stochastic processes and their applications

ELSEVIER Stochastic Processes and their Applications 64 (1996) 103–125

Large deviations results for subexponential tails, with applications to insurance risk

Søren Asmussen[a], Claudia Klüppelberg[b,*]

[a] Department of Mathematics, University of Lund, Box 118, S-22100 Lund, Sweden

[b] Department of Mathematics, Johannes Gutenberg-University, D-55099 Mainz, Germany

Received 23 April 1995; revised 14 May 1996

Abstract

Consider a random walk or Lévy process $\{S_t\}$ and let $\tau(u) = \inf\{t \geqslant 0 : S_t > u\}$, $\mathbb{P}^{(u)}(\cdot) = \mathbb{P}(\cdot \mid \tau(u) < \infty)$. Assuming that the upwards jumps are heavy–tailed, say subexponential (e.g. Pareto, Weibull or lognormal), the asymptotic form of the $\mathbb{P}^{(u)}$–distribution of the process $\{S_t\}$ up to time $\tau(u)$ is described as $u \to \infty$. Essentially, the results confirm the folklore that level crossing occurs as result of one big jump. Particular sharp conclusions are obtained for downwards skip-free processes like the classical compound Poisson insurance risk process where the formulation is in terms of total variation convergence. The ideas of the proof involve excursions and path decompositions for Markov processes. As a corollary, it follows that for some deterministic function $a(u)$, the limiting $\mathbb{P}^{(u)}$–distribution of $\tau(u)/a(u)$ is either Pareto or exponential, and corresponding approximations for the finite time ruin probabilities are given.

Keywords: Conditioned limit theorem; Downwards skip-free process; Excursion; Extreme value theory; Insurance risk; Integrated tail; Maximum domain of attraction; Path decomposition; Random walk; Regular variation; Ruin probability; Subexponential distribution; Total variation convergence

AMS classification : 60F10, 60K10

1. Introduction and statement of results

Let $\{S_t\}$ be a random walk in discrete time or a Lévy process in continuous time. Assume throughout that the drift $-\mu = \mathbb{E}S_1$ is negative. Then $M = \max_{t \geqslant 0} S_t$ is finite, and the event $\{M > u\} = \{\tau(u) < \infty\}$ where $\tau(u) = \inf\{t > 0 : S_t > u\}$ is rare when u is large, i.e. $\psi(u) = \mathbb{P}(\tau(u) < \infty)$ is small. Thus, typical problems for large deviations theory (e.g. Bucklew, 1990; Dembo and Zeitouni, 1993; or Deuschel and Stroock, 1989) are the study

(i) of the asymptotic form of $\psi(u)$ as $u \to \infty$, and

* Corresponding author.

PII S0304-4149(96)00087-7

(ii) letting $\mathbb{P}^{(u)}(\cdot) = \mathbb{P}(\cdot \mid \tau(u) < \infty)$, of the $\mathbb{P}^{(u)}$–distribution of the path

$$S_{[0,\tau(u))} = \{S_t\}_{0 \leqslant t < \tau(u)} \tag{1.1}$$

leading to the occurrence of the rare event.

Assuming $\kappa(\alpha) = \log \mathbb{E}e^{\alpha S_1}$ to be finite for sufficiently large $\alpha > 0$, the solutions to these problems are in fact well known: Let γ be the positive solution of $\kappa(\gamma) = 0$, then as $u \to \infty$ ($\sim$ means that the quotient of lhs and rhs tends to 1 in the indicated sense),

(i) $-\log \psi(u) \sim \gamma u$, and

(ii) $\tau(u) \sim u/\kappa'(\gamma)$ and $S_{\lfloor t\tau(u) \rfloor} \sim \kappa'(\gamma) t\tau(u)$, $0 < t < 1$, in $\mathbb{P}^{(u)}$–distribution.

The form of these statements are the typical ones of large deviations theory (being applicable to much more general settings like Markov processes and different types of rare events), but in fact rather much sharper results can be derived in this specific setting:

(i) $\psi(u) \sim Ce^{-\gamma u}$ for some positive constant C (Cramér, 1930; Feller, 1971),

(ii) the $\mathbb{P}^{(u)}$–distribution of $S_{[0,\tau(u))}$ is in an appropriate sense the same as the unconditional distribution w.r.t. the Lévy process obtained by the exponential change of measure corresponding to replacing $\kappa(\alpha)$ by

$$\kappa_\gamma(\alpha) = \kappa(\alpha + \gamma) - \kappa(\gamma) \tag{1.2}$$

(Asmussen, 1982).

The present paper is concerned with the same type of problem in the heavy–tailed case where $\kappa(\alpha) = \infty$ for all $\alpha > 0$. The precise set–up is stated later for the various models we consider, but basically 'heavy–tailed' means that positive jumps have the same tail behaviour as a certain subexponential distribution B. We recall that B is a *subexponential distribution function* ($B \in \mathscr{S}$) if for positive iid rv's $Y_1, \ldots, Y_n$ with common distribution B

$$\mathbb{P}(\max(Y_1, \ldots, Y_n) > x) \sim \mathbb{P}(Y_1 + \cdots + Y_n > x), \quad x \to \infty,$$

holds for all $n \in \mathbb{N}$ (Embrechts and Goldie, 1980). This definition immediately classifies subexponential distributions as heavy–tailed distributions: for large x the maximum dominates the sum. Moreover, it is this characterisation which is fundamental to the results of this paper. It is well-known that Pareto–, lognormal – and certain Weibull distributions are subexponential; see e.g. Klüppelberg (1987) and references therein.

In this setting, the solution to problem (i) is known (Embrechts and Veraverbeke, 1982 and references there) and stated in (1.3), (1.9) below: $\psi(u)$ is asymptotically proportional to $\overline{B}_0(u)$ where B_0 is the integrated tail distribution,

$$\overline{B}_0(u) = \frac{1}{\mu_B} \int_u^\infty \overline{B}(x)\, \mathrm{d}x$$

(here $\overline{B}(x) = 1 - B(x)$ and μ_B is the mean of B). One of the main contributions of this paper is to present a solution to problem (ii) in the heavy–tailed case where a common folklore asserts that rare events of the type we consider occur as consequence of one big jump. We make here precise for various examples, how big this jump is, what the asymptotic distribution of the time $\tau(u)$, when it occurs, looks like, and what it means that the process evolves in its 'typical' way up to time $\tau(u)$. We work in two different though closely related settings.

1.1. Random walks and Lévy processes

Assume first that $\{S_n\} = \{X_1 + \cdots + X_n\}$ is a discrete time random walk with increment distribution F, such that the mean of F is $-\mu < 0$ and that $\overline{F}(x) \sim \overline{B}(x)$ as $x \to \infty$ where B is a subexponential distribution on $(0, \infty)$. In this setting, the result of Embrechts and Veraverbeke (1982) can be written as

$$\mathbb{P}(\tau(u) < \infty) \sim \frac{1}{\mu} \int_u^\infty \overline{F}(x) \mathrm{d}x, \quad u \to \infty. \tag{1.3}$$

Our main result for the random walk case uses classical extreme value theory. This makes it necessary to distinguish between two classes of subexponential distributions, corresponding to the maximum domain of attraction of a Frèchet distribution $\Phi_\alpha, \alpha > 0$, and the Gumbel distribution Λ, respectively. We write MDA(Φ_α) and MDA(Λ). MDA(Φ_α) consists of df's with regularly varying tail, i.e. $\overline{B}(x) = x^{-\alpha}L(x)$, where $\alpha > 0$ and L is a slowly varying function. We write $\overline{B} \in \mathscr{R}(-\alpha)$ (see Bingham et al., 1987 for definitions and properties of regularly varying functions). MDA(Λ) includes the Weibull- and lognormal cases. We refer to Section 3 for more details.

Let V_α, where $0 < \alpha \leqslant \infty$, be a rv with a *generalised Pareto distribution* G_α, i.e. V_α is positive and its tail is given by

$$\overline{G}_\alpha(x) = \mathbb{P}(V_\alpha > x) = \begin{cases} (1 + x/\alpha)^{-\alpha} & \alpha < \infty, \\ \mathrm{e}^{-x} & \alpha = \infty, \end{cases} \quad x > 0. \tag{1.4}$$

Let T_α be defined on the same probability space, such that $\mathbb{P}(V_\alpha > x, T_\alpha > y) = \overline{G}_\alpha(x + y)$ (then T_α has marginal distribution G_α as well and, when $\alpha = \infty$, V_α and T_α are just independent).

Define further $-Z(u) = S_{\tau(u)-1}$ as the level of the random walk just before the upcrossing of level u, $Y(u) = S_{\tau(u)}$ as the level just after, and $s_0(t) = -\mu t$ for $0 \leqslant t < 1$.

Theorem 1.1. *Assume that either $\overline{B} \in \mathscr{R}(-\alpha - 1)$ for $\alpha \in (0, \infty)$ or $B \in$ MDA(Λ) and $B_0 \in \mathscr{S}$, and let $a(u) \sim \int_u^\infty \overline{B}(x)\mathrm{d}x / \overline{B}(u)$. Then*

$$\left(\frac{Z(u)}{a(u)}, \frac{\tau(u)}{a(u)}, \left\{ \frac{S_{\lfloor t\tau(u) \rfloor}}{\tau(u)} \right\}_{0 \leqslant t < 1}, \frac{Y(u) - u}{a(u)} \right) \to (V_\alpha, V_\alpha/\mu, s_0, T_\alpha), \quad u \to \infty, \tag{1.5}$$

in $\mathbb{P}^{(u)}$–distribution in $\mathbb{R} \times \mathbb{R}_+ \times D[0, 1) \times \mathbb{R}_+$.

