this property remain valid for  $\mathbf{E}X_\nu$  if  $\nu$  is a stopping time? From Wald's identity we know that this is the case for the martingale from Example 15.1.1. In the general case one has the following.

**Theorem 15.2.1** (Doob) *Let  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  be a martingale (submartingale) and  $\nu_1, \nu_2$  be stopping times such that*

$$\mathbf{E}|X_{\nu_i}| < \infty, \quad i = 1, 2, \quad (15.2.1)$$

$$\liminf_{n \rightarrow \infty} \mathbf{E}(|X_n|; \nu_2 \geq n) = 0. \quad (15.2.2)$$

*Then, on the set  $\{\nu_2 \geq \nu_1\}$ ,*

$$\mathbf{E}(X_{\nu_2} | \mathfrak{F}_{\nu_1}) = X_{\nu_1} \quad (\geq X_{\nu_1}). \quad (15.2.3)$$

This theorem extends the martingale (submartingale) property to random time.

**Corollary 15.2.1** *If  $\nu_2 = \nu \geq 0$  is an arbitrary stopping time, then putting  $\nu_1 = n$  (also a stopping time) we have that, on the set  $\nu \geq n$ ,*

$$\mathbf{E}(X_\nu | \mathfrak{F}_n) = X_n, \quad \mathbf{E}X_\nu = \mathbf{E}X_0,$$

*or, which is the same, for any  $A \in \mathfrak{F}_n \cap \{\nu \geq n\}$ ,*

$$\mathbf{E}(X_\nu; A) = \mathbf{E}(X_n; A).$$

For submartingales substitute “=” by “ $\geq$ ”.

*Proof of Theorem 15.2.1* To prove (15.2.3) it suffices to show that, for any  $A \in \mathfrak{F}_{\nu_1}$ ,

$$\mathbf{E}(X_{\nu_2}; A \cap \{\nu_2 \geq \nu_1\}) = \mathbf{E}(X_{\nu_1}; A \cap \{\nu_2 \geq \nu_1\}). \quad (15.2.4)$$

Since the random variables  $\nu_i$  are discrete, we just have to establish (15.2.4) for sets  $A_n = A \cap \{\nu_1 = n\} \in \mathfrak{F}_n$ ,  $n = 0, 1, \dots$ , i.e. to establish the equality

$$\mathbf{E}(X_{\nu_2}; A_n \cap \{\nu_2 \geq n\}) = \mathbf{E}(X_n; A_n \cap \{\nu_2 \geq n\}). \quad (15.2.5)$$

Thus the proof is reduced to the case  $\nu_1 = n$ . We have

$$\begin{aligned} \mathbf{E}(X_n; A_n \cap \{\nu_2 \geq n\}) &= \mathbf{E}(X_n; A_n \cap \{\nu_2 = n\}) + \mathbf{E}(X_n; A_n \cap \{\nu_2 \geq n+1\}) \\ &= \mathbf{E}(X_{\nu_2}; A_n \cap \{\nu_2 = n\}) + \mathbf{E}(X_{n+1}; A_n \cap \{\nu_2 \geq n+1\}). \end{aligned}$$

Here we used the fact that  $\{\nu_2 \geq n+1\} \in \mathfrak{F}_n$  and the martingale property (15.1.3).

Applying this equality  $m - n$  times we obtain that

$$\begin{aligned} &\mathbf{E}(X_{\nu_2}; A_n \cap \{n \leq \nu_2 < m\}) \\ &= \mathbf{E}(X_n; A_n \cap \{\nu_2 \geq n\}) - \mathbf{E}(X_m; A_n \cap \{\nu_2 \geq m\}). \end{aligned} \quad (15.2.6)$$

By (15.2.2) the last expression converges to zero for some sequence  $m \rightarrow \infty$ .

Since

$$A_{n,m} := A_n \cap \{n \leq \nu_2 < m\} \uparrow B_n = A_n \cap \{n \leq \nu_2\},$$

by the property of integrals and by virtue of (15.2.6),

$$\mathbf{E}(X_{\nu_2}; A_n \cap \{n \leq \nu_2\}) = \lim_{m \rightarrow \infty} \mathbf{E}(X_{\nu_2}; A_{n,m}) = \mathbf{E}(X_n; A_n \cap \{\nu_2 \geq n\}).$$

Thus we proved equality (15.2.5) and hence Theorem 15.2.1 for martingales. The proof for submartingales can be obtained by simply changing the equality signs in certain places to inequalities. The theorem is proved.  $\square$

The conditions of Theorem 15.2.1 are far from always being met, even in rather simple cases. Consider, for instance, a fair game (see Examples 4.2.3 and 4.4.5) versus an infinitely rich adversary, in which  $z + S_n$  is the fortune of the first gambler after  $n$  plays (given he has not been ruined yet). Here  $z > 0$ ,  $S_n = \sum_{k=1}^n \xi_k$ ,  $\mathbf{P}(\xi_k = \pm 1) = 1/2$ ,  $\eta(z) = \min\{k : S_k = -z\}$  is obviously a Markov (stopping) time, and the sequence  $\{S_n; n \geq 0\}$ ,  $S_0 = 0$ , is a martingale, but  $S_{\eta(z)} = -z$ . Hence  $\mathbf{E}S_{\eta(z)} = -z \neq \mathbf{E}S_n = 0$ , and equality (15.2.5) does not hold for  $v_1 = 0$ ,  $v_2 = \eta(z)$ ,  $z > 0$ ,  $n > 0$ . In this example, this means that condition (15.2.2) is not satisfied (this is related to the fact that  $\mathbf{E}\eta(z) = \infty$ ).

Conditions (15.2.1) and (15.2.2) of Theorem 15.2.1 can, generally speaking, be rather hard to verify. Therefore the following statements are useful in applications.

Put for brevity

$$\xi_n := X_n - X_{n-1}, \quad \xi_0 := X_0, \quad Y_n := \sum_{k=0}^n |\xi_k|, \quad n = 0, 1, \dots$$

**Lemma 15.2.1** *The condition*

$$\mathbf{E}Y_v < \infty \tag{15.2.7}$$

*is sufficient for (15.2.1) and (15.2.2) (with  $v_i = v$ ).*

*The Proof* is almost evident since  $|X_v| \leq Y_v$  and

$$\mathbf{E}(|X_n|; v > n) \leq \mathbf{E}(Y_v; v > n).$$

Because  $\mathbf{P}(v > n) \rightarrow 0$  and  $\mathbf{E}Y_v < \infty$ , it remains to use the property of integrals by which  $\mathbf{E}(\eta; A_n) \rightarrow 0$  if  $\mathbf{E}|\eta| < \infty$  and  $\mathbf{P}(A_n) \rightarrow 0$ .  $\square$

We introduce the following notation:

$$a_n := \mathbf{E}(|\xi_n| | \mathfrak{F}_{n-1}), \quad \sigma_n^2 := \mathbf{E}(\xi_n^2 | \mathfrak{F}_{n-1}), \quad n = 0, 1, 2, \dots,$$

where  $\mathfrak{F}_{-1}$  can be taken to be the trivial  $\sigma$ -algebra.

**Theorem 15.2.2** *Let  $\{X_n; n \geq 0\}$  be a martingale (submartingale) and  $v$  be a stopping time (with respect to  $\{\mathfrak{F}_n = \sigma(X_0, \dots, X_n)\}$ ).*

(1) *If*

$$\mathbf{E}v < \infty \tag{15.2.8}$$

*and, for all  $n \geq 0$ , on the set  $\{v \geq n\} \in \mathfrak{F}_{n-1}$  one has*

$$a_n \leq c = \text{const}, \tag{15.2.9}$$

*then*

$$\mathbf{E}|X_v| < \infty, \quad \mathbf{E}X_v = \mathbf{E}X_0 \quad (\geq \mathbf{E}X_0). \tag{15.2.10}$$



(2) If, in addition,  $\mathbf{E}\sigma_n^2 = \mathbf{E}\xi_n^2 < \infty$  then

$$\mathbf{E}X_v^2 = \mathbf{E} \sum_{k=1}^v \sigma_k^2. \quad (15.2.11)$$

*Proof* By virtue of Theorem 15.2.1, Corollary 15.2.1 and Lemma 15.2.1, to prove (15.2.10) it suffices to verify that conditions (15.2.8) and (15.2.9) imply (15.2.7). Quite similarly to the proof of Theorem 4.4.1, we have

$$\mathbf{E}|Y_v| = \sum_{n=0}^{\infty} \left( \sum_{k=0}^n \mathbf{E}(|\xi_k|; \nu = n) \right) = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \mathbf{E}(|\xi_k|; \nu = n) = \sum_{k=0}^{\infty} \mathbf{E}(|\xi_k|; \nu \geq k).$$

Here  $\{\nu \geq k\} = \Omega \setminus \{\nu \leq k-1\} \in \mathfrak{F}_{k-1}$ . Therefore, by condition (15.2.9),

$$\mathbf{E}(|\xi_k|; \nu \geq k) = \mathbf{E}(\mathbf{E}(|\xi_k| | \mathfrak{F}_{k-1}); \nu \geq k) \leq c \mathbf{P}(\nu \geq k).$$

This means that

$$\mathbf{E}Y_v \leq c \sum_{k=0}^{\infty} \mathbf{P}(\nu \geq k) = c \mathbf{E}\nu < \infty.$$

Now we will prove (15.2.11). Set  $Z_n := X_n^2 - \sum_0^n \sigma_k^2$ . One can easily see that  $Z_n$  is a martingale, since

$$\mathbf{E}(X_{n+1}^2 - X_n^2 - \sigma_{n+1}^2 | \mathfrak{F}_n) = \mathbf{E}(2X_n\xi_{n+1} + \xi_{n+1}^2 - \sigma_{n+1}^2 | \mathfrak{F}_n) = 0.$$

It is also clear that  $\mathbf{E}|Z_n| < \infty$  and  $\nu(n) = \min(\nu, n)$  is a stopping time. By virtue of Lemma 15.2.1, conditions (15.2.1) and (15.2.2) always hold for the pair  $\{Z_k, \nu(n)\}$ . Therefore, by the first part of the theorem,

$$\mathbf{E}Z_{\nu(n)} = 0, \quad \mathbf{E}X_{\nu(n)}^2 = \mathbf{E} \sum_{k=1}^{\nu(n)} \sigma_k^2. \quad (15.2.12)$$

It remains to verify that

$$\lim_{n \rightarrow \infty} \mathbf{E}X_{\nu(n)}^2 = \mathbf{E}X_v^2, \quad \lim_{n \rightarrow \infty} \mathbf{E} \sum_{k=1}^{\nu(n)} \sigma_k^2 = \mathbf{E} \sum_{k=1}^v \sigma_k^2. \quad (15.2.13)$$

The second equality follows from the monotone convergence theorem ( $\nu(n) \uparrow \nu$ ,  $\sigma_k^2 \geq 0$ ). That theorem implies the former equality as well, for  $X_{\nu(n)}^2 \xrightarrow{a.s.} X_v^2$  and  $X_{\nu(n)}^2 \uparrow$ . To verify the latter claim, note that  $\{X_n^2, \mathfrak{F}_n; n \geq 0\}$  is a martingale, and therefore, for any  $A \in \mathfrak{F}_n$ ,

$$\begin{aligned} \mathbf{E}(X_{\nu(n)}^2; A) &= \mathbf{E}(X_v^2; A \cap \{\nu \leq n\}) + \mathbf{E}(X_n^2; A \cap \{\nu > n\}) \\ &\leq \mathbf{E}(X_v^2; A \cap \{\nu \leq n\}) + \mathbf{E}(\mathbf{E}(X_{n+1}^2 | \mathfrak{F}_n); A \cap \{\nu > n\}) \\ &= \mathbf{E}(X_v^2; A \cap \{\nu < n+1\}) + \mathbf{E}(X_{n+1}^2; A \cap \{\nu \geq n+1\}) \\ &= \mathbf{E}(X_{\nu(n+1)}^2; A). \end{aligned}$$

Thus (15.2.12) and (15.2.13) imply (15.2.11), and the theorem is completely proved.  $\square$

The main assertion of Theorem 15.2.2 for martingales (submartingales):

$$\mathbf{E}X_\nu = \mathbf{E}X_0 \quad (\geq \mathbf{E}X_0) \tag{15.2.14}$$

was obtained as a consequence of Theorem 15.2.1. However, we could get it directly from some rather transparent relations which, moreover, enable one to extend it to improper stopping times  $\nu$ .

