VERAVERBEKE’S THEOREM AT LARGE

ON THE MAXIMUM OF SOME PROCESSES
WITH NEGATIVE DRIFT
AND HEAVY TAIL INNOVATIONS

Ph. Barbe(1) and W.P. McCormick(2)
(1)CNRS, France, (2)University of Georgia

Abstract. Veraverbeke’s (1977) theorem relates the tail of the distribution of the supremum of a random walk with negative drift to the tail of the distribution of its increments, or equivalently, the probability that a centered random walk with heavy-tail increments hits a moving linear boundary. We study similar problems for more general processes. In particular, we derive an analogue of Veraverbeke’s theorem for fractional integrated ARMA models without prehistoric influence, when the innovations have regularly varying tails. Furthermore, we prove some limit theorems for the trajectory of the process, conditionally on a large maximum. Those results are obtained by using a general scheme of proof which we present in some detail and should be of value in other related problems.

AMS 2000 Subject Classifications: Primary: 60G50; Secondary: 60F99, 60G99, 60K30, 62P05, 62M10, 26A12, 26A33.

Keywords: maximum of random walk, heavy tail, fractional ARIMA process, long range dependence, boundary crossing probability, nonlinear renewal theory.
Contents

1. Introduction

2. \((g, F)\)-processes, their maximum and sample paths
   2.1. \((g, F)\)-processes and their maximum
   2.2. Removing the positivity of the coefficients
   2.3. Typical trajectories leading to a large maximum
   2.4. Examples
   2.5. Note on the quantiles of the maximum of \((g, F)\)-processes
   2.6. Concluding remarks

3. Veraverbeke’s theorem at large
   3.1. The single large jump heuristic
   3.2. From the heuristic to a theorem
   3.3. Analysis of the paths leading to a large maximum

4. A large deviation inequality and a Karamata type theorem
   4.1. A large deviation inequality
   4.2. A Karamata type theorem

5. Some asymptotic analysis related to analytic functions
   5.1. Preliminaries
   5.2. The functions \(\Psi_n\) and their inverses
   5.3. Approximation of \(\Psi_n^{-1}\) when \((g_n)_{n\geq 0}\) tends to 0
   5.4. Asymptotic analysis of \(\Psi_n^{-1}\) when \((g_n)_{n\geq 0}\) tends to infinity
   5.5. Approximation of \(\Psi_n^{-1}\) when \((g_n)_{n\geq 0}\) has a positive and finite limit

6. Proof of the results of section 2
   6.1. Proof of Theorem 2.1.1 – upper bound
   6.2. Proof of Theorem 2.1.1 – lower bound
   6.3. Proof of Theorem 2.1.2 – upper bound
   6.4. Proof of Theorem 2.1.2 – lower bound
   6.5. Proof of Theorem 2.1.3
   6.6. Proof of Theorem 2.2.1
   6.7. Proof of Theorem 2.3.1

References
1. Introduction. Veraverbeke’s (1977) theorem relates the tail behavior of the maximum of some random walks with negative drift to the tail behavior of their increments. The purpose of this paper is to show that such relation holds for a much larger class of discrete time stochastic processes which encompass some nonstationary FARIMA ones.

Before going further, let us recall Veraverbeke’s theorem. Let \( X_i, i \geq 1, \) be a sequence of independent and identically distributed random variables, having a negative mean \( \mu. \) Define the random walk \( S_n \) by \( S_0 = 0 \) and \( S_n = S_{n-1} + X_n \) for all positive \( n. \) Since the increments \( X_i \) have negative mean, the maximum of the walk, \( M = \max_{n \geq 0} S_n, \) is almost surely finite.

Let \( F \) be the common distribution function of the \( X_i, \) and let \( F = 1 - F \) be its tail. This tail is regularly varying (see e.g. Bingham, Goldie, Teugels, 1989) if there exists a nonnegative \( \alpha \) such that

\[
\lim_{t \to \infty} \frac{F(\lambda t)}{F(t)} = \lambda^{-\alpha}
\]

for any positive \( \lambda. \) The number \( -\alpha \) is called the index of regular variation.

Veraverbeke’s theorem asserts that if the distribution of the increments of the random walk has negative mean and regularly varying tail with index \( -\alpha \) less than \(-1,\) then

\[
P\{ M > t \} \sim \frac{1}{-\mu} \int_t^\infty F(u) \, du,
\]

as \( t \) tends to infinity; or, equivalently, using Karamata’s theorem,

\[
P\{ M > t \} \sim \frac{1}{-\mu} \frac{t \overline{F}(t)}{\alpha - 1}
\]

as \( t \) tends to infinity.

The original question which motivated us to write this paper is the following simple one: if the increment of the ‘random walk’ are themselves a random walk, should we replace \( t\overline{F}(t) \) in (1.2) by \( t^2 \overline{F}(t)? \) The answer to this question is given in subsection 2.4 of this paper.

Veraverbeke’s theorem has been extended in several directions. Clearly, one can seek to prove that it holds for a larger class of distributions for the increments. In that line of investigation,
Veraverbeke’s (1977) original result asserts in fact that \((1.2)\) holds whenever \(F\) is subexponential. Later, Korshunov (1997) obtained necessary and sufficient condition for \((1.1)\) to hold, building upon Borovkov (1971), Pakes (1975), Veraverbeke (1977), Embrechts and Veraverbeke (1982). We also mention that it is not necessary that the increments of the random walk have negative mean for the maximum of the process to be almost surely finite. An analysis of the tail distribution for the global maximum of a random walk with heavy tail increments without mean and with a left tail dominance can be found in Borovkov (2003).

In a different direction, Mikosch and Samorodnitsky (2000) replace the independent increments of the random walk by an infinite order moving average process; they consider \(X_n = \mu + \sum_{j \in \mathbb{Z}} \phi_n^{-j} \epsilon_j\), where \(\mu\) is negative, the \(\epsilon_j\) are centered, independent, equidistributed with a common distribution having regularly varying tail. They assume further a natural tail balance condition so that the distribution of \(X_n\) is itself regularly varying even if some of the \(\phi_j\) are negative. They prove that if \(\sum_{j \in \mathbb{Z}} |j\phi_j|\) is finite, then an analogue of Veraverbeke’s theorem holds, that is, there exists a constant \(c\) such that

\[
P\{M > t\} \sim c \frac{t \overline{F}(t)}{\alpha - 1}
\]

as \(t\) tends to infinity. Interestingly, the constant \(c\) is explicit — but its value is irrelevant here — and is, in general, different than the factor \(1/(-\mu)\) involved in \((1.1)\) or \((1.2)\). However, Mikosch and Samorodnitsky’s (2000) result shows that the decay in Veraverbeke’s theorem remains unchanged in their more general setting. They also note that their assumptions exclude some fractional integrated ARMA models. We remark that such models are considered in section 2.4 of the current paper.

Yet in another direction, Konstantinides and Mikosch (2005) make a study of the tail behavior of the global maximum of partial sums of dependent heavy tailed summands with negative drift. More specifically, they consider the stationary solution \(Y_n\) of a stochastic recurrence equation \(Y_n = A_n Y_{n-1} + B_n\) and consider the random walk \(S_n = Y_1 + \cdots + Y_n\). Considering a negative real number \(\mu\), they provide an asymptotic equivalent for the tail distribution of the maximum of the process \((S_n - ES_n + \mu n)_{n \geq 1}\). One of their findings is that, with their assumptions on the coefficients \(A_n\) and \(B_n\), the order of decay of the classical Veraverbeke result is preserved but
the constant in the asymptotic expression changes. They interpret this change of constant as a measure of clustering of extremes for the stationary \((Y_n)_{n \geq 1}\) sequence.

As indicated, a purpose of this paper is to describe the tail behavior of the maximum of some processes which generalize in a natural way the random walk model, which are nonstationary and exhibit long range dependence — for instance, in the sense that the series of the correlations between the process at a fixed time and time \(n\) is not summable in \(n\). Of Veraverbeke’s original result, only that there is a relation between the tail of the maximum and the tail of the innovation will be preserved; neither the constant \(1/(−\mu)\) nor, in contrast to the models studied by Mikosch and Samorodnitsky (2000) or Konstantinides and Mikosch (2005), the order of decay, \(t^{-1}\), will be preserved in general.

To understand the motivation for the class of processes which we are going to introduce and to which we will extend Veraverbeke’s result, as well as to frame the contribution of this paper in a larger context, it is necessary to recall two results on random walks and fractional ARIMA processes; such presentation requires defining some notation related to the latter. Toward this end, for any nonpositive integer \(n\), we set \(X_n\) to be 0. The backward shift \(B\) acts on the sequence \((X_n)_{n \in \mathbb{Z}}\) by \(BX_n = X_{n-1}\). As usual, this operator can be raised to a nonnegative power, with \(B^0\) being the identity and \(B^n\) being defined inductively as \(BB^{n-1}\). For any positive real number \(d\) and any polynomials \(\Phi\) and \(\Theta\) with both \(\Phi(1)\) and \(\Theta(1)\) nonzero, the nonstationary FARIMA\((\Phi,d,\Theta)\) process \((Y_n)_{n \geq 0}\) with innovations \((X_n)_{n \geq 1}\) is defined by the formula

\[
Y_n = (1 - B)^{-d}\Phi(B)^{-1}\Theta(B)X_n;
\]

the actual meaning of this expression is obtained by expanding the function

\[
g(x) = (1 - x)^{-d}\Phi(x)^{-1}\Theta(x)
\]

as a Taylor series \(\sum_{i \geq 0} g_i x^i\) and setting \(Y_n = g(B)X_n\), that is,

\[
Y_n = \sum_{0 \leq i < n} g_i X_{n-i}, \quad n \geq 0.
\]

Note that \(Y_0\) vanishes. One sees that the random walk is obtained for \(d = 1\) and the polynomials \(\Phi\) and \(\Theta\) being constant equal to 1. The classical ARMA processes are obtained for \(d = 0\).
The two results alluded to — which we will not use but put the present paper in a broader perspective — are that as $n$ tends to infinity, a random walk up to time $n$, suitably rescaled and under the proper moment conditions on the increments (the exact assumptions are irrelevant to this discussion)

(i) converges in distribution to a Wiener process (see e.g. Billingsley, 1968);

(ii) obeys a large deviation principle with rate function involving a derivative (Varadhan, 1966; see e.g. Dembo and Zeitouni, 1992).

Extending those two results with possibly stronger assumptions on the increments, some FARIMA processes, properly rescaled

(i) converge in distribution to a fractional Brownian motion (Akomom and Gouriéroux, 1987; see also Wang, Lin and Gulati, 2002)

(ii) obey a large deviation principle with rate function involving a fractional derivative (Barbe and Broniatowski, 1998).

Thus, at a broad level, underlying these results is the idea that some statements valid for the partial sum processes may be extended to some FARIMA processes, replacing integrals or derivative by their fractional analogue (see Oldham and Spanier, 2006, for fractional calculus). This suggests that for some FARIMA processes, an analogue of Veraverbeke’s theorem might be true, replacing the integrated tail by a fractional integrated tail.

There are further general motivations for results of this paper, which are related to the disparate reasons for studying the maximum of random walks with negative drifts and FARIMA processes. One area where the interest in this type of result is clear is that of insurance risk. For the classical model of claims arriving according to a homogeneous Poisson process and constant premium rate, the surplus claim process viewed at lattice time points forms such a random walk and the assumption of profitability ensures that the increments have negative mean. The distribution of the global maximum of the walk describes the ruin probability over an infinite horizon. This was one of the motivation of Embrechts and Veraverbeke (1982). For dependent heavy tail claims, asymptotic bounds on ruin probability have been given by Nyrhinen (2005). A good reference for ruin probability calculations under a wide variety of model assumptions can be found in Asmussen (2000). From this perspective, our
results allow calculation of ruin probability when the claim process is a nonstationary FARIMA one.

In queuing theory, for a GI/G/1 queue with traffic intensity less than 1, the stationary distribution of the waiting time is given by the distribution of the global maximum of a random walk with negative drift and if the service time distribution has a heavy tail, we are precisely in the situation governed by Veraverbeke’s result; see Pakes (1975) and Asmussen (1987). For extension of the theory to a dependent setting, see Asmussen, Schmidli, Schmidt (1999). For related information in the case of queuing networks, we refer to Baccelli and Foss (2004) and Baccelli, Foss and Lelarge (2005). Again, our result could be converted into statements on waiting time distribution for some queue.

From a modeling perspective, the processes which we will study extend the FARIMA ones. FARIMA processes possess the desirable property that both short- and long-memory components of a time series can be accounted for. For example, in hydrology, Montanari, Rosso and Taqqu (1997) use a FARIMA$(1,d,1)$ process to model deseasonalized daily flows into a lake. A value for the parameter $d$ in the range $0 < d < 1/2$ corresponds to a long-memory process. By way of illustration, a value of $d = 0.26$ was obtained for the lake inflow data studied by Montanari, Rosso and Taqqu (1997), indicating that long memory models are of value in this type of application. Further discussion of applications of those models can be found in Samorodnitsky and Taqqu (1992). For the estimation and theoretical properties of FARIMA processes with heavy tails, we refer to Kokoszka and Taqqu (1995). Resnick (2007) is also a source of information concerning long-range dependence and heavy-tailed modeling.

FARIMA processes have also been of much use in econometric and time series analysis, in part because their occurrence in aggregation of light-tailed time series — see Granger (1980) and the clear exposition in Beran (1994) — and also in connection with the problem of testing for unit root (Akonon and Gouriéroux, 1987; Phillips, 1987; Tanaka, 1999).

2. $(g,F)$-processes, their maximum and sample paths. This section contains our main concrete results. A more abstract formulation is presented in section 3. In the first subsection we define the $(g,F)$-processes, which generalize in a natural way the FARIMA
processes, and we state our tail equivalent of the distribution of their supremum. This is done under some positivity assumption which we remove in the second subsection. In the third subsection we analyze the likely paths for such processes to reach a high level. In the fourth subsection we discuss two examples.

2.1. \((g, F)\)-processes and their maximum. Let \(g\) be a real analytic function on the segment \((-1, 1)\). Its Taylor series expansion

\[
g(x) = \sum_{i \geq 0} g_i x^i
\]

allows one to define the nonstationary process \(S_n = g(B)X_n\). We call such a process a \((g, F)\)-process, \(F\) being the common distribution of the \(X_n\) with \(n\) positive, and with the convention that \(X_n = 0\) if \(n\) is nonpositive. Thus, \(S_n = \sum_{0 \leq i < n} g_i X_{n-i}\). We see that if all the \(g_i\) are equal to 1, that is if \(g(x) = 1/(1 - x)\), then \(S_n\) is a random walk. In this section we give an analogue of Veraverbeke’s theorem for the maximum of some \((g, F)\)-processes.

Defining for any nonnegative integers \(k\) and \(n\) with \(k\) less than \(n\),

\[
g[k,n] = \sum_{k \leq i < n} g_i,
\]

and keeping the notation \(\mu\) for the mean of the \(X_j\), the mean of \(S_n\) is \(\mu g[0,n]\). For the maximum \(M\) of this process to be finite in a setting which extends that of a random walk, it is natural to require \(\mu\) to be negative and \(\lim_{n \to \infty} g[0,n] = +\infty\). In particular, this latter requirement suggests that \(g\) should have a singularity at 1. Because we will need to have some estimation on the decay of the expectation of \(S_n\) toward minus infinity, because it is sufficient to encompass the FARIMA processes, and because it yields a nice mathematical theory, we restrict the singularity of \(g\) by requiring \(g\) to be regularly varying at 1, meaning the existence of some \(\gamma\) such that

\[
\lim_{\epsilon \to 0} \frac{g(1 - \lambda \epsilon)}{g(1 - \epsilon)} = \lambda^{-\gamma}.
\]

For \(g\) to have a singularity at 1, it is then necessary that \(\gamma\) is nonnegative. We will only consider positive \(\gamma\).

If \((g_n)_{n \geq 0}\) is asymptotically equivalent to a monotone sequence as \(n\) tends to infinity, then a straightforward variant of Karamata’s
Tauberian theorem for power series (Bingham, Goldie and Teugels, 1989, Corollary 1.7.3) shows that regular variation of $g$ with index $\gamma$ is equivalent to that of the sequence $(g_n)_{n \geq 0}$ with index $\gamma - 1$; furthermore, in this case, writing $\Gamma(\cdot)$ for the gamma function,

$$g_n \sim \frac{g(1 - 1/n)}{n\Gamma(\gamma)}$$ (2.1.1)

as $n$ tends to infinity. Thus, when applying our results, either the function $g$ or its coefficients can be given.

Note that if the sequence $(g_n)_{n \geq 0}$ is regularly varying with index $\gamma - 1$, then, whenever $\gamma$ is positive and different than 1, the sequence $(g_n)_{n \geq 0}$ is asymptotically equivalent to a monotone sequence; indeed, this follows from Bojanic and Seneta’s theorem (Bingham, Goldie and Teugels, 1989, Theorem 1.5.3). An alternative point of view, replacing any assumption on the coefficients by assumptions solely on the function $g$, is given by Braaksma and Stark’s (1997) complex variable analogue of Karamata’s power series theorem.

**Note.** In the remainder of this section, whenever we use an analytic function $g$, we assume that its Taylor coefficients at 0 are nonnegative. This assumption will be dropped in section 2.2.