Notice that $a(u) \sim \int_u^\infty \overline{F}(x)\mathrm{d}x / \overline{F}(u)$, providing another normalising function in (1.5).

Note furthermore that by Karamata's theorem $a(u) \sim u/\alpha$ when $B \in \mathscr{R}(-\alpha-1)$. For the case $B \in \mathrm{MDA}(\Lambda)$, see Section 3.

The statement that $S_{\lfloor t\tau(u)\rfloor} \sim -\mu t\tau(u)$ for $0 < t < 1$ is an intuitive support that the random walk evolves 'typically' up to time $\tau(u)$ because $S_t \sim -\mu t$ for large t.

A further substantiation of this fact is the following.

Theorem 1.2. *Let*

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} I(X_k \leqslant x)$$

be the empirical df. Then under the conditions of Theorem 1.1

$$\mathbb{P}^{(u)}\left(\|F_{\tau(u)} - F\| > \varepsilon\right) \to 0, \quad u \to \infty. \tag{1.6}$$

Here $\|\cdot\|$ denotes the supremum norm. Note that this formulation is similar to the ones of Asmussen (1982) for the case of exponentially bounded tails (then the limit of $F_{\tau(u)}$ is not F but the exponentially twisted distribution given by (1.2)).

Using a discrete skeleton argument, it is straightforward to give parallels of the above results for Lévy processes (processes with stationary-independent increments in continuous time) when the tail of the Lévy measure is subexponential. We omit the details, but proceed to a special case.

1.2. Insurance risk processes

Let now S be the claim surplus process of a classical compound risk process R,

$$S_t = \sum_{\{n\in\mathbb{N}:\ 0<T_n\leqslant t\}} U_n - ct \quad \text{and} \quad R_t = u - S_t, \quad t \geqslant 0, \tag{1.7}$$

where the T_n are the epochs of a Poisson process with intensity β and the $\{U_n\}_{n\in\mathbb{N}}$ are iid with common distribution B (the claim size distribution). The positive constant c is the premium rate, which w.l.o.g we assume to be equal to 1. Then $\tau(u)$ is the ruin time,

$$\psi(u,T) = \mathbb{P}(\tau(u) \leqslant T) \quad \text{and} \quad \psi(u) = \psi(u,\infty) = \mathbb{P}(\tau(u) < \infty) \tag{1.8}$$

are the probability of ruin before time T and in infinite time, respectively. In earlier notation, $\mu = -\mathbb{E}S_1 = 1 - \rho$ where $\rho = \beta\mathbb{E}U_1$ is the expected claim amount per unit time; the assumption $\mu > 0$ is thus equivalent to the net profit condition $\rho < 1$. The approximation of Embrechts and Veraverbeke (1982) for the probability of ruin in infinite time is

$$\psi(u) = \mathbb{P}(\tau(u) < \infty) \sim \frac{\rho}{1-\rho}\overline{B}_0(u). \tag{1.9}$$

The motivation for looking at (1.7) in particular is twofold. First, it has mathematically convenient features, in particular the property that S is downwards skip-free. This lies behind the classical property of an explicit Wiener–Hopf factorization (Proposition

2.1(a) below) which is important for us as well, but has also other consequences. In particular, we obtain a limit of $S_{[0,\tau(u))}$ in a strong 'local' total variation sense rather than subject to normalizing as in Theorem 1.1, with a form of the limit that makes only sense in the downwards skip-free case (not even for discrete–time random walks). For further studies into downwards skip-freeness in related problems, see Bertoin (1993). The second motivation for (1.7) is from applications: our analysis leads here to the solution of a problem which has long been open, to find the asymptotic $\mathbb{P}^{(u)}$–distribution of $\tau(u)$ and thereby approximations for the finite time ruin probability $\psi(u,T)$ in the heavy–tailed case (Corollary 1.6 below). For the case of light tails, the classical result in this direction is due to Segerdahl (1955), who obtained a CLT for $\tau(u)$ (properly scaled and normalised). As suggested by Theorem 1.1, we will show that for heavy tails there typically exists a deterministic normalising function $a(u)$ such that $\tau(u)/a(u)$ has a weak limit which is either Pareto or exponential.

In order to state our results, we first introduce (or update) our notation. Let $(Z(u), Y(u))$ be a random vector having the $\mathbb{P}^{(u)}$-distribution of $(-S_{\tau(u)-}, S_{\tau(u)})$ and write $(Z,Y) = (Z(0), Y(0))$. Furthermore, recall that $s_0(t) = -\mu t$ for $0 \leqslant t < 1$. We note that it is well-known (Proposition 2.1 below) that the marginal distribution of either of Z, Y is B_0.

Defining

$$\mathbb{P}^{(u,z)}(\cdot) = \mathbb{P}^{(u)}(\cdot \mid Z(u) = z) = \mathbb{P}(\cdot \mid \tau(u) < \infty, Z(u) = z), \tag{1.10}$$

it is easy to see that the $\mathbb{P}^{(u,z)}$–distribution of $Y(u) - u$ is $B^{(u+z)}$, the distribution of the overshoot of a claim over $u + z$,

$$B^{(u+z)}(x) = \frac{B(u+z+x) - B(u+z)}{\overline{B}(u+z)} = 1 - \frac{\overline{B}(u+z+x)}{\overline{B}(u+z)}.$$

(Similar notation is used later for the overshoot distribution corresponding to B_0.) Thus, the structure of $\mathbb{P}^{(u)}$ is completely described by the marginal distribution of $Z(u)$ and the $\mathbb{P}^{(u,z)}$–distribution of $S_{[0,\tau(u))}$, cf. (1.1). Let $\|\cdot\|$ denote the total variation (t.v.) distance and define

$$\delta(z) = \sup\{t > 0 : S_t = -z\}$$

as the time of the last downcrossing of level $-z$ (which is finite w.r.t. $\mathbb{P}$ since the drift is negative and the process downwards skip-free).

Theorem 1.3. *Assume that $B_0 \in \mathscr{S}$. Then as $u \to \infty$:*

(a) $\|\mathbb{P}^{(u)}(Z(u) \in \cdot) - B_0^{(u)}\| \to 0$;

(b) $g(u, Z(u)) \xrightarrow{\mathbb{P}^{(u)}} 0$, *where*

$$g(u,z) = \left\|\mathbb{P}^{(u,z)}\left(S_{[0,\tau(u))} \in \cdot\right) - \mathbb{P}\left(S_{[0,\delta(z))} \in \cdot\right)\right\|.$$

(In (b), one can formally view $S_{[0,\tau(u))}$ as defined in (1.1) as a random element of the space D^* of $D[0,\infty)$ functions with finite lifelength, see e.g. Williams (1979, III.14).) A weaker result of the same form as Theorem 1.1 is

Corollary 1.4. *The conclusion of Theorem* 1.1 *holds under the same conditions as there, provided one replaces* V_α/μ *by* $V_\alpha/(1-\rho)$.

Denote by $W(u) = Z(u) + Y(u)$ the size of the claim causing ruin and let $\psi(u,T)$ and $\psi(u)$ be the ruin events as defined in (1.8).

Corollary 1.5. (a) *If* $\overline{B} \in \mathscr{R}(-\alpha-1)$ *for* $\alpha \in (0,\infty)$, *then*

$$\lim_{u\to\infty} \mathbb{P}^{(u)}\left(\frac{W(u)}{u} > x\right) = (1+\alpha(1-x^{-1}))x^{-\alpha}.$$

(b) *If* $B \in \mathrm{MDA}(\Lambda)$ *and* $B_0 \in \mathscr{S}$, *then*

$$\lim_{u\to\infty} \mathbb{P}^{(u)}\left(\frac{W(u)-u}{a(u)} > x\right) = (1+x)\mathrm{e}^{-x}$$

with $a(u) \sim \mu\overline{B}_0(u)/\overline{B}(u)$.

Corollary 1.6. (a) *If* $\overline{B} \in \mathscr{R}(-\alpha-1)$ *for* $\alpha \in (0,\infty)$, *then*

$$\lim_{u\to\infty} \frac{\psi(u,uT)}{\psi(u)} = 1-(1+(1-\rho)T)^{-\alpha}.$$

(b) *If* $B \in \mathrm{MDA}(\Lambda)$ *and* $B_0 \in \mathscr{S}$, *then*

$$\lim_{u\to\infty} \frac{\psi(u,a(u)T)}{\psi(u)} = 1-\mathrm{e}^{-(1-\rho)T}$$

with $a(u) \sim \mu\overline{B}_0(u)/\overline{B}(u)$.

The rest of the paper contains the proofs of the above results as well as certain corollaries. We start by the proofs for the insurance risk model in Sections 2 and 3, with the conditioned limit theorems for the path being given in Section 2 and the ruin probability approximations in Section 3, which also studies problems such as the asymptotic distributions of the claim leading to ruin and the largest claim before then. The discussion in Section 3 is based upon studies of the asymptotic form of $B_0^{(u)}$ by Balkema and de Haan (1974), see also Geluk and de Haan (1987); an important feature is the classification of B_0 according to extreme value theory. Section 4 then contains the proofs for the discrete time random walk case which basically are just simplifications of arguments from Sections 2 and 3; this is not surprising since the results are somewhat weaker from a mathematical point of view, being given in terms of weak convergence of a suitable normalized version of the process.