A stopping time  $\nu$  is called *improper* if  $0 < \mathbf{P}(\nu < \infty) = 1 - \mathbf{P}(\nu = \infty) < 1$ . To give an example of an improper stopping time, consider independent identically distributed random variables  $\xi_k$ ,  $a = \mathbf{E}\xi_k < 0$ ,  $X_n = \sum_{k=1}^n \xi_k$ , and put

$$\nu = \eta(x) := \min\{k \geq 1 : X_k > x\}, \quad x \geq 0.$$

Here  $\nu$  is finite only for such trajectories  $\{X_k\}$  that  $\sup_k X_k > x$ . If the last inequality does not hold, we put  $\nu = \infty$ . Clearly,

$$\mathbf{P}(\nu = \infty) = \mathbf{P}\left(\sup_k X_k \leq x\right) > 0.$$

Thus, for an arbitrary (possibly improper) stopping time, we have

$$\mathbf{E}(X_\nu; \nu < \infty) = \sum_{k=0}^{\infty} \mathbf{E}(X_k; \nu = k) = \sum_{k=0}^{\infty} [\mathbf{E}(X_k; \nu \geq k) - \mathbf{E}(X_k; \nu \geq k + 1)]. \tag{15.2.15}$$

Assume now that changing the order of summation is justified here. Then, by virtue of the relation  $\{\nu \geq k + 1\} \in \mathfrak{F}_k$ , we get

$$\begin{aligned} \mathbf{E}(X_\nu; \nu < \infty) &= \mathbf{E}X_0 + \sum_{k=0}^{\infty} \mathbf{E}(X_{k+1} - X_k; \nu \geq k + 1) \\ &= \mathbf{E}X_0 + \sum_{k=0}^{\infty} \mathbf{E}I(\nu \geq k + 1)\mathbf{E}(X_{k+1} - X_k | \mathfrak{F}_k). \end{aligned} \tag{15.2.16}$$

Since for martingales (submartingales) the factors  $\mathbf{E}(X_{k+1} - X_k | \mathfrak{F}_k) = 0$  ( $\geq 0$ ), we obtain the following.

**Theorem 15.2.3** *If the change of the order of summation in (15.2.15) and (15.2.16) is legitimate then, for martingales (submartingales),*

$$\mathbf{E}(X_\nu; \nu < \infty) = \mathbf{E}X_0 \quad (\geq \mathbf{E}X_0). \tag{15.2.17}$$

Assumptions (15.2.8) and (15.2.9) of Theorem 15.2.2 are nothing else but conditions ensuring the absolute convergence of the series in (15.2.15) (see the proof of Theorem 15.2.2) and (15.2.16), because the sum of the absolute values of the terms in (15.2.16) is dominated by

$$\sum_{k=1}^{\infty} a_k \mathbf{P}(\nu \geq k + 1) \leq a \mathbf{E}\nu < \infty,$$

where, as before,  $a_k = \mathbf{E}(|\xi_k| \mid \mathfrak{F}_{k-1})$  with  $\xi_k = X_k - X_{k-1}$ . This justifies the change of the order of summation.

There is still another way of proving (15.2.17) based on (15.2.15) specifying a simple condition ensuring the required justification. First note that identity (15.2.17) assumes that the expectation  $\mathbf{E}(X_\nu; \nu < \infty)$  exists, i.e. both values  $\mathbf{E}(X_\nu^\pm; \nu < \infty)$  are finite, where  $x^\pm = \max(\pm x, 0)$ .

**Theorem 15.2.4** 1. *Let  $\{X_n, \mathfrak{F}_n\}$  be a martingale. Then the condition*

$$\lim_{n \rightarrow \infty} \mathbf{E}(X_n; \nu > n) = 0 \quad (15.2.18)$$

*is necessary and sufficient for the relation*

$$\lim_{n \rightarrow \infty} \mathbf{E}(X_n; \nu \leq n) = \mathbf{E}X_0. \quad (15.2.19)$$

*A necessary and sufficient condition for (15.2.17) is that (15.2.18) holds and at least one of the values  $\mathbf{E}(X_\nu^\pm; \nu < \infty)$  is finite.*

2. *If  $\{X_n, \mathfrak{F}_n\}$  is a supermartingale and*

$$\liminf_{n \rightarrow \infty} \mathbf{E}(X_n; \nu > n) \geq 0, \quad (15.2.20)$$

*then*

$$\limsup_{n \rightarrow \infty} \mathbf{E}(X_n; \nu \leq n) \leq \mathbf{E}X_0.$$

*If, in addition, at least one of the values  $\mathbf{E}(X_\nu^\pm; \nu < \infty)$  is finite then*

$$\mathbf{E}(X_\nu; \nu < \infty) \leq \mathbf{E}X_0.$$

3. *If, in conditions (15.2.18) and (15.2.20), we replace the quantity  $\mathbf{E}(X_n; \nu > n)$  with  $\mathbf{E}(X_n; \nu \geq n)$ , the first two assertions of the theorem will remain true.*

*The corresponding symmetric assertions hold for submartingales.*

*Proof* As we have already mentioned, for martingales,  $\mathbf{E}(\xi_k; \nu \geq k) = 0$ . Therefore, by virtue of (15.2.18)

$$\mathbf{E}X_0 = \lim_{n \rightarrow \infty} \left[ \mathbf{E}X_0 + \sum_{k=1}^n \mathbf{E}(\xi_k; \nu \geq k) - \mathbf{E}(X_n; \nu \geq n+1) \right].$$

Here

$$\begin{aligned} \sum_{k=1}^n \mathbf{E}(\xi_k; \nu \geq k) &= \sum_{k=1}^n \mathbf{E}(X_k; \nu \geq k) - \sum_{k=1}^n \mathbf{E}(X_{k-1}; \nu \geq k) \\ &= \sum_{k=1}^n \mathbf{E}(X_k; \nu \geq k) - \sum_{k=1}^{n-1} \mathbf{E}(X_k; \nu \geq k+1). \end{aligned}$$

Hence

$$\begin{aligned} \mathbf{E}X_0 &= \lim_{n \rightarrow \infty} \sum_{k=0}^n [\mathbf{E}(X_k; \nu \geq k) - \mathbf{E}(X_k; \nu \geq k+1)] \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^n \mathbf{E}(X_k; \nu = k) = \lim_{n \rightarrow \infty} \mathbf{E}(X_\nu; \nu \leq n). \end{aligned}$$

These equalities also imply the necessity of condition (15.2.18).

If at least one of the values  $\mathbf{E}(X_\nu^\pm; \nu < \infty)$  is finite, then by the monotone convergence theorem

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{E}(X_n; \nu \leq n) &= \lim_{n \rightarrow \infty} \mathbf{E}(X_n^+; \nu \leq n) - \lim_{n \rightarrow \infty} \mathbf{E}(X_n^-; \nu \leq n) \\ &= \mathbf{E}(X_\nu^+; \nu < \infty) - \mathbf{E}(X_\nu^-; \nu < \infty) = \mathbf{E}(X_\nu; \nu < \infty). \end{aligned}$$

The third assertion of the theorem follows from the fact that the stopping time  $\nu(n) = \min(\nu, n)$  satisfies the conditions of the first part of the theorem (or those of Theorems 15.2.1 and 15.2.3), and therefore, for the martingale  $\{X_n\}$ ,

$$\mathbf{E}X_0 = \mathbf{E}X_{\nu(n)} = \mathbf{E}(X_\nu; \nu < n) + \mathbf{E}(X_\nu; \nu \geq n),$$

so that (15.2.19) implies the convergence  $\mathbf{E}(X_n; \nu \geq n) \rightarrow 0$  and vice versa.

The proof for semimartingales is similar. The theorem is proved.  $\square$

That assertions (15.2.17) and (15.2.19) are, generally speaking, not equivalent even when (15.2.18) holds (i.e.,  $\lim_{n \rightarrow \infty} \mathbf{E}(X_\nu; \nu \leq n) = \mathbf{E}(X_\nu; \nu < \infty)$  is not always the case), can be illustrated by the following example. Let  $\xi_k$  be independent random variables with

$$\mathbf{P}(\xi_k = 3^k) = \mathbf{P}(\xi_k = -3^k) = 1/2,$$

$\nu$  be independent of  $\{\xi_k\}$ , and  $\mathbf{P}(\nu = k) = 2^{-k}$ ,  $k = 1, 2, \dots$ . Then  $X_0 = 0$ ,  $X_k = X_{k-1} + \xi_k$  for  $k \geq 1$  is a martingale,

$$\mathbf{E}X_n = 0, \quad \mathbf{P}(\nu < \infty) = 1, \quad \mathbf{E}(X_n; \nu > n) = \mathbf{E}X_n \mathbf{P}(\nu > n) = 0$$

by independence, and condition (15.2.18) is satisfied. By virtue of (15.2.19), this means that  $\lim_{n \rightarrow \infty} \mathbf{P}(X_\nu; \nu \leq n) = 0$  (one can also verify this directly). On the other hand, the expectation  $\mathbf{E}(X_\nu; \nu < \infty) = \mathbf{E}X_\nu$  is not defined, since  $\mathbf{E}X_\nu^+ = \mathbf{E}X_\nu^- = \infty$ . Indeed, clearly

$$\begin{aligned} X_{k-1} \geq -\frac{3^k - 3}{2}, \quad \{\xi_k = 3^k\} \subset \left\{ X_k \geq \frac{3^k + 3}{2} \right\}, \quad \mathbf{P}\left( X_k \geq \frac{3^k + 3}{2} \right) \geq \frac{1}{2}, \\ \mathbf{E}X_k^+ \geq \frac{3^k + 3}{4}, \quad \mathbf{E}X_\nu^+ = \sum_{k=1}^{\infty} 2^{-k} \mathbf{E}X_k^+ \geq \sum_{k=1}^{\infty} 2^{-k-2} 3^k = \infty. \end{aligned}$$

By symmetry, we also have  $\mathbf{E}X_\nu^- = \infty$ .

**Corollary 15.2.2** 1. *If  $\{X_n, \mathfrak{F}_n\}$  is a nonnegative martingale, then condition (15.2.18) is necessary and sufficient for (15.2.17).*

2. If  $\{X_n, \mathfrak{F}_n\}$  is a nonnegative supermartingale and  $\nu$  is an arbitrary stopping time, then

$$\mathbf{E}(X_\nu; \nu < \infty) \leq \mathbf{E}X_0. \tag{15.2.21}$$

*Proof* The assertion follows in an obvious way from Theorem 15.2.4 since one has  $\mathbf{E}(X_\nu^-; \nu < \infty) = 0$ .  $\square$

Theorem 15.2.2 implies the already known Wald's identity (see Theorem 4.4.3) supplemented with another useful statement.

**Theorem 15.2.5** (Wald's identity) *Let  $\zeta_1, \zeta_2, \dots$  be independent identically distributed random variables,  $S_n = \zeta_1 + \dots + \zeta_n$ ,  $S_0 = 0$ , and assume  $\mathbf{E}\zeta_1 = a$ . Let, further,  $\nu$  be a stopping time with  $\mathbf{E}\nu < \infty$ . Then*

$$\mathbf{E}S_\nu = a\mathbf{E}\nu. \tag{15.2.22}$$

If, moreover,  $\sigma^2 = \text{Var } \zeta_k < \infty$ , then

$$\mathbf{E}[S_\nu - \nu a]^2 = \sigma^2 \mathbf{E}\nu. \tag{15.2.23}$$

*Proof* It is clear that  $X_n = S_n - na$  forms a martingale and conditions (15.2.8) and (15.2.9) are met. Therefore  $\mathbf{E}X_\nu = \mathbf{E}X_0 = 0$ , which is equivalent to (15.2.22), and  $\mathbf{E}X_\nu^2 = \mathbf{E}\nu\sigma^2$ , which is equivalent to (15.2.23).  $\square$

*Example 15.2.1* Consider a generalised renewal process (see Sect. 10.6)  $S(t) = S_{\eta(t)}$ , where  $S_n = \sum_{j=1}^n \xi_j$  (in this example we follow the notation of Chap. 10 and change the meaning of the notation  $S_n$  from the above),  $\eta(t) = \min\{k : T_k > t\}$ ,  $T_n = \sum_{j=1}^n \tau_j$  and  $(\tau_j, \xi_j)$  are independent vectors distributed as  $(\tau, \xi)$ ,  $\tau > 0$ . Set  $a_\xi = \mathbf{E}\xi$ ,  $a = \mathbf{E}\tau$ ,  $\sigma_\xi^2 = \text{Var } \xi$  and  $\sigma^2 = \text{Var } \tau$ . As we know from Wald's identity in Sect. 4.4,

$$\mathbf{E}\eta(t) = \frac{t + \mathbf{E}\chi(t)}{a}, \quad \mathbf{E}S(t) = a_\xi \mathbf{E}\eta(t),$$

where  $\mathbf{E}\chi(t) = o(t)$  as  $t \rightarrow \infty$  (see Theorem 10.1.1) and, in the non-lattice case,  $\mathbf{E}\chi(t) \rightarrow \frac{\sigma^2 + a^2}{2a^2}$  if  $\sigma^2 < \infty$  (see Theorem 10.4.3).