Since we are only interested in situations where the drift pushes the $(g, F)$-process toward minus infinity, it is natural to assume also that $\lim_{\epsilon \to 0} g(1 - \epsilon) = +\infty$. Thus, if $g$ is regularly varying, there exists a regularly varying function $U$, of index $1/\gamma$, unique up to an asymptotic equivalence, such that

$$g\left(1 - \frac{1}{U(t)}\right) \sim t$$

as $t$ tends to infinity. This function $U$ appears in our results.

Our first result gives an asymptotic equivalent of the tail of the distribution of $M$ when the sequence $(g_n)_{n \geq 0}$ converges to 0. In this case, the number

$$g^* = \sup_{i \geq 0} g_i$$

is well defined, and in fact the supremum is even a maximum.

Recall the beta integral,

$$B(p, q) = \int_0^1 u^{p-1}(1 - u)^{q-1} \, du = \int_0^\infty u^{p-1}(1 + u)^{-p-q} \, du.$$
Theorem 2.1.1. Let $g$ be a real analytic function on $(-1, 1)$ whose Taylor coefficients $(g_n)_{n \geq 0}$ are nonnegative, regularly varying of index $\gamma - 1$, not summable, and tend to 0 at infinity. If $\gamma = 1$, assume further that $(g_n)_{n \geq 0}$ is asymptotically equivalent to a monotone sequence.

Let $F$ be a distribution function with negative mean and whose tail is regularly varying with index $-\alpha$. If $1/\alpha < \gamma < 1$ then the maximum $M$ of the corresponding $(g, F)$-process satisfies

$$P\{ M > t \} \sim \frac{g^{*\alpha}}{\gamma} B\left(\frac{1}{\gamma}, \frac{1}{\gamma} \right) \left( \frac{\Gamma(1 + \gamma)}{-\mu} \right)^{1/\gamma} (UF)(t), \quad (2.1.2)$$

as $t$ tends to infinity.

In the case of Theorem 2.1.1, the distribution tail of $M$ decays at rate $UF$ which, because of (2.1.1) and the convergence of $(g_n)_{n \geq 0}$ to 0, is slower than that in Veraverbeke’s theorem. Note that $UF$ is regularly varying of index $(1/\gamma) - \alpha$, which can assume any value between $1 - \alpha$ and 0 by a proper choice of $\gamma$.

We next consider the case where the sequence $(g_n)_{n \geq 0}$ tends to infinity, which forces $\gamma$ to be at least 1. This case is more involved in particular when $\gamma$ is 1. Indeed, in this latter case our proof involves some more refined asymptotic analysis which requires more precise assumptions; a general result in this situation remains elusive. We limit ourself to models neighboring the classical random walk and find sufficient conditions for preserving the result obtained when $\gamma$ exceeds 1. The analysis requires some extra notions related to the theory of regularly varying functions, notions which we now introduce.

Recall that a slowly varying function $\ell$ has a Karamata representation (Bingham, Goldie and Teugels, 1989, Theorem 1.3.1) which asserts the existence of a function $\varepsilon(\cdot)$ with limit 0 at infinity and a function $a(\cdot)$ with a finite positive limit at infinity, such that

$$\ell(x) = a(x) \exp\left( \int_1^x \frac{\varepsilon(u)}{u} \, du \right)$$

ultimately. When $\gamma$ is equal to 1, it is natural to assume that

$$g_k \sim \ell(k) \quad (2.1.3)$$
where $\ell$ is slowly varying and tends to infinity at infinity. It then follows from Corollary 1.3.5 in Bingham, Goldie and Teugels (1989) that we can take $\varepsilon(\cdot)$ to be nonnegative. In fact, by the same argument used to prove their Corollary 1.3.5, the function $\varepsilon(\cdot)$ can be taken positive if $\ell$ is increasing. Note that in the Karamata representation of $\ell$, if the function $\varepsilon(\cdot)$ is equal to $1/\log p$ for some $p$ greater than 1, then $\ell$ has a finite limit at infinity. Thus, to fix the ideas, under some extra conditions which we will not assume, we could force $\varepsilon(\cdot)$ to tend to 0 at a rate slower than $1/\log p$ for any $p$ greater than 1. The point of this remark is to suggest that, actually, it is natural to assume that $\varepsilon(\cdot)$ is itself slowly varying. We will assume in fact a weak form of super-slow variation (see Bingham, Goldie and Teugels, 1989, §3.12.2) for $\varepsilon(\cdot)$, namely that

$$
\lim_{t \to \infty} \frac{\varepsilon(\lambda t \varepsilon(t))}{\varepsilon(t)} = 1 \quad (2.1.4)
$$

uniformly in $\lambda$ in any compact subset of the positive half-line. Following the terminology given in Bingham, Goldie and Teugels (1989, §2.11) we will also assume further that

$$
\log \varepsilon(e^t) \text{ is self-neglecting,} \quad (2.1.5)
$$

meaning that the function $\varphi(t) = \log \varepsilon(e^t)$ satisfies locally uniformly, that is, uniformly for $\lambda$ in any fixed compact set,

$$
\lim_{t \to \infty} \frac{\varphi(t + \lambda \varphi(t))}{\varphi(t)} = 1.
$$

These unfortunately technical looking conditions still allow a wide array of interesting examples. For instance, if $g_k = \log^p k$ for some positive $p$, then $\ell(x) = \log^p x$ and $\varepsilon(x) = p/\log x$ satisfies both (2.1.4) and (2.1.5). Similarly, $\ell(x) = \exp(\log^p x)$ with $p$ less than 1 yields $\varepsilon(x) = p \log^{p-1} x$, and that latter function satisfies (2.1.4) and (2.1.5).

Note that the function (1.3) is regularly varying at 1, with index $d$.

For $\gamma$ positive we need to introduce the functions

$$
\rho_\gamma(u) = \min_{y > 0} \frac{\Gamma(1 + \gamma) + (u + y)\gamma}{\gamma y^{\gamma-1}},
$$

11
whose argument \( u \) is nonnegative. In general, it does not seem possible to obtain an explicit form of the minimum. However, one can easily see that \( \rho_1(u) = 1 + u \) and \( \rho_2(u) = u + \sqrt{2 + u^2} \). We will see that \( \rho_\gamma^{-\alpha} \) is integrable on the positive half-line whenever \( \alpha \) is greater than 1 — see Lemma 5.4.2.

The following result is an analogue of Veraverbeke’s theorem for \((g, F)\)-processes when the sequence \((g_n)_{n \geq 0}\) diverges to infinity. Note that by (2.1.1) this condition necessitates that \( \gamma \) is at least 1. We write \( \text{Id} \) for the identity function on the real line.

\[ \text{Theorem 2.1.2.} \quad \text{Let } g \text{ be a real analytic function on } (-1, 1) \text{ whose Taylor coefficients } (g_n)_{n \geq 0} \text{ tend to infinity and are regularly varying of index } \gamma - 1. \text{ If } \gamma = 1 \text{ assume further that } (g_n)_{n \geq 0} \text{ is asymptotically equivalent to a monotone sequence and that (2.1.3), (2.1.4) and (2.1.5) hold.}

Let \( F \) be a distribution function with negative mean and whose tail is regularly varying with index \(-\alpha\) less than \(-1\). The maximum \( M \) of the corresponding \((g, F)\)-process satisfies

\[ P\{ M > t \} \sim (-\mu)^{\alpha(1-1/\gamma)-1/\gamma}(\text{Id} F) \circ U(t) \int_0^\infty \rho_\gamma^{-\alpha}(u) \, du \quad (2.1.6) \]

as \( t \) tends to infinity.

When Theorem 2.1.2 applies, the tail probability of \( M \) decays like \((\text{Id} F) \circ U\). Under the assumptions of Theorem 2.1.2, \( U \) tends to infinity at a rate slower than that of the identity. Then, the rate of decay of \((\text{Id} F) \circ U\) is slower than that involved in Veraverbeke’s (1977) theorem. By a suitable choice of \( U \), hence of \( g \), the index of regular variation of \((\text{Id} F) \circ U\) can assume any value between \( 1 - \alpha \) and 0.

The last result of this subsection somewhat fills the main gap left by Theorems 2.1.1 and 2.1.2, assuming now that the sequence \((g_n)_{n \geq 0}\) converges to a positive and finite limit. Recall that \( g^* \) is defined as the supremum of the sequence \((g_n)_{n \geq 0}\).

\[ \text{Theorem 2.1.3.} \quad \text{Let } g \text{ be a real analytic function on } (-1, 1) \text{ whose Taylor coefficients } (g_n)_{n \geq 0} \text{ converge to a finite and positive limit } g_\infty.

Let \( F \) be a distribution function with negative mean and whose tail is regularly varying with index \(-\alpha\) less than \(-1\). Then, the maximum
\( M \) of the associated \((g,F)\)-process satisfies

\[
P\{ M > t \} \sim \frac{g^\gamma t^{\overline{F}(t)}}{-\mu g^\infty} \alpha - 1
\]
as \( t \) tends to infinity.

Under the assumptions of Theorem 2.1.3, it is easy to see that \( U(t) \sim t/g^\infty \) at infinity. Thus, the result of Theorem 2.1.3 is formally that of Theorem 2.1.1 when \( \gamma = 1 \) and \( U(t) \sim t/g^\infty \). This is no coincidence, and the proof of Theorem 2.1.3 builds upon that of Theorem 2.1.1.

Clearly, if \( g(x) = 1/(1-x) \), then all the Taylor coefficients \( g_i \) are equal to 1, and Theorem 2.1.3 implies Veraverbeke’s (1977) result.

It is interesting to compare the index of regular variation of the tail probability of \( M \) in Theorems 2.1.1, 2.1.2 and 2.1.3. We see that it varies between \( 1 - \alpha \) and 0, and that within the class of \((g,F)\)-processes considered, the classical random walk of Veraverbeke’s original result is an extreme case where the index is equal to \( 1 - \alpha \).

2.2. Removing the positivity of the coefficients. For application to FARIMA processes without too restrictive assumptions on the polynomials \( \Theta \) and \( \Phi \), it is necessary to extend the results of the previous section to the case where some coefficients \( g_i \) may be negative. Our technique allows this extension. It is particularly simple if only a finite number of \( g_i \) are negative. It is still simple if \( \sum_{0 \leq i < n} g_i \) is of the same order as \( \sum_{0 \leq i < n} |g_i| \) as \( n \) tends to infinity. However, the discussion is more delicate if \( \sum_{0 \leq i < n} g_i = o(\sum_{0 \leq i < n} |g_i|) \) as \( n \) tends to infinity — this is related to the asymptotic behavior of the function \( \psi_n \) to be defined in section 3 and analyzed in some detail in section 4. Since this latter case does not seem to have much bearing to applications, we limit ourselves to the situation where only a finite number of \( g_i \) have negative sign. As we will see, this is sufficient to cover the FARIMA processes.

If \( \lim_{n \to \infty} g_n = +\infty \), then it is easy to see from its proof that Theorem 2.1.2 remains valid without any change, even if finitely many coefficients are negative.

If \( \lim_{n \to \infty} g_n = 0 \), then the lower tail of the distribution may play a role in the tail behavior of \( M \). The usual assumption is the following tail balance condition. Let \( F_* \) be the distribution function
of $|X_1|$. The tail balance condition asserts that

$$\lim_{t \to \infty} \frac{F(t)}{F^*(t)} \in (0, \infty) \quad \text{and} \quad \lim_{t \to \infty} \frac{F(-t)}{F^*(t)} \in [0, \infty). \tag{2.2.1}$$

Writing $g_*$ for the smallest negative Taylor coefficient of $g$ if it exists and for 0 otherwise, that is

$$g_* = 0 \land \inf_{i \geq 0} g_i,$$

we have the following extension of Theorem 2.1.1. Recall that, by our convention, a regularly varying sequence is ultimately positive; in particular this forces $g^*$ to be positive.

**Theorem 2.2.1.** Assume that $(g_n)_{n \geq 0}$ is regularly varying, not summable and tends to 0 at infinity. Let $F$ be a distribution function with negative mean, satisfying the tail balance condition (2.2.1), and whose upper tail is regularly varying with index $-\alpha$. If $1/\alpha < \gamma < 1$, the maximum $M$ of the corresponding $(g, F)$-process satisfies

$$P\{M > t \} \sim \frac{1}{\gamma} B\left(\frac{1}{\gamma}, \frac{1}{\gamma} - \frac{1}{\gamma}\right) \left(\frac{\Gamma(1+\gamma)}{-\mu}\right)^{1/\gamma} U(t)\left(g^* F(t) + (-g_*)^\alpha F(-t)\right)$$

as $t$ tends to infinity.

Note that with regard to the previous theorem, writing $\widetilde{F}$ for the distribution function $x \mapsto F(-x-)$ — that is, if $X$ has distribution function $F$ then $\widetilde{F}$ is the distribution function of $-X$ — a $(g, F)$-process has the same distribution as a $(-g, \widetilde{F})$ one. However, the theorem does not allow such a substitution for it is assumed that $F$ has negative mean and $g$ tends to $+\infty$ at 1.

### 2.3. Typical trajectories leading to a large maximum.

The purpose of this subsection is to describe the most likely trajectories of $(g, F)$-processes leading to a large value of their maximum. For simplicity, we consider only the case where all the $g_n$ are positive; the extension to the setting of the previous subsection poses no real difficulties and does not seem to bring any further understanding. Our result can also be viewed as an extension of those of Asmussen and Klüppelberg (1996). For the classical random walk with negative
drift and heavy-tail increments, they prove that conditionally on having the maximum of the process larger than \( t \), the process properly normalized, run through the proper time scale and up to the time at which it reaches its maximum, converges to a straight line with slope equal to the mean of the increments — indicating that the process behaves as expected in this time frame — whereas the overshoot at the jump time properly normalized converges to a Pareto distribution. Further information on the time of jump and conditional path behavior in this context is given in Asmussen (2000). This section addresses a similar problem and carries the analysis further by providing a description of \((g,F)\)-processes both before and after the jump time.

To analyse those trajectories, we write \( N_t \) for the first passage time of the process \((S_n)_{n \geq 0}\) over the threshold \( t \), that is

\[
N_t = \min \{ n : S_n > t \},
\]

with the convention that the minimum of the empty set is \(+\infty\). Clearly, \( M \) exceeds \( t \) if and only if \( N_t \) is finite. The single large jump heuristic described in detail in section 3 suggests that \( M \) exceeds \( t \) because, most likely, one of the \( X_i, 1 \leq i \leq N_t \), is large. Thus, it is natural to consider the index \( J_t \) of occurrence of the 'big jump', that is, the integer between 0 and \( N_t \) such that

\[
X_{J_t} = \max \{ X_i : 1 \leq i \leq N_t \}.
\]

In case of ties, we take \( J_t \) to be the smallest such index. Furthermore, recalling that the function \( U \) is defined by the asymptotic equivalence \( g(1 - 1/U) \sim Id \) at infinity, we consider the rescaled process

\[
\mathcal{E}_t(\lambda) = S_{\lfloor \lambda U(t) \rfloor}/t
\]

as well as the rescaled random variables

\[
\tau_t = J_t/U(t) \quad \text{and} \quad Y_t = X_{J_t}/U(t).
\]

The rescaled process \( \mathcal{E}_t \) is right continuous with left limit, and therefore is viewed here in the space \( D[0, \infty) \) of all càdlàg functions equipped with the Skorohod topology (see Billingsley, 1968; Pollard, 1984). In order to obtain a pleasing result, we will assume that the tail balance condition (2.2.1) holds. We will restrict ourselves to what happens under the assumptions of Theorem 2.1.2.
Theorem 2.3.1. Let $\gamma$ be greater than 1 and assume that the hypotheses of Theorem 2.1.2 hold as well as the tail balance condition (2.2.1). If $\mu = -1$, then the conditional distribution of $(S_t, \tau_t, Y_t)$ conditional on $M > t$ converges weakly* to that of $(\hat{S}, \tau, \rho_\gamma(\tau)Y)$ where $Y$ and $\tau$ are independent and

(i) $Y$ has a Paréto distribution on $[1, \infty)$ with parameter $\alpha$,

(ii) $\tau$ has density proportional to $\rho_\gamma^{-\alpha}$,

(iii) $\hat{S}(\lambda) = \frac{1}{\Gamma(1 + \gamma)}(-\lambda^\gamma + \mathbb{1}\{\lambda \geq \tau\} \gamma(\lambda - \tau)^{\gamma-1}\rho_\gamma(\tau)Y)$.

When $\mu$ is an arbitrary negative number, we will prove by a rescaling argument that the limiting triple is

$$\left(\hat{S}\left((\mu)^{1/\gamma}.\right), (-\mu)^{-1/\gamma}\tau, (-\mu)^{-1-1/\gamma}\rho_\gamma((-\mu)^{-1/\gamma}\tau)Y\right).$$

It is not difficult to adapt our proof to the case of the random walk with negative drift, for which $\gamma = 1$, and show that the same result remains true.

One can check, starting from the definition of $\rho_\gamma$ and the fact that $Y$ is at least 1 almost surely, that the maximum of the process $\hat{S}$ is at least 1 almost surely. Note that the trajectories of the limiting process $\hat{S}$ are infinitely differentiable on the positive half-line, except at $\tau$. If $\gamma < 2$, the trajectories are not differentiable at $\tau+$ but are Hölderian of index $\gamma - 1$. Thus, on the right of the random time $\tau$ they have a vertical tangent going upward. If $\gamma > 2$, the trajectories are differentiable. If $\gamma = 2$, the trajectories are not differentiable at $\tau$ but admit left and right tangents.