We conclude this section by mentioning some further relevant literature. In connection with (1.9), see also Klüppelberg (1988) and Asmussen et al. (1994). Large deviations for the insurance risk model have been studied by Asmussen (1984), Asmussen and Nielsen (1995), Barndorff-Nielsen and Schmidli (1995), Djehiche (1993), Martin-Löf (1983, 1986) and Slud and Hoesman (1989); all these authors assume light tails. Asmussen and Teugels (1997) study the tail of the $\mathbb{P}^{(u)}$-distribution of $\tau(u)$ for a fixed u, considering the case of regular variation only. Klüppelberg & Mikosch

(1996) prove large deviation results for the total claim amount process, where the claim size distribution has regularly varying tail. As for conditioned limit theorems in a heavy-tailed setting, some important references are Durrett (1980) and Anantharam (1988), who look at problems of a different type such as describing $S_{[0,t]}$ given $S_t > t(\varepsilon - \mu)$, assuming regular variation. A general reference in the context is Asmussen (1996).

2. Total variation limits for the insurance risk process

We start the proofs of our results by some exact (non-asymptotic) facts. In order to state the first, it will be convenient to let the arrival process have doubly infinite time such that we can represent it by the marked point process $A = (T_k, U_k)_{k=0,\pm 1,\pm 2,\dots}$ where

$$\cdots < T_{-1} < T_0 < 0 \leqslant T_1 < T_2 < \cdots$$

and $(U_n)_{n\in\mathbb{Z}}$ are iid with common distribution B. We let A_t denote the arrival process prior to t, i.e.

$$A_t = \left(T_{K(t)-k}, U_{K(t)-k}\right)_{k=0,1,\dots},$$

where $K(t) = \sup\{k : T_k \leqslant t\}$.

Recall that $Y = S_{\tau(0)}$, $Z = -S_{\tau(0)-}$. Then $W = Y + Z$ is the size of the claim leading to ruin with initial reserve 0.

Proposition 2.1. *Subject to the probability measure $\mathbb{P}^{(0)} = \mathbb{P}(\cdot|\tau(0) < \infty)$, it holds that:*

(a) *The marginal distributions of Y,Z are both B_0.*

(b) *The distribution of W is the distribution with density x/μ w.r.t. B. Further, the pair (Y,Z) is distributed as $(UW,(1-U)W)$ where U is uniform on $(0,1)$ and independent of W.*

(c) *The conditional distribution of $A_{\tau(0)}$ given (Y,Z) is the same as the unconditional $\mathbb{P}$–distribution of A_0.*

Proof. The statement concerning Y in (a) is classical (Cramér, 1930 or Feller, 1971, XI.4), whereas the one concerning Z in (a) and part (b) were proved by Dufresne and Gerber (1988). Part (c) follows from Theorem 2 of Asmussen and Schmidt (1995) by considering the marked point process (T_k, U_k, A_{T_k}) and noting that the Palm distribution of A_{T_k} is just the unconditional $\mathbb{P}$–distribution of A_0. □

Note that analytically, the content of (b) is that

$$\mathbb{P}(Z \geqslant z, Y \geqslant u) = \frac{1}{\mu_B}\int_{u+z}^{\infty} \overline{B}(y)\,\mathrm{d}y; \tag{2.1}$$

this is the way in which the result is stated in Dufresne and Gerber (1988). A probabilistic interpretation is from renewal theory: in a stationary renewal process with interarrival distribution B, we can think of Z as the backwards recurrence time and of Y as the forward recurrence time at 0 (see e.g. Feller, 1971, XI.4). In particular, Z has marginal distribution B_0 and

$$\mathbb{P}(Y > y \mid Z = z) = \overline{B}(z+y)/\overline{B}(z).$$

From (2.1) we obtain

$$\mathbb{P}(Z > z | Y > u) = \frac{\mu_B^{-1}\int_{u+z}^{\infty}\overline{B}(y)\,\mathrm{d}y}{\mu_B^{-1}\int_{u}^{\infty}\overline{B}(y)\,\mathrm{d}y} = 1 - B_0^{(u)}(z). \tag{2.2}$$

Note that since stationary renewal processes are time–reversible, we can interchange Y and Z in this description.

Proposition 2.1 provides the following 'simulation algorithm' for a single ladder segment $S_{[0,\tau(0))}$ with distribution $\mathbb{P}^{(0)}$: first generate (Y,Z) as in (b), and next run the risk process backwards in time starting from $-Z$; by (c), the time the risk process hits 0 has then the same distribution as the ruin time $\tau(0)$. In particular, letting

$$\sigma(z) = \inf\{t > 0 : R_t = z \mid R_0 = 0\} = \inf\{t > 0 : S_t = -z\},$$

we obtain the following representation for the distribution of $\tau(0)$:

$$\mathbb{P}^{(0)}(\tau(0) \in \cdot) = \int_0^{\infty} \mathbb{P}(\sigma(z) \in \cdot)\mathbb{P}(Z \in \mathrm{d}z) = \int_0^{\infty} \mathbb{P}(\sigma(z) \in \cdot)B_0(\mathrm{d}z). \tag{2.3}$$

An alternative approach to Proposition 2.1(c) is via excursion theory for Markov processes. We thereby view $\{S_t\}$ as a Markov process, allowing any given initial value x. Consider an open subset O of $\mathbb{R}$ and let

$$\overline{O}_+ = \{x \in \mathbb{R} : (x-\delta, x) \subseteq O \text{ for some } \delta > 0\},$$
$$\overline{O}_- = \{x \in \mathbb{R} : (x, x+\delta) \subseteq O \text{ for some } \delta > 0\}$$

(note that $O \subseteq \overline{O}_+$, $O \subseteq \overline{O}_-$ since O is open). We define

$$S_{[0,\zeta)}^{(x,y)} = \left\{S_t^{(x,y)}\right\}_{0 \leqslant t < \zeta},$$

an excursion of length $\zeta = \zeta_S$ in O of $\{S_t\}$, conditioned to start in $x \in \overline{O}_+$ and end in $y \in \overline{O}_-$, as the random element of D^* with distribution

$$\mathbb{P}\left(S_{[0,\zeta)} \in \cdot \mid S_0 = x, S_{\zeta-} = y\right),$$

where $\zeta = \inf\{t > 0 : S_t \in O^{\mathrm{c}}\}$. The following result is a special case of Proposition 2.14 of Fitzsimmons (1987) (see also Kaspi, 1985); to obtain it from that reference, note that the claim surplus process $\{S_t\}$ and the risk process $\{R_t\}$ are in classical duality w.r.t. Lebesgue measure (in the case of Lebesgue reference measure and processes with independent increments, this means just sign reversion). Without risk of ambiguity, use the same letter $\zeta = \zeta_R$ for the excursions of R in O.

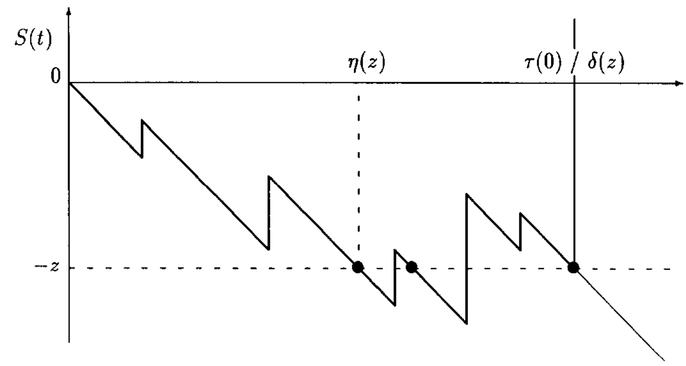

Fig. 1.

Proposition 2.2. *For* $x \in \overline{O}_+$, $y \in \overline{O}_-$,

$$S^{(x,y)}_{[0,\zeta)} = \left\{S^{(x,y)}_t\right\}_{0 \leqslant t < \zeta} \stackrel{\mathscr{D}}{=} \left\{R^{(y,x)}_{(\zeta - t)-}\right\}_{0 \leqslant t < \zeta}.$$

Note that the backwards description in Proposition 2.1(c) can alternatively be viewed as the case $O = (-\infty, 0)$, $x = 0$, $y = z$ of Proposition 2.2. Further special cases occur in the proof of the following lemma. Define $\mathbb{P}^{(\flat)}(\cdot) = \mathbb{P}(\cdot \mid \tau(0) = \infty)$ (see Bertoin, 1993 for a thorough discussion of $\mathbb{P}^{(\flat)}$).

Lemma 2.3. $\mathbb{P}^{(0,z)}(S_{[0,\tau(0))} \in \cdot) = \mathbb{P}^{(\flat)}(S_{[0,\delta(z))} \in \cdot)$.

Proof. We use the path decomposition in Fig. 1.

If we read the path backwards from $(\tau(0), -z)$ to $(0,0)$, this is just $\{R_t\}$ started from $-z$ and stopped when hitting 0. This path starts with a number of excursions (separated by •'s on the Figure) in $O = (-\infty, z) \cup (z, 0)$ with $x = y = -z$; the number is a geometric rv with parameter p_R, the probability that an excursion of R in $O = (-\infty, z) \cup (z, \infty)$ with $x = y = -z$ does not hit 0. The segment from $(\eta(z), -z)$ to $(0,0)$ is an excursion in $O = (-z, 0)$ starting at $y = -z$ and ending in $x = 0$.

Similarly, read the $\mathbb{P}^{(\flat)}$–path of S forwards from $(0,0)$ to $(\delta(z), -z)$. This starts with an excursion of S in $O = (-z, 0))$ starting at $x = 0$ and ending in $y = -z$, followed by a geometric number of excursions in $O = (-\infty, z) \cup (z, 0)$ with $x = y = -z$; the number is a geometric rv with parameter p_S, the probability that an excursion of S in $O = (-\infty, z) \cup (z, \infty)$ with $x = y = -z$ does not hit 0.