We now find  $\text{Var } \eta(t)$  and  $\text{Var } S(t)$ . Omitting for brevity's sake the argument  $t$ , we can write

$$\begin{aligned} a^2 \text{Var } \eta(t) &= a^2 \text{Var } \eta = \mathbf{E}(a\eta - a\mathbf{E}\eta)^2 = \mathbf{E}(a\eta - T_\eta + T_\eta - a\mathbf{E}\eta)^2 \\ &= \mathbf{E}(T_\eta - a\eta)^2 + \mathbf{E}(T_\eta - a\mathbf{E}\eta)^2 - 2\mathbf{E}(T_\eta - a\eta)(T_\eta - a\mathbf{E}\eta). \end{aligned}$$

The first summand on the right-hand side is equal to

$$\sigma^2 \mathbf{E}\eta = \frac{\sigma^2 t}{a} + O(1)$$

by Theorem 15.2.3. The second summand equals, by (10.4.8) ( $\chi(t) = T_{\eta(t)} - t$ ),

$$\mathbf{E}(t + \chi(t) - a\mathbf{E}\eta)^2 = \mathbf{E}(\chi(t) - \mathbf{E}\chi(t))^2 \leq \mathbf{E}\chi^2(t) = o(t).$$

The last summand, by the Cauchy–Bunjakovsky inequality, is also  $o(t)$ . Finally, we get

$$\text{Var } \eta(t) = \frac{\sigma^2 t}{a^3} + o(t).$$

Consider now (with  $r = a_\xi/a$ ;  $\zeta_j = \xi_j - r\tau_j$ ,  $\mathbf{E}\zeta_j = 0$ )

$$\begin{aligned} \text{Var } S(t) &= \mathbf{E}(S_\eta - a_\xi \mathbf{E}\eta)^2 = \mathbf{E}[S_\eta - rT_\eta + r(T_\eta - a\mathbf{E}\eta)]^2 \\ &= \mathbf{E}\left(\sum_{j=1}^{\eta} \zeta_j\right)^2 + r^2 \mathbf{E}(T_\eta - a\mathbf{E}\eta)^2 + 2r \mathbf{E}\left(\sum_{j=1}^{\eta} \zeta_j\right)(T_\eta - a\mathbf{E}\eta). \end{aligned}$$

The first term on the right-hand side is equal to

$$\mathbf{E}\eta \text{Var } \zeta = \frac{t \text{Var } \zeta}{a} + O(1)$$

by Theorem 15.2.3. The second term has already been estimated above. Therefore, as before, the sum of the last two terms is  $o(t)$ . Thus

$$\text{Var } S(t) = \frac{t}{a} \mathbf{E}(\xi - r\tau)^2 + o(t).$$

This corresponds to the scaling used in Theorem 10.6.2.

*Example 15.2.2* Examples 4.4.4 and 4.5.5 referring to the fair game situation with  $\mathbf{P}(\zeta_k = \pm 1) = 1/2$  and  $v = \min\{k : S_k = z_2 \text{ or } S_k = -z_1\}$  ( $z_1$  and  $z_2$  being the capitals of the gamblers) can also illustrate the use of Theorem 15.2.5.

Now consider the case  $p = \mathbf{P}(\zeta_k = 1) \neq 1/2$ . The sequence  $X_n = (q/p)^{S_n}$ ,  $n \geq 0$ ,  $q = 1 - p$  is a martingale, since

$$\mathbf{E}(q/p)^{\zeta_k} = p(q/p) + q(p/q) = 1.$$

By Theorem 15.2.5 (the probabilities  $P_1$  and  $P_2$  were defined in Example 4.4.5),

$$\mathbf{E}X_v = \mathbf{E}X_0 = 1, \quad P_1(q/p)^{z_2} + P_2(q/p)^{z_1} = 1.$$

From this relation and equality  $P_1 + P_2 = 1$  we have

$$P_1 = \frac{(q/p)^{z_1} - 1}{(q/p)^{z_1} - (q/p)^{z_2}}, \quad P_2 = 1 - P_1.$$

Using Wald's identity again, we also obtain that

$$\mathbf{E}v = \frac{\mathbf{E}S_v}{\mathbf{E}\zeta_1} = \frac{P_1 z_2 - P_2 z_1}{p - q}.$$

Note that these equalities could have been obtained by elementary methods<sup>1</sup> but this would require lengthy calculations.

---

<sup>1</sup>See, e.g., [12].

In the cases when the nature of  $S_\nu$  is simple enough, the assertions of the type of Theorems 15.2.1–15.2.2 enable one to obtain (or estimate) the distribution of the random variable  $\nu$  itself. In such situations, the following assertion is rather helpful.

Suppose that the conditions of Theorem 15.2.5 are met, but, instead of conditions on the moments of  $\zeta_n$ , the Cramér condition (cf. Chap. 9) is assumed to be satisfied:

$$\psi(\lambda) := \mathbf{E}e^{\lambda\zeta} < \infty$$

for some  $\lambda \neq 0$ .

In other words, if

$$\lambda_+ := \sup\{\lambda : \psi(\lambda) < \infty\} \geq 0, \quad \lambda_- := \inf\{\lambda : \psi(\lambda) < \infty\} \leq 0,$$

then  $\lambda_+ - \lambda_- > 0$ . Everywhere in what follows we will only consider the values

$$\lambda \in B := \{\psi(\lambda) < \infty\} \subseteq [\lambda_-, \lambda_+]$$

for which  $\psi'(\lambda) < \infty$ . For such  $\lambda$ , the positive martingale

$$X_n = \frac{e^{\lambda S_n}}{\psi^n(\lambda)}, \quad X_0 = 1,$$

is well-defined so that  $\mathbf{E}X_n = 1$ .

**Theorem 15.2.6** *Let  $\nu$  be an arbitrary stopping time and  $\lambda \in B$ . Then*

$$\mathbf{E}\left(\frac{e^{\lambda S_\nu}}{\psi(\lambda)^\nu}; \nu < \infty\right) \leq 1 \quad (15.2.24)$$

and, for any  $s > 1$  and  $r > 1$  such that  $1/r + 1/s = 1$ ,

$$\mathbf{E}(e^{\lambda S_\nu}; \nu < \infty) \leq \{\mathbf{E}[\psi^{r\nu/s}(\lambda S_\nu); \nu < \infty]\}^{1/r}. \quad (15.2.25)$$

A necessary and sufficient condition for

$$\mathbf{E}\left(\frac{e^{\lambda S_\nu}}{\psi(\lambda)^\nu}; \nu < \infty\right) = 1 \quad (15.2.26)$$

is that

$$\lim_{n \rightarrow \infty} \mathbf{E}\left(\frac{e^{\lambda S_n}}{\psi(\lambda)^n}; \nu > n\right) = 0. \quad (15.2.27)$$

*Remark 15.2.1* Relation (15.2.26) is known as the *fundamental Wald identity*. In the literature it is usually considered for a.s. finite  $\nu$  (when  $\mathbf{P}(\nu < \infty) = 1$ ) being in that case an extension of the obvious equality  $\mathbf{E}e^{\lambda S_n} = \psi^n(\lambda)$  to the case of random  $\nu$ . Originally, identity (15.2.26) was established by A. Wald in the special case where  $\nu$  is the exit time of the sequence  $\{S_n\}$  from a finite interval (see Corollary 15.2.3), and was accompanied by rather restrictive conditions. Later, these conditions were removed (see e.g. [13]). Below we will obtain a more general assertion for the problem on the first exit of the trajectory  $\{S_n\}$  from a strip with curvilinear boundaries.

*Remark 15.2.2* The fundamental Wald identity shows that, although the nature of a stopping time could be quite general, there exists a stiff functional constraint (15.2.26) on the joint distribution of  $\nu$  and  $S_\nu$  (the distribution of  $\zeta_k$  is assumed to be known). In the cases where one of these variables can somehow be “computed” or “eliminated” (see Examples 15.2.2–15.2.4) Wald’s identity turns into an explicit formula for the Laplace transform of the distribution of the other variable. If  $\nu$  and  $S_\nu$  prove to be independent (which rarely happens), then (15.2.26) gives the relationship

$$\mathbf{E}e^{\lambda S_\nu} = [\mathbf{E}\psi(\lambda)^{-\nu}]^{-1}$$

between the Laplace transforms of the distributions of  $\nu$  and  $S_\nu$ .

*Proof of Theorem 15.2.6* As we have already noted, for

$$X_n = e^{\lambda S_n} \psi^{-n}(\lambda), \quad \mathfrak{F}_n = \sigma(\zeta_1, \dots, \zeta_n),$$

$\{X_n, \mathfrak{F}_n; n \geq 0\}$  is a positive martingale with  $X_0 = 1$  and  $\mathbf{E}X_n = 1$ . Corollary 15.2.2 immediately implies (15.2.24).

Inequality (15.2.25) is a consequence of Hölder’s inequality and (15.2.24):

$$\begin{aligned} \mathbf{E}(e^{(\lambda/s)S_\nu}; \nu < \infty) &= \mathbf{E}\left[\left(\frac{e^{\lambda S_\nu}}{\psi^\nu(\lambda)}\right)^{1/s} \psi^{\nu/s}(\lambda); \nu < \infty\right] \\ &\leq [\mathbf{E}(\psi^{\nu r/s}(\lambda); \nu < \infty)]^{1/r}. \end{aligned}$$

The last assertion of the theorem (concerning the identity (15.2.26)) follows from Theorem 15.2.4. □

We now consider several important special cases. Note that  $\psi(\lambda)$  is a convex function ( $\psi''(\lambda) > 0$ ),  $\psi(0) = 1$ , and therefore there exists a unique point  $\lambda_0$  at which  $\psi(\lambda)$  attains its minimum value  $\psi(\lambda_0) \leq 1$  (see also Sect. 9.1).

**Corollary 15.2.3** *Assume that we are given a sequence  $g(n)$  such that*

$$g^+(n) := \max(0, g(n)) = o(n) \quad \text{as } n \rightarrow \infty.$$

*If  $S_n \leq g(n)$  holds on the set  $\{\nu > n\}$ , then (15.2.26) holds for  $\lambda \in (\lambda_0, \lambda_+] \cap B$ ,  $B = \{\lambda : \psi(\lambda) < \infty\}$ .*

The random variable  $\nu = \nu_g = \inf\{k \geq 1 : S_k > g(k)\}$  for  $g(k) = o(k)$  obviously satisfies the conditions of Corollary 15.2.3. For stopping times  $\nu_g$  one could also consider the case  $g(n)/n \rightarrow c \geq 0$  as  $n \rightarrow \infty$ , which can be reduced to the case  $g(n) = o(n)$  by introducing the random variables

$$\zeta_k^* := \zeta_k - c, \quad S_k^* := \sum_{j=1}^k \zeta_j^*,$$

for which  $\nu_g = \inf\{k \geq 1 : S_k^* > g(k) - ck\}$ .



*Proof of Corollary 15.2.3* For  $\lambda > \lambda_0$ ,  $\lambda \in B$ , we have

$$\begin{aligned} \mathbf{E}\left(\frac{e^{\lambda S_n}}{\psi^n(\lambda)}; v > n\right) &\leq \psi^{-n}(\lambda) \mathbf{E}(e^{\lambda S_n}; S_n \leq g(n)) \\ &= \psi^{-n}(\lambda) \mathbf{E}(e^{(\lambda-\lambda_0)S_n} \cdot e^{\lambda_0 S_n}; S_n \leq g(n)) \\ &\leq \psi^{-n}(\lambda) e^{(\lambda-\lambda_0)g(n)} \mathbf{E}(e^{\lambda_0 S_n}; S_n \leq g(n)) \\ &\leq \psi^{-n}(\lambda) e^{(\lambda-\lambda_0)g^+(n)} \mathbf{E}e^{\lambda_0 S_n} = \left(\frac{\psi(\lambda_0)}{\psi(\lambda)}\right)^n e^{(\lambda-\lambda_0)g^+(n)} \rightarrow 0 \end{aligned}$$

as  $n \rightarrow \infty$ , because  $(\lambda - \lambda_0)g^+(n) = o(n)$ . It remains to use Theorem 15.2.6. The corollary is proved.  $\square$

We now return to Theorem 15.2.6 for arbitrary stopping times. It turns out that, based on the Cramér transform introduced in Sect. 9.1, one can complement its assertions without using any martingale techniques.

Together with the original distribution  $\mathbf{P}$  of the sequence  $\{\zeta_k\}_{k=1}^\infty$  we introduce the family of distributions  $\mathbf{P}_\lambda$  of this sequence in  $(\mathbb{R}^\infty, \mathfrak{B}^\infty)$  (see Sect. 5.5) generated by the finite-dimensional distributions

$$\begin{aligned} \mathbf{P}_\lambda(\zeta_k \in dx_k) &= \frac{e^{\lambda x_k}}{\psi(\lambda)} \mathbf{P}(\zeta_k \in dx_k), \\ \mathbf{P}_\lambda(\zeta_k \in dx_1, \dots, \zeta_n \in dx_n) &= \prod_{k=1}^n \mathbf{P}_\lambda(\zeta_k \in dx_k). \end{aligned}$$

This is the *Cramér transform* of the distribution  $\mathbf{P}$ .

**Theorem 15.2.7** *Let  $v$  be an arbitrary stopping time. Then, for any  $\lambda \in B$ ,*

$$\mathbf{E}\left(\frac{e^{\lambda S_v}}{\psi^v(\lambda)}; v < \infty\right) = \mathbf{P}_\lambda(v < \infty). \quad (15.2.28)$$

*Proof* Since  $\{v = n\} \in \sigma(\zeta_1, \dots, \zeta_n)$ , there exists a Borel set  $D_n \subset \mathbb{R}^n$ , such that

$$\{v = n\} = \{(\zeta_1, \dots, \zeta_n) \in D_n\}.$$

Further,

$$\mathbf{E}\left(\frac{e^{\lambda S_v}}{\psi^v(\lambda)}; v < \infty\right) = \sum_{n=0}^{\infty} \mathbf{E}\left(\frac{e^{\lambda S_n}}{\psi^n(\lambda)}; v = n\right),$$

where

$$\begin{aligned} \mathbf{E}\left(\frac{e^{\lambda S_n}}{\psi^n(\lambda)}; v = n\right) &= \int_{(x_1, \dots, x_n) \in D_n} \frac{e^{\lambda(x_1 + \dots + x_n)}}{\psi^n(\lambda)} \mathbf{P}(\zeta_1 \in dx_1, \dots, \zeta_n \in dx_n) \\ &= \int_{(x_1, \dots, x_n) \in D_n} \mathbf{P}_\lambda(\zeta_1 \in dx_1, \dots, \zeta_n \in dx_n) = \mathbf{P}_\lambda(v = n). \end{aligned}$$

This proves the theorem.  $\square$

For a given function  $g(n)$ , consider now the stopping time

$$v = v_g = \inf\{k : S_k \geq g(k)\}$$

(cf. Corollary 15.2.3). The assertion of Theorem 15.2.7 can be obtained in that case in the following way. Denote by  $\mathbf{E}_\lambda$  the expectation with respect to the distribution  $\mathbf{P}_\lambda$ .