The following pictures show typical paths of the limiting process $\hat{S}$ for different values of $\gamma$. 

\begin{center}
\begin{tikzpicture}
\draw[->] (0,0) -- (5,0) node[right] {$\tau$};
\draw[->] (0,0) -- (0,3) node[above] {$\gamma=1$};
\end{tikzpicture}
\hspace{1cm}
\begin{tikzpicture}
\draw[->] (0,0) -- (5,0) node[right] {$\tau$};
\draw[->] (0,0) -- (0,3) node[above] {$\gamma=1.5$};
\end{tikzpicture}
\end{center}
It is noticeable that for $\gamma > 2$, the limiting process keeps decreasing for some time after $\tau$, before increasing to reach its maximum and finally regains a path asymptotic to that given by the law of large number. Thus there is a delay not only in reaching the maximum but also in changing from the negative drift to a positive one which will lead to the maximum.

It is not difficult to adapt the proof of Theorem 2.3.1 to study more precisely what happens near the jump when $\gamma$ is equal to 1 and the sequence $(g_n)_{n \geq 0}$ tends to infinity. Similarly, one can see with our proof that when $\gamma$ is less than 1, the processes $\mathcal{E}_t$ do not converge in distribution in the space $D[0, \infty)$ — one of the many reasons is that the function $h$ defined in section 3.3.1 is not continuous from $[0,1)^2$ to $D[0,\infty)$. The phenomenological reason is that in order for those processes to converge, our proof shows that the limiting function needs to be

$$\frac{-1}{\Gamma(1+\gamma)} \left(-\lambda^\gamma + 1 \{ \lambda = \tau \} \gamma \rho_\gamma(\tau) Y \right)$$

which, because of the jump at $\tau$, is not a càdlàg function. It is not particularly difficult, though somewhat lengthy, to adapt our proof to study the process rescaled with different time scales, $U(t)$ and 1, before and after $J_\lambda$, and show that with these different scalings, both parts converge.

Define the random variables $\mathcal{N}$ and $\mathcal{M}$ by

$$\mathcal{N} = \inf \{ \lambda : \mathcal{E}(\lambda) > 1 \},$$

and

$$\mathcal{M} = \sup_{\lambda > 0} \mathcal{E}(\lambda) = \frac{1}{\Gamma(1+\gamma)} \max_{\lambda \geq \tau} -\lambda^\gamma + \gamma(\lambda - \tau)^{\gamma-1} \rho_\gamma(\tau) Y.$$
By the continuous mapping theorem (Billingsly, 1968, §5; Pollard, 1984, §VI.1, example 2), we deduce from Theorem 2.3.1 that when \( \gamma \) is greater than 1, the conditional distribution of \( (N_t/U(t), M/t) \) given \( M > t \) converge to that of \( ((-\mu)^{-1}/\gamma N, M) \). It follows that the conditional limiting distribution of the overshoot, \( (M - t)/t \), given \( M > t \) converges to that of \( M - 1 \). In the same spirit, write \( L \) for the first time that the process \((S_n)_{n \geq 0}\) attains it maximum. Defining \( L \) to be the largest solution in \( [\tau, \infty) \) of the equation
\[
-\mathcal{L}^{-1} + (\gamma - 1)(\mathcal{L} - \tau)\gamma^{-2}Y \rho_{\gamma}(\tau) = 0,
\]
the conditional distribution of \( L/U(t) \) given \( M > t \) converges to the distribution of \( (-\mu)^{-1}/\gamma \mathcal{L} \).

The same conclusions hold when \( \gamma \) is 1 and the hypotheses of Theorem 2.1.2 are satisfied or in the case of a random walk.

2.4. Examples. For a FARIMA(\( \Phi, d, \Theta \)) process, \( g \) is given by (1.3). Let \( c = \Phi(1)/\Theta(1) \). To calculate the associated function \( U \), note that if \( u \) tends to infinity, then
\[
g\left(1 - \frac{1}{u}\right) = u^{\alpha}\frac{\Theta}{\Phi}\left(1 - \frac{1}{u}\right) \sim \frac{u^d}{c}.
\]
We assume that \( c \) is positive; if this is not the case, we should replace \( g \) by \(-g\) and \( X_n \) by \(-X_n \) and permute the upper and lower tails in what follows. Since \( c \) is positive, we obtain
\[
U(t) \sim (ct)^{1/d}
\]
as \( t \) tends to infinity. For the FARIMA processes, Akonom and Gouriéroux (1987) showed directly that
\[
g_n \sim \frac{n^{d-1}}{c\Gamma(d)},
\]
as \( n \) tends to infinity. Thus, if \( d \) is larger than 1, Theorem 2.1.2 yields
\[
P\{ M > t \} \sim (-\mu)^{\alpha(1-1/d)-1/d} (\text{IdF})_{(c^{1/d}t^{1/d})} \int_0^\infty \rho_{d}(v)^{-\alpha} \, dv \\
\sim (-\mu)^{\alpha(1-1/d)-1/d} c^{(1-\alpha)/d} t^{1/d} \text{F}(t^{1/d}) \int_0^\infty \rho_{d}(v)^{-\alpha} \, dv
\]
as $t$ tends to infinity. In contrast, if $d$ is less than 1 and the tail balance condition (2.2.1) holds, then, writing $p$ for $\lim_{t \to \infty} F(t)/F_*(t)$ and $q$ for $\lim_{t \to \infty} F(-t)/F_*(t)$, Theorem 2.2.1 yields

$$P\{ M > t \} \sim \frac{1}{d} B \left( \frac{1}{d}, \alpha - \frac{1}{d} \right) \left( \frac{(1 + d)}{-\mu} \right)^{1/d} c^{1/d} \left( p g^* + q (-g_*)^{\alpha} \right) t^{1/d} F_*(t)$$

as $t$ tends to infinity. There is no explicit expression for $g^*$ and $g_*$ but those can be calculated numerically if needed.

The above asymptotic equivalent sheds further light on the parallel mentioned in the introduction between FARIMA process and the classical random walk. Indeed, one may consider the fractional integrated tail of order $1/d$,

$$I_{1/d} F_*(t) = \int_t^\infty (x - t)^{(1/d) - 1} F_*(x) \, dx.$$  

A change of variable $x = (1 + \lambda)t$ shows that

$$I_{1/d} F_*(t) \sim B \left( \frac{1}{d}, \alpha - \frac{1}{d} \right) t^{1/d} F_*(t)$$

as $t$ tends to infinity. Therefore, when $d$ is less than 1 we have

$$P\{ M > t \} \sim \left( \frac{(1 + d)}{-\mu} \right)^{1/d} c^{1/d} \left( p g^* + q (-g_*)^{\alpha} \right) I_{1/d} F_*(t)$$

as $t$ tends to infinity, continuing the similarity mentioned in the introduction between FARIMA processes and the usual random walk, namely that an integrated tail is replaced with a fractionally integrated tail.

The ‘random walk’ whose increments are themselves a random walk, corresponds to $g(x) = (1 - x)^{-2}$. To apply Theorem 2.1.2, we need to consider $\rho_2(u) = u + \sqrt{2 + u^2}$. The change of variable $s = u + \sqrt{2 + u^2}$ shows that

$$\int_0^\infty \rho_2(u)^{-\alpha} \, du = \frac{1}{2\alpha + 1} \frac{3\alpha + 1}{\alpha^2 - 1}.$$  

Hence, in this case, we obtain

$$P\{ M > t \} \sim \frac{(-\mu)^{(\alpha-1)/2} 3\alpha + 1}{2^{\alpha+1} \alpha^2 - 1} \sqrt{t} F_*(\sqrt{t}).$$
as \( t \) tends to infinity.

More generally, a ‘random walk’ whose increments are a \((g,F)\)-process is a \(((1-Id)^{-1}g,F)\)-process. Provided the drift is negative, Theorem 2.1.2 or 2.2.1 yield an asymptotic estimate on the tail of the distribution of its maximum.

2.5. Note on the quantiles of the maximum of \((g,F)\)-processes. Motivated by the last paragraph of the previous section, the purpose of this section is to describe how the tail of the maximum of a \((g,F)\)-process changes when the increments derive from a process of the same type. Interestingly, we will see that various quantile functions are asymptotically related. Though this is not our main purpose, this connection between high-order quantiles is of potential interest in the theory of value at risk (see e.g. McNeil, Frey and Embrechts, 2005). Indeed, motivated by the Basel II regulatory framework, there has been some studies of how extreme quantiles behave under addition of random variables, with the most recent research emphasizing cases where the random variables are dependent (Barbe, Fougères, Genest, 2006; Embrechts, Nešlehová, Wütrich, 2008; Embrechts, Lambrigger, Wütrich, 2008, and references therein). The following result gives corresponding results when aggregation is made according to some moving average scheme and the quantity of interest is the global maximum of the process.

To state our next results, let \( H_+ \) be the set of all analytic functions on \((-1,1)\) whose sequence of Taylor coefficients at 0 tends to infinity and are regularly varying of positive index. This set of functions is a semi-group both under addition and multiplication. Let \( F \) be a fixed distribution function with regularly varying tail of index less than \(-1\) and with negative mean. Let \( M_g \) be the maximum of the corresponding \((g,F)\)-process. We write \( q_g \) for the function,

\[
q_g(s) = \inf \{ t : P\{ M_g \leq t \} \geq 1 - 1/s \}.
\]

This is the quantile function of \( M_g \) evaluated at \( 1 - 1/s \). We also write

\[
c_g = (-\mu)^{\alpha(1-1/\gamma)-1/\gamma} \int_0^{\infty} \rho_u^{-\alpha}(u) \, du
\]

for the constant involved in the statement of Theorem 2.1.2. The quantile function of \( M_g/c_g^{\gamma/(1-\alpha)} \) evaluated at \( 1 - 1/s \) is

\[
\tilde{q}_g(s) = c_g^{-\gamma/(1-\alpha)} q_g(s).
\]
We then have the following asymptotic relations showing that the map $g \in H_+ \mapsto \tilde{q}_g$ is a linear morphism of semigroups in an asymptotic sense.

**Proposition 2.5.1.** Let $g$, $g_1$ and $g_2$ be some functions in $H_+$. Then for any positive $\lambda$, the following asymptotic equivalences hold at infinity.

(i) $\tilde{q}_\lambda g \sim \lambda \tilde{q}_g$.

(ii) $\tilde{q}_{g_1 + g_2} \sim \tilde{q}_{g_1} + \tilde{q}_{g_2}$.

(iii) $\tilde{q}_{g_1 g_2} \sim \tilde{q}_{g_1} \tilde{q}_{g_2}$.

**Proof.** Note that with some obvious notations, the relation defining $U_g$, that is, $g(1 - 1/U_g) \sim \text{Id}$, implies $U_g^{-} \sim g(1 - 1/\text{Id})$ at infinity. Consequently

$$U_{g\lambda}^{-} \sim \lambda U_g^{-}, \quad U_{g_1 + g_2}^{-} \sim U_{g_1}^{-} + U_{g_2}^{-}, \quad \text{and} \quad U_{g_1 g_2}^{-} \sim U_{g_1}^{-} U_{g_2}^{-}.$$  \hspace{1cm} (2.5.1)

Theorem 2.1.2 shows that

$$c_g(\text{IdF}) \circ U_g \circ q_g \sim 1/\text{Id}.$$  

Thus,

$$q_g \sim U_g^{-} \circ (\text{IdF})^{-1}/c_g(\text{Id}) \circ (1/\text{Id}).$$

In particular, $q_g$ is regularly varying of index $\gamma_g/(\alpha - 1)$ and

$$q_g \sim c_g^\gamma_g/(\alpha - 1) U_g^{-} \circ (\text{IdF})^{-1}/(1/\text{Id}).$$

Hence, $\tilde{q}_g \sim U_g^{-} \circ (\text{IdF})^{-1}/(1/\text{Id})$. The result then follows from (2.5.1).

**2.6. Concluding remarks.** The next sections will show the technique used to prove the results of the current section. As it should be clear at the end of this paper, this technique can be used to give extensions of Veraverbeke’s result in different directions.

For instance, Veraverbeke’s theorem can be interpreted as a statement on the probability that a centered random walk crosses a linear moving boundary. Indeed, let $Z_i$ be the centered random variable $X_i - \mu$. The maximum of the random walk based on the $X_i$ exceeds $t$ if the random walk based on the $Z_i$ crosses the boundary $t - \mu i$, which is called a moving boundary since $t$ translates it
upward. In nonlinear renewal theory (see, e.g., Woodroofe, 1982), calculations of crossing probabilities of nonlinear moving boundaries are questions of importance. Our technique yields estimates for the crossing of some nonlinear boundaries. However, it does not allow one to recover all known results and in particular our technique does not work for boundaries near the range of the law of the iterated logarithm — compare to the remarkably general Theorem 1 in Foss, Palmowski and Zachary (2005).

In a similar spirit, but further away from Veraverbeke’s theorem, our technique can be used to evaluate the probability that at least one of the $X_i$ exceeds $t + b(i)$ where $b$ is an increasing function. Clearly, this probability can be evaluated directly, and so the purpose of this remark is only to delineate further the range of usefulness of our technique. Our technique applies when $b$ is regularly varying of index greater $1/\alpha$, but not if $b$ is regularly varying of index $1/\alpha$. Therefore, it does not yield estimates as sharp as extreme value theory.

On a more positive note, our technique seems useful when some form of dependence is present, provided that one has a good representation of the random variables involved. The $(g, F)$-processes provide a nontrivial example. Another interesting example is as follows. Let $(Z_i)_{i \geq 1}$ be a sequence of independent random variables, equidistributed and centered. Let $p$ be a fixed integer, let $\mu$ be a negative real number, and let further $X_i = Z_i Z_{i+1} \ldots Z_{i+p} + \mu$. Consider the ‘random walk’ $(S_n)_{n \geq 0}$ associated to the $X_i$ and defined by $S_0 = 0$ and $S_i = S_{i-1} + X_i$ if $i \geq 1$. The increments of this random walk are the $p$-dependent sequence $(X_i)_{i \geq 1}$. Such process has been considered in the context of large deviations by Choi, Cover and Csizár (1987) as well as Bolthausen (1993). Our technique allows one to show that if $X_i$ has a distribution function $F$ whose tail is regularly varying of index $\alpha$, then Veraverbeke’s result remains valid, namely that

$$P\{ \max_{n \geq 0} S_n > t \} \sim \frac{1}{-\mu \alpha - 1} \frac{tF(t)}{F(t)}$$  \hspace{1cm} (2.6.1)$$

as $t$ tends to infinity. To sketch the proof of this assertion requires the notation to be developed in the next section, and, perhaps, the next paragraph can only be understood after reading the remainder of this paper; however, the shortness of the following sketch seems
a compelling argument in favor of the general framework which we will develop in the next section.

To prove (2.6.1), we evaluate

\[ \psi_{i,n}(x) = E(S_i \mid X_n = x) = \begin{cases} 
  i\mu & \text{if } i < n \\
  (i-1)\mu + x & \text{if } i \geq n 
\end{cases} \]

Thus, \( \psi_n(x) = \max_{i \geq 1} \psi_{i,n}(x) \) is invertible on the preimage of some interval \((t_0, \infty)\) and \( \psi^{-1}_n(t) = t - (n-1)\mu \) on this preimage. It follows that we can take \( U \) and \( \chi \) to be the identity function, and \( \rho(x) = 1 - \mu x \). From this, the result can be guessed using formula (3.1.2). To actually prove the result, steps 3–7 of subsection 3.1 are established as follows. We first split \( S_n \) as a sum of the \( p \) random walks with independent increments

\[ S_{n,j} = \sum_{\substack{1 \leq i \leq n \\
i \equiv j \mod p}} X_i, \quad 0 \leq j < p. \]

Then, whenever we need to estimate a probability involving the event \( S_n > t \) we note that this event is included in \( \bigcup_{0 \leq j < p} \{ S_{n,j} > t/p \} \) and use Bonferroni’s inequality. The result then follows by the estimates of section 6.5.

3. Veraverbeke’s theorem at large. The proofs of our theorems are conceptually simple, but this simplicity is somewhat lost in the many steps needed in its execution. In order to make this simplicity more obvious and intuitive, as well as in order to make our technique easy to adapt to different problems, we first describe a general scheme for how to prove the type of results which we are aiming for. In the second subsection we prove a theorem which asserts that, indeed, whenever this general scheme can be applied, it yields the correct result. It will be used to prove the results of section 2.

In this section only, we consider a stochastic process \((S_n)_{n \geq 1}\) built through functions \( S_n \) mapping a sequence of independent random variables \((X_n)_{n \geq 1}\) into the real line. We are seeking some tail estimate for the maximum of \((S_n)_{n \geq 1}\). Clearly, nothing useful can be said with that level of generality, but our purpose is to describe a technique at a conceptual level. We make no claim that this technique yields the correct result in general — as a matter of fact, it is very easy to find counterexamples — but the remainder of this paper will show that this description can be most useful.
We will use the following notation.