The lemma now follows by combining these two descriptions with Proposition 2.2. One needs also that $p_R = p_S$, which follows by noting that the probability that either of R, S will return to $-z$, if started there, is ρ and using Proposition 2.2 once more. □

We now turn to the asymptotics in the limit $u \to \infty$. We shall use repeatedly the following lemma (where Δ denotes the symmetric difference of two sets). Though we have no direct reference for parts (a) and (b), we consider these to be well known (the proofs are straightforward). Part (c) follows from Scheffé's theorem (Billingsley, 1968, p. 224).

Lemma 2.4. (a) *If* $A(u)$, $\tilde{A}(u)$ *are events such that* $\mathbb{P}(A(u)\Delta\tilde{A}(u)) = \mathrm{o}(\mathbb{P}(\tilde{A}(u)))$, *then*

$$\|\mathbb{P}(\cdot \mid A(u)) - \mathbb{P}(\cdot \mid \tilde{A}(u))\| \;\to\; 0,$$

(b) *If* $\mathbb{P}_u$, Q_u *are probability measures such that* $\|\mathbb{P}_u - Q_u\| \to 0$ *and if* $K(\omega, F)$ *is a Markov kernel (conditional probability), then*

$$\left\| \int K(\omega,\cdot)\mathbb{P}_u(\mathrm{d}\omega) - \int K(\omega,\cdot)Q_u(\mathrm{d}\omega) \right\| \to 0,$$

(c) *If* $M(u)$, M *are discrete rv's such that* $\mathbb{P}(M(u) = n) \to \mathbb{P}(M = n)$ *for all* n, *then*

$$\|\mathbb{P}(M(u) \in \cdot) - \mathbb{P}(M \in \cdot)\| \;\to 0.$$

We also use the standard consequence

$$\lim_{u\to\infty} \overline{B}_0^{(u)}(x) = \lim_{u\to\infty} \frac{\overline{B}_0(x+u)}{\overline{B}_0(u)} = 1\,, \quad \forall x \in \mathbb{R}, \tag{2.4}$$

of $B_0 \in \mathscr{S}$. In the following, let $Y_1, Y_2, \ldots$ be i.i.d. with distribution B_0 and let $\otimes$ denote product measure.

Lemma 2.5. *Let* n *be fixed and* $A(u) = \{Y_1 + \cdots + Y_{n-1} \leqslant u,\ Y_1 + \cdots + Y_n > u\}$. *Then*

$$\left\| \mathbb{P}\Big((Y_1,\ldots,Y_{n-1},Y_n - u) \in \cdot \mid A(u)\Big) \;-\; B_0^{\otimes(n-1)} \otimes B_0^{(u)} \right\| \;\to\; 0.$$

Proof. Letting $\tilde{A}(u) = \{Y_n > u\}$, the condition of Lemma 2.4(a) follows easily from the definition of subexponential distributions, which implies that

$$\mathbb{P}\Big(\{Y_1 + \cdots + Y_n > u\}\,\Delta\,\{\max(Y_1,\ldots,Y_n) > u\}\Big) = \mathrm{o}(\overline{B}_0(u)).$$

Now just note that

$$\mathbb{P}\Big((Y_1,\ldots,Y_{n-1},Y_n - u) \in \cdot \mid \tilde{A}(u)\Big) = B_0^{\otimes(n-1)} \otimes B_0^{(u)}. \quad \square$$

Define the ladder epochs by $\tau_+(0) = 0$,

$$\tau_+ = \tau_+(1) = \tau(0) = \inf\{t > 0 : S_t > 0\}\,,$$
$$\tau_+(k+1) = \inf\{t > \tau_+(k) : S_t > S_{\tau_+(k)}\}\,,$$

and let $N(u) = \inf\{n : S_{\tau_+(n)} > u\}$. Then $\tau(u) = \tau_+(N(u))$.

Lemma 2.6. $\mathbb{P}^{(u)}(N(u) = n) \to (1-\rho)\rho^{n-1} \quad \forall n \in \mathbb{N}$.

Proof. From the proof of Lemma 2.5,

$$\mathbb{P}(A(u)) = \mathbb{P}(\tilde{A}(u))(1+o(1)) = \overline{B}_0(u)(1+o(1)).$$

Hence,

$$\mathbb{P}(N(u) = n) = \mathbb{P}(\tau_+(n) < \infty)\,\mathbb{P}(A(u)) = \rho^n \overline{B}_0(u)(1+o(1)).$$

Dividing by $\psi(u)$ and using (1.9), the result follows. □

Note that the situation in Lemma 2.6 is in contrast to the light–tailed case, where it is easy to see (Asmussen, 1982) that $N(u)$ is of the order of magnitude $u/\kappa'(\gamma)$ w.r.t. $\mathbb{P}^{(u)}$.

Lemma 2.7. *For any distribution G in $(0,\infty)$, $\|B_0^{(u)} * G - B_0^{(u)}\| \to 0$.*

Proof. Let

$$b_0^{(u)}(x) = \frac{\overline{B}(x+u)}{\int_u^\infty \overline{B}(y)\,\mathrm{d}y}$$

be the density of $B_0^{(u)}$, $0 \leqslant x < \infty$ ($b_0^{(u)}(x) = 0$, $x < 0$). Then for $A \subseteq (0,\infty)$,

$$\begin{aligned}
\left|B_0^{(u)}(A) - B_0^{(u)} * G(A)\right| &= \left|\int_0^\infty G(\mathrm{d}a)\int_A \left(b_0^{(u)}(x) - b_0^{(u)}(x-a)\right)\mathrm{d}x\right| \\
&\leqslant \int_0^\infty G(\mathrm{d}a)\int_{\mathbb{R}} \left|b_0^{(u)}(x) - b_0^{(u)}(x-a)\right|\mathrm{d}x \\
&= \int_0^\infty G(\mathrm{d}a)\left\{\int_0^a b_0^{(u)}(x)\mathrm{d}x\right. \\
&\qquad \left. + \int_a^\infty \left(b_0^{(u)}(x-a) - b_0^{(u)}(x)\right)\mathrm{d}x\right\} \\
&= 2\int_0^\infty B_0^{(u)}(a)G(\mathrm{d}a),
\end{aligned}$$

where we used the monotonicity of $b_0^{(u)}$ in the third step. The result now follows since the rhs does not depend on A and converges to 0 as $u \to \infty$ by (2.4) and dominated convergence. □

Proof of Theorem 1.3. We use the path decomposition in Fig. 2.
Define the nth ladder segment as

$$\mathscr{L}_n = \{S_{\tau_+(n-1)+t} - S_{\tau_+(n-1)}\}_{0 \leqslant t \leqslant \tau_+(n) - \tau_+(n-1)},$$

identify Y_n with $S_{\tau_+(n)} - S_{\tau_+(n-1)}$ and write $-Z_n = S_{\tau_+(n)-} - S_{\tau_+(n-1)}$. Then

$$Z(u) = Z_{N(u)} - Y_1 - \cdots - Y_{N(u)-1}.$$

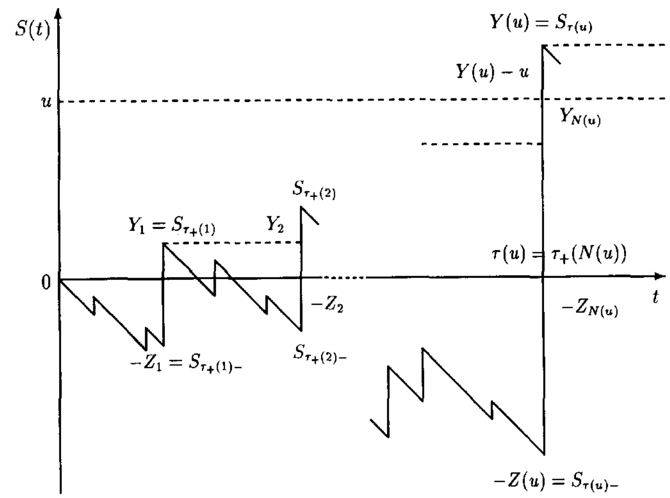

Fig. 2.

Recall the definition of $A(u)$ in Lemma 2.5. Then $\{N(u) = n\} = \{\tau_+(n) < \infty\} \cap A(u)$. By Lemma 2.5 and Eq. (2.2),

$$\left\| \mathbb{P}\Big((Y_1, \ldots, Y_{n-1}, Z_n) \in \cdot \mid A(u)\Big) - B_0^{\otimes(n-1)} \otimes B_0^{(u)} \right\| \to 0.$$

Hence,

$$\left\| \mathbb{P}\Big(Z_n - Y_1 - \cdots - Y_{n-1} \in \cdot \mid A(u)\Big) - \mathbb{P}\Big(Z^{(u)} - Y_1 - \cdots - Y_{n-1} \in \cdot\Big) \right\| \to 0.$$

From this it follows that

$$\left\| \mathbb{P}^{(u)}(Z(u) \in \cdot) - \mathbb{P}\left(Z^{(u)} - Y_1 - \cdots - Y_{N-1} \in \cdot\right) \right\| \to 0, \tag{2.5}$$

where $N, Z^{(u)}, Y_1, Y_2, \ldots$ are independent rv's such that N is geometric with parameter ρ, $Z^{(u)}$ has distribution $B_0^{(u)}$ and the Y_i have distribution B_0. To see this, appeal to Lemma 2.4(b), with $\mathbb{P}_u = \mathbb{P}^{(u)}(N(u) \in \cdot)$, $Q_u = \mathbb{P}(N \in \cdot)$ and K the conditional distribution given $N(u)$; the assumption $\|\mathbb{P}_u - Q_u\| \to 0$ is satisfied because of Lemmas 2.6 and 2.4(c).