**Corollary 15.2.4** 1. *If  $g^+(n) = \max(0, g(n)) = o(n)$  as  $n \rightarrow \infty$  and  $\lambda \in (\lambda_0, \lambda_+) \cap B$ , then one has  $\mathbf{P}_\lambda(v_g < \infty) = 1$  in relation (15.2.28).*

2. *If  $g(n) \geq 0$  and  $\lambda < \lambda_0$ , then  $\mathbf{P}_\lambda(v_g < \infty) < 1$ .*

3. *For  $\lambda = \lambda_0$ , the distribution  $\mathbf{P}_{\lambda_0}$  of the variable  $v$  can either be proper (when one has  $\mathbf{P}_{\lambda_0}(v_g < \infty) = 1$ ) or improper ( $\mathbf{P}_{\lambda_0}(v_g < \infty) < 1$ ). If  $\lambda_0 \in (\lambda_-, \lambda_+)$ ,  $g(n) < (1 - \varepsilon)\sigma(2 \log \log n)^{1/2}$  for all  $n \geq n_0$ , starting from some  $n_0$ , and  $\sigma^2 = \mathbf{E}_{\lambda_0} \zeta_1^2$ , then  $\mathbf{P}_\lambda(v_g < \infty) = 1$ .*

*But if  $\lambda \in (\lambda_-, \lambda_+)$ ,  $g(n) \geq 0$ , and  $g(n) \geq (1 + \varepsilon)\sigma(2 \log \log n)^{1/2}$  for  $n \geq n_0$ , then  $\mathbf{P}_\lambda(v_g < \infty) < 1$  (we exclude the trivial case  $\zeta_k \equiv 0$ ).*

*Proof* Since  $\mathbf{E}_\lambda \zeta_k = \frac{\psi'(\lambda)}{\psi(\lambda)}$ , the expectation  $\mathbf{E}_\lambda \zeta_k$  is of the same sign as the difference  $\lambda - \lambda_0$ , and  $\mathbf{E}_{\lambda_0} \zeta_k = 0$  ( $\psi'(\lambda_0) = 0$  if  $\lambda_0 \in (\lambda_-, \lambda_+)$ ). Hence the first assertion follows from the relations

$$\mathbf{P}_\lambda(v = \infty) = \mathbf{P}_\lambda(X_n < g(n) \text{ for all } n) < \mathbf{P}(X_n < g^+(n)) \rightarrow 0$$

as  $n \rightarrow \infty$  by the law of large numbers for the sums  $X_n = \sum_{k=1}^n \zeta_k$ , since  $\mathbf{E}_\lambda \zeta_k > 0$ .

The second assertion is a consequence of the strong law of large numbers since  $\mathbf{E}_\lambda \zeta_k < 0$  and hence  $\mathbf{P}_\lambda(v = \infty) = \mathbf{P}(\sup_n X_n \leq 0) > 0$ .

The last assertion of the corollary follows from the law of the iterated logarithm which we prove in Sect. 20.2. The corollary is proved.  $\square$

The condition  $g(n) \geq 0$  of part 2 of the corollary can clearly be weakened to the condition  $g(n) = o(n)$ ,  $\mathbf{P}(v > n) > 0$  for any  $n > 0$ . The same is true for part 3.

An assertion similar to Corollary 15.2.4 is also true for the (stopping) time  $v_{g_-, g_+}$  of the first passage of one of the two boundaries  $g_\pm(n) = o(n)$ :

$$v_{g_-, g_+} := \inf\{k \geq 1 : S_k \geq g_+(k) \text{ or } S_k \leq g_-(k)\}.$$

**Corollary 15.2.5** *For  $\lambda \in B \setminus \{\lambda_0\}$ , we have  $\mathbf{P}_\lambda(v_{g_-, g_+} < \infty) = 1$ .*

*If  $\lambda = \lambda_0 \in (\lambda_-, \lambda_+)$ , then the  $\mathbf{P}_\lambda$ -distribution of  $v$  may be either proper or improper.*

*If, for some  $n_0 > 2$ ,*

$$g_\pm(n) \leq \pm(1 - \varepsilon)\sigma\sqrt{2 \ln \ln n}$$

*for  $n \geq n_0$  then  $\mathbf{P}_{\lambda_0}(v_{g_-, g_+} < \infty) = 1$ .*

*If  $g_\pm(n) \geq 0$  and, additionally,*

$$g_\pm(n) \geq \pm(1 + \varepsilon)\sigma\sqrt{2 \ln \ln n}$$

*for  $n \geq n_0$  then  $\mathbf{P}_{\lambda_0}(v_{g_-, g_+} < \infty) < 1$ .*

*Proof* The first assertion follows from Corollary 15.2.4 applied to the sequences  $\{\pm X_n\}$ . The second is a consequence of the law of the iterated logarithm from Sect. 20.2.  $\square$

We now consider several relations following from Corollaries 15.2.3, 15.2.4 and 15.2.5 (from identity (15.2.26)) for the random variables  $\nu = \nu_g$  and  $\nu = \nu_{g-,g+}$ .

Let  $a < 0$  and  $\psi(\lambda_+) \geq 1$ . Since  $\psi'(0) = a < 0$  and the function  $\psi(\lambda)$  is convex, the equation  $\psi(\lambda) = 1$  will have a unique root  $\mu > 0$  in the domain  $\lambda > 0$ . Setting  $\lambda = \mu$  in (15.2.26) we obtain the following.

**Corollary 15.2.6** *If  $a < 0$  and  $\psi(\lambda_+) \geq 1$  then, for the stopping times  $\nu = \nu_g$  and  $\nu = \nu_{g-,g+}$ , we have the equality*

$$\mathbf{E}(e^{\mu S_\nu}; \nu < \infty) = 1.$$

*Remark 15.2.3* For an  $x > 0$ , put (as in Chap. 10)  $\eta(x) := \inf\{k : S_k > 0\}$ . Since  $S_{\eta(x)} = x + \chi(x)$ , where  $\chi(x) := S_{\eta(x)} - x$  is the value of overshoot over the level  $x$ , Corollary 15.2.6 implies

$$\mathbf{E}(e^{\mu(x+\chi(x))}; \eta(x) < \infty) = 1. \tag{15.2.29}$$

Note that  $\mathbf{P}(\eta(x) < \infty) = \mathbf{P}(S > x)$ , where  $S = \sup_{k \geq 0} S_k$ . Therefore, Theorem 12.7.4 and (15.2.29) imply that, as  $x \rightarrow \infty$ ,

$$e^{\mu x} \mathbf{P}(\eta(x) < \infty) = [\mathbf{E}(e^{\mu \chi(x)} | \eta(x) < \infty)]^{-1} \rightarrow c. \tag{15.2.30}$$

The last convergence relation corresponds to the fact that the limiting conditional distribution (as  $x \rightarrow \infty$ )  $\mathbf{G}$  of  $\chi(x)$  exists given  $\eta(x) < \infty$ . If we denote by  $\chi$  a random variable with the distribution  $\mathbf{G}$  then (15.2.30) will mean that  $c = [\mathbf{E} e^{\mu \chi}]^{-1} < 1$ . This provides an interpretation of the constant  $c$  that is different from the one in Theorem 12.7.4.

In Corollary 15.2.6 we “eliminated” the “component”  $\psi^\nu(\lambda)$  in identity (15.2.26). “Elimination” of the other component  $e^{\lambda S_\nu}$  is possible only in some special cases of random walks, such as the so-called skip-free walks (see Sect. 12.8) or walks with exponentially (or geometrically) distributed  $\zeta_k^+ = \max(0, \zeta_k)$  or  $\zeta_k^- = -\min(0, \zeta_k)$ . We will illustrate this with two examples.

*Example 15.2.3* We return to the ruin problem discussed in Example 15.2.2. In that case, Corollary 15.2.4 gives, for  $g_-(n) := -z_1$  and  $g_+(n) = z_2$ , that

$$e^{\lambda z_2} \mathbf{E}(\psi(\lambda)^{-\nu}; S_\nu = z_2) + e^{-\lambda z_1} \mathbf{E}(\psi(\lambda)^{-\nu}; S_\nu = -z_1) = 1.$$

In particular, for  $z_1 = z_2 = z$  and  $p = 1/2$ , we have by symmetry that

$$\mathbf{E}(\psi(\lambda)^{-\nu}; S_\nu = z) = \frac{1}{e^{\lambda z} + e^{-\lambda z}}, \quad \mathbf{E}(\psi(\lambda)^{-\nu}) = \frac{2}{e^{\lambda z} + e^{-\lambda z}}. \tag{15.2.31}$$

Let  $\lambda(s)$  be the unique positive solution of the equation  $s\psi(\lambda) = 1$ ,  $s \in (0, 1)$ . Since here  $\psi(\lambda) = \frac{1}{2}(e^\lambda + e^{-\lambda})$ , solving the quadratic equation yields

$$e^{\lambda(s)} = \frac{1 + \sqrt{1 - s^2}}{s}.$$

Identity (15.2.31) now gives

$$\mathbf{E}s^\nu = 2(e^{\lambda(s)z} + e^{-\lambda(s)z}).$$

We obtain an explicit form of the generating function of the random variable  $\nu$ , which enables us to find the probabilities  $\mathbf{P}(\nu = n)$ ,  $n = 1, 2, \dots$  by expanding elementary functions into series.

*Example 15.2.4* Simple explicit formulas can also be obtained from Wald’s identity in the problem with one boundary, where  $\nu = \nu_g$ ,  $g(n) = z$ . In that case, the class of distributions of  $\zeta_k$  could be wider than in Example 15.2.3. Suppose that one of the two following conditions holds (cf. Sect. 12.8).

1. *The transform walk is arithmetic and skip-free, i.e.  $\zeta_k$  are integers,  $\mathbf{P}(\xi_k = 1) > 0$  and  $\mathbf{P}(\zeta_k \geq 2) = 0$ .*
2. *The walk is right exponential, i.e.*

$$\mathbf{P}(\zeta_k > t) = ce^{-\alpha t} \tag{15.2.32}$$

*either for all  $t > 0$  or for  $t = 0, 1, 2, \dots$  if the walk is integer-valued (the geometric distribution).*

The random variable  $\nu_g$  will be proper if and only if  $\mathbf{E}\xi_k = \psi'(0) \geq 0$  (see Chaps. 9 and 12). For skip-free random walks, Wald’s identity (15.2.26) yields  $(g(n) = z > 0, S_\nu = z)$

$$e^{\lambda z} \mathbf{E}\psi^{-\nu}(\lambda) = 1, \quad \lambda > \lambda_0. \tag{15.2.33}$$

For  $s \leq 1$ , the equation  $\psi(\lambda) = s^{-1}$  (cf. Example 15.2.3) has in the domain  $\lambda > \lambda_0$  a unique solution  $\lambda(s)$ . Therefore identity (15.2.33) can be written as

$$\mathbf{E}s^\nu = e^{-z\lambda(s)}. \tag{15.2.34}$$

This statement implies a series of results from Chaps. 9 and 12. Many properties of the distribution of  $\nu := \nu_z$  can be derived from this identity, in particular, the asymptotics of  $\mathbf{P}(\nu_z = n)$  as  $z \rightarrow \infty, n \rightarrow \infty$ . We already know one of the ways to find this asymptotics. It consists of using Theorem 12.8.4, which implies

$$\mathbf{P}(\nu_z = n) = \frac{x}{n} \mathbf{P}(S_n = z), \tag{15.2.35}$$

and the local Theorem 9.3.4 providing the asymptotics of  $\mathbf{P}(S_n = z)$ . Using relation (15.2.34) and the inversion formula is an alternative approach to studying the asymptotics of  $\mathbf{P}(\nu_z = n)$ . If we use the inversion formula, there will arise an integral of the form

$$\int_{|s|=1} s^{-n} e^{-z\mu(s)} ds, \tag{15.2.36}$$

where the integrand  $s^{-n}e^{-z\mu(s)}$ , after the change of variable  $\mu(s)=\lambda$  (or  $s = \psi(\lambda)^{-1}$ ), takes the form

$$\exp - \{z\lambda - n \ln \psi(\lambda)\}.$$

The integrand in the inversion formula for the probability  $\mathbf{P}(S_n = z)$  has the same form. This probability has already been studied quite well (see Theorem 9.3.4); its exponential part has the form  $e^{-n\Lambda(\alpha)}$ , where  $\alpha = z/n$ ,  $\Lambda(\alpha) = \sup_{\lambda} (\alpha\lambda - \ln \psi(\lambda))$  is the large deviation rate function (see Sect. 9.1 and the footnote for Definition 9.1.1). A more detailed study of the inversion formula (15.2.36) allows us to obtain (15.2.35).