**Definition.** A function $f$ mapping a neighborhood of infinity to a neighborhood of infinity has an asymptotic inverse if it is asymptotically equivalent to a monotone function and there exists a function $f^{-}$ such that $f \circ f^{-} \sim f^{-} \circ f \sim \text{Id}$ at infinity.

### 3.1. The single large jump heuristic.

Our purpose in this subsection is not to do rigorous mathematics but to give some useful intuitions. The basic idea underlying the single large jump heuristic is that a single large $X_n$ is what is likely to make the maximum $M$ of the process $(S_n)_{n \geq 1}$ to be large, and that the other $X_i$ contribute to the process in an average way, having in mind some form of law of large numbers. This leads us to consider the conditional expectations functions

$$
\psi_{i,n}(x) = E(S_i | X_n = x), \quad i, n \geq 1.
$$

If $X_n$ is large, we expect the process $(S_k)_{k \geq 0}$ to reach the level $\psi_{i,n}(X_n)$ at time $i$. Therefore, defining

$$
\psi_n(x) = \max_{i \geq 1} \psi_{i,n}(x),
$$

the maximum of the process, given $X_n$, is expected to reach the level $\psi_n(X_n)$. It exceeds $t$ if $\psi_n(X_n)$ does, that is, if $\psi_n$ is increasing and invertible, if $X_n > \psi_n^{-1}(t)$. So, we anticipate that, as $t$ tends to infinity,

$$
P\{ M > t \} \sim P\{ \exists n \geq 1 : X_n > \psi_n^{-1}(t) \} \sim \sum_{n \geq 1} F \circ \psi_n^{-1}(t).
$$

In what follows, let $r$ be a function such that

$$
r(t) \sim \sum_{n \geq 1} F \circ \psi_n^{-1}(t)
$$

at infinity. It is of course assumed that this function tends to 0 at infinity, that is $M$ is almost surely finite. Thus, $r$ is a tentative asymptotic equivalent for the probability that $M$ exceeds $t$. A general scheme to turn this tentative equivalent into an actual one is as follows. It has two parts, an analytical one and a probabilistic one.
Analytical part. The purpose of this part is to obtain useful information on the function $\psi_n^{-1}$.

**Step 1.** Find two regularly varying functions $U$ and $\chi$ with limit infinity at infinity and a function $\rho$ continuous on the positive half-line, bounded away from 0 on any compact subset of the positive half-line, such that $\rho^{-\alpha}$ is Lebesgue integrable on the nonnegative half-line, and such that the asymptotic factorization

$$
\psi_{[xU(t)]}^{-1}(t) \sim \chi(t)\rho(x)
$$

holds as $t$ tends to infinity, uniformly in $x$ in any compact subset of the positive half-line. Implicit in this assertion is that for $t$ large enough, $\psi_{[xU(t)]}$ is invertible on the preimage of $[t, \infty)$.

The functions $\chi$ and $\rho$ are not unique. However, (S1) implies that for another such pair, say $(\chi_1, \rho_1)$, we have, for any positive $x$,

$$
\frac{X_1}{\chi_1}(t) \sim \frac{\rho_1}{\rho}(x)
$$

as $t$ tends to infinity. This forces $\rho/\rho_1$ to be constant, equal to some $c$ say, and then $\chi_1 \sim c\chi$. The constant $c$ is positive for both $\chi$ and $\chi_1$ are assumed to tend to infinity at infinity. It follows that even though the functions $\chi$ and $\rho$ are not unique, they are asymptotically unique up to a positive multiplicative constant. Our results do not depend on the choice of the constant, and in applications we will choose whatever constant makes the calculation less cumbersome.

In what follows, to any positive real number $\epsilon$ less than 1 we associate the set

$$
I_{\epsilon,t} = \{ n \in \mathbb{N} : \epsilon U(t) \leq n \leq U(t)/\epsilon \},
$$

and we write

$$
N_{\epsilon,t} = \lfloor U(t)/\epsilon \rfloor
$$

for its largest element. Often we will drop the subscripts, writing $I$ and $N$ for $I_{\epsilon,t}$ and $N_{\epsilon,t}$.

**Step 2.** Prove that the asymptotic behavior of $\sum_{n \geq 1} F \circ \psi_n^{-1}(t)$ as $t$ tends to infinity is driven by the terms for which $n$ is of order $U(t)$, that is,

$$
1 = \lim_{\epsilon \to 0} \liminf_{t \to \infty} \frac{1}{r(t)} \sum_{n \in I} F \circ \psi_n^{-1}(t) \\
\leq \lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{1}{r(t)} \sum_{n \in I} F \circ \psi_n^{-1}(t) = 1. \quad (S2)
$$
In particular, this step suggests that in relation to step 1, we should also obtain order of magnitudes or rather crude bounds for \( \psi^{-1} \) when \( n \) is outside \( I \).

Note that the completion of step 2 implies that the function \( r \) can be identified as follows. Since \( F \) is regularly varying, steps 1 and 2 imply

\[
\sum_{n \in I} F \circ \psi^{-1}(t) \sim \sum_{n \in I} F(\chi(t)\rho\left(\frac{n}{U(t)}\right)) \\
\sim \int_{U(t)/\epsilon}^{U(t)} F \circ \chi(t)\rho\left(\frac{u}{U(t)}\right)^{-\alpha} du \\
\sim (U \circ \chi)(t) \int_{\epsilon}^{1/\epsilon} \rho(u)^{-\alpha} du.
\]

Thus, completion of steps 1 and 2 implies that

\[
r(t) = (U \circ \chi)(t) \int_{0}^{\infty} \rho(v)^{-\alpha} dv.
\]

In particular \( r \) is regularly varying. Given how \( r \) was initially defined, this suggests that one could easily guess the tail behavior of \( M \) by a simple examination of \( \psi^{-1} \), just guessing what \( U, \chi \) and \( \rho \) are. This has been illustrated in the last example discussed in section 2.5.

**Probabilistic part.** This part consists in proving that if the process reaches the level \( t \), then it is unlikely to occur at a time too small or too large, and that some form of law of large numbers holds.

**Step 3.** Prove that the process is unlikely to reach the level \( t \) at a time of smaller order than \( U(t) \), that is

\[
\lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{1}{r(t)} P\{ \exists n \leq \epsilon U(t) : S_n > t \} = 0.
\]

The proof of such result is sometimes made easier by the following remark. For any fixed \( \theta \), consider the events

\[
B_i = \{ X_i \leq \theta \chi(t) \},
\]

whose notation does not keep track of the dependence on \( \theta \) and \( t \). We write \( B_i^c \) for the complement of \( B_i \). Since

\[
P\{ \bigcup_{i \leq \epsilon U(t)} B_i^c \} \sim \epsilon U(t) F(\theta \chi(t)) \sim \epsilon \theta^{-\alpha} (U \circ \chi)(t),
\]

(3.1.1)
it suffices to prove that
\[
\lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{1}{r(t)} P \{ \exists n \leq \epsilon U(t) : S_n > t; \bigcap_{i \leq \epsilon U(t)} B_i \} = 0.
\] (3.1.3)

The advantage of this formulation is that on the event \( \bigcap_{i \leq \epsilon U(t)} B_i \) the random variables are bounded, and many more inequalities exist for bounded random variables than for unbounded ones.

**Step 4.** Prove that the process is unlikely to reach the level \( t \) at a time of larger order than \( U(t) \), that is,
\[
\lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{1}{r(t)} P \{ \exists n \geq U(t)/\epsilon : S_n > t \} = 0. \tag{S4}
\]

As in step 3, it is sometimes useful to replace the unbounded variables \( X_i \) by bounded ones. Let \( \kappa \) be the index of regular variation of \( \chi \) and assume that \( \alpha \kappa \gamma \) is greater than 1. We have
\[
P \{ \exists n \geq 0 : X_n > \theta (\chi \circ U^- (n) + \chi (t)/\eta) \} \sim \sum_{n \geq 1} F \left( \theta (\chi \circ U^- (n) + \chi (t)/\eta) \right).
\]

Replacing this series by the corresponding Riemann integral,
\[
\int_{1}^{\infty} F \left( \theta (\chi \circ U^- (u) + \chi (t)/\eta) \right) du,
\]
making the change of variable \( u = \lambda U(t) \) and using the regular variation of \( F \), the series is asymptotically equivalent to
\[
\theta^{-\alpha} (U F \circ \chi) (t) \int_{0}^{\infty} (\lambda^{\kappa \gamma} + 1/\eta)^{-\alpha} d\lambda
\]
as \( t \) tends to infinity. Since \( (\lambda^{\kappa \gamma} + 1/\eta)^{-\alpha} \) is at most \( \eta^\alpha \) when \( \lambda \) is in \([0,1]\) and at most \( \lambda^{-\alpha \kappa \gamma} \) when \( \lambda \) is at least 1, by dominated convergence,
\[
\lim_{\eta \to 0} \limsup_{t \to \infty} \frac{1}{r(t)} P \{ \exists n \geq 0 : X_n > \theta (\chi \circ U^- (n) + \chi (t)/\eta) \} = 0.
\]

Therefore, we can replace the original problem of this step by that of proving
\[
\lim_{\eta \to 0} \limsup_{t \to \infty} \frac{1}{r(t)} P \{ \exists n \geq U(t)/\eta : S_n > t; \bigcap_{i \geq 1} \{ X_i \leq \theta (\chi \circ U^- (i) + \chi (t)/\eta) \} \} = 0. \tag{3.1.4}
\]
The next step consists in formalizing the single large jump heuristic, proving that for $M$ to exceed $t$ then at least one large jump had likely occurred.

**Step 5.** Prove that for the process to exceed $t$ at a time in $I$, we need at least one variable prior to that time to exceed $\theta \chi(t)$, that is, for any positive $\epsilon$, there exists some positive $\theta$ such that

$$P\left( \bigcup_{i \in I} \left( \{ S_i > t \} \cap \bigcap_{n \leq i} B_n \right) \right) = o(r(t)). \quad (S5)$$

Note that if this holds for some $\theta$ then it holds for any smaller one, because the sets $B_i$ are decreasing in $\theta$.

Interestingly, completion of steps 1, 2 and 5 are enough to show that it is unlikely that $M$ reaches $t$ because two of the $X_i$’s are large. Indeed, we have

$$P\left( \bigcup_{i,j \leq N} B_i^c \cap B_j^c \right) \sim \frac{N}{2} F(\theta \chi(t))^2 \sim \frac{1}{2\theta^2 \epsilon^2} (UF \circ \chi)^2(t),$$

and since $r \sim UF \circ \chi$ tends to 0, this implies

$$\lim_{t \to \infty} \frac{1}{r(t)} P\left( \bigcup_{i,j \leq N} B_i^c \cap B_j^c \right) = 0.$$

We can then move on to the next step, showing that when a single $X_n$ is large, then the process can be approximated by its conditional expectation given that large random variable. Recall that $N = \lfloor U(t)/\epsilon \rfloor$. It is convenient for what follows to introduce the events that all $X_i$, $1 \leq i \leq N$, are at most $\theta \chi(t)$ except perhaps $X_n$, that is,

$$C_n = \bigcap_{1 \leq i \leq N, i \neq n} B_i,$$

and the events that the $S_i$, $1 \leq i \leq n$, are well approximated by $\psi_{i,n}(X_n)$, that is,

$$D_n = \bigcap_{1 \leq i \leq N} \{ |S_i - \psi_{i,n}(X_n)| \leq \delta t \}.$$

**Step 6.** Prove that if $X_n$ exceeds $\theta \chi(t)$ and all the other $X_i$, $i \leq N$, are at most $\theta \chi(t)$, then each $S_i$ is about $\psi_{i,n}(X_n)$; more precisely, prove that for any positive $\delta$,

$$P\left( \bigcup_{n \in I} B_n^c \cap C_n \cap D_n^c \right) = o(r(t)). \quad (S6)$$
This is in fact a little stronger than what we need, and sometimes a one-sided bound, replacing $|S_i - \psi_{i,n}(X_n)|$ by $S_i - \psi_{i,n}(X_n)$, may suffice.

A naive and yet effective way to prove such law of large numbers is to show first that

$$\max_{n \in I} P(B^c_n \cap C_n \cap D^c_n) = o\left(\frac{r(t)}{U(t)}\right),$$

and then use Bonferroni’s inequality, upon noting that the cardinality of $I$ is of order $U(t)$.

The combination of all these steps suggests that $M > t$ occurs most likely because a single $X_n$ is large, that conditionally on this event, $S_i$ is about $\psi_{i,n}(X_n)$, and that the maximum of the process will indeed exceed $t$ if some $\psi_{i,n}(X_n)$ exceeds $t$, that is if $\psi_n(X_n)$ does. In fact, we will prove rigorously in the next subsection that completion of steps 1–6, in other words (S1)–(S6), implies the upper bound

$$\limsup_{t \to \infty} \frac{1}{r(t)} P\{ M > t \} \leq 1.$$

To obtain a matching lower bound, additional knowledge seems needed for the following reason. Let $\delta$ be a positive real number. If $\psi_n(X_n)$ exceeds $(1 + 2\delta)t$, there exists an integer $i$ such that

$$\psi_{i,n}(X_n) > (1 + \delta)t. \quad (3.1.5)$$

We would like to use step 6 to prove that, perhaps up to intersecting with a further set,

$$S_i \geq \psi_{i,n}(X_n) - \delta t > t$$

and so $M$ exceeds $t$, suggesting that $\sum_{n \in I} F^\circ \psi_n^{-1}(t)$ is an asymptotic lower bound for the probability that the process reaches the level $t$ at some time. The problem with this approach is that nothing guarantees that the $i$ involved in (3.1.5) stays of order $U(t)$, and, therefore, that step 6 gives the needed law of large numbers on the proper range of $i$. Various assumptions could be made to remove this difficulty. In some cases we may adapt an argument due to Zachary (2004) while in others the following may do.

**Step 7.** Let $i(n, x)$ be an integer which maximizes $\psi_{i,n}(x)$. Prove that for any positive $\delta$ and $\theta$ less than 1, there exists a positive $\eta$ such that for any $t$ large enough,

$$\{ i(n, x) : \theta \chi(t) \leq x \leq \chi(t)/\theta, \ n \in I_{\delta, t} \} \subset I_{\eta, t}. \quad (S7)$$
We will show in the next subsection that if in addition to the previous steps this last one can be completed then

$$\liminf_{t \to \infty} \frac{1}{r(t)} P\{ M > t \} \geq 1.$$ 

Therefore, once all seven steps have been verified we obtain

$$P\{ M > t \} \sim U(t) \bar{F} \circ \chi(t) \int_0^\infty \rho(u)^{-\alpha} \, du$$

as $t$ tends to infinity.

**Remarks.** While we defined $\psi_{i,n}$ as a conditional expectation, and from there $\psi_n$, we could as well have defined those two functions in a more axiomatic way, with no connection to conditional expectation, as functions which allows us to carry out steps 1–6 if we are seeking only an upper bound, or 1–7 if we are seeking an asymptotic equivalent for the tail probability of $M$.

In step 1, we assumed that $\psi_n$ is invertible on the preimage of some interval $(t_0, \infty)$. This assumption could be replaced by a weaker one using asymptotic inverse; however, that requires some form of uniformity with respect to $n$ in the asymptotic inversion. While technically possible, such refinement does not seem relevant in applications.

### 3.2. From the heuristic to a theorem.

The previous subsection sketched a possible path to obtain an asymptotic equivalent of the probability that the process reaches a large level at some time. In this subsection, we prove rigorously that this scheme, if it can be completed, indeed yields an asymptotic equivalent of the probability that $M$ exceeds $t$. Given its unsightly assumptions, one may be skeptical that the following theorem is of any value, but the next sections will demonstrate that its virtue is to break somewhat complicated problems into bits far more tractable. In particular, this theorem will be used to prove the results given in section 2.

**Theorem 3.2.1.** Referring to the previous subsection, if steps 1–7 have been completed, that is, if (S1)–(S7) hold, then

$$P\{ M > t \} \sim r(t)$$

as $t$ tends to infinity.

Needless to say that the function $r$ in this statement refers to that defined in (3.1.2).
Remark. The proof of Theorem 3.2.1 shows that under (S1)–(S5) and the one-sided version of (S6) with $D_n$ replaced by

$$D_n = \bigcap_{i \leq N} \{ S_i - \psi_{i,n} (X_n) \leq \delta t \}$$

then

$$\limsup_{t \to \infty} P\{ M > t \} / r(t) \leq 1.$$ 

If in addition the two sided version of (S6) holds as well as (S7) then

$$\liminf_{t \to \infty} P\{ M > t \} / r(t) \geq 1.$$ 

Proof. We first derive an upper bound for the probability that $M$ exceeds $t$ under (S1)–(S6), and, with the addition of (S7), a matching lower bound.