Part (a) of the theorem now follows from (2.5) combined with Lemma 2.7. The proof of (b) is almost the same. For $z > 0$, we have the following obvious description of $S_{[0,\delta(z))}$.

(i) The process starts with N ladder segments where $\mathbb{P}(N = n) = (1-\rho)\rho^n$, $n = 0, 1, 2, \ldots$ (Lemma 2.6).

(ii) Given $N = n$, the n ladder segments are iid each distributed as L, i.e. their joint distribution is $\mathbb{P}^{(0)}(L \in \cdot)^{\otimes n}$.

(iii) Given $N = n$, $Y_1 = y_1, \ldots, Y_n = y_n$, the segment of S from $(\tau_+(n), y_1 + \cdots + y_n)$ to $(\delta(z), z)$ has distribution $\mathbb{P}^{(b)}\left(S_{[0,\delta(z+y_1+\cdots+y_n))} \in \cdot\right)$.

For $\mathbb{P}^{(u,z)}(S_{[0,\tau(u))} \in \cdot)$, the description is:

(i) The process starts with $N(u)$ ladder segments where $\mathbb{P}(N(u) = n+1) \to (1-\rho)\rho^n$, $n = 0, 1, \ldots$ (Lemma 2.6).

(ii) Given $N(u) = n+1$, the distribution of $Y_1,\dots,Y_n$) converges in t.v. to $B_0^{\otimes n}$ (Lemma 2.6). Therefore, the distribution of the first n ladder segments converges in t.v. to $\mathbb{P}^{(0)}(L \in \cdot)^{\otimes n}$ (Lemma 2.4(b)).

(iii) Given $N(u) = n+1$, $Y_1 = y_1,\dots,Y_n = y_n$, the segment of S from $(\tau_+(n), y_1 + \cdots + y_n)$ to $(\tau(u), z)$ has distribution $\mathbb{P}^{(b)}\left(S_{[0,\delta(z+y_1+\cdots+y_n))}\right)$ (Lemma 2.3).

Putting these two descriptions together and using Lemma 2.4 as in the proof of part (a) shows that $g(u,z) \to 0$. Hence $g(u,Z(u))I(Z(u)>0) \to 0$ and it only remains to show that $\mathbb{P}^{(u)}(Z(u)>0) \to 1$. But by (2.4), $Z^{(u)} \xrightarrow{\mathbb{P}} \infty$, and hence also $Z(u) \xrightarrow{\mathbb{P}^{(u)}} \infty$ by part (a) of the theorem. □

3. Ruin probability approximations

We start with the result of Theorem 1.3(a) that for $u,x>0$

$$\mathbb{P}^{(u)}(Z^{(u)} \leqslant x) = B_0^{(u)}(x) = \mathbb{P}^{(u)}(Y - u \leqslant x \mid Y > u),$$

where Y has df B_0. Now $B_0 \in \mathscr{S}$ implies that $Z^{(u)} \to \infty$ in distribution, cf. (2.4). This is a rather weak statement, but more information can be obtained if we allow for a scale function $a(u)$, i.e. if we consider $Z^{(u)}/a(u)$. This, besides giving more insight in the asymptotic behaviour of $Z^{(u)}$, also provides a link to extreme value theory. Here $B_0^{(u)}$ is called the *excess distribution* referring to the fact that it is the df of the excess of the rv Y over a threshold u.

We recall some facts from extreme value theory; see e.g. de Haan (1970), Resnick (1987) or in particular in the insurance context, Embrechts et al. (1997). For iid rv's $X, X_1,\dots,X_n$ with df F we say $F \in \mathrm{MDA}(H)$ (F belongs to the maximum domain of attraction of H) if there exist norming constants $c(n)>0$ and $d(n) \in \mathbb{R}$ such that

$$(c(n))^{-1}(\max(X_1,\dots,X_n) - d(n)) \xrightarrow{\mathscr{L}} H, \quad n \to \infty.$$

H is called *extreme value distribution.*

In the present paper, only the case where F has infinite right endpoint, i.e. $F(x) < 1$ for all $x \in \mathbb{R}$ is of interest. Then either

$$H(x) = \Phi_\alpha(x) = \begin{cases} \exp\left\{-x^{-\alpha}\right\}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

for some $\alpha > 0$, or

$$H(x) = \Lambda(x) = \exp\left\{-e^{-x}\right\}, \quad x \in \mathbb{R}.$$

Proposition 3.1. (Balkema and de Haan, 1974; Geluk and de Haan 1987). *Assume that F has infinite right endpoint. Then $F \in \mathrm{MDA}(H)$ if and only if there exists some measurable function $a : \mathbb{R} \to \mathbb{R}^+$ such that*

$$\lim_{u\to\infty} \mathbb{P}\left(\frac{X-u}{a(u)} > x \,\middle|\, X > u\right) = \overline{G}_\alpha(x). \tag{3.1}$$

Here $\alpha < \infty$ corresponds to $H = \Phi_\alpha$ and $\alpha = \infty$ to $H = \Lambda$. The function a can be chosen as $a(u) \sim \int_u^\infty \overline{F}(y)\,\mathrm{d}y/\overline{F}(u)$. □

Recall that G_α is the distribution of the generalised Pareto rv V_α which has been defined in (1.4).

The exponential tail ($\alpha = \infty$) can be considered as the limiting case for $\alpha \to \infty$. Notice that for $\alpha \in (0,\infty)$ G_α is just a reparametrisation of the Pareto distribution.

Now we assume that the claim size distribution B is heavy-tailed in the sense that B_0 is subexponential. In order to apply Proposition 3.1 we need that $B_0 \in \mathrm{MDA}(H)$. Since B_0 has decreasing density $\mu_B^{-1}\overline{B}$, this is equivalent to $B \in \mathrm{MDA}(\widetilde{H})$ for extreme value distributions $H, \widetilde{H}$.

More precisely, $B \in \mathrm{MDA}(\Phi_{\alpha+1}), \alpha > 0$, is equivalent to $B_0 \in \mathrm{MDA}(\Phi_\alpha)$, i.e. $\overline{B}_0 \in \mathscr{R}(-\alpha)$. This is ensured by the monotone density theorem (Bingham et al., 1987, Theorem 1.7.2). Moreover, for $B \in \mathrm{MDA}(\Phi_{\alpha+1}), \alpha > 0$, Karamata's theorem gives

$$a(u) \sim \frac{u}{\alpha}, \quad u \to \infty .$$

$B \in \mathrm{MDA}(\Lambda)$ is equivalent to $B_0 \in \mathrm{MDA}(\Lambda)$; moreover, Proposition 3.1 applies to B and B_0 with the same normalising function a (see e.g. Resnick, 1987, Proposition 1.17). Furthermore, B_0 has the representation

$$\overline{B}_0(u) = \exp\left\{-\int_0^u \frac{1}{a(t)}\,\mathrm{d}t\right\}, \quad u > 0$$

with $a(u) = \mu_B \overline{B}_0(u)/\overline{B}(u)$.

For $B \in \mathrm{MDA}(\Lambda)$ and $B_0 \in \mathscr{S}$, Goldie and Resnick (1988) have shown that $a(u) \to \infty$ as $u \to \infty$. They furthermore derive conditions on the function a such that $B_0 \in \mathscr{S}$. The following applies to the claim size distributions we have in mind.

Proposition 3.2. ((Goldie and Resnick, 1988)). *Let B be a distribution function with $a(u) = \mu_B \overline{B}_0(u)/\overline{B}(u) \to \infty, a'(u) \to 0$ as $u \to \infty$ and a is eventually non-decreasing. Furthermore, assume that for some $t > 1$*

$$\liminf_{u\to\infty} \frac{a(tu)}{a(u)} > 1$$

holds. Then $B \in \mathrm{MDA}(\Lambda)$ and $B_0 \in \mathscr{S}$.

We show that $Z^{(u)}/a(u)$ converges in $\mathbb{P}^{(u)}$-distribution to a generalised Pareto distribution. Obviously, the function $a(u)$ is unique only up to asymptotic equivalence. Hence, any function $a(u) \sim \mu_B \overline{B}_0(u)/\overline{B}(u)$ may be chosen as normalising function.

Theorem 3.3. (a) *If $\overline{B} \in \mathscr{R}(-\alpha - 1)$ for $\alpha \in (0,\infty)$, then*

$$\alpha\frac{Z^{(u)}}{u} \to V_\alpha$$

in $\mathbb{P}^{(u)}$–distribution.

(b) *If* $B \in \mathrm{MDA}(\Lambda)$ *and* $B_0 \in \mathscr{S}$, *then*

$$\frac{Z^{(u)}}{a(u)} \to V_\infty$$

in $\mathbb{P}^{(u)}$*–distribution, where* $a(u) \sim \mu \overline{B}_0(u)/\overline{B}(u) \to \infty$.