Similar relations can be obtained for random walks with exponential right distribution tails. Let, for example, (15.2.32) hold for all  $t > 0$ . Then the conditional distribution  $\mathbf{P}(S_\nu > t | \nu = n, S_{n-1} = x)$  coincides with the distribution

$$\mathbf{P}(\zeta_n > z - x + t | \zeta_n > z - x) = e^{-\alpha t}$$

and clearly depends neither on  $n$  nor on  $x$ . This means that  $\nu$  and  $S_\nu$  are independent,  $S_\nu = z + \gamma$ ,  $\gamma \in \Gamma_\alpha$ ,

$$\mathbf{E}\psi(\lambda)^{-\nu} = \frac{1}{\mathbf{E}e^{(z+\gamma)\lambda}} = e^{-\lambda z} \frac{\alpha - \lambda}{\alpha}, \quad \lambda_0 < \lambda < \alpha; \quad \mathbf{E}S^\nu = e^{-z\lambda(s)} \frac{\alpha - \lambda(s)}{\alpha},$$

where  $\lambda(s)$  is, as before, the only solution to the equation  $\psi(\lambda) = s^{-1}$  in the domain  $\lambda > \lambda_0$ . This implies the same results as (15.2.34).

If  $\mathbf{P}(\zeta_k > t) = c_1 e^{-\alpha t}$  and  $\mathbf{P}(\zeta_k < -t) = c_2 e^{-\beta t}$ ,  $t > 0$ , then, in the problem with two boundaries, we obtain for  $\nu = \nu_{g_-, g_+}$ ,  $g_+(n) = z_2$  and  $g_-(n) = -z_1$  in exactly the same way from (15.2.26) that

$$\frac{\alpha e^{\lambda z_2}}{\alpha - \lambda} \mathbf{E}(\psi^{-\nu}(\lambda); S_\nu \geq z_2) + \frac{\beta e^{-\lambda z_1}}{\beta + \lambda} \mathbf{E}(\psi^{-\nu}(\lambda); S_\nu \leq -z_1) = 1, \quad \lambda \in (-\beta, \alpha).$$

## 15.3 Inequalities

### 15.3.1 Inequalities for Martingales

First of all we note that the property  $\mathbf{E}X_n \leq 1$  of the sequence  $X_n = e^{\lambda S_n} \psi_0(\lambda)^{-n}$  forming a supermartingale for an appropriate function  $\psi_0(\lambda)$  remains true when we replace  $n$  with a stopping time  $\nu$  (an analogue of inequality (15.2.24)) in a much more general case than that of Theorem 15.2.6. Namely,  $\zeta_k$  may be dependent.

Let, as before,  $\{\mathfrak{F}_n\}$  be an increasing sequence of  $\sigma$ -algebras, and  $\zeta_n$  be  $\mathfrak{F}_n$ -measurable random variables. Suppose that a.s.

$$\mathbf{E}(e^{\lambda \zeta_n} | \mathfrak{F}_{n-1}) \leq \psi_0(\lambda). \tag{15.3.1}$$

This condition is always met if a.s.

$$\mathbf{P}(\zeta_n \geq x | \mathfrak{F}_{n-1}) \leq G(x), \quad \psi_0(\lambda) = - \int e^{\lambda x} dG(x) < \infty.$$

In that case the sequence  $X_n = e^{\lambda S_n} \psi_0^{-n}(\lambda)$  forms a supermartingale:

$$\mathbf{E}(X_n | \mathfrak{F}_{n-1}) \leq X_{n-1}, \quad \mathbf{E}X_n \leq 1.$$

**Theorem 15.3.1** *Let (15.3.1) hold and  $\nu$  be a stopping time. Then inequalities (15.2.24) and (15.2.25) will hold true with  $\psi$  replaced by  $\psi_0$ .*

*The Proof* of the theorem repeats almost verbatim that of Theorem 15.2.6. □

Now we will obtain inequalities for the distribution of

$$\bar{X}_n = \max_{k \leq n} X_k \quad \text{and} \quad X_n^* = \max_{k \leq n} |X_k|,$$

$X_n$  being an arbitrary submartingale.

**Theorem 15.3.2** (Doob) *Let  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  be a nonnegative submartingale. Then, for all  $x \geq 0$  and  $n \geq 0$ ,*

$$\mathbf{P}(\bar{X}_n > x) \leq \frac{1}{x} \mathbf{E}X_n.$$

*Proof* Let

$$\nu = \eta(x) := \inf\{k \geq 0 : X_k > x\}, \quad \nu(n) := \min(\nu, n).$$

It is obvious that  $n$  and  $\nu(n)$  are stopping times,  $\nu(n) \leq n$ , and therefore, by Theorem 15.2.1 (see (15.2.3) for  $\nu_2 = n, \nu_1 = \nu(n)$ ),

$$\mathbf{E}X_n \geq \mathbf{E}X_{\nu(n)}.$$

Observing that  $\{\bar{X}_n > x\} = \{X_{\nu(n)} > x\}$ , we have from Chebyshev's inequality that

$$\mathbf{P}(\bar{X}_n > x) = \mathbf{P}(X_{\nu(n)} > x) \leq \frac{1}{x} \mathbf{E}X_{\nu(n)} \leq \frac{1}{x} \mathbf{E}X_n.$$

The theorem is proved. □

Theorem 15.3.2 implies the following.

**Theorem 15.3.3** (The second Kolmogorov inequality) *Let  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  be a martingale with a finite second moment  $\mathbf{E}X_n^2$ . Then  $\{X_n^2, \mathfrak{F}_n; n \geq 0\}$  is a submartingale and by Theorem 15.3.2*

$$\mathbf{P}(X_n^* > x) \leq \frac{1}{x^2} \mathbf{E}X_n^2.$$

Originally A.N. Kolmogorov established this inequality for sums  $X_n = \xi_1 + \dots + \xi_n$  of independent random variables  $\xi_n$ . Theorem 15.3.3 extends Kolmogorov's proof to the case of submartingales and refines Chebyshev's inequality.

The following generalisation of Theorem 15.3.3 is also valid.

**Theorem 15.3.4** *If  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  is a martingale and  $\mathbf{E}|X_n|^p < \infty$ ,  $p \geq 1$ , then  $\{|X_n|^p, \mathfrak{F}_n; n \geq 0\}$  forms a nonnegative submartingale and, for all  $x > 0$ ,*

$$\mathbf{P}(X_n^* \geq x) \leq \frac{1}{x^p} \mathbf{E}|X_n|^p.$$

*If  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  is a submartingale,  $\mathbf{E}e^{\lambda X_n} < \infty$ ,  $\lambda > 0$ , then  $\{e^{\lambda X_n}, \mathfrak{F}_n; n \geq 0\}$  also forms a nonnegative submartingale,*

$$\mathbf{P}(\bar{X}_n \geq x) \leq e^{-\lambda x} \mathbf{E}e^{\lambda X_n}.$$

Both Theorem 15.3.4 and Theorem 15.3.3 immediately follow from Lemma 15.1.3 and Theorem 15.3.2.

If  $X_n = S_n = \sum_{k=1}^n \zeta_k$ , where  $\zeta_k$  are independent, identically distributed and satisfy the Cramér condition:  $\lambda_+ = \sup\{\lambda : \psi(\lambda) < \infty\} > 0$ , then, with the help of the fundamental Wald identity, one can obtain sharper inequalities for  $\mathbf{P}(\bar{X}_n > x)$  in the case  $a = \mathbf{E}\xi_k < 0$ .

Recall that, in the case  $a = \psi'(0) < 0$ , the function  $\psi(\lambda) = \mathbf{E}e^{\lambda \zeta_k}$  decreases in a neighbourhood of  $\lambda = 0$ , and, provided that  $\psi(\lambda_+) \geq 1$ , the equation  $\psi(\lambda) = 1$  has a unique solution  $\mu$  in the domain  $\lambda > 0$ .

Let  $\zeta$  be a random variable having the same distribution as  $\zeta_k$ . Put

$$\psi_+ := \sup_{t>0} \mathbf{E}(e^{\mu(\zeta-t)} | \zeta > t), \quad \psi_- := \inf_{t>0} \mathbf{E}(e^{\mu(\zeta-t)} | \zeta > t).$$

If, for instance,  $\mathbf{P}(\zeta > t) = ce^{-\alpha t}$  for  $t > 0$  (in this case necessarily  $\alpha > \mu$  in (15.2.32)), then

$$\mathbf{P}(\zeta - t > v | \zeta > t) = \frac{\mathbf{P}(\zeta > t + v)}{\mathbf{P}(\zeta > t)} = e^{-\alpha v}, \quad \psi_+ = \psi_- = \frac{\alpha}{\alpha - \mu}.$$

A similar equality holds for integer-valued  $\xi$  with a geometric distribution.

For other distributions, one has  $\psi_+ > \psi_-$ .

Under the above conditions, one has the following assertion which supplements Theorem 12.7.4 for the distribution of the random variable  $S = \sup_k S_k$ .

**Theorem 15.3.5** *If  $a = \mathbf{E}\zeta < 0$  then*

$$\psi_+^{-1} e^{-\mu x} \leq \mathbf{P}(S > x) \leq \psi_-^{-1} e^{-\mu x}, \quad x > 0. \tag{15.3.2}$$

This theorem implies that, in the case of exponential right tails of the distribution of  $\zeta$  (see (15.2.32)), inequalities (15.3.2) become the exact equality

$$\mathbf{P}(S > x) = \frac{\alpha - \mu}{\alpha} e^{-\mu x}.$$

(The same result was obtained in Example 12.5.1.) This means that inequalities (15.3.2) are unimprovable. Since  $\bar{S}_n = \max_{k \leq n} S_k \leq S$ , relation (15.3.2) implies that, for any  $n$ ,

$$\mathbf{P}(\bar{S}_n > x) \leq \psi_-^{-1} e^{-\mu x}.$$

*Proof of Theorem 15.3.5* Set  $v := \infty$  if  $S = \sup_{k \geq 0} S_k \leq x$ , and put  $v := \eta(x) = \min\{k : S_k > x\}$  otherwise. Further, let  $\chi(x) := S_{\eta(x)} - x$  be the excess of the level  $x$ . We have

$$\begin{aligned} \mathbf{P}(\chi(x) > v; v < \infty) &= \sum_{k=1}^{\infty} \int_{-\infty}^x \mathbf{P}(\bar{S}_{k-1} \leq x, S_{k-1} \in du, \zeta_k > x - u + v) \\ &= \sum_{k=1}^{\infty} \int_{-\infty}^x \mathbf{P}(\bar{S}_{k-1} \leq x, S_{k-1} \in du, \zeta_k > x - u) \\ &\quad \times \mathbf{P}(\zeta_k > x - u + v | \zeta_k > x - u), \\ \mathbf{E}(e^{\mu\chi(x)}; v < \infty) &\leq \sum_{k=1}^{\infty} \int_{-\infty}^x \mathbf{P}(\bar{S}_{k-1} \leq x, S_{k-1} \in du, \zeta_k > x - u) \psi_+ \\ &= \psi_+ \sum_{k=1}^{\infty} \mathbf{P}(v = k) = \psi_+ \mathbf{P}(v < \infty). \end{aligned}$$

Similarly,

$$\mathbf{E}(e^{\mu\chi(x)}; v < \infty) \geq \psi_- \mathbf{P}(v < \infty).$$

Next, by Corollary 15.2.6,

$$1 = \mathbf{E}(e^{\mu S_v}; v < \infty) = e^{\mu x} \mathbf{E}(e^{\mu\chi(x)}; v < \infty) \leq e^{\mu x} \psi_+ \mathbf{P}(v < \infty).$$

Because  $\mathbf{P}(v < \infty) = \mathbf{P}(S > x)$ , we get from this the right inequality of Theorem 15.3.5. The left inequality is obtained in the same way. The theorem is proved. □

*Remark 15.3.1* We proved Theorem 15.3.5 with the help of the fundamental Wald identity. But there is a direct proof based on the following relations:

$$\begin{aligned} \psi^n(\lambda) = \mathbf{E}e^{\lambda S_n} &\geq \sum_{k=1}^n \mathbf{E}(e^{(S_k + S_n - S_k)\lambda}; v = k) \\ &= \sum_{k=1}^n \mathbf{E}(e^{(x + \chi(x))\lambda} e^{(S_n - S_k)\lambda}; v = k). \end{aligned} \tag{15.3.3}$$

Here the random variables  $e^{\lambda\chi(x)} \mathbf{I}(v = k)$  and  $S_n - S_k$  are independent and, as before,

$$\mathbf{E}(e^{\lambda\chi(x)}; v = k) \geq \psi_- \mathbf{P}(v = k).$$

Therefore, for all  $\lambda$  such that  $\psi(\lambda) \leq 1$ ,

$$\psi^n(\lambda) \geq e^{\lambda x} \psi_- \sum_{k=1}^n \psi^{-k}(\lambda) \mathbf{P}(v = k) \geq \psi_- e^{\lambda x} \psi^n(\lambda) \mathbf{P}(v \leq n).$$

Hence we obtain

$$\mathbf{P}(\bar{S}_n > x) = \mathbf{P}(v \leq n) \leq \psi_-^{-1} e^{-\lambda x}.$$



Since the right-hand side does not depend on  $n$ , the same inequality also holds for  $\mathbf{P}(S > x)$ . The lower bound is obtained in a similar way. One just has to show that, in the original equality (cf. (15.3.3))

$$\psi^n(\lambda) = \sum_{k=1}^n \mathbf{E}(e^{\lambda S_n}; \nu = k) + \mathbf{E}(e^{\lambda S_n}; \nu > n),$$

one has  $\mathbf{E}(e^{\lambda S_n}; \nu > n) = o(1)$  as  $n \rightarrow \infty$  for  $\lambda = \mu$ , which we did in Sect. 15.2.