Upper bound. We set

$$A_i = \{ S_i > t \}.$$ 

Let $\delta$ be a positive real number. Steps 3 and 4 show that we can find a positive $\epsilon$ such that, ultimately in $t$,

$$P\{ M > t \} \leq P \left( \bigcup_{i \in I} A_i \right) + \delta r(t).$$ 

Using step 5, find $\theta$ such that, for any $t$ large enough,

$$P \left( \bigcup_{i \in I} \left( A_i \cap \bigcap_{n \leq i} B_i \right) \right) \leq \delta r(t).$$ 

Then, the equality

$$\bigcup_{i \in I} A_i = \left( \bigcup_{i \in I} \left( A_i \cap \bigcap_{n \leq i} B_n \cap \bigcap_{n \neq i} B_i \right) \right) \cup \left( \bigcup_{i \in I} \left( A_i \cap \bigcap_{n \leq i} B_n \cap \bigcap_{n \neq i} B_i \right) \right)$$ 

yields, for any $t$ large enough,

$$P\{ M > t \} \leq P \left( \bigcup_{i \in I} A_i \cap \bigcap_{n \leq N} B_n^c \right) + 2\delta r(t). \quad (3.2.1)$$ 

Consider the event that all random variables before $N$, except perhaps $X_n$, are at most $\theta \chi(t)$, that is

$$C_n = \bigcap_{1 \leq i \leq N, i \neq n} B_i.$$ 

31
The event $B_n^c \cap C_n$ expresses that all the $X_i, i \leq N$, but $X_n$ are at most $\theta \chi(t)$. We have the identity

$$A_i \cap \bigcup_{n \leq N} B_n^c = A_i \cap \bigcup_{n \leq N} \left( (B_n^c \cap C_n) \cup (B_n^c \cap C_n^c) \right) = \bigcup_{n \leq N} (A_i \cap B_n^c \cap C_n) \cup \bigcup_{n \leq N} (A_i \cap B_n^c \cap C_n^c).$$

But

$$B_n^c \cap C_n^c = B_n^c \cap \bigcup_{i \neq n} B_i^c = \bigcup_{i \neq n} B_n^c \cap B_i^c,$$

and we saw after step 5 that $\bigcup_{n,i \leq N} B_n^c \cap B_i^c$ has probability $o(r(t))$.

Thus, for $t$ large enough, (3.2.1) shows that

$$P \{ M > t \} \leq P \left( \bigcup_{i \in I} \bigcup_{n \leq N} A_i \cap B_n^c \cap C_n \right) + 3\delta r(t). \quad (3.2.2)$$

We then consider the events

$$D_n = \bigcap_{i \leq N} \{ S_i - \psi_{i,n}(X_n) \leq \delta t \}.$$

Completion of step 6 — in fact, the one-sided version would suffice here — ensures that

$$P \left( \bigcup_{n \in I} B_n^c \cap C_n \cap D_n^c \right) = o(r(t)) \quad (3.2.3)$$

as $t$ tends to infinity. Thus, (3.2.2) implies that for any $t$ large enough,

$$P \{ M > t \} \leq P \left( \bigcup_{i \in I} \bigcup_{n \leq N} A_i \cap B_n^c \cap C_n \cap D_n \right) + 4\delta r(t).$$

On $A_i \cap D_n$,

$$t < S_i < \delta t + \psi_{i,n}(X_n).$$

Thus, we proved that, ultimately in $t$,

$$P \{ M > t \} \leq P \left( \bigcup_{i \in I} \bigcup_{n \leq N} \{ \psi_{i,n}(X_n) > (1 - \delta) t \} \right) + 4\delta r(t)$$

$$\leq \sum_{n \geq 1} P \{ \max_{i \geq 1} \psi_{i,n}(X_n) > (1 - \delta) t \} + 4\delta r(t)$$

$$\leq \sum_{n \geq 1} F \circ \psi_n^{-1}((1 - \delta)t) + 4\delta r(t)$$

$$\leq r((1 - \delta)t) + 5\delta r(t).$$

32
Since $\delta$ is arbitrary and $r$ is regularly varying, it follows that
\[
\limsup_{t \to \infty} \frac{1}{r(t)} P\{ M > t \} \leq 1.
\]

**Lower bound.** Let $\epsilon$ be a positive real number. Using step 2, let $\delta$ be such that
\[
(1 - \epsilon)r(t) \leq \sum_{n \in I_{\delta,t}} F \circ \psi_n^{-1}(t) \leq (1 + \epsilon)r(t)
\]
ultimately. Let $\theta$ be small enough so that $\theta^{\alpha}/\delta \leq \epsilon$. Consider the events
\[
F_n = \{ \psi_n(X_n) > (1 + \delta)t; \theta\chi(t) < X_n \leq \chi(t)/\theta \},
\]
and let $B_n$, $C_n$ be the same events as defined previously, with $C_n$ defined with reference to the set $I_{\eta,t}$ obtained from step 7. Let $N = [U(t)/\eta]$ be the largest element of that $I_{\eta,t}$ and redefine $D_n$ to be
\[
D_n = \bigcap_{i \leq N} \{ S_i - \psi_{i,n}(X_n) \geq -\delta t \}.
\]
Recall the notation $i(n, x)$ introduced in step 7. For $n$ in $I_{\delta,t}$ and on $F_n \cap D_n$,
\[
\psi_{i(n, X_n), n}(X_n) > (1 + \delta)t
\]
and
\[
S_{i(n, X_n)} \geq \psi_{i(n, X_n), n}(X_n) - \delta t > t.
\]
Thus,
\[
P\{ M > t \} \geq P\{ \bigcup_{n \in I_{\delta,t}} F_n \cap B_n^c \cap C_n \cap D_n \}
\geq P\{ \bigcup_{n \in I_{\delta,t}} F_n \} - P\{ \bigcup_{n \in I_{\delta,t}} F_n \cap (B_n^c \cap C_n \cap D_n)^c \}.
\]
(3.2.4)

We consider the event
\[
(B_n^c \cap C_n \cap D_n)^c = B_n \cup (B_n^c \cap C_n^c) \cup (B_n^c \cap C_n \cap D_n^c).
\]
Note that $F_n \cap B_n = \emptyset$ for $\theta\chi(t) < X_n$ in $F_n$ while $X_n \leq \theta\chi(t)$ on $B_n$. Recall that the event $C_n$ is defined with reference to the set $I_{\eta,t}$ obtained from step 7, while (3.2.4) involves the different set $I_{\delta,t}$.  

33
Taking $\eta$ to be at most $\delta$, which can be done without any loss of generality, guarantees

$$\bigcup_{n \in I_{\delta,t}} (B_n \cap C_n^c) \subset \bigcup_{n \in I_{\eta,t}} (B_n \cap C_n^c)$$

and, as mentioned after step 5, the event in the right hand side of this inclusion had probability $o(r(t))$ as $t$ tends to infinity. Finally, by step 6, one has

$$P\left\{ \bigcup_{n \in I_{\eta,t}} (B_n \cap C_n \cap D_n) \right\} = o(r(t))$$
as $t$ tends to infinity. Therefore, (3.2.4) yields

$$P\{ M > t \} \geq P\left\{ \bigcup_{n \in I_{\delta,t}} F_n \right\} + o(r(t)). \quad (3.2.5)$$

Given step 1, we can also take $\theta$ small enough so that $\psi_n(X_n) > (1 + \delta)t$ and $n \in I_{\delta,t}$ guarantees $X_n > \theta \chi(t)$. Then (3.2.5) implies

$$P\{ M > t \} \geq \sum_{n \in I_{\delta,t}} F \circ \psi_n^{-1}\left( (1 + \delta)t + r I_{\delta,t} \right) \chi(t)/\theta + o(r(t)).$$

Thus, given our choice of $\delta$, we obtain for any $t$ large enough

$$P\{ M > t \} \geq (1 - \epsilon)r((1 + \delta)t) - \theta^\alpha\left( \frac{1}{\delta} - \delta \right) U(t) \chi(t) (1 + o(1)) + o(r(t)).$$

Since $\theta^\alpha/\delta$ is at most $\epsilon$ and $\epsilon$ is arbitrary, regular variation of $r$ yields

$$\liminf_{t \to \infty} \frac{1}{r(t)} P\{ M > t \} \geq 1.$$

### 3.3. Analysis of the paths leading to a large maximum.

The purpose of this subsection is to examine the likely trajectories of the process which lead to a large maximum, in the same formal framework as in the previous subsection. That is, we are seeking for the limiting distribution of the process $(S_n)_{n \geq 0}$ conditionally on $M$ exceeding $t$, as $t$ tends to infinity. Clearly the process needs to be rescaled to avoid degeneracy. The right rescaling is suggested by the proof of Theorem 3.2.1 and that proof also suggests introducing other random variables of interests. We define $N_t$ to be the first time
that $S_n$ exceeds $t$ and $J_t$ the index of the largest random variable $X_i$ among $X_1, \ldots, X_{N_t}$, that is
\[ X_{J_t} = \max_{1 \leq i \leq N_t} X_i, \]
with the convention that $J_t$ is minimal in case of ties. The proof of Theorem 3.2.1 suggests that $J_t$ is of order $U(t)$ while $X_{J_t}$ is of order $\chi(t)$. Thus, it is natural to introduce the random variables
\[ \tau_t = J_t/U(t) \quad \text{and} \quad Y_t = X_{J_t}/\chi(t), \]
as well as the rescaled process
\[ \mathfrak{S}_t(\lambda) = S_{\lfloor \lambda U(t) \rceil}/t. \]
This process belongs to the space $D[0, \infty)$ of all real-valued càdlàg functions endowed with the projective topology inherited from the Skorokhod topology on $D[0, 1/\epsilon]$ for any positive $\epsilon$ (see e.g. Billingsley, 1968, chapter 3; Pollard, 1984, chapter 6). Step 6 of subsection 3.1 also suggests that $S_{\lfloor \lambda U(t) \rceil}$ should be about $\psi_{\lfloor \lambda U(t) \rceil, \lfloor \tau U(t) \rceil}(\chi(t)Y_t)$. Therefore, for the process to converge it is natural to assume that there is a function $h$ on $[0, \infty)^3$ such that for any $\lambda, \tau$ and $y$,
\[ \lim_{t \to \infty} t^{-1} \psi_{\lfloor \lambda U(t) \rceil, \lfloor \tau U(t) \rceil}(\chi(t)y) = h(\lambda, \tau, y). \]
This pointwise convergence is not sufficient to guarantee the convergence in distribution of the process $\mathfrak{S}_t$ in $D[0, \infty)$. To strengthen it, set
\[ h_t(\lambda, \tau, y) = t^{-1} \psi_{\lfloor \lambda U(t) \rceil, \lfloor \tau U(t) \rceil}(\chi(t)y). \]
We assume that
\[ (\tau, y) \mapsto h_t(\cdot, \tau, y) \quad \text{and} \quad (\tau, y) \mapsto h(\cdot, \tau, y) \]
are measurable, and, for Lebesgue almost all $(\tau, y)$ in $[0, \infty)^2$, the functions $h_t(\cdot, \tau, y)$ converge to $h(\cdot, \tau, y)$ in $D[0, \infty)$ as $t$ tends to infinity.

Though this is not important for our purpose, assumption (3.3.1) is not independent of (S1) and there is a somewhat complicated though explicit relation between the functions $h$ and $\rho$. 

35
The following result describes the most likely trajectories of the process leading to a large maximum.

**Theorem 3.3.1.** Under (S1)–(S7) and (3.3.1), the conditional distributions of \((\xi_t, \tau_t, Y_t)\) given \(M > t\) converges weakly* to the distribution of \((\xi, \tau, \rho(\tau)Y)\) where \(\tau\) and \(Y\) are independent and

(i) \(Y\) has a Paréto distribution on \([1, \infty)\) with parameter \(\alpha\),

(ii) \(\tau\) has density proportional to \(\rho^{-\alpha}\),

(iii) \(\xi(\lambda) = h(\lambda, \tau, \rho(\tau)Y)\).

**Proof.** The proof requires establishing a couple of lemmas, describing the limiting behavior of \((\tau_t, Y_t)\) given that \(M\) exceeds \(t\), as \(t\) tends to infinity.

**Lemma 3.3.2.** Under the assumption of Theorem 3.3.1, for any nonnegative \(u\) and \(y\)

\[
\lim_{t \to \infty} P\{\tau_t \leq u; Y_t > y \mid M > t\} = \frac{\int_0^u (y \lor \rho(v))^{-\alpha} \, dv}{\int_0^\infty \rho(v)^{-\alpha} \, dv}.
\]

**Proof.** Let \(\epsilon\) and \(\eta\) be two positive real numbers. Referring to the sets introduced in subsection 3.1 and 3.2 and with \(N = [U(t)/\epsilon]\), consider the event

\[
F = \bigcup_{i \in I} \bigcup_{1 \leq n \leq N} A_i \cap B_n^c \cap C_n \cap D_n,
\]

whose dependence on \(t\), \(\epsilon\) and \(\theta\) is not kept track of. Note that if \(\theta\) is small enough and \(t\) is large enough and if \(F\) occurs, the proof of the upper bound of Theorem 3.2.1 shows that

\[
P(\{M > t\} \setminus F) \leq 4\eta \rho(t).
\]

Therefore, ultimately,

\[
P\{\tau_t \leq u; Y_t > y; M > t\} \leq P\{\tau_t \leq u; Y_t > y; M > t; F\} + 4\eta \rho(t). \tag{3.3.2}
\]

We can also write \(F\) as

\[
F = \bigcup_{1 \leq n \leq N} \left(\bigcup_{i \in I} A_i \cap B_n^c \cap C_n \cap D_n\right).
\]
Take $\theta$ smaller than $y$ and $\epsilon$ sufficiently small so that $u$ lies between $\epsilon$ and $1/\epsilon$. If $\tau_t \leq u$ and $Y_t > y$ and $B_n^c$ occur and if $J_t \neq n$, then the two distinct random variables $X_{J_t}$ and $X_n$ exceed $\theta \chi(t)$ and both $J_t$ and $n$ are at most $N$. But we have seen after (S5) that the probability for two distinct $X_i$ with $1 \leq i \leq N$ to exceed $\theta \chi(t)$ is $o(r(t))$ as $t$ tends to infinity. Therefore, (3.3.2) is ultimately at most $5\eta r(t)$ plus

$$P\left\{ \bigcup_{1 \leq n \leq N} \left( \{ \tau_t \leq u; J_t = n; Y_t > y; M > t \} \cap \bigcup_{i \in I} A_i \cap B_n^c \cap C_n \cap D_n \right) \right\}.$$

Using Bonferroni’s inequality, ultimately, this is at most

$$\sum_{1 \leq n \leq uU(t)} P\{ X_n > y \chi(t) ; \bigcup_{i \in I} A_i \cap D_n \} + 5\eta r(t).$$

Next, on $A_i \cap D_n$,

$$t < S_i \leq \delta t + \psi_{i,n}(X_n).$$

Therefore, taking $\epsilon$ small enough so that $u$ is between $\epsilon$ and $1/\epsilon$, the right hand side in (3.3.2) is at most

$$\sum_{1 \leq n \leq uU(t)} P\left\{ X_n > y \chi(t) ; \max_{i \in I} \psi_{i,n}(X_n) > (1 - \delta)t \right\} + 5\eta r(t).$$

Since

$$\sum_{1 \leq n \leq uU(t)} P\{ X_n > y \chi(t) \} \sim c y^{-\alpha}(U \mathcal{T} \circ \chi)(t),$$

as $t$ tends to infinity, we see that, provided $\epsilon$ is small enough, (3.3.2) is at most

$$\sum_{\epsilon U(t) \leq n \leq uU(t)} P\left\{ X_n > y \chi(t) ; \max_{i \in I} \psi_{i,n}(X_n) > (1 - \delta)t \right\} + 6\eta(U \mathcal{T} \circ \chi)(t),$$

that is, at most

$$\sum_{\epsilon U(t) \leq n \leq uU(t)} P\left\{ X_n > y \chi(t) ; \psi_n(X_n) > (1 - \delta)t \right\} + 6\eta(U \mathcal{T} \circ \chi)(t). \quad (3.3.3)$$
Note that (S1) implies
\[ \psi^{-1}_n((1 - \delta)t) \sim \chi((1 - \delta)t) \rho((1 - \delta)^{-1/\gamma}u). \]
Recall we set \( \kappa \) for the index of regular variation of \( \chi \). The same argument used to derive (3.1.2), that is, regular variation and comparison to a Riemann integral, shows that the sum in (3.3.3) is asymptotically equivalent to
\[ (UF \circ \chi)(t) \int_{\epsilon}^u (y \vee (1 - \delta)^\kappa \rho((1 - \delta)^{-1/\gamma}v))^\alpha dv. \]
We let \( \epsilon \), then \( \delta \) and then \( \eta \) tend to 0 to obtain
\[ \limsup_{t \to \infty} \frac{P\{ \tau_t \leq u; Y_t > y; M > t \}}{(UF \circ \chi)(t)} \leq \int_{0}^u (y \vee \rho(v))^{-\alpha} dv. \]
To obtain a matching lower bound, we refer to how we proved the lower bound of Theorem 3.2.1. In particular, keeping the notation of that proof and remembering that the sets \( B_n^c \cap C_n \) are disjoint for different values of \( n \), we see that
\[ P\{ \tau_t \leq u; Y_t > y; M > t \} \]
is at least (cf. (3.2.4))
\[
\sum_{n \in I_{\delta,t}} \sum_{n \leq uU(t)} P\{ J_t = n; X_n > y; F_n \cap B_n^c \cap C_n \cap D_n \} \geq \sum_{n \in I_{\delta,t}} P\{ X_n > y; \psi^{-1}_n((1 + \delta)t) \} - \#I_{\delta,t} F(\chi(t)/\theta) + o(r(t)) \tag{3.3.4}
\]
as \( t \) tends to infinity. Since (S1) implies that for \( n \) in \( I_{\delta,t} \)
\[ \psi^{-1}_n((1 + \delta)t) \sim \chi((1 + \delta)t) \rho(n/U((1 + \delta)t)) \]
and \( \chi \) is regularly varying with index \( \kappa \), the sum in (3.3.4) is asymptotically equivalent to
\[
\int_{\delta U(t)}^{uU(t)} F\left(\chi((1 + \delta)t) \rho\left(\frac{s}{U((1 + \delta)t)}\right) \vee \chi(t)y\right) ds
\sim \int_{\delta}^{u} F\left(\chi(t)((1 + \delta)^\kappa \rho((1 + \delta)^{-1/\gamma}v) \vee y)\right) U(t) dv
\sim (UF \circ \chi)(t) \int_{\delta}^{u} ((1 + \delta)^\kappa \rho((1 + \delta)^{-1/\gamma}v) \vee y)^{-\alpha} dv
\]
38
as $t$ tends to infinity. Since $\delta$ can be made as small as desired, it follows that

$$\liminf_{t \to \infty} P\{ \tau_t \leq u ; Y_t > y ; M > t \} \geq \int_0^u (y \vee \rho(v))^{-\alpha} \, dv$$

and this completes the proof.