Proof. (a) Assume that

$$\overline{B}(u) = u^{-\alpha-1}L(u), \quad u > 0 \quad (\alpha \in (0,\infty),\ L \text{ slowly varying}).$$

Then by Karamata's theorem

$$\overline{B}_0(u) \sim \frac{1}{\mu_B \alpha} u^{-\alpha} L(u), \quad u \to \infty .$$

Furthermore, $B_0 \in \mathscr{S}$ and $a(u) \sim u/\alpha$ as $u \to \infty$. Regular variation implies

$$\mathbb{P}^{(u)}\left(\frac{\alpha Z^{(u)}}{u} > x\right) = \overline{B}_0^{(u)}\left(\frac{ux}{\alpha}\right) = \frac{\overline{B}_0(u(1+x/\alpha))}{\overline{B}_0(u)} \to \left(1+\frac{x}{\alpha}\right)^{-\alpha} .$$

(b) $B \in \mathrm{MDA}(\Lambda)$ implies $B_0 \in \mathrm{MDA}(\Lambda)$ with the same normalising function a. Hence, by Proposition 3.1

$$\mathbb{P}^{(u)}\left(\frac{Z^{(u)}}{a(u)} > x\right) = \overline{B}_0^{(u)}(a(u)x) = \frac{\overline{B}_0(a(u)x+u)}{\overline{B}_0(u)}$$
$$= \mathbb{P}\left(\frac{Y-u}{a(u)} > x \,\middle|\, Y > u\right) \to \mathrm{e}^{-x} . \qquad \square$$

Example 3.4. The Pareto model

$$\overline{B}(x) = \left(1+\frac{x}{\theta}\right)^{-\alpha-1}, \quad x > 0, \quad (\theta > 0, \alpha > 0).$$

Then

$$\overline{B}_0(x) = \frac{1}{\mu_B \alpha}\left(1+\frac{x}{\theta}\right)^{-\alpha}, \quad x > 0 .$$

This implies

$$\overline{B}_0^{(u)}(x) = \frac{\overline{B}_0(x+u)}{\overline{B}_0(u)} = \left(1+\frac{x}{\theta+u}\right)^{-\alpha}, \quad x > 0,$$

i.e. $\overline{B}_0^{(u)}$ is again a Pareto distribution with index α and the parameter θ has been transformed into $\theta+u$. Hence immediately, we see that by scaling with $a(u) = (\theta+u)/\alpha$

$$\overline{B}_0^{(u)}(a(u)x) = \overline{B}_0^{(u)}\left(\frac{\theta+u}{\alpha}x\right) = \left(1+\frac{x}{\alpha}\right)^{-\alpha}, \quad x > 0,$$

which is of the form (1.4) for all u. This means that

$$\frac{\alpha Z^{(u)}}{\theta+u} \stackrel{\mathscr{D}}{=} V_\alpha \quad \forall u > 0 .$$

Example 3.5. The Weibull model

$$\overline{B}(x) = \exp\{-x^\alpha\}, \quad x > 0 \quad (\alpha \in (0,1)).$$

Then by Proposition 3.2, $B \in \mathrm{MDA}(\Lambda)$ and $B_0 \in \mathscr{S}$. By partial integration,

$$\overline{B}_0(x) \sim \frac{1}{\mu_B \alpha} x^{1-\alpha} \exp\{-x^\alpha\}, \quad x \to \infty,$$

and hence

$$a(u) = \mu_B \frac{\overline{B}_0(u)}{\overline{B}(u)} \sim \frac{1}{\alpha} u^{1-\alpha}, \quad u \to \infty.$$

Then we obtain

$$\frac{\alpha Z^{(u)}}{u^{1-\alpha}} \to V_\infty$$

in $\mathbb{P}^{(u)}$-distribution.

Example 3.6. The lognormal model

$$B(x) = N\left(\frac{\ln x - b}{a}\right), \quad x > 0 \quad (a > 0, b > 0),$$

where N denotes the standard normal distribution. It is well-known that the lognormal distribution is in $\mathrm{MDA}(\Lambda)$ (see e.g. Resnick, 1987) and also that $B_0 \in \mathscr{S}$ (see e.g. Klüppelberg, 1987). B is absolutely continuous with density B' which can be expressed in terms of the standard normal density n. Using l'Hospital and Mill's ratio,

$$a(u) \sim \frac{\overline{B}(u)}{B'(u)} = \frac{\overline{N}((\ln u - b)/a)\, a\, u}{n((\ln u - b)/a)} \sim \frac{a^2 u}{\ln u - b}.$$

Hence,

$$\frac{\ln u\, Z^{(u)}}{a^2 u} \to V_\infty$$

in $\mathbb{P}^{(u)}$-distribution.

Proof of Corollary 1.4. Consider the joint distribution of

$$\left(\frac{Z^{(u)}}{a(u)}, \frac{Y(u) - u}{a(u)}\right).$$

First assume that $\overline{B} \in \mathscr{R}(-\alpha - 1)$ for $\alpha > 0$. Then by (1.10) and (2.5) we obtain

$$\mathbb{P}^{(u)}\left(\frac{\alpha(Y(u) - u)}{u} > t \,\middle|\, \frac{\alpha Z^{(u)}}{u} = v\right) = \mathbb{P}^{(u)}\left(Y(u) - u > \frac{tu}{\alpha} \,\middle|\, Z^{(u)} = \frac{vu}{\alpha}\right)$$

$$= \mathbb{P}^{(u, vu/\alpha)}\left(Y(u) - u > \frac{tu}{\alpha}\right)(1 + o(1))$$

$$= \frac{\overline{B}\left(u\left(1+\frac{v+t}{\alpha}\right)\right)}{\overline{B}\left(u\left(1+\frac{v}{\alpha}\right)\right)}(1+o(1))$$

$$\to \frac{\left(1+\frac{v+t}{\alpha}\right)^{-\alpha-1}}{\left(1+\frac{v}{\alpha}\right)^{-\alpha-1}}.$$

Then we obtain for the joint distribution

$$\mathbb{P}^{(u)}\left(\frac{\alpha(Y(u)-u)}{u} > t, \frac{\alpha Z^{(u)}}{u} > v\right) \to \int_v^\infty \frac{\left(1+\frac{s+t}{\alpha}\right)^{-\alpha-1}}{\left(1+\frac{s}{\alpha}\right)^{-\alpha-1}}\, \mathrm{d}G_\alpha(s)$$

$$= \left(1+\frac{v+t}{\alpha}\right)^{-\alpha}$$

$$= \overline{G}_\alpha(v+t).$$

Now assume that $B \in \mathrm{MDA}(\Lambda)$ and $B_0 \in \mathscr{S}$. Then by Proposition 3.1,

$$\mathbb{P}^{(u)}\left(\frac{Y(u)-u}{a(u)} > t \;\middle|\; \frac{Z^{(u)}}{a(u)} = v\right) = \mathbb{P}^{(u,va(u))}(Y(u)-u > ta(u))$$

$$= \frac{\overline{B}(u+(v+t)a(u))}{\overline{B}(u+va(u))}$$

$$\to \mathrm{e}^{-t}.$$

This implies

$$\mathbb{P}^{(u)}\left(\frac{Y(u)-u}{a(u)} > t\,, \frac{Z^{(u)}}{a(u)} > v\right) \to \mathrm{e}^{-(t+v)}.$$

To get the four–variate joint limit, just invoke the law of large numbers in the form $S_t/t \to 1-\rho$ uniformly on bounded intervals, which in particular implies $\delta(z)/z \to (1-\rho)^{-1}$. □

Using Corollary 1.4, the proof of Corollary 1.6 on finite time ruin probability is immediate.

Proof of Corollary 1.5. Immediately by Corollary 1.4 we obtain for all $\alpha \in (0,\infty]$,

$$\frac{W(u)-u}{a(u)} \to V_\alpha + T_\alpha$$

in $\mathbb{P}^{(u)}$-distribution. V_α and T_α are absolutely continuous with joint density

$$f_{V_\alpha,T_\alpha}(v,t) = \begin{cases} \dfrac{\alpha+1}{\alpha}\left(1+\dfrac{v+t}{\alpha}\right)^{-\alpha-2} & \text{if } \alpha < \infty, \\ \mathrm{e}^{-(v+t)} & \text{if } \alpha = \infty. \end{cases}$$

From this we obtain the density of $V_\alpha + T_\alpha$ as

$$f_{V_\alpha+T_\alpha}(x) = \begin{cases} \dfrac{\alpha+1}{\alpha} x \left(1+\dfrac{x}{\alpha}\right)^{-\alpha-2} & \text{if } \alpha < \infty, \\ x\mathrm{e}^{-x} & \text{if } \alpha = \infty. \end{cases}$$

(b) follows immediately by integration as (a) does by setting $a(u) = u/\alpha$. □

We now investigate the asymptotic distribution of the largest claim before the one leading to ruin in order to substantiate that it actually is smaller. Assume that $B \in \mathrm{MDA}(H)$. Denote

$$M(u) = \max_{1 \leqslant i \leqslant N(\tau(u))} U_i,$$

where $N(\tau(u)) = \sup\{n \in \mathbb{N} : T_n < \tau(u)\}$. Furthermore, let the functions $c(t)$ and $d(t)$ for $t > 0$ be the usual interpolations of the sequences $(c(n))$ and $(d(n))$. Then the following result holds.