### 15.3.2 Inequalities for the Number of Crossings of a Strip

We now return to arbitrary submartingales  $X_n$  and prove an inequality that will be necessary for the convergence theorems of the next section. It concerns the number of crossings of a strip by the sequence  $X_n$ . Let  $a < b$  be given numbers. Set  $\nu_0 = 0$ ,

$$\begin{aligned} \nu_1 &:= \min\{n > 0 : X_n \leq a\}, & \nu_2 &:= \min\{n > \nu_1 : X_n \geq b\}, \\ &\dots\dots\dots & &\dots\dots\dots \\ \nu_{2k-1} &:= \min\{n > \nu_{2k-2} : X_n \leq a\}, & \nu_{2k} &:= \min\{n > \nu_{2k-1} : X_n \geq b\}. \end{aligned}$$

We put  $\nu_m := \infty$  if the path  $\{X_n\}$  for  $n \geq \nu_{m-1}$  never crosses the corresponding level. Using this notation, one can define the number of upcrossings of the strip (interval)  $[a, b]$  by the trajectory  $X_0, \dots, X_n$  as the random variable

$$\nu(a, b; n) := \begin{cases} \max\{k : \nu_{2k} \leq n\} & \text{if } \nu_2 \leq n, \\ 0 & \text{if } \nu_2 > n. \end{cases}$$

Set  $(a)^+ = \max(0, a)$ .

**Theorem 15.3.6** (Doob) *Let  $\{X_n, \mathfrak{F}_n; n \geq 0\}$  be a submartingale. Then, for all  $n$ ,*

$$\mathbf{E}\nu(a, b; n) \leq \frac{\mathbf{E}(X_n - a)^+}{b - a}. \tag{15.3.4}$$

It is clear that inequality (15.3.4) assumes by itself that only the submartingale  $\{X_n, \mathfrak{F}_n; 0 \leq k \leq n\}$  is given.

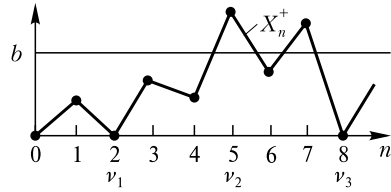
*Proof* The random variable  $\nu(a, b; n)$  coincides with the number of upcrossings of the interval  $[0, b - a]$  by the sequence  $(X_n - a)^+$ . Now  $\{(X_n - a)^+, \mathfrak{F}_n; n \geq 0\}$  is a nonnegative submartingale (see Example 15.1.4) and therefore, without loss of generality, one can assume that  $a = 0$  and  $X_n \geq 0$ , and aim to prove that

$$\mathbf{E}\nu(0, b; n) \leq \frac{\mathbf{E}X_n}{b}.$$

Let

$$\eta_j := \begin{cases} 1 & \text{if } \nu_k < j \leq \nu_{k+1} \text{ for some odd } k, \\ 0 & \text{if } \nu_k < j \leq \nu_{k+1} \text{ for some even } k. \end{cases}$$

**Fig. 15.1** Illustration to the proof of Theorem 15.3.6 showing the locations of the random times  $\nu_1, \nu_2,$  and  $\nu_3$  (here  $a = 0$ )



In Fig. 15.1,  $\nu_1 = 2, \nu_2 = 5, \nu_3 = 8; \eta_j = 0$  for  $j \leq 2, \eta_j = 1$  for  $3 \leq j \leq 5$  etc. It is not hard to see (using the Abel transform) that (with  $X_0 = 0, \eta_0 = 1$ )

$$\eta_0 X_0 + \sum_1^n \eta_j (X_j - X_{j-1}) = \sum_0^{n-1} X_j (\eta_j - \eta_{j+1}) + \eta_n X_n \geq b\nu(0, b; n).$$

Moreover (here  $\mathcal{N}_1$  denotes the set of odd numbers),

$$\{\eta_j = 1\} = \bigcup_{k \in \mathcal{N}_1} \{\nu_k < j \leq \nu_{k+1}\} = \bigcup_{k \in \mathcal{N}_1} [\{\nu_k \leq j - 1\} - \{\nu_{k+1} \leq j - 1\}] \in \mathfrak{F}_{j-1}.$$

Therefore, by virtue of the relation  $\mathbf{E}(X_j | \mathfrak{F}_{j-1}) - X_{j-1} \geq 0$ , we obtain

$$\begin{aligned} b\mathbf{E}\nu(0, b; n) &\leq \mathbf{E} \sum_1^n \eta_j (X_j - X_{j-1}) = \sum_1^n \mathbf{E}(X_j - X_{j-1}; \eta_j = 1) \\ &= \sum_1^n \mathbf{E}[\mathbf{E}(X_j - X_{j-1} | \mathfrak{F}_{j-1}); \eta_j = 1] = \sum_1^n \mathbf{E}[\mathbf{E}(X_j | \mathfrak{F}_{j-1}) - X_{j-1}; \eta_j = 1] \\ &\leq \sum_1^n \mathbf{E}[\mathbf{E}(X_j | \mathfrak{F}_{j-1}) - X_{j-1}] = \sum_1^n \mathbf{E}(X_j - X_{j-1}) = \mathbf{E}X_n. \end{aligned}$$

The theorem is proved. □

### 15.4 Convergence Theorems

**Theorem 15.4.1** (Doob's martingale convergence theorem) *Let*

$$\{X_n, \mathfrak{F}_n; -\infty < n < \infty\}$$

*be a submartingale. Then*

- (1) *The limit  $X_{-\infty} := \lim_{n \rightarrow -\infty} X_n$  exists a.s.,  $\mathbf{E}X_{-\infty}^+ < \infty$ , and the process  $\{X_n, \mathfrak{F}_n; -\infty \leq n < \infty\}$  is a submartingale.*
- (2) *If  $\sup_n \mathbf{E}X_n^+ < \infty$  then  $X_{\infty} := \lim_{n \rightarrow \infty} X_n$  exists a.s. and  $\mathbf{E}X_{\infty}^+ < \infty$ . If, moreover,  $\sup_n \mathbf{E}|X_n| < \infty$  then  $\mathbf{E}|X_{\infty}| < \infty$ .*
- (3) *The random sequence  $\{X_n, \mathfrak{F}_n; -\infty \leq n \leq \infty\}$  forms a submartingale if and only if the sequence  $\{X_n^+\}$  is uniformly integrable.*

*Proof* (1) Since

$$\{\limsup X_n > \liminf X_n\} = \bigcup_{\substack{\text{rational} \\ a, b}} \{\limsup X_n > b > a > \liminf X_n\}$$

(here the limits are taken as  $n \rightarrow \infty$ ), the assumption on divergence with positive probability

$$\mathbf{P}(\limsup X_n > \liminf X_n) > 0$$

means that there exist rational numbers  $a < b$  such that

$$\mathbf{P}(\limsup X_n > b > a > \liminf X_n) > 0. \quad (15.4.1)$$

Let  $v(a, b; m)$  be the number of upcrossings of the interval  $[a, b]$  by the sequence  $Y_1 = X_{-m}, \dots, Y_m = X_{-1}$  and  $v(a, b) = \lim_{m \rightarrow \infty} v(a, b; m)$ . Then (15.4.1) means that

$$\mathbf{P}(v(a, b) = \infty) > 0. \quad (15.4.2)$$

By Theorem 15.3.6 (applied to the sequence  $Y_1, \dots, Y_m$ ),

$$\mathbf{E}v(a, b; m) \leq \frac{\mathbf{E}(X_{-1} - a)^+}{b - a} \leq \frac{\mathbf{E}X_{-1}^+ + |a|}{b - a}, \quad (15.4.3)$$

$$\mathbf{E}v(a, b) \leq \frac{\mathbf{E}X_{-1}^+ + |a|}{b - a}. \quad (15.4.4)$$

Inequality (15.4.4) contradicts (15.4.2) and hence proves that

$$\mathbf{P}(\limsup X_n = \liminf X_n) = 1.$$

Moreover, by the Fatou–Lebesgue theorem ( $X_{-\infty}^+ := \liminf X_n^+$ ),

$$\mathbf{E}X_{-\infty}^+ \leq \liminf \mathbf{E}X_n^+ \leq \mathbf{E}X_{-1}^+ < \infty. \quad (15.4.5)$$

Here the second inequality follows from the fact that  $\{X_n^+, \mathfrak{F}_n\}$  is also a submartingale (see Lemma 15.1.3) and therefore  $\mathbf{E}X_n^+ \uparrow$ .

By Lemma 15.1.2, to prove that  $\{X_n, \mathfrak{F}_n; -\infty \leq n < \infty\}$  is a submartingale, it suffices to verify that, for any  $A \in \mathfrak{F}_{-\infty} \subset \mathfrak{F}$ ,

$$\mathbf{E}(X_{-\infty}; A) \leq \mathbf{E}(X_n; A). \quad (15.4.6)$$

Set  $X_n(a) := \max(X_n, a)$ . By Lemma 15.1.4,  $\{X_n(a), \mathfrak{F}_n; n \leq 0\}$  is a uniformly integrable submartingale. Therefore, for any  $-\infty < k < n$ ,

$$\begin{aligned} \mathbf{E}(X_k(a); A) &\leq \mathbf{E}(X_n(a); A), \\ \mathbf{E}(X_{-\infty}(a); A) &= \lim_{k \rightarrow -\infty} \mathbf{E}(X_k(a); A) \leq \mathbf{E}(X_n(a); A). \end{aligned} \quad (15.4.7)$$

Letting  $a \rightarrow -\infty$  we obtain (15.4.6) from the monotone convergence theorem.

(2) The second assertion of the theorem is proved in the same way. One just has to replace the right-hand sides of (15.4.3) and (15.4.4) with  $\mathbf{E}X_n^+$  and  $\sup_n \mathbf{E}X_n^+$ , respectively. Instead of (15.4.5) we get (the limits here are as  $n \rightarrow \infty$ )

$$\mathbf{E}X_{\infty}^+ \leq \liminf \mathbf{E}X_n^+ < \infty,$$

and if  $\sup_n \mathbf{E}|X_n| < \infty$  then

$$\mathbf{E}|X_\infty| \leq \liminf \mathbf{E}|X_n| < \infty.$$

(3) The last assertion of the theorem is proved in exactly the same way as the first one—the uniform integrability enables us to deduce along with (15.4.7) that, for any  $A \in \mathfrak{F}_n$ ,

$$\mathbf{E}(X_\infty(a); A) = \lim_{k \rightarrow \infty} \mathbf{E}(X_k(a); A) \geq \mathbf{E}(X_n(a); A).$$

The converse part of the third assertion of the theorem follows from Lemma 15.1.4. The theorem is proved.  $\square$

Now we will obtain some consequences of Theorem 15.4.1.

So far (see Sect. 4.8), while studying convergence of conditional expectations, we dealt with expectations of the form  $\mathbf{E}(X_n|\mathfrak{F})$ . Now we can obtain from Theorem 15.4.1 a useful theorem on convergence of conditional expectations of another type.

**Theorem 15.4.2** (Lévy) *Let a nondecreasing family  $\mathfrak{F}_1 \subseteq \mathfrak{F}_2 \subseteq \dots \subseteq \mathfrak{F}$  of  $\sigma$ -algebras and a random variable  $\xi$ , with  $\mathbf{E}|\xi| < \infty$ , be given on a probability space  $(\Omega, \mathfrak{F}, \mathbf{P})$ . Let, as before,  $\mathfrak{F}_\infty := \sigma(\bigcup_n \mathfrak{F}_n)$  be the  $\sigma$ -algebra generated by events from  $\mathfrak{F}_1, \mathfrak{F}_2, \dots$ . Then, as  $n \rightarrow \infty$ ,*

$$\mathbf{E}(\xi|\mathfrak{F}_n) \xrightarrow{a.s.} \mathbf{E}(\xi|\mathfrak{F}_\infty). \tag{15.4.8}$$

*Proof* Set  $X_n := \mathbf{E}(\xi|\mathfrak{F}_n)$ . We already know (see Example 15.1.3) that the sequence  $\{X_n, \mathfrak{F}_n; 1 < n \leq \infty\}$  is a martingale and therefore, by Theorem 15.4.1, the limit  $\lim_{n \rightarrow \infty} X_n = X_{(\infty)}$  exists a.s. It remains to prove that  $X_{(\infty)} = \mathbf{E}(\xi|\mathfrak{F}_\infty)$  (i.e., that  $X_{(\infty)} = X_\infty$ ). Since  $\{X_n, \mathfrak{F}_n; 1 \leq n \leq \infty\}$  is by Lemma 15.1.4 a uniformly integrable martingale,

$$\mathbf{E}(X_{(\infty)}; A) = \lim_{n \rightarrow \infty} \mathbf{E}(X_n; A) = \lim_{n \rightarrow \infty} \mathbf{E}(\mathbf{E}(\xi|\mathfrak{F}_n); A) = \mathbf{E}(\xi; A)$$

for  $A \in \mathfrak{F}_k$  and any  $k = 1, 2, \dots$ . This means that the left- and right-hand sides of the last relation, being finite measures, coincide on the algebra  $\bigcup_{n=1}^\infty \mathfrak{F}_n$ . By the theorem on extension of a measure (see Appendix 1), they will coincide for all  $A \in \sigma(\bigcup_{n=1}^\infty \mathfrak{F}_n) = \mathfrak{F}_\infty$ . Therefore, by the definition of conditional expectation,

$$X_{(\infty)} = \mathbf{E}(\xi|\mathfrak{F}_\infty) = X_\infty.$$

The theorem is proved.  $\square$

We could also note that the uniform integrability of  $\{X_n, \mathfrak{F}_n; 1 \leq n \leq \infty\}$  implies that  $\xrightarrow{a.s.}$  in (47) can be replaced by  $\xrightarrow{(1)}$ .