The next lemma gives a simple representation of random variables having the limiting distribution written in Lemma 3.3.2.

**Lemma 3.3.3.** Let $\tau$ be a random variable having density proportional to $\rho^{-\alpha}$, and let $Y$ be a random variable independent of $\tau$, having a Pareto distribution of index $\alpha$ on $[1, \infty)$. Then

$$P\{ \tau \leq u ; \rho(\tau) Y > y \} = \frac{\int_0^u (y \vee \rho(v))^{-\alpha} \, dv}{\int_0^\infty \rho(v)^{-\alpha} \, dv}.$$

**Proof.** We simply write

$$P\{ \tau \leq u ; \rho(\tau) Y > y \} \int_0^\infty \rho(v)^{-\alpha} \, dv$$

$$= \int_0^u P\{ Y > y/\rho(v) \} \rho(v)^{-\alpha} \, dv$$

$$= \int_0^u \left( \frac{y}{\rho(v)} \vee 1 \right)^{-\alpha} \rho(v)^{-\alpha} \, dv$$

$$= \int_0^u (y \vee \rho(v))^{-\alpha} \, dv. \quad \blacksquare$$

We can now conclude the proof of Theorem 3.3.1. Let $\epsilon$ be a positive real number and consider the event

$$G = \{ \exists \lambda \leq 1/\epsilon : |\xi_t(\lambda) - h_t(\lambda, \tau_t, Y_t)| > \delta \}.$$

Recall that the event $F$ introduced in the proof of the upper bound related to Lemma 3.3.2 depends on a parameter $\theta$ through the events $B_n$ and $C_n$. As we saw in that proof, for any fixed $\eta$ and for any $\theta$ small enough

$$P(G ; M > t) \leq P(G \cap F ; M > t) + 4\eta r(t)$$

39
ultimately. If $F$ occurs, then $N_t$ is at most $N$ and, also, $\bigcup_{1 \leq n \leq N} B_n^c \cap C_n \cap D_n$ occurs. Again, this union is disjoint for the sets $B_n^c \cap C_n$ are disjoint for different values of $n$. If $B_n^c \cap C_n$ occurs, it is tempting to conclude that $J_t = n$. Since $J_t$ is at most $N_t$, this is true provided that $N_t$ is at least $n$. We now show that assumption (S5) guarantees that this is likely to be the case. Clearly, from the definition of the event $F$,

$$F \subset \bigcup_{1 \leq n \leq N} \bigcup_{i \leq n} A_i \cap B_n^c \cap C_n \cap D_n \quad \bigcup \bigcup_{1 \leq n \leq N} \bigcup_{n \leq i \leq N} \left( A_i \cap B_n^c \cap C_n \cap D_n \cap \left( \bigcup_{j < n} A_j \right)^c \right).$$

Note that if $A_i \cap B_n^c \cap C_n \cap D_n$ occurs, then so does $A_i \cap C_n$. In this case, if $i$ is less than $n$ then the definition of $C_n$ shows that $A_i \cap \bigcap_{1 \leq j \leq i} B_j$ occurs as well. But (S5) combined with (S3) imply that $\bigcup_{1 \leq i \leq N} (A_i \cap \bigcap_{1 \leq j \leq i} B_j)$ occurs with probability at most $2\delta r(t)$ provided $\epsilon$ is small enough and $t$ is large enough. Therefore,

$$P(G; M > t) \leq P \left( G \cap \bigcup_{1 \leq n \leq N} \bigcup_{n \leq i \leq N} \left( A_i \cap B_n^c \cap C_n \cap D_n \cap \left( \bigcup_{j < n} A_j \right)^c \right) \right) + 6\eta r(t).$$

Considering the events involved in this upper bound, if $A_i \cap (\bigcup_{j < n} A_j)^c$ occurs and $i$ is at least $n$, then $N_t$ is at least $n$ and, as announced, $J_t = n$; moreover, if $D_n$ occurs, then $|S_i - \psi_{i,n}(X_n)|$ is at most $\delta t$ for any $i$ at most $N$. In that case, uniformly in $\lambda$ in $[0, 1/\epsilon]$,

$$|t^{-1} S_{[\lambda U(t)]} - t^{-1} \psi_{[\lambda U(t)]}, J_t(X_J_t)| \leq \delta,$$

that is,

$$|\mathcal{S}_t(\lambda) - h_t(\lambda, \tau_t, Y_t)| \leq \delta,$$

and so $G$ does not occur. Therefore,

$$\lim_{t \to \infty} P(G \mid M > t) = 0. \quad (3.3.5)$$

Recall that the Skorohod topology is metric. Combining Lemmas 3.3.2, 3.3.3 and Theorem 5.5 in Billingsley (1968) upon using (3.3.1), imply that the conditional distribution of $h_t(\cdot, \tau_t, Y_t)$ given $M > t$ converges weakly to that of $h(\cdot, \tau, \rho(\tau) Y)$ as $t$ tends to infinity. The result then follows from (3.3.5) which asserts that, on compact sets, $\mathcal{S}_t - h_t(\cdot, \tau_t, Y_t)$ converges uniformly to 0 in probability under
the conditional probability given \( M > t \) as \( t \) tends to infinity, and uniform convergence on compact sets implies convergence in \( D[0, \infty] \) under the Skorohod metric.

4. A large deviation inequality and a Karamata type theorem. The folklore attributes to Kolmogorov that behind every limit theorem there is an inequality. The purpose of this short section is to derive the large deviation inequality behind some of our theorems as well as to state a Karamata type theorem which we will be needing.

4.1. A large deviation inequality. The result of this subsection is of a more technical nature. It provides a bound on the moment generating function of a centered random variable truncated from above. Its use is explained after its proof, and it will be instrumental to show that some probabilities tend to 0. It is inspired from a technique used in Cline and Hsing (1991) as well as Ng, Tang, Yan and Yang (2004).

**Lemma 4.1.1.** Let \( Z \) be a centered random variable with distribution function \( H \). For any positive \( \lambda \) and \( a \), for any positive \( \eta \) less than 1,

\[
\log E \exp(\lambda Z \mathbb{1}\{ Z \leq a \}) \leq \eta \lambda E |Z| - \lambda E Z \mathbb{1}\{ \lambda Z \leq \log(1 - \eta) \} + e^{\lambda a H'(\frac{\log(1 + \eta)}{\lambda})}.
\]

**Proof.** Let \( M(\lambda) \) be the expected value of \( \exp(\lambda Z \mathbb{1}\{ Z \leq a \}) \). The inequality \( 1 + x \leq e^x \) yields

\[
M(\lambda) = 1 + \int_{-\infty}^{a} e^{\lambda z} - 1 \, dH(z) \\
\leq \exp \left( \int_{-\infty}^{a} e^{\lambda z} - 1 \, dH(z) \right).
\]

Since \( e^x - 1 \) is nonnegative and at most \( (1 + \eta)x \) on \( [0, \log(1 + \eta)] \)
and is nonpositive and at most \((1 - \eta)x\) on \([\log(1 - \eta), 0]\), we have
\[
\log M(\lambda) \leq \int_{-\infty}^{\log(1 + \eta)} e^{\lambda z} - 1 \, dH(z)
\]
\[
\leq (1 + \eta) \int_{0}^{\infty} \lambda z \mathbb{1}\{\lambda z \leq \log(1 + \eta)\} \, dH(z)
\]
\[
+ (1 - \eta) \int_{-\infty}^{0} \lambda z \mathbb{1}\{\log(1 - \eta) \leq \lambda z\} \, dH(z)
\]
\[
+ e^{\lambda a} \frac{\log(1 + \eta)}{\lambda}.
\]
Since \(Z\) is centered,
\[
\int_{\mathbb{R}} \lambda z \mathbb{1}\{\log(1 - \eta) \leq \lambda z \leq \log(1 + \eta)\} \, dH(z)
\]
\[
\leq -\int_{-\infty}^{\log(1 - \eta)/\lambda} \lambda z \, dH(z),
\]
and the result follows.

We will use Lemma 4.1.1 in the following situation. Consider a sequence \((a_j)_{j \geq 1}\) of positive real numbers and a sequence of independent and equidistributed and centered random variables \((Z_j)_{j \geq 1}\). We write \(Z\) for a random variable having the same distribution as \(Z_1\). Substituting \(\lambda\) with \(\lambda g_j\) and \(a\) with \(a_{n-j}\) in Lemma 4.1.1 we obtain the Chernoff type inequality, valid for all positive \(\eta\) less than 1, any positive \(\lambda\), all sequences of positive reals \(a_j\), nonnegative \(g_j\), and any real numbers \(t\) and \(s_n\) such that \(t - s_n\) is positive,
\[
\log P\left\{ \sum_{0 \leq j < n} g_j Z_{n-j} \mathbb{1}\{Z_{n-j} \leq a_{n-j}\} > t - s_n \right\}
\]
\[
\leq -\lambda(t - s_n) + \eta \lambda g_{[0,n)} E|Z| - \lambda \sum_{0 \leq j < n} g_j EZ \mathbb{1}\{\lambda g_j Z \leq \log(1 - \eta)\}
\]
\[
+ \sum_{0 \leq j < n} e^{\lambda g_j a_{n-j}} \frac{\log(1 + \eta)}{\lambda g_j}.
\]
(4.1.1)
We will take \(\lambda\) small in this bound. The a priori strange formulation of this inequality, using a rather mysterious \(t - s_n\) instead of a single variable is on purpose and designed to make the remainder of this paper easier to read.
4.2. A Karamata type theorem. The following result is an easy extension of the direct half of Karamata’s theorem (see Bingham, Goldie and Teugels, 1989, Proposition 1.5.10). Recall that if $b$ is a function with limit infinity at infinity, then, if it exists, $\overleftarrow{b}$ is an asymptotic inverse of $b$, that is, a function such that $b \circ \overleftarrow{b} \sim \overleftarrow{b} \circ b \sim \text{Id}$ at infinity.

**Lemma 4.2.1.** Let $b$ be a regularly varying function of positive index $\beta$. If $\alpha > 1 \lor (1/\beta)$ then for any positive real number $r$

$$\int_r^\infty \overline{F}(t + b(u)) \, du \sim \frac{1}{\beta} B\left(\frac{1}{\beta}, \alpha - \frac{1}{\beta}\right)(\overleftarrow{b} \circ \overline{F})(t)$$

as $t$ tends to infinity.

**Proof.** Let $\epsilon$ be a positive real number. The change of variable $u = \lambda \overleftarrow{b}(t)$ and regular variation of $\overline{F}$ and $b$ show that

$$\int_{\epsilon \overleftarrow{b}(t)}^{\epsilon \overleftarrow{b}(t)/\epsilon} \overline{F}(t + b(u)) \, du \sim (\overleftarrow{b} \circ \overline{F})(t) \int_\epsilon^{1/\epsilon} (1 + \lambda^\beta)^{-\alpha} \, d\lambda.$$

Next, since $b$ is ultimately positive, tends to infinity at infinity, so does $\overleftarrow{b}$. Therefore, by the monotonicity of $\overline{F}$,

$$\int_r^{\epsilon \overleftarrow{b}(t)} \overline{F}(t + b(u)) \, du \leq \epsilon (\overleftarrow{b} \circ \overline{F})(t) (1 + o(1))$$

as $t$ tends to infinity. Furthermore, by Karamata’s theorem and monotonicity of $\overline{F}$,

$$\int_{b^{-}(t)/\epsilon}^{\infty} \overline{F}(t + b(u)) \, du \leq \int_{b^{-}(t)/\epsilon}^{\infty} \overline{F} \circ b(u) \, du \sim \frac{\epsilon^{\alpha \beta - 1}}{\alpha \beta - 1} (\overleftarrow{b} \circ \overline{F})(t).$$

Since $\alpha \beta > 1$,

$$\int_r^{\infty} \overline{F}(t + b(u)) \, du \sim (\overleftarrow{b} \circ \overline{F})(t) \int_0^{\infty} (1 + \lambda^\beta)^{-\alpha} \, d\lambda$$

as $t$ tends to infinity. The change of variable $x = \lambda^\beta$ then yields the result. 

43
5. Some asymptotic analysis related to analytic functions. The purpose of this section is prove some purely analytical results related to analytic functions which will be needed to prove our results on the maximum of \((g, F)\)-processes and their trajectories.

In the first subsection we restate known results in a form suitable for our purpose. The second subsection introduces a family of functions, \(\Psi_n, n \geq 1\), associated to an analytic function. The notation is not fortuitous, for if \(S_n\) is a \((g, F)\)-process, then the function \(\psi_n(x) = \max_{i \geq 0} E(S_i|X_n = x)\) involved in our heuristic will be related to the function \(\Psi_n\) in a simple way. In order to apply the methodology presented in section 3, we need to have some information on \(\psi_n^{-1}\), and for this reason, we will obtain some basic information on \(\Psi_n^{-1}\). This will be done in the third subsection when the sequence of Taylor coefficients \((g_n)_{n \geq 0}\) tends to 0, in the fourth subsection when that sequence diverges toward infinity, and in the fifth subsection when \((g_n)_{n \geq 0}\) has a positive finite limit.

Throughout this subsection, \(g(x) = \sum_{i \geq 0} g_i x^i\) is a real analytic function on \((-1, 1)\), regularly varying at 1 with positive index \(\gamma\). In particular, \(\lim_{x \to 1; x < 1} g(x) = +\infty\). Recall that for any positive \(n\), we write \(g(0, n)\) for \(\sum_{0 \leq j < n} g_j\).

The following notation will save some unsightly \(\epsilon, t_0, n_0\) as well as various quantifiers.

**Notation.** Throughout this section, if \((a_n)\) and \((b_n)\) are two sequences, we say that ‘\(a_n\) is bounded from above by an equivalent of \(b_n\)’ and write \(a_n \preceq b_n\) if \(a_n \leq b_n (1 + o(1))\) as \(n\) tends to infinity, or, equivalently, if \(\limsup_{n \to \infty} a_n / b_n \leq 1\). Similarly we define in an obvious manner what it is to be asymptotically bounded from below, and write \(\succeq\) for this relation. Both relations are transitive.

**5.1. Preliminaries.** In this section we restate some known result in a form suitable for our analysis.

Our first lemma essentially restates Karamata’s Tauberian theorem for power series (Bingham, Goldie and Teugels, 1989, Corollary 1.7.3), and adds some uniformity to it. Recall that throughout this paper we assume that \((g_n)_{n \geq 0}\) is asymptotically equivalent to a monotone sequence.
Lemma 5.1.1. The following asymptotic equivalences hold as $n$ tends to infinity, uniformly in $x$ in any compact subset of the positive half-line,

(i) $g_{\lceil nx \rceil} \sim \frac{x^{\gamma-1} g(1-1/n)}{\Gamma(\gamma) n}$,

(ii) $g_{(0, nx)} \sim \frac{x^{\gamma}}{\Gamma(1+\gamma)} g(1-1/n)$.

Proof. For a fixed $x$, the result is Corollary 1.7.3 in Bingham, Goldie and Teugels (1989). Uniformity follows by the same proof or the following one. We note that for any $\epsilon$ positive,

$$\sup_{\epsilon \leq x \leq 1/\epsilon} \left| \frac{g_{\lceil nx \rceil} \lceil nx \rceil \Gamma(\gamma)}{g(1-1/\lceil nx \rceil)} - 1 \right| \leq \sup_{m \geq n} \left| \frac{g_m m \Gamma(\gamma)}{g(1-1/m)} - 1 \right|.$$

The pointwise version implies that this upper bound tends to 0 as $n$ tends to infinity. Since $\lceil nx \rceil/nx$ tends to 1 uniformly in $x$ in any compact set of the positive half-line, and by the uniform convergence theorem (Bingham, Goldie and Teugels, 1989, Theorem 1.5.2) $g(1-1/\lceil nx \rceil)/g(1-1/n)$ tends to $x^{\gamma}$ uniformly as well, it follows that

$$\lim_{n \to \infty} \sup_{\epsilon \leq x \leq 1/\epsilon} \left| \frac{ng_{\lceil nx \rceil}}{g(1-1/n)} - \frac{x^{\gamma-1}}{\Gamma(\gamma)} \right| = 0.$$

A similar argument proves the uniformity in the convergence of $g_{(0, nx)}/g(1-1/n)$.