Theorem 3.7. (a) *If $\overline{B} \in \mathscr{R}(-\alpha-1)$ for $\alpha \in (0,\infty)$, then*

$$\lim_{u\to\infty} \mathbb{P}^{(u)}(u^{-1/\alpha}M(u) \leqslant x) = \int_0^\infty \exp\left\{-\frac{\beta}{(\alpha-1)(1-\rho)}x^{-\alpha}z\right\}\left(1+\frac{z}{\alpha}\right)^{-\alpha-1} \mathrm{d}z, \quad x \geqslant 0.$$

(b) *If $B \in \mathrm{MDA}(\Lambda)$ and $B_0 \in \mathscr{S}$, then*

$$\lim_{u\to\infty} \mathbb{P}^{(u)}\left(\frac{M(u)-d(a(u))}{c(a(u))} \leqslant x\right) = \left(1+\frac{\beta}{1-\rho}\mathrm{e}^{-x}\right)^{-1}, \quad x \in \mathbb{R}.$$

Proof. By Corollary 1.4 we have for $\alpha \in (0,\infty]$,

$$\frac{N(\tau(u))}{a(u)} = \frac{N(\tau(u))}{\tau(u)}\frac{\tau(u)}{a(u)} \to \frac{\beta}{1-\rho}V_\alpha$$

in $\mathbb{P}^{(u)}$–distribution, where β is the intensity of the claim arrival process. We consider

$$M_n = \max\{U_1, \ldots, U_n\},$$

where (U_n) are the claim sizes with df B. The normalising function $a(u)$ can in the case of (a) be chosen as $u/(\alpha-1)$, in the case of (b) it can be chosen to be the same as for B_0. This implies in the case of (a) with $a(u) = u/(\alpha-1)$ and $\gamma = ((\alpha-1)(1-\rho))^{-1}\beta$,

$$\begin{aligned}
&\lim_{u\to\infty} \mathbb{P}u(u^{-1/\alpha}M(u) \leqslant x) \\
&= \lim_{u\to\infty}\int_0^\infty \mathbb{P}\left(M(u) \leqslant u^{1/\alpha}x \,\middle|\, \frac{\tau(u)}{a(u)} = z\right) \mathrm{d}G_\alpha(z) \\
&= \lim_{u\to\infty}\int_0^\infty \mathbb{P}\left(\max_{1\leqslant i\leqslant N(\tau(u))} U_i \leqslant u^{1/\alpha}x \,\middle|\, \frac{\tau(u)}{a(u)} = z\right) \mathrm{d}G_\alpha(z) \\
&= \lim_{u\to\infty}\int_0^\infty \mathbb{P}\left(M_{[\gamma z u]} \leqslant u^{1/\alpha}x\right) \mathrm{d}G_\alpha(z),
\end{aligned}$$

where $[y]$ denotes the integer part of y. Set $n = [\gamma z u]$, then the rhs is equal to

$$\lim_{n\to\infty} \int_0^\infty \mathbb{P}\left(M_n \leqslant \left(\frac{n}{\gamma z}\right)^{1/\alpha} x\right) \mathrm{d}G_\alpha(z)$$

$$= \int_0^\infty \Phi_{\alpha-1}\left(x(\gamma z)^{-1/\alpha}\right) \mathrm{d}G_\alpha(z)$$

$$= \int_0^\infty \exp\left\{-\gamma z x^{-\alpha}\right\} \left(1+\frac{z}{\alpha}\right)^{-\alpha-1} \mathrm{d}z .$$

In the case of (b) we obtain with $\gamma = (1-\rho)^{-1}\beta$ as above

$$\lim_{u\to\infty} \mathbb{P}^{(u)}\Big((M(u) - d(a(u)))/c(a(u)) \leqslant x\Big)$$

$$= \lim_{u\to\infty} \int_0^\infty \mathbb{P}\left(M_{[\gamma z a(u)]} \leqslant c(a(u))x + d(a(u))\right) \mathrm{d}G_\infty(z)$$

$$= \lim_{n\to\infty} \int_0^\infty \mathbb{P}\left(M_n \leqslant c\left(n/(\gamma z)\right)x + d\left(n/(\gamma z)\right)\right) \mathrm{d}G_\infty(z)$$

$$= \lim_{n\to\infty} \int_0^\infty \mathbb{P}\left(\frac{M_n - d(n)}{c(n)} \leqslant \frac{c\left(n/(\gamma z)\right)}{c(n)}x + \frac{d\left(n/(\gamma z)\right) - d(n)}{c(n)}\right) \mathrm{d}G_\infty(z).$$

Now $c(t)$ is a slowly varying function, hence $c\left(n/(\gamma z)\right)/c(n) \to 1$ as $n\to\infty$. Furthermore, by Proposition 3.1,

$$B \in \Lambda \iff \lim_{n\to\infty} \frac{P(U > u + xa(u))}{P(U > u)} = \mathrm{e}^{-x}, \quad x \in \mathbb{R}.$$

Since $\overline{B}(d(t)) \sim t^{-1}$ and $c(t) \sim d(a(t))$, Bingham et al. (1987, Theorem 3.10.4) (see also Geluk and de Haan, 1987) yields

$$\frac{d\left(\frac{n}{\gamma z}\right) - d(n)}{c(n)} \to -\ln \gamma z .$$

By continuity the rhs of the limit above is equal to

$$\lim_{n\to\infty} \int_0^\infty \mathbb{P}\big((M_n - d(n))/c(n) \leqslant x - \ln\gamma z\big)\mathrm{e}^{-z}\mathrm{d}z$$

$$= \int_0^\infty \exp\left\{-\mathrm{e}^{-x}\gamma z\right\}\mathrm{e}^{-z}\mathrm{d}z = \left(1+\gamma\mathrm{e}^{-x}\right)^{-1}, \quad x\in\mathbb{R}. \quad \square$$

Remark 3.8. The analogue of Theorem 3.7 for the light-tailed case does not seem to be stated in the literature but follows easily along the lines of Asmussen (1982): $M(u)$ has the same asymptotic properties w.r.t. $\mathbb{P}^{(u)}$ as $\max_{i\leqslant[ku]} U_i$ w.r.t. the exponentially changed measure (1.2) for $k = 1/\kappa'(\gamma)$ (cf. Introduction).

4. Proofs for the random walk case

When unambiguous, all notation introduced for the risk process will be used also for the random walk (say τ_+, $Z = -S_{\tau_+-1}$, $\tau(u)$, $Z(u) = -S_{\tau(u)-1}$ etc.).

We let G_+, G_- denote the strict ascending (defective), resp. weak descending (proper), ladder height distribution and $-\nu_-$ the mean of G_-. It is well known (Asmussen, 1987, p. 169) that

$$\mathbb{P}(\tau_+ < \infty) \;=\; \|G_+\| = 1 - \frac{1}{\mathbb{E}\tau_-} = 1 + \frac{\mu}{\nu_-}.$$

A main difficulty when considering random walks which are not skip-free in one direction is the fact that the Wiener–Hopf factorization is not explicit, i.e. G_+, G_-, etc., cannot in general be found in closed form. However, many properties remain valid asymptotically; e.g. compare Proposition 2.1(a):

Lemma 4.1. *For large y it holds that*

$$\mathbb{P}(Y \geqslant y) \sim \frac{1}{\nu_-}\int_y^\infty \overline{F}(x)\,\mathrm{d}x, \qquad \mathbb{P}(Z \geqslant y) \sim \frac{1}{\nu_-}\int_y^\infty \overline{F}(x)\,\mathrm{d}x.$$

Here the statement on Y is well known (Veraverbeke, 1977) and the key in the proof of (1.3). The statement on Z follows along the lines of the proof of Proposition 4.2 below (Lemma 4.1 is not used in the following but stated because of its independent interest).

Proposition 4.2. $\mathbb{P}^{(0)}(Z \geqslant za(u) \mid Y > u) \to \overline{G}_\alpha(z)$.

Proof. The fundamental tool is the representation

$$G_+(A) = \int_{-\infty}^0 F(A-x)R(\mathrm{d}x), \qquad A \subseteq (0,\infty), \tag{4.1}$$

of G_+ where

$$R(B) = \mathbb{E}\sum_{n=0}^{\tau_+-1} I(S_n \in B)$$

is the pre-τ_+ occupation measure which in turn can be identified with the renewal measure $U_- = \sum_0^\infty G_-^{*n}$ (see Feller, 1971, Ch. XII; Asmussen, 1987, Ch. VII, 1989). In the following, we assume w.l.o.g. that $\overline{B}(x) = \overline{F}(x)/\overline{F}(0)$, $x > 0$. Defining

$$c(u,h) = \sup_{y \leqslant -za(u)} R([y-h,y]) = \sup_{y \leqslant -za(u)} U_-([y-h,y]),$$

it follows from (4.1) by conditioning upon S_{τ_+-1} that

$$\mathbb{P}(Y > u,\, Z > za(u),\, \tau_+ < \infty) = \int_{-\infty}^{-za(u)} \overline{F}(u-x)R(\mathrm{d}x)$$

$$
\begin{aligned}
&= \overline{F}(0)\int_{-\infty}^{-za(u)} \overline{B}(u-x)R(\mathrm{d}x)\\
&\le \frac{c(u,h)\overline{F}(0)}{h} h \sum_{n=0}^{\infty} \overline{B}(u+za(u)+nh)\\
&= \frac{c(u,h)\overline{F}(0)}{h}\left\{\int_{u+za(u)}^{\infty} \overline{B}(x)\,\mathrm{d}x + \varepsilon(u,h)\right\},
\end{aligned}
$$

where $|\varepsilon(u,h)| \leqslant h\overline{B}(u+za(u))$. Since $c(u,h) \to h/v_-$ by Blackwell's renewal theorem, it follows by invoking the standard estimate $\overline{B}(x) = \mathrm{o}(\overline{B}_0(x))$ (see Section 3) that

$$
\begin{aligned}
\limsup_{u\to\infty} \frac{\mathbb{P}(Y>u,\, Z>za(u),\, \tau_+<\infty)}{\overline{B}_0(u)} &\le \frac{\overline{F}(0)}{v_-}\limsup_{u\to\infty} \frac{\int_{u+za(u)}^{\infty}\overline{B}(x)\,\mathrm{d}x}{\overline{B}_0(u)}\\
&= \frac{\overline{F}(0)\mu_B}{v_-}\limsup_{u\to\infty}\frac{\overline{B}_0(u+za(u))}{\overline{B}_0(u)}\\
&= \frac{\overline{F}(0)\mu_B\overline{G}_\alpha(z)}{v_-}.
\end{aligned}
$$