Theorem 15.4.1 implies the strong law of large numbers. Indeed, turn to our Example 15.1.4. By Theorem 15.4.1, the limit  $X_{-\infty} = \lim_{n \rightarrow -\infty} X_n = \lim_{n \rightarrow \infty} n^{-1} S_n$  exists a.s. and is measurable with respect to the tail (trivial)  $\sigma$ -algebra, and therefore it is constant with probability 1. Since  $\mathbf{E}X_{-\infty} = \mathbf{E}\xi_1$ , we have  $n^{-1} S_n \xrightarrow{a.s.} \mathbf{E}\xi_1$ .

One can also obtain some extensions of the theorems on series convergence of Chap. 11 to the case of dependent variables. Let

$$X_n = S_n = \sum_{k=1}^n \xi_k$$

and  $X_n$  form a submartingale ( $\mathbf{E}(\xi_{n+1} | \mathfrak{F}_n) \geq 0$ ). Let, moreover,  $\mathbf{E}|X_n| < c$  for all  $n$  and for some  $c < \infty$ . Then the limit  $S_\infty = \lim_{n \rightarrow \infty} S_n$  exists a.s. (As well as Theorem 15.4.1, this assertion is a generalisation of the monotone convergence theorem. The crucial role is played here by the condition that  $\mathbf{E}|X_n|$  is bounded.) In particular, if  $\xi_k$  are independent,  $\mathbf{E}\xi_k = 0$ , and the variances  $\sigma_k^2$  of  $\xi_k$  are such that  $\sum_{k=1}^{\infty} \sigma_k^2 < \sigma^2 < \infty$ , then

$$\mathbf{E}|X_n| \leq (\mathbf{E}X_n^2)^{1/2} \leq \left( \sum_{k=1}^n \sigma_k^2 \right)^{1/2} \leq \sigma < \infty,$$

and therefore  $S_n \xrightarrow{a.s.} S_\infty$ . Thus we obtain, as a consequence, the Kolmogorov theorem on series convergence.

*Example 15.4.1* Consider a *branching process*  $\{Z_n\}$  (see Sect. 7.7). We know that  $Z_n$  admits a representation

$$Z_n = \zeta_1 + \cdots + \zeta_{Z_{n-1}},$$

where the  $\zeta_k$  are identically distributed integer-valued random variables independent of each other and of  $Z_{n-1}$ ,  $\zeta_k$  being the number of descendants of the  $k$ -th particle from the  $(n-1)$ -th generation. Assuming that  $Z_0 = 1$  and setting  $\mu := \mathbf{E}\zeta_k$ , we obtain

$$\mathbf{E}(Z_n | Z_{n-1}) = \mu Z_{n-1}, \quad \mathbf{E}Z_n = \mu \mathbf{E}Z_{n-1} = \mu^n.$$

This implies that  $X_n = Z_n / \mu^n$  is a martingale, because

$$\mathbf{E}(X_n | X_{n-1}) = \mu^{1-n} Z_{n-1} = X_{n-1}.$$

For branching processes we have the following.

**Theorem 15.4.3** *The sequence  $X_n = \mu^{-n} Z_n$  converges almost surely to a proper random variable  $X$  with  $\mathbf{E}X < \infty$ . The ch.f.  $\varphi(\lambda)$  of the random variable  $X$  satisfies the equation*

$$\varphi(\mu\lambda) = p(\varphi(\lambda)),$$

where  $p(v) = \mathbf{E}v^{\zeta_k}$ .

Theorem 15.4.3 means that  $\mu^{-n} Z_n$  has a proper limiting distribution as  $n \rightarrow \infty$ .

*Proof* Since  $X_n \geq 0$  and  $\mathbf{E}X_n = 1$ , the first assertion follows immediately from Theorem 15.4.1.

Since  $\mathbf{E}Z^{Z_n}$  is equal to the  $n$ -th iteration of the function  $f(z)$ , for the ch.f. of  $Z_n$  we have  $(\varphi_n(\lambda) := \mathbf{E}e^{i\lambda\eta})$

$$\begin{aligned} \varphi_{Z_n}(\lambda) &= p(\varphi_{Z_{n-1}}(\lambda)), \\ \varphi_{X_n}(\lambda) &= \varphi_{Z_n}(\mu^{-n}\lambda) = p(\varphi_{Z_{n-1}}(\mu^{-n}\lambda)) = p\left(\varphi_{X_{n-1}}\left(\frac{\lambda}{\mu}\right)\right). \end{aligned}$$

Because  $X_n \Rightarrow X$  and the function  $p$  is continuous, from this we obtain the equation for the ch.f. of the limiting distribution  $X$ :

$$\varphi(\lambda) = p\left(\varphi\left(\frac{\lambda}{\mu}\right)\right).$$

The theorem is proved. □

In Sect. 7.7 we established that in the case  $\mu \leq 1$  the process  $Z_n$  becomes extinct with probability 1 and therefore  $\mathbf{P}(X = 0) = 1$ . We verify now that, for  $\mu > 1$ , the distribution of  $X$  is nondegenerate (not concentrated at zero). It suffices to prove that  $\{X_n, 0 \leq n \leq \infty\}$  forms a martingale and consequently

$$\mathbf{E}X = \mathbf{E}X_n \neq 0.$$

By Theorem 15.4.1, it suffices to verify that the sequence  $X_n$  is uniformly integrable. To simplify the reasoning, we suppose that  $\text{Var}(\zeta_k) = \sigma^2 < \infty$  and show that then  $\mathbf{E}X_n^2 < c < \infty$  (this certainly implies the required uniform integrability of  $X_n$ , see Sect. 6.1). One can directly verify the identity

$$Z_n^2 - \mu^{2n} = \sum_{k=1}^n [Z_k^2 - (\mu Z_{k-1})^2] \mu^{2n-2k}.$$

Since  $\mathbf{E}[Z_k^2 - (\mu Z_{k-1})^2 | Z_{k-1}] = \sigma^2 Z_{k-1}$  (recall that  $\text{Var}(\eta) = \mathbf{E}(\eta^2 - (\mathbf{E}\eta)^2)$ ), we have

$$\begin{aligned} \text{Var}(Z_n) &= \mathbf{E}(Z_n^2 - \mu^{2n}) = \sum_{k=1}^n \mu^{2n-2k} \sigma^2 \mathbf{E}Z_{k-1} \\ &= \mu^{2n} \sigma^2 \sum_{k=1}^n \mu^{-k-1} = \frac{\sigma^2 \mu^n (\mu^n - 1)}{\mu(\mu - 1)}, \\ \mathbf{E}X_n^2 &= \mu^{-2n} \mathbf{E}Z_n^2 = 1 + \frac{\sigma^2(1 - \mu^{-n})}{\mu(\mu - 1)} \leq 1 + \frac{\sigma^2}{\mu(\mu - 1)}. \end{aligned}$$

Thus we have proved that  $X$  is a nondegenerate random variable,

$$\mathbf{E}X = 1, \quad \text{Var}(X_n) \rightarrow \frac{\sigma^2}{\mu(\mu - 1)}.$$

From the last relation one can easily obtain that  $\text{Var}(X) = \frac{\sigma^2}{\mu(\mu-1)}$ . To this end one can, say, prove that  $X_n$  is a Cauchy sequence in mean quadratic and hence (see Theorem 6.1.3)  $X_n \xrightarrow{(2)} X$ .

## 15.5 Boundedness of the Moments of Stochastic Sequences

When one uses convergence theorems for martingales, conditions ensuring boundedness of the moments of stochastic sequences  $\{X_n, \mathfrak{F}_n\}$  are of significant interest (recall that the boundedness of  $\mathbf{E}X_n$  is one of the crucial conditions for convergence of submartingales). The boundedness of the moments, in turn, ensures that  $X_n$  is stochastically bounded, i.e., that  $\sup_n \mathbf{P}(X_n > N) \rightarrow 0$  as  $N \rightarrow \infty$ . The last boundedness is also of independent interest in the cases where one is not able to prove, for the sequence  $\{X_n\}$ , convergence or any other ergodic properties.

For simplicity's sake, we confine ourselves to considering nonnegative sequences  $X_n \geq 0$ . Of course, if we could prove convergence of the distributions of  $X_n$  to a limiting distribution, as was the case for Markov chains or submartingales in Theorem 15.4.1, then we would have a more detailed description of the asymptotic behaviour of  $X_n$  as  $n \rightarrow \infty$ . This convergence, however, requires that the sequence  $X_n$  satisfies stronger constraints than will be used below.

The basic and rather natural elements of the boundedness conditions to be considered below are: the boundedness of the moments of  $\xi_n = X_n - X_{n-1}$  of the respective orders and the presence of a negative "drift"  $\mathbf{E}(\xi_n | \mathfrak{F}_{n-1})$  in the domain  $X_{n-1} > N$  for sufficiently large  $N$ . Such a property has already been utilised for Markov chains; see Corollary 13.7.1 (otherwise the trajectory of  $X_n$  may go to  $\infty$ ).

Let us begin with exponential moments. The simplest conditions ensuring the boundedness of  $\sup_n \mathbf{E}e^{\lambda X_n}$  for some  $\lambda > 0$  are as follows: for all  $n \geq 1$  and some  $\lambda > 0$  and  $N < \infty$ ,

$$\mathbf{E}(e^{\lambda \xi_n} | \mathfrak{F}_{n-1}) \mathbf{I}(X_{n-1} > N) \leq \beta(\lambda) < 1, \quad (15.5.1)$$

$$\mathbf{E}(e^{\lambda \xi_n} | \mathfrak{F}_{n-1}) \mathbf{I}(X_{n-1} \leq N) \leq \psi(\lambda) < \infty. \quad (15.5.2)$$

**Theorem 15.5.1** *If conditions (15.5.1) and (15.5.2) hold then*

$$\mathbf{E}(e^{\lambda X_n} | \mathfrak{F}_0) \leq \beta(\lambda) e^{\lambda X_0} + \frac{\psi(\lambda) e^{\lambda N}}{1 - \beta(\lambda)}. \quad (15.5.3)$$

*Proof* Denote by  $A_n$  the left-hand side of (15.5.3). Then, by virtue of (15.5.1) and (15.5.2), we obtain

$$\begin{aligned} A_n &= \mathbf{E}\left\{\mathbf{E}\left[e^{\lambda X_n} (\mathbf{I}(X_{n-1} > N) + \mathbf{I}(X_{n-1} \leq N)) \mid \mathfrak{F}_{n-1}\right] \mid \mathfrak{F}_0\right\} \\ &\leq \mathbf{E}\left[e^{\lambda X_{n-1}} (\beta(\lambda) \mathbf{I}(X_{n-1} > N) + \psi(\lambda) \mathbf{I}(X_{n-1} \leq N)) \mid \mathfrak{F}_0\right] \\ &\leq \beta(\lambda) A_{n-1} + e^{\lambda N} \psi(\lambda). \end{aligned}$$

This immediately implies that

$$A_n \leq A_0 \beta^n(\lambda) + e^{\lambda N} \psi(\lambda) \sum_{k=0}^{n-1} \beta^k(\lambda) \leq A_0 \beta^n(\lambda) + \frac{e^{\lambda N} \psi(\lambda)}{1 - \beta(\lambda)}.$$

The theorem is proved.  $\square$

The conditions

$$\mathbf{E}(\xi_n | \mathfrak{F}_{n-1}) \leq -\varepsilon < 0 \quad \text{on the } \omega\text{-set } \{X_{n-1} > N\}, \tag{15.5.4}$$

$$\mathbf{E}(e^{\lambda|\xi_n|} | \mathfrak{F}_{n-1}) \leq \psi_1(\lambda) < \infty \quad \text{for some } \lambda > 0 \tag{15.5.5}$$

are sufficient for (15.5.1) and (15.5.2).

The first condition means that  $Y_n := (X_n + \varepsilon n) \mathbf{I}(X_{n-1} > N)$  is a supermartingale.