5.2. The functions $\Psi_n$ and their inverses. In this subsection we develop some asymptotic estimates for some functions derived from a real analytic function on $(-1, 1)$ with a singularity at 1. Some of the results presented may be of independent interest and fit in the rich corpus of Tauberian theorems in the realm of analytic functions.

Recall that $U$ is a function which satisfies $g(1 - 1/U(t)) \sim t$ as $t$ tends to infinity, and that $g_n, n \geq 0$, is the sequence of the Taylor coefficients of $g$ at the origin. In this section, we do not assume that these coefficients are nonnegative, but that only finitely many of them may be negative. We also assume that

$$g^* = \sup_{n \geq 0} g_n \quad \text{is positive.}$$
We consider the following functions,

\[ \Psi_n(x) = 0 \lor \max_{k \geq 0} (g_k x - g_{[0,n+k]}), \quad n \geq 1. \]

Since Lemma 5.1.1 implies that \( \lim_{k \to \infty} g_{[0,n+k]} / g_k = \infty \), these functions are defined for all \( n \) and \( x \) nonnegative, that is to say, since the maximum is attained, it is proper to write a maximum instead of a supremum. Lemma 5.1.1 implies that there exists \( n_0 \) such that for any \( n \) at least \( n_0 \), both \( g_n \) and \( g_{[0,n]} \) are nonnegative. Clearly, if all the coefficients \( g_n \) are nonnegative, we can take \( n_0 \) to be 0. For \( n \) at least \( n_0 \), our first lemma gives another expression for \( \Psi_n \) and shows that this function is increasing and convex on the half-line where it is positive. For this purpose, we define

\[ K_+ = \{ k \in \mathbb{N} : g_k > 0 \}, \]

and

\[ x_n = \min_{k \in K_+} g_{[0,n+k]} / g_k. \]

**Lemma 5.2.1.** For any \( n \) at least \( n_0 \),

(i) \( \Psi_n \) vanishes on \([0,x_n]\);

(ii) \( \Psi_n \) is positive and increasing on \((x_n, \infty)\). Moreover, on the half-line \((x_n, \infty)\),

\[ \Psi_n(x) = \max_{k \in K_+} g_k x - g_{[0,n+k]}; \]

(iii) \( \Psi_n \) is continuous and convex on the nonnegative half-line;

(iv) If \( n_0 \leq n \leq m \), then \( \Psi_m \leq \Psi_n \) and \( x_n \leq x_m \).

**Proof.** (i) Let \( x \) be a nonnegative real number at most equal to \( x_n \). The definition of \( x_n \) implies that \( g_k x - g_{[0,n+k]} \) is nonpositive for any \( k \) in \( K_+ \). If \( k \) does not belong to \( K_+ \) then \( g_k x - g_{[0,n+k]} \) is nonpositive, for both \( g_k \) and \( -g_{[0,n+k]} \) are nonpositive. Therefore, \( \Psi_n \) vanishes at \( x \).

(ii) If \( x \) is larger than \( x_n \), then \( g_k x - g_{[0,n+k]} \) is positive for some \( k \) in \( K_+ \) and so is \( \Psi_n(x) \). Since \( g_k x - g_{[0,n+k]} \) is nonpositive for \( k \) not in \( K_+ \), this proves that \( \Psi_n \) has the representation given in (ii). This representation shows that \( \Psi_n \) is increasing on \( (x_n, \infty) \).

(iii) The representation obtained in (ii) and the proof of (i) show that \( \Psi_n(x) = 0 \lor \max_{k \in K_+} g_k x - g_{[0,n+k]} \) on the nonnegative half-line. As
the supremum of nondecreasing linear functions, \( \Psi_n \) is convex and therefore continuous.

(iv) If \( n_0 \leq n \leq m \) then \( g_{[0,n+k]} \leq g_{[0,m+k]} \) for any \( k \) and the result follows.  

For \( n \) at least \( n_0 \), the expression for \( \Psi_n \) in Lemma 5.2.1.ii shows that \( \lim_{x \to \infty} \Psi_n(x) = +\infty \). Lemma 5.2.1.ii–iii imply that \( \Psi_n \) is invertible as a map from \((x_n, \infty)\) to the positive half-line. Therefore, for \( n \) at least \( n_0 \), it is meaningful to define the inverse \( \Psi_n^{-1} \) on the positive half-line. We extend it to 0 by continuity, defining \( \Psi_n^{-1}(0) = x_n \). The following lemma provides an expression for that inverse.

**Lemma 5.2.2.** Let \( n \) be at least \( n_0 \). For any nonnegative \( t \),

\[
\Psi_n^{-1}(t) = \min_{k \in K_+} \frac{t + g_{[0,n+k]}}{g_k}.
\]

**Proof.** Since \( \Psi_n \) is invertible, for any positive \( t \) and any \( k \) in \( K_+ \),

\[
t = \Psi_n \circ \Psi_n^{-1}(t) \geq g_k \Psi_n^{-1}(t) - g_{[0,n+k]}.
\]

Therefore,

\[
\Psi_n^{-1}(t) \leq \min_{k \in K_+} \frac{t + g_{[0,n+k]}}{g_k}.
\]

To prove that this upper bound is sharp, assume that it is not, so that there exists a positive \( \epsilon \) such that for any positive \( g_k \),

\[
\Psi_n^{-1}(t) \leq \frac{t + g_{[0,n+k]}}{g_k} - \epsilon.
\]

Since \( \Psi_n \) is onto, there exists \( x \) such that \( \Psi_n(x) = t \). Then

\[
x = \Psi_n^{-1} \circ \Psi_n(x) \leq \frac{\Psi_n(x) + g_{[0,n+k]}}{g_k} - \epsilon,
\]

and therefore

\[
\Psi_n(x) \geq x g_k - g_{[0,n+k]} + \epsilon g_k.
\]

In particular, considering this inequality for a value of \( k \) which maximizes \( x g_k - g_{[0,n+k]} \), this last quantity being then equal to \( \Psi_n(x) \), we obtain

\[
\Psi_n(x) \geq \Psi_n(x) + \epsilon g_k.
\]

(5.2.1)
If $g_k$ were equal to 0 then $t = \Psi_n(x) = -g_{(0,n+k)}$ would be negative, for $n$ is at least $n_0$. Therefore, $g_k$ is positive and (5.2.1) yields $\Psi_n(x) > \Psi_n(x)$ which is a contradiction.

Given Lemma 5.2.2, we write $k_n(t)$ for an integer such that

$$\Psi^{-1}_n(t) = \frac{t + g_{(0,n+k_n(t))}}{g_{k_n(t)}}.$$  

Such an integer may not be unique, but whatever statement we will make about it will not depend on its particular choice.

5.3. Approximation of $\Psi_n^{-1}$ when $(g_n)_{n \geq 0}$ tends to 0. When the sequence $(g_n)_{n \geq 0}$ tends to 0 at infinity, the minimization involved in Lemma 5.2.2 can be made explicit for large argument $t$. For this purpose, let $k^*$ be the smallest integer at which the sequence $(g_n)_{n \geq 0}$ achieves its maximum; thus $k^*$ is the smallest integer for which $g_k = g^*$.

**Lemma 5.3.1.** There exists a nonnegative $t_0$ such that for any $n$ at least $n_0$ and any $t$ at least $t_0$,

$$\Psi^{-1}_n(t) = \frac{t + g_{(0,n+k^*)}}{g^*}.$$  

**Proof.** Given Lemma 5.2.2, it is clear that the proposed expression is an upper bound for $\Psi_n^{-1}$. If $k$ is in $K_+$ and larger than $k^*$, then $g^*/g_k \geq 1$ and $g_{(n+k^*,n+k)} \geq 0$. Therefore,

$$\frac{t + g_{(0,n+k)}}{g_k} = \frac{t + g_{(0,n+k^*)} g^*}{g_k} + \frac{g_{(n+k^*,n+k)}}{g_k} \geq \frac{t + g_{(0,n+k^*)}}{g^*}.$$  

If $k$ is in $K_+$ and less than $k^*$ we write $(t + g_{(0,n+k)})/g_k$ as the sum of $(t + g_{(0,n+k^*)})/g^*$ and

$$\frac{t + g_{(0,n+k^*)}}{g^*} \left(\frac{g^*}{g_k} - 1\right) = \frac{g_{(n+k,n+k^*)}}{g_k}.$$  

This quantity is positive if

$$t > g_{(n+k,n+k^*)} \frac{g^*}{g^* - g_k} - g_{(0,n+k^*)}.$$  

48
Since $k^*$ is minimal, the maximum of this lower bound over all $k$ in $K_+$ and less than $k^*$ and all $n \geq n_0$ is finite. Call $t_0$ its maximum.

It follows from Lemmas 5.1.1.ii and 5.3.1 that
\[ \Psi_n^{-1}(t) \sim \frac{t + g_{[0,n]}}{g^*} \] (5.3.1)
as $n$ tends to infinity, uniformly in $t$ in $[t_0, \infty)$. We then deduce the following asymptotic equivalence.

**Lemma 5.3.2.** For any $x$ and $c$ in any compact subset of the positive half-line, as $t$ tends to infinity,
\[ \Psi_{[xU(t)]}^{-1}(t) \sim \frac{t}{g^*} \left(1 + \frac{x^\gamma}{\Gamma(1 + \gamma)}\right), \]
and, for any $n$ at least $n_0$ and $t$ at least $t_0$, the equality $k_n(t) = k^*$ holds.

**Proof.** This is immediate from Lemmas 5.1.1 and 5.3.1.

Our next two lemmas provide some bounds for $\Psi_n^{-1}$.

**Lemma 5.3.3.** For any $n$ at least $n_0$ and any $t$ at least $t_0$,
\[ \Psi_n^{-1}(t) \geq \frac{t}{g^*}. \]

**Proof.** The result follows from the formula in Lemma 5.3.1 since $g_{[0,n+k^*]}$ is nonnegative.

**Lemma 5.3.4.** As $n$ tends to infinity
\[ \inf_{t \geq t_0} \Psi_n^{-1}(t) \geq \frac{g(1 - 1/n)}{g^* \Gamma(1 + \gamma)}. \]

**Proof.** This follows from Lemmas 5.1.1 and 5.3.1, since $t$ is at least $t_0$ and hence positive.

5.4. **Asymptotic analysis of $\Psi_n^{-1}$ when $(g_n)_{n \geq 0}$ tends to infinity.** The purpose of this subsection is to derive an asymptotic
equivalent for $\Psi_n^{-1}$ and $k_n(\cdot)$ when $n$ is of order $U(t)$ and the argument is of order $t$, as well as some bounds for this function when the sequence of coefficients, $(g_n)_{n \geq 0}$ tends to infinity.

When $\gamma$ is 1, we assume that

$$(g_n)_{n \geq 0} \text{ is asymptotically equivalent to an increasing}$$

sequence with limit $+\infty$ \hfill (5.4.1)

and moreover that

the Karamata representation of $g$ satisfies (2.1.4) and

(2.1.5). \hfill (5.4.2)

For any positive real number $a$, we introduce the functions

$$\xi_a(x) = \min_{y > 0} \frac{a + (x + y)\gamma}{\gamma y^{\gamma-1}}.$$ 

Since $\gamma$ is fixed throughout this section, the notation $\xi_a$ does not keep track of the dependence of this function on $\gamma$. In general an explicit form for the minimum cannot be found, however, if $\gamma$ is 1, then $\xi_a(x) = a + x$. With respect to Theorem 2.1.2, note that the function $\rho_\gamma$ is equal to $\xi_{\Gamma(1+\gamma)}$.

The following shows that the minimum involved in the definition of $\xi_a(x)$ is achieved at a unique point and that all the functions $\xi_a$ can be recovered from a knowledge $\xi_1$.

**Lemma 5.4.1.** (i) The function

$$y \mapsto \frac{a + (x + y)\gamma}{\gamma y^{\gamma-1}}$$

has a unique minimum on the positive half-line.  

(ii) If $a \leq b$, then $\xi_a \leq \xi_b$.

(iii) The identity $\xi_a(x) = a^{1/\gamma} \xi_1(a^{-1/\gamma} x)$ holds.

**Proof.** (i) The derivative of the function vanishes at the minimum. Thus, the minimizer $y$ satisfies

$$\left(\frac{x + y}{y}\right)^{\gamma-1} - \frac{\gamma - 1}{\gamma} \frac{a + (x + y)^\gamma}{y^{\gamma}} = 0.$$
Setting \( s = y/(x + y) \), this equation asserts that

\[
s = \frac{\gamma - 1}{\gamma} \left( a \frac{(1 - s)^{\gamma}}{x^{\gamma}} + 1 \right).
\]

The left hand side of this equality is increasing in \( s \) in \((0, 1)\] while the right hand side is decreasing. Therefore, the equality is achieved for a unique \( s \).

(ii) This is obvious.

(iii) We substitute \( y \) with \( a^{-1/\gamma} y \) in the definition of \( \xi_1(a^{-1/\gamma} x) \).

Knowledge of the behavior of \( \xi_a \) at the origin and at infinity will also be useful and some information is given now.

**Lemma 5.4.2.** The following hold for \( \gamma \) greater than 1.

(i) \( \xi_a(0) = a^{1/\gamma}(\gamma - 1)^{(1/\gamma) - 1} \).

(ii) \( \xi_0(1) = \left( \frac{\gamma - 1}{\gamma} \right)^{\gamma - 1} \).

(iii) \( \xi_a(x) \sim x\xi_0(1) \) as \( x \) tends to infinity.

Moreover, if \( \gamma = 1 \), then \( \xi_a(x) = a + x \) and (i)–(iii) hold provided they are extended by continuity as \( \gamma \) tends to 1.

**Proof.** Assume that \( \gamma \) is greater than 1. Standard calculus shows that the function to minimize to calculate \( \xi_a(0) \) achieves its minimum at \( y = a^{1/\gamma}(\gamma - 1)^{1/\gamma} \), while that to calculate \( \xi_0(1) \) achieves its minimum at \( y = \gamma - 1 \). Parts (i) and (ii) follow.

To prove (iii), Lemma 5.4.1.iii yields, with \( xa^{1/\gamma} \) in place of \( x \) and setting \( a = x^{-\gamma} \), the identity

\[
\xi_1(x) = a^{-1/\gamma} \xi_a(xa^{1/\gamma}) = x\xi_{x^{-\gamma}}(1).
\]

It suffices to prove that \( \xi_a(1) \) tends to \( \xi_0(1) \) as \( a \) tends to 0 from above. On the one hand, if \( a \) is positive, Lemma 5.4.1.ii shows that \( \xi_a \geq \xi_0 \), and on the other hand, with \( y = \gamma - 1 \),

\[
\xi_a(1) \leq \frac{a + \gamma}{\gamma(\gamma - 1)^{\gamma - 1}}
\]

and this upper bound tends to \( \xi_0(1) \) as \( a \) tends to 0. 

We can now derive an asymptotic equivalent for \( \Psi_n^{-1} \) when \( n \) is of order \( U(t) \) and the argument is of order \( t \). To proceed, we note that
Given Lemma 5.4.1.i, it is legitimate to define $\kappa_a(x)$ as the unique positive real number such that

$$\xi_a(x) = \frac{a + (x + \kappa_a(x))^\gamma}{\gamma \kappa_a(x)^{\gamma - 1}}.$$ 

If $\gamma$ is greater than 1, then

$$\lim_{\epsilon \to 0} \frac{a + (x + \epsilon)^\gamma}{\gamma \epsilon^{\gamma - 1}} = +\infty.$$ 

Thus, for $\gamma$ greater than 1, the function $\kappa_a$ maps compact subsets of the positive half-line to compact subsets of the positive half-line. In particular, on any compact subset of the positive half-line, $\kappa_a$ is lower bounded by a positive constant.

We remark that if $\gamma = 1$, taking $0^0$ to be 1, we have $\kappa_a(x) = 0$.

In the next result, the function $\varepsilon$ is that involved in the Karamata representation of $g$. Recall that (5.4.2) holds, that is (2.1.4) and (2.1.5) are assumed to hold when $\gamma$ is 1.

**Lemma 5.4.3.** The following asymptotic equivalents hold uniformly in $x$ and $c$ in any compact subset of the positive half-line as $t$ tends to infinity,

$$\Psi_{\lfloor xU(t) \rfloor}^{-1}(ct) \sim U(t)\xi_c \Gamma(1+\gamma)(x),$$

and

$$k_{\lfloor xU(t) \rfloor}(ct) \sim \begin{cases} \kappa_c \Gamma(1+\gamma)(x)U(t) & \text{if } \gamma > 1, \\ (c + x)(\text{Id } \varepsilon) \circ U(t) & \text{if } \gamma = 1. \end{cases}$$

**Proof.** We distinguish two cases, according to whether $\gamma$ is greater than 1 or not.

**Case $\gamma > 1$.** Write $n = \lfloor xU(t) \rfloor$. In the minimization defining $\Psi_n^{-1}(ct)$, we consider three ranges of $k$. First, if $k \sim yU(t)$ for $y$ in some compact subset of the positive half-line, Lemma 5.1.1 shows that $(ct + g(0, n + k))/g_k$ is equal to

$$ct + g\left(1 - \frac{1}{(x + y)U(t)}\right)\frac{1 + o(1)}{\Gamma(1 + \gamma)} yU(t) \Gamma(\gamma) (1 + o(1))$$

$$g \left(1 - \frac{1}{yU(t)}\right)$$

$$= U(t) \frac{c \Gamma(1 + \gamma) + (x + y)^\gamma(1 + o(1))}{\gamma y^{\gamma - 1}} (1 + o(1)).$$
with the $o(1)$ terms being uniform in $x$ and $y$ in any compact subset of the positive half-line. Therefore, by Lemma 5.4.1, provided $\epsilon$ is small enough, the minimum of $(c t + g_{[0,n+k]})/g_k$ over $k$ in the range $\epsilon U(t) \leq k \leq U(t)/\epsilon$ is asymptotically equivalent to $\xi_{c t \Gamma(1+\gamma)}(x) U(t)$ as $t$ tends to infinity and any $k$ minimizing in this range is asymptotically equivalent to $U(t) \kappa_{c t \Gamma(1+\gamma)}(x)$.