Combining this with a similar lower bound gives

$$
\frac{\mathbb{P}(Y>u,\, Z>za(u),\, \tau_+<\infty)}{\overline{B}_0(u)} \to \frac{\overline{F}(0)\mu_B\overline{G}_\alpha(z)}{v_-}.
$$

Taking $z=0$ and dividing, the result follows. □

Proof of Theorem 1.1. Given Proposition 4.2, the proof is basically just an easier version of the proof of Theorem 1.3. As there, we get

$$
\left\|\mathbb{P}^{(u)}(Z(u)\in\cdot) - \mathbb{P}\left(Z^{(u)} - Y_1 - \cdots - Y_{N-1}\in\cdot\right)\right\| \to 0,
$$

where $N, Z^{(u)}, Y_1, Y_2, \ldots,$ are independent rv's such that N is geometric with parameter $\|G_+\|$, $Z^{(u)}$ has the conditional distribution of Z given $Y>u$ and the Y_i have distribution $G_+/\|G_+\|$. From this and Proposition 4.2, it follows immediately that $Z(u)/a(u) \to V_\alpha$ in $\mathbb{P}^{(u)}$–distribution. For the following, we also note that

$$
\tau_+(N(u)-1) = \mathrm{O}(1), \tag{4.2}
$$

$$
\sup_{0\leqslant t\leqslant 1} S_{t\tau_+(N(u)-1)} = \mathrm{O}(1). \tag{4.3}
$$

We next need to invoke the discrete time analogue of Proposition 2.2. The dual process is the sign–reversed random walk $\{-S_n\}$, and $-S_n/n \to \mu$ combined with $Z(u)\to\infty$ in $\mathbb{P}^{(u)}$–distribution therefore yields

$$
\tau_+(N(u)) - \tau_+(N(u)-1) \sim \frac{Z_1^{(u)}}{\mu}, \tag{4.4}
$$

$$
\frac{S_{t(\tau_+(N(u))-1)} - S_{\tau_+(N(u)-1)}}{\tau_+(N(u)) - \tau_+(N(u)-1)} \sim -\mu t, \tag{4.5}
$$

where $Z_1^{(u)} = S_{\tau_+(N(u)-1)} - S_{\tau_+(N(u))-1} \xrightarrow{\mathbb{P}^{(u)}} \infty$, hence $Z(u)/Z_1^{(u)} \xrightarrow{\mathbb{P}^{(u)}} 1$ by (4.2). Combining (4.2) and (4.4) then yields $\tau(u)/Z(u) = \tau_+(N(u))/Z(u) \to 1/\mu$. That $(Y(u)-u)/a(u) \to T_\alpha$ follows by conditioning upon $Z(u)$ as in Section 3, and finally

$$\left\{ \frac{S_{\lfloor t\tau(u) \rfloor}}{\tau(u)} \right\}_{0 \leqslant t < 1} \to s_0$$

follows by combining (4.3) and (4.5). □

Proof of Theorem 1.2. Define $\sigma(z) = \inf\{n : -S_n \geqslant z\}$. Then by (4.2),

$$F_{\tau(u)}(x) \sim \frac{1}{\tau_+(N(u)) - \tau_+(N(u)-1)} \sum_{n=\tau_+(N(u)-1)+1}^{\tau_+(N(u))-1} I(X_n \leqslant x)$$
$$= F_{\sigma(Z_1^{(u)})}(x),$$

where the last equality is in $\mathbb{P}^{(u)}$–distribution. Here $\sigma(Z_1^{(u)}) \to \infty$ (because $Z_1^{(u)} \to \infty$ and $\sigma(z)/z \to 1/\mu$) and hence $F_{\tau(u)}(x) \xrightarrow{\mathbb{P}^{(u)}} F(x)$ because $F_n(x) \xrightarrow{\text{a.s.}} F(x)$. It is standard that the truth of this for each x implies uniformity in x. □

Acknowledgements

We started this project during the 'Conference on Risk Theory' at the Research Institute in Oberwolfach in September 1994. We take pleasure to thank the Institute for its excellent hospitality. We are furthermore indebted to Haya Kaspi for pointing out some relevant references.

References

V. Anantharam, How large delays build up in a GI/GI/1 queue, Queueing Systems 5 (1988) 345–368.
S. Asmussen, Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the GI/G/1 queue, Adv. Appl. Probab. 14 (1982) 143–170.
S. Asmussen, Approximations for the probability of ruin within finite time, Scand. Act. J. (1984) 31–57; ibid. (1985) 57.
S. Asmussen, Applied Probability and Queues (Wiley, Chichester, 1987).
S. Asmussen, Aspects of matrix Wiener-Hopf factorisation in applied probability, Math. Scientist 14 (1989) 101–116.
S. Asmussen, Ruin Probabilities (World Scientific Singapore, 1997), to appear.
S. Asmussen, L. Fløe Henriksen and C. Klüppelberg, Large claims approximations for risk processes in a Markovian environment, Stochastic Processes Appl. 54 (1994) 29–43.
S. Asmussen and H.M. Nielsen, Ruin probabilities via local adjustment coefficients, J. Appl. Probab. 33 (1995) 736–755.
S. Asmussen and V. Schmidt, Ladder height distributions with marks, Stochastic Processes Appl. 58 (1995) 105–119.
S. Asmussen and J.L. Teugels, Convergence rates for M/G/1 queues and ruin problems with heavy tails, J. Appl. Probab. 35 (1997).
A.A. Balkema and L. de Haan, Residual life-time at great age, Ann. Probab. 2 (1974) 792–804.
O. Bardorff-Nielsen and H. Schmidli, Saddlepoint approximations for the probability of ruin in finite time, Scand. Actuarial J. (1995) 169–186.

J. Bertoin, Splitting at the infimum and excursions in half-lines for random walks and Lévy processes, Stochastic Processes Appl. 47 (1994) 17–35.
P. Billingsley, Convergence of Probability Measures (Wiley, New York, 1968).
N.H. Bingham, C.M. Goldie and J.L. Teugels, Regular Variation (Cambridge University Press, Cambridge, 1987).
J.A. Bucklew, Large Deviation Techniques in Decision, Simulation and Estimation (Wiley, New York, 1990).
H. Cramér, On the Mathematical Theory of Risk, in: A. Martin-Löf, ed., Skandia Jubilee Volume, Stockholm. Reprinted in H. Cramér, Collected Works (Springer, Berlin 1994).
A. Dembo and O. Zeitouni, Large Deviations. Techniques and Applications (Jones and Bartlett, Boston, 1993).
J.-D. Deuschel and D.W. Stroock, Large Deviations (Academic Press, Duluth, London, 1989).
B. Djehiche, A large deviation estimate for ruin probabilities, Scand. Act. J. (1993) 42–59.
F. Dufresne and H.U. Gerber, The surpluses immediately before and at ruin, and the amount of claim causing ruin, Insurance: Math. Economics 7 (1988) 193–199.
R. Durrett, Conditioned limit theorems for random walks with negative drift, Z. Wahrscheinlichkeitsth. Verw. Geb. 52 (1980) 277–287.
P. Embrechts and C.M. Goldie, On closure and factorization properties of subexponential and related distributions, J. Austral. Math. Soc. A 29 (1980) 243–256.
P. Embrechts and N. Veraverbeke, Estimates for the probability of ruin with special emphasis on the possibility of large claims, Insurance: Math. Economics 1 (1982) 55–72.
P. Embrechts, C. Klüppelberg and T. Mikosch, Modelling Extremal Events for Insurance and Finance, (Springer, Berlin, 1997, to appear).
W. Feller, An Introduction to Probability Theory and its Applications II (Wiley, New York, 2nd ed., 1971).
P.J. Fitzsimmons, On the excursions of Markov processes in classical duality, Probab. Theory Related Fields 75 (1987) 159–178.
J.L. Geluk and L. de Haan, Regular Variation, Extensions and Tauberian Theorems, CWI Tract 40 (CWI, Amsterdam 1987).
C.M. Goldie and S.I. Resnick, Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution, Adv. Appl. Probab. 20 (1988) 706–718.
L. de Haan, On Regular Variation and Its Application to Weak Convergence of Sample Extremes, Mathematical Center Tract 32 (Mathematisch Centrum, Amsterdam, 1970).
H. Kaspi, Excursions laws of Markov processes in classical duality, Ann. Probab. 13 (1985) 492–518.
C. Klüppelberg, Estimation of ruin probabilities by means of hazard rates. Insurance: Math. Economics 8 (1987) 279–285.
C. Klüppelberg, Subexponential distributions and integrated tails, J. Appl. Probab. 25 (1988) 132–141.
C. Klüppelberg and T. Mikosch, Large deviations of heavy-tailed random sums with applications in insurance and finance, J. Appl. Probab. 34, to appear.
A. Martin-Löf, Entropy estimates for ruin probabiliites, in: A. Gut and J. Holst, eds., Probability and Mathematical Statistics (1983) 29–39.
A. Martin-Löf, Entropy, a useful concept in risk theory, Scand. Act. J. (1986) 223–235.
S.I. Resnick, Extreme Values, Regular Variation, and Point Processes (Springer, New York, 1987).
C.-O. Segerdahl, When does ruin occur in the collective theory of risk? Skand. Aktuar Tidsskr. (1955) 22–36.
E. Slud and C. Hoesman, Moderate- and large-deviation probabilities in actuarial risk theory, Adv. Appl. Probab. 21 (1989) 725–741.
N. Veraverbeke, Asymptotic behaviour of Wiener-Hopf factors of a random walk, Stochastic Processes Appl. 5 (1977) 27–37.
D. Williams, Diffusions, Markov Processes, and Martingales (Wiley, Chichester, 1979).