We now prove sufficiency of (15.5.4) and (15.5.5). That (15.5.2) holds is clear. Further, make use of the inequality

$$e^x \leq 1 + x + \frac{x^2}{2} e^{|x|},$$

which follows from the Taylor formula for  $e^x$  with the remainder in the Cauchy form:

$$e^x = 1 + x + \frac{x^2}{2} e^{\theta x}, \quad \theta \in [0, 1].$$

Then, on the set  $\{X_{n-1} > N\}$ , one has

$$\mathbf{E}(e^{\lambda\xi_n} | \mathfrak{F}_{n-1}) \leq 1 - \lambda\varepsilon + \frac{\lambda^2}{2} \mathbf{E}(\xi_n^2 e^{\lambda|\xi_n|} | \mathfrak{F}_{n-1}).$$

Since  $x^2 < e^{\lambda x/2}$  for sufficiently large  $x$ , by the Hölder inequality it follows that, together with (15.5.5), we will have

$$\mathbf{E}(\xi_n^2 e^{\lambda|\xi_n|/2} | \mathfrak{F}_{n-1}) \leq \psi_2(\lambda) < \infty.$$

This implies that, for sufficiently small  $\lambda$ , one has on the set  $\{X_{n-1} > N\}$  the inequality

$$\mathbf{E}(e^{\lambda\xi_n} | \mathfrak{F}_{n-1}) \leq 1 - \lambda\varepsilon + \frac{\lambda^2}{2} \psi_2(\lambda) =: \beta(\lambda) \leq 1 - \frac{\lambda\varepsilon}{2} < 1.$$

This proves (15.5.1). □

**Corollary 15.5.1** *If, in addition to the conditions of Theorem 15.5.1, the distribution of  $X_n$  converges to a limiting distribution:  $\mathbf{P}(X_n < t) \Rightarrow \mathbf{P}(X < t)$ , then*

$$\mathbf{E}e^{\lambda X} \leq \frac{e^{\lambda N} \psi(\lambda)}{1 - \beta(\lambda)}.$$

The corollary follows from the Fatou–Lebesgue theorem (see also Lemma 6.1.1):

$$\mathbf{E}e^{\lambda X} \leq \liminf_{n \rightarrow \infty} \mathbf{E}e^{\lambda X_n}. \tag{□}$$

We now obtain bounds for “conventional” moments. Set

$$\begin{aligned} M^l(n) &:= \mathbf{E}X_n^l, \\ m(0) &:= 1, \quad m(1) := \sup_{n \geq 1} \sup_{\omega \in \{X_{n-1} > N\}} \mathbf{E}(\xi_n | \mathfrak{F}_{n-1}), \\ m(l) &:= \sup_{n \geq 1} \sup_{\omega} \mathbf{E}(|\xi_n|^l | \mathfrak{F}_{n-1}), \quad l > 1. \end{aligned}$$



**Theorem 15.5.2** Assume that  $\mathbf{E}X_0^s < \infty$  for some  $s > 1$  and there exist  $N \geq 0$  and  $\varepsilon > 0$  such that

$$m(1) \leq -\varepsilon, \quad (15.5.6)$$

$$m(s) < c < \infty. \quad (15.5.7)$$

Then

$$\liminf_{n \rightarrow \infty} M^{s-1}(n) < \infty. \quad (15.5.8)$$

If, moreover,

$$M^s(n+1) > M^s(n) - c_1 \quad (15.5.9)$$

for some  $c_1 > 0$ , then

$$\sup_n M^{s-1}(n) < \infty. \quad (15.5.10)$$

**Corollary 15.5.2** If conditions (15.5.6) and (15.5.7) are met and the distribution of  $X_n$  converges weakly to a limiting distribution:  $\mathbf{P}(X_n < t) \Rightarrow \mathbf{P}(X < t)$ , then  $\mathbf{E}X^{s-1} < \infty$ .

This assertion follows from the Fatou–Lebesgue theorem (see also Lemma 6.1.1), which implies

$$\mathbf{E}X^{s-1} \leq \liminf_{n \rightarrow \infty} \mathbf{E}X_n^{s-1}. \quad \square$$

The assertion of Corollary 15.5.2 is unimprovable. One can see this from the example of the sequence  $X_n = (X_{n-1} + \zeta_n)^+$ , where  $\zeta_k \stackrel{d}{=} \zeta$  are independent and identically distributed. If  $\mathbf{E}\zeta_k < 0$  then the limiting distribution of  $X_n$  coincides with the distribution of  $S = \sup_k S_k$  (see Sect. 12.4). From factorisation identities one can derive that  $\mathbf{E}S^{s-1}$  is finite if and only if  $\mathbf{E}(\zeta^+)^s < \infty$ . An outline of the proof is as follows. Theorem 12.3.2 implies that  $\mathbf{E}S^k = c \mathbf{E}(\chi_+^k; \eta_+ < \infty)$ ,  $c = \text{const} < \infty$ . It follows from Corollary 12.2.2 that

$$1 - \mathbf{E}(e^{i\lambda\chi_+}; \eta_+ < \infty) = (1 - \mathbf{E}e^{i\lambda\zeta}) \int_0^\infty e^{-i\lambda x} dH(x),$$

where  $H(x)$  is the renewal function for the random variable  $-\chi_-^0 \geq 0$ . Since

$$a_1 + b_1x \leq H(x) \leq a_2 + b_2x$$

(see Theorem 10.1.1 and Lemma 10.1.1;  $a_i, b_i$  are constants), integrating the convolution

$$\mathbf{P}(\chi_+ > x, \eta_+ < \infty) = \int_0^\infty \mathbf{P}(\zeta > v + x) dH(v)$$

by parts we verify that, as  $x \rightarrow \infty$ , the left-hand side has the same order of magnitude as  $\int_0^\infty \mathbf{P}(\zeta > v + x) dv$ . Hence the required statement follows.

We now return to Theorem 15.5.2. Note that in all of the most popular problems the sequence  $M^{s-1}(n)$  behaves “regularly”: either it is bounded or  $M^{s-1}(n) \rightarrow \infty$ . Assertion (15.5.8) means that, under the conditions of Theorem 15.5.2, the second possibility is excluded. Condition (15.5.9) ensuring (15.5.10) is also rather broad.

*Proof of Theorem 15.5.2* Let for simplicity’s sake  $s > 1$  be an integer. We have

$$\begin{aligned} \mathbf{E}(X_n^s; X_{n-1} > N) &= \int_N^\infty \mathbf{E}((x + \xi_n)^s; X_{n-1} \in dx) \\ &= \sum_{l=0}^s \binom{s}{l} \int_N^\infty x^l \mathbf{E}(\xi_n^{s-l}; X_{n-1} \in dx). \end{aligned}$$

If we replace  $\xi_n^{s-l}$  for  $s - l \geq 2$  with  $|\xi_n|^{s-l}$  then the right-hand side can only increase. Therefore,

$$\mathbf{E}(X_n^s; X_{n-1} > N) \leq \sum_{l=0}^s \binom{s}{l} m(s-l) M_N^l(n-1),$$

where

$$M_N^l(n) = \mathbf{E}(X_n^l; X_n > N).$$

The moments  $M^s(n) = \mathbf{E}X_n^s$  satisfy the inequalities

$$\begin{aligned} M^s(n) &\leq \mathbf{E}[(N + |\xi_n|)^s; X_{n-1} \leq N] + \sum_{l=0}^s \binom{s}{l} m(s-l) M_N^l(n-1) \\ &\leq 2^s [N^s + c] + \sum_{l=0}^s \binom{s}{l} m(s-l) M_N^l(n-1). \end{aligned} \tag{15.5.11}$$

Suppose now that (15.5.8) does not hold:  $M^{s-1}(n) \rightarrow \infty$ . Then all the more  $M^s(n) \rightarrow \infty$  and there exists a subsequence  $n'$  such that  $M^s(n') > M^s(n' - 1)$ . Since  $M^l(n) \leq [M^{l+1}(n)]^{l/l+1}$ , we obtain from (15.5.6) and (15.5.11) that

$$\begin{aligned} M^s(n') &\leq \text{const} + M^s(n' - 1) + sM^{s-1}(n' - 1)m(1) + o(M^{s-1}(n' - 1)) \\ &\leq M^s(n' - 1) - \frac{1}{2}s\varepsilon M^{s-1}(n' - 1) \end{aligned}$$

for sufficiently large  $n'$ . This contradicts the assumption that  $M^s(n) \rightarrow \infty$  and hence proves (15.5.8).

We now prove (15.5.10). If this relation is not true then there exists a sequence  $n'$  such that  $M^{s-1}(n') \rightarrow \infty$  and  $M^s(n') > M^s(n' - 1) - c_1$ . It remains to make use of the above argument.

We leave the proof for a non-integer  $s > 1$  to the reader (the changes are elementary). The theorem is proved.  $\square$

*Remark 15.5.1* (1) The assertions of Theorems 15.5.1 and 15.5.2 will remain valid if one requires inequalities (15.5.4) or  $\mathbf{E}(\xi_n + \varepsilon | \mathfrak{F}_{n-1}) \mathbf{I}(X_{n-1} > N) \leq 0$  to hold not for all  $n$ , but only for  $n \geq n_0$  for some  $n_0 > 1$ .

(2) As in Theorem 15.5.1, condition (15.5.6) means that the sequence of random variables  $(X_n + \varepsilon n) \mathbf{I}(X_{n-1} > N)$  forms a supermartingale.

(3) The conditions of Theorems 15.5.1 and 15.5.2 may be weakened by replacing them with “averaged” conditions. Consider, for instance, condition (15.5.1). By integrating it over the set  $\{X_{n-1} > x > N\}$  we obtain

$$\mathbf{E}(e^{\lambda \xi_n}; X_{n-1} > x) \leq \beta(\lambda) \mathbf{P}(X_{n-1} > x)$$

or, which is the same,

$$\mathbf{E}(e^{\lambda \xi_n} | X_{n-1} > x) \leq \beta(\lambda). \tag{15.5.12}$$

The converse assertion that (15.5.12) for all  $x > N$  implies relation (15.5.1) is obviously false, so that condition (15.5.12) is weaker than (15.5.1). A similar remark is true for condition (15.5.4).

One has the following generalisations of Theorems 15.5.1 and 15.5.2 to the case of “averaged conditions”.

**Theorem 15.5.1A** *Let, for some  $\lambda > 0$ ,  $N > 0$  and all  $x \geq N$ ,*

$$\mathbf{E}(e^{\lambda \xi_n} | X_{n-1} > x) \leq \beta(\lambda) < 1, \quad \mathbf{E}(e^{\lambda \xi_n}; X_{n-1} \leq N) \leq \psi(\lambda) < \infty.$$

*Then*

$$\mathbf{E}e^{\lambda X_n} \leq \beta^n(\lambda) \mathbf{E}e^{\lambda X(0)} + \frac{e^{\lambda N} \psi(\lambda)}{1 - \beta(\lambda)}.$$

Put

$$\bar{m}(1) := \sup_{n \geq 1} \sup_{x \geq N} \mathbf{E}(\xi_n | X_{n-1} > x),$$

$$\bar{m}(l) := \sup_{n \geq 1} \sup_{x \geq N} \mathbf{E}(|\xi_n|^l | X(n) > x), \quad l > 1.$$

**Theorem 15.5.2A** *Let  $\mathbf{E}X_0^s < \infty$  and there exist  $N \geq 0$  and  $\varepsilon > 0$  such that*

$$\bar{m}(1) \leq -\varepsilon, \quad \bar{m}(s) < \infty, \quad \mathbf{E}(|\xi_n|^s; X_{n-1} \leq N) < c < \infty.$$

*Then (15.5.8) holds true. If, in addition, (15.5.9) is valid, then (15.5.10) is true.*

*The proofs of Theorems 15.5.1A and 15.5.2A are quite similar to those of Theorems 15.5.1 and 15.5.2. The only additional element in both cases is integration by parts. We will illustrate this with the proof of Theorem 15.5.1A. Consider*

$$\begin{aligned}
\mathbf{E}(e^{\lambda X_n}; X_{n-1} > N) &= \int_N^\infty e^{\lambda x} \mathbf{E}(e^{\lambda \xi_n}; X_{n-1} \in dx) \\
&= \mathbf{E}(e^{\lambda(N+\xi_n)}; X_{n-1} > N) + \int_N^\infty \lambda e^{\lambda x} \mathbf{E}(e^{\lambda \xi_n}; X_{n-1} > x) dx \\
&\leq \mathbf{E}(e^{\lambda(N+\xi_n)}; X_{n-1} > N + \beta(\lambda)) \int_N^\infty \lambda e^{\lambda x} \mathbf{P}(X_{n-1} > x) dx \\
&= e^{\lambda N} \mathbf{E}(e^{\lambda \xi_n} - \beta(\lambda); X_{n-1} > N) + \beta(\lambda) \int_N^\infty e^{\lambda x} \mathbf{P}(X_{n-1} \in dx) \\
&\leq \beta(\lambda) \mathbf{E}(e^{\lambda X_{n-1}}; X_{n-1} > N).
\end{aligned}$$

From this we find that

$$\begin{aligned}
\beta_n(\lambda) &:= \mathbf{E}e^{\lambda X_n} \leq \mathbf{E}(e^{\lambda(X_{n-1}+\xi_n)}; X_{n-1} \leq N) + \mathbf{E}(e^{\lambda X_n}; X_{n-1} > N) \\
&\leq e^{\lambda N} \psi(\lambda) + \beta(\lambda) \mathbf{E}(e^{\lambda X_{n-1}}; X_{n-1} > N) \\
&\leq e^{\lambda N} \psi(\lambda) - \mathbf{P}(X_{n-1} \leq N) \beta(\lambda) + \beta(\lambda) \beta_n(\lambda); \\
\beta_n(\lambda) &\leq \beta^n(\lambda) \beta_0(\lambda) + \frac{e^{\lambda N} \psi(\lambda)}{1 - \beta(\lambda)}. \quad \square
\end{aligned}$$

Note that Theorem 13.7.2 and Corollary 13.7.1 on “positive recurrence” can also be referred to as theorems on boundedness of stochastic sequences.