If $0 \leq k \leq \epsilon U(t)$ and $g_k$ is positive, then for $t$ large enough and $\epsilon$ small enough, $(c t + g_{[0,n+k]})/g_k$ is at least

$$\frac{c t + g_{[0,n]}}{2 g_{[\epsilon U(t)]}} \sim U(t) \frac{c \Gamma(1 + \gamma) + x^\gamma}{2 \gamma \epsilon^{\gamma-1}} \geq 2 U(t) \xi_{c t \Gamma(1+\gamma)}(x),$$

where the last inequality uses our earlier observation that when $\gamma$ is greater than 1, the function $\kappa_a$ is lower bounded by a positive constant on any compact subset of the positive half-line.

Finally, if $k \geq U(t)/\epsilon$, then $(c t + g_{[0,n+k]})/g_k$ is asymptotically bounded from below by an equivalent of

$$\frac{g_{[0,k]}}{g_k} \sim \frac{k}{\gamma} \geq \frac{U(t)}{\epsilon \beta},$$

which is asymptotically greater than $U(t) \xi_{c t \Gamma(1+\gamma)}(x)$, again provided that $\epsilon$ is small enough.

**Case $\gamma = 1$.** This case is more involved, and for clarity of the argument, we split the proof into several steps. The first one consists in proving the result for a continuous analogue of the minimization problem involved in the variational form of $\Psi_n^{-1}$. Recall that $g(1 - 1/t) = t \ell(t)$ for some slowly varying function $\ell$ having the Karamata representation

$$\ell(x) = d(x) \exp \int_1^x \frac{\varepsilon(u)}{u} \, du$$

ultimately and where $\varepsilon$ satisfies (2.1.4) and (2.1.5).

**Step 1.** Assume that the function $d(\cdot)$ is constant. For any fixed positive $c$, consider the function

$$\phi_t(y) = \frac{c t + g(1 - \frac{1}{x U(t)} + y)}{\ell(y)}.$$

Asymptotically in $t$, this function is a continuous analogue of the function $y \mapsto (c t + g_{[0,x U(t)+y]})/g_{[y]}$. In view of Lemma 5.2.2, we are
seeking the minimum value of $\phi_t$ as well as its minimizing argument. First, the minimizer has to tend to infinity with $t$, for if $y$ stays bounded then $\phi_t(y) \sim (c + x^\gamma)t/\ell(y)$ and, since we assume that (5.4.1) holds so that $\ell$ tends to infinity, this asymptotic equivalent can be made smaller by increasing $y$. Second, the minimizer has to be of smaller order than $U(t)$ because if $y \sim \theta U(t)$ for some $\theta$ in a compact subset of the positive half-line, then

$$\phi_t(y) \sim \frac{t}{\ell \circ U(t)}(c + x + \theta)$$

which attains its minimum for $\theta$ vanishing; and, moreover, if $y$ is of order larger than $U(t)$, then the same argument as in the case $\gamma > 1$ show that $\phi_t(y)$ cannot be minimum.

Next, differentiating $\phi_t$ and after substitution of $\text{Id} \ell$ for $g(1 - 1/\text{Id})$, the minimizer satisfies

$$0 = \frac{\ell(xU(t) + y) + (xU(t) + y)\ell'(xU(t) + y)}{\ell(y)} - \frac{\ell'(y)}{\ell^2(y)}\left(ct + g\left(1 - \frac{1}{xU(t) + y}\right)\right).$$

Since the Karamata representation of $\ell$ with a constant function $d$ implies $\text{Id} \ell'/\ell = o(1)$, and since the minimizer is $o(U(t))$, it follows that, after factoring $1/\ell(y)$ and simplifying,

$$0 = \ell(xU(t))(1 + o(1)) - \frac{\varepsilon(y)}{y}\left(ct + g\left(1 - \frac{1}{xU(t) + y}\right)\right).$$

Consequently, since $\gamma = 1$ and

$$g\left(1 - \frac{1}{xU(t) + y}\right) \sim g\left(1 - \frac{1}{xU(t)}\right) \sim xt,$$

the minimizer satisfies

$$\frac{\varepsilon(y)}{y} \sim \frac{\ell \circ U(t)}{t} \frac{1}{c + x}.$$

Furthermore,

$$\ell \circ U \sim g(1 - 1/U)/U \sim \text{Id}/U,$$

at infinity, and therefore the minimizer satisfies

$$\frac{\varepsilon(y)}{y} \sim \frac{1}{U(t)(c + x)}.$$
Assumption (2.1.4) then implies \( y \sim (c + x)U(t)\varepsilon \circ U(t) \) — compare with the value for \( k_{\lfloor xU(t) \rfloor}(ct) \) given in the statement of the lemma. For such value of \( y \), we have

\[
\phi_t(y) \sim \frac{t}{\ell(U(t)\varepsilon \circ U(t))}(c + x).
\]

Since \( \log \varepsilon(e^t) \) is self-neglecting and therefore self-controlled, Theorem 3.12.5 in Bingham, Goldie and Teugels (1989) shows that \( \ell(U\varepsilon(U)) \sim \ell(U) \) as \( U \) tends to infinity. Combined with (5.4.3), this yields

\[
\phi_t(y) \sim \frac{t}{\ell \circ U(t)}(c + x) \sim U(t)(c + x)
\]

as \( t \) tends to infinity.

**Step 2.** This step consists in showing that the minimum of \( \phi_t(y) \) has some form of continuity with respect to the asymptotic behavior of \( g \). Assume now that we have another function, \( g_1 \) asymptotically equivalent to \( g \) at \( 1^- \). With obvious notation, this new function gives rise to the corresponding functions \( U_1 \) and \( \ell_1 \). Then, for any positive \( \eta \),

\[
ct + g \left( 1 - \frac{1}{xU(t) + y} \right) \leq \frac{ct + (1 + \eta)g_1 \left( 1 - \frac{1}{xU_1(t) + y} \right)}{(1 - \eta)\ell_1(y)} \leq \frac{ct + g_1 \left( 1 - \frac{1}{xU_1(t) + y} \right)}{\ell_1(y)}
\]

as \( t \) tends to infinity, and uniformly in the range \( y \) nonnegative.

Step 1 of this proof shows that even though we do not assume \( \ell_1 \) to be smooth, the minimizer of the corresponding \( \phi_{1,t}(y) \) function is \( o(U_1(t)) \). Therefore, (5.4.4) and the analogous lower bound obtained by permuting \( g \) and \( g_1 \) show that the minimizer of \( \phi_{1,t} \) is asymptotically equivalent to that of \( \phi_t \); moreover, \( \phi_{1,t} \) and \( \phi_t \) have asymptotically equivalent minimum values. It follows that the conclusion of step 1 remains valid if we only assume that the function \( d(\cdot) \) in the Karamata representation of \( \ell \) has a limit and that the function \( \varepsilon(\cdot) \) in that representation satisfies (2.1.4) and (2.1.5).

**Step 3.** Recall that \( \gamma = 1 \) here, so that both \( \Gamma(\gamma) \) and \( \Gamma(1 + \gamma) \) are 1 as well. Going back to the problem of evaluating \( \Psi_{[xU(t)]}^{-1}(ct) \), we have \( g_k \sim \ell(k) \) and, by Karamata’s theorem, \( g_{[0,k]} \sim k\ell(k) \), this
equivalent being uniform in the range of \( k \) of order \( U(t) \). This allows us to replace the discrete minimization to calculate \( \Psi_{[xU(t)]}(ct) \) by the continuous one solved in the first step. Using that \( \xi_a(x) = a + x \) when \( \gamma \) is 1, this proves the lemma.

In the preceding lemma, writing \( n \) for \( \lfloor xU(t) \rfloor \), we obtain

\[
\Psi_n^{-1}(t) \sim U(t) \xi_{\Gamma(1+\gamma)}(n/U(t))
\]

as \( t \) tends to infinity. Lemma 5.4.2 asserts that if \( x \) is large, then \( \xi_{\Gamma(1+\gamma)}(x) \) is about \( x\xi_0(1) \). Consequently, we expect that if \( n/U(t) \) is large then \( \Psi_n^{-1} \) is about \( n\xi_0(1) \). The following bounds show that in some sense this is indeed the case.

**Lemma 5.4.4.** Let \( \epsilon \) be a positive real number less than 1.

(i) For \( n \) large enough, for any positive \( t \),

\[
\Psi_n^{-1}(t) \geq (1 - \epsilon)n\xi_0(1) .
\]

(ii) For \( n \) at least \( U(t)/\epsilon \) and \( t \) large enough,

\[
\Psi_n^{-1}(t) \leq (1 + \epsilon)n\xi_0(1) .
\]

**Proof.** (i) Clearly, \( \Psi_n^{-1}(t) \) is at least \( \min_{k \in K} g_{[0,n+k]} / g_k \). Setting \( k = y\delta \) with \( y \) in a compact set of the positive half-line, we obtain, as \( n \) tends to infinity,

\[
\frac{g_{[0,n+k]}}{g_k} \sim \frac{1}{\gamma} \frac{g(1 - \frac{1}{n(1+y)})}{g(1 - \frac{1}{ny})} ny \sim \frac{n}{\gamma} \frac{(1+y)^{1+y}}{y^{\gamma-1}} \geq n\xi_0(1) .
\]

Next, let \( \delta \) be a positive real number. For any positive \( k \) at most \( \delta n \) and for any \( g_k \) positive, Lemma 5.1.1 yields

\[
\frac{g_{[0,n+k]}}{g_k} \geq \frac{g_{[0,n]}}{g_{[0,n+k]}} \sim \frac{n}{\gamma} \frac{1}{\delta^{\gamma-1}} ,
\]

while for any \( k \) at least \( n/\delta \), it yields

\[
\frac{g_{[0,n+k]}}{g_k} \geq \frac{g_{[0,k]}}{g_k} \geq \frac{n}{\gamma} \frac{1}{\delta} .
\]
The result follows by taking $\delta$ such that $\delta^{1-\gamma} \land \delta^{-1} \geq \xi_0(1)$. Note that when $\gamma$ is 1, the result still holds because then $\xi_0(1) = 1$.

(ii) Recall that $g(1 - 1/\cdot)$ is regularly varying with nonvanishing index. Hence, it is asymptotically equivalent to a monotone function (Bingham, Goldie and Teugels, 1989, Theorem 1.5.3). Since $U(t) \leq \epsilon n$, for $t$ large enough,

$$t \sim g\left(1 - \frac{1}{U(t)}\right) \lesssim g\left(1 - \frac{1}{\epsilon n}\right) \sim g\left(1 - \frac{1}{n}\right)\epsilon^\gamma,$$

as $n$ tends to infinity. Therefore, taking $k = yn$ with $y$ fixed, when $\gamma$ is greater than 1, we obtain that for $t$ large enough, $\Psi_n^{-1}(t)$ is at most

$$g\left(1 - \frac{1}{n}\right)\epsilon^\gamma + g\left(1 - \frac{1}{n(1+y)}\right)\frac{(1 + o(1))}{\Gamma(1 + \gamma)}ny\Gamma(\gamma)(1 + o(1))$$

$$\sim n\frac{\Gamma(1 + \gamma)\epsilon^\gamma + (1 + y)^\gamma}{\gamma y^{\gamma-1}}.$$

If $\gamma$ is greater than 1, the result follows by taking $y$ minimizing $(1 + y)^\gamma / y^{\gamma-1}$, upon noting that the minimizing value is positive. If $\gamma$ is 1, the result follows by taking $y = \epsilon$.

While the previous lemma gives valuable information on $\Psi_n^{-1}$ when $n$ is large, it does not give any estimate for $\Psi_n^{-1}$ when $n$ is moderate, say, and $t$ is large. The next result fills this gap. It should be compared to (5.4.5).

**Lemma 5.4.5.** There exists $n_1$ such that

$$\min_{n \geq n_1} \Psi_n^{-1}(t)/U(t) \gtrsim \xi_{\Gamma(1+\gamma)}(0).$$

as $t$ tends to infinity. Moreover, if all the $g_n$ are nonnegative, we can take $n_1$ to be 0.

**Proof.** We first prove the following claim.

**Claim.** There exists $n_1$ such that for any $n$ at least $n_1$ and any $k$ nonnegative, $g_{[0,n+k]} \geq g_{[0,k]}$. 

57
Indeed, for any fixed $k$, define

$$m_k = \min\{ n \geq 1 : \forall i \geq n, g_{[k,i+k]} > 0 \}.$$  

This integer exists since the partial sums $g_{[0,n)}$ diverge as $n$ tends to infinity. Recall that $n_0$ is the smallest integer such that both $g_n$ and $g_{[0,n)}$ are positive if $n$ is at least $n_0$. If $k$ is at least $n_0$, then $m_k = 1$. Therefore the sequence $(m_k)_{k \geq 1}$ is bounded and admits a maximum element which we call $\tilde{n}_1$. We then set $n_1$ to be the maximum of $n_0$ and $\tilde{n}_1$.

Having proved the claim, for $n$ at least $n_1$,

$$\Psi^{-1}_n (t) \geq \min_{k \in K_{+}} \frac{t + g_{[0,k)}}{g_k}.$$  

If $k = [yU(t)]$ for some $y$ in a compact subset of the positive half-line, then

$$\frac{t + g_{[0,k)}}{g_k} \sim U(t) \frac{\Gamma(1+\gamma) + y^\gamma}{\gamma y^\gamma - 1} \geq U(t) \xi \Gamma(1+\gamma)(0)$$

as $t$ tends to infinity.

If $k$ is at most $\delta U(t)$, then

$$\frac{t + g_{[0,k)}}{g_k} \geq \frac{t}{g_{[\delta U(t)]}} \sim U(t) \Gamma(\gamma) \delta^{1-\gamma},$$

while if $k$ is at least $U(t)/\delta$, then

$$\frac{t + g_{[0,k)}}{g_k} \geq \frac{g_{[0,k)}}{g_k} \sim \frac{k}{\gamma} \geq \frac{U(t)}{\gamma \delta}.$$  

The result follows by choosing $\delta$ small enough so that $\Gamma(1+\gamma) \delta^{1-\gamma}$ and $\delta^{-1} \geq \gamma \xi \Gamma(1+\gamma)(0)$ which as noted previously holds trivially when $\gamma$ is 1.

5.5. Approximation of $\Psi^{-1}_n$ when $(g_n)_{n \geq 0}$ has a positive and finite limit. In this subsection we consider the case where the sequence $(g_n)_{n \geq 0}$ has a positive and finite limit. The minimization involved in Lemma 5.2.2 may or may not be made explicit, according to whether the supremum of the sequence is achieved or not. To be more precise, recall that $g^* = \sup_{n \geq 0} g_n$, and, if it exists let $k^*$ be
the smallest integer such that $g_{k^*} = g^*$. Note that such number does not exists for a sequence such as $(1 - (n+1)^{-1})_{n \geq 0}$.

**Lemma 5.5.1.** If $g^*$ is attained, that is $k^*$ is well defined, then there exists a nonnegative real number $t_0$ such that for any $n$ at least $n_0$ and any $t$ at least $t_0$,

$$
\Psi_n^{-1}(t) = \frac{t + g_{[0,n+k^*)}}{g^*}.
$$

Otherwise, for any $n$ at least $n_0$, and any positive $t$,

$$
\Psi_n^{-1}(t) \geq \frac{t + g_{[0,n)}}{g^*},
$$

and for any positive $\epsilon$ there exists $k$ such that for any positive $t$,

$$
\Psi_n^{-1}(t) \leq (1 + \epsilon)\frac{t + g_{[0,n)}}{g^*} + (1 + \epsilon)\frac{g_{[n,n+k)}}{g^*}.
$$

**Proof.** If $k^*$ exists, the proof of Lemma 5.3.1 is still valid and yields the result. Hence, we assume that all the $g_n$ are less than their limit $g^*$. For $n$ at least $n_0$, the inequality $g_{[0,n+k)} \geq g_{[0,n)}$ holds and, since all the $g_n$ are less than $g^*$, the formula for $\Psi_n^{-1}$ in Lemma 5.2.2 implies the given lower bound for $\Psi_n^{-1}$.

To prove the upper bound, let $k$ be any integer such that $g_k \geq g^*/(1 + \epsilon)$. Then, the formula for $\Psi_n^{-1}$ in Lemma 5.2.2 shows that

$$
\Psi_n^{-1}(t) \leq (1 + \epsilon)\frac{t + g_{[0,n+k)}}{g^*},
$$

and the result follows by writing $g_{[0,n+k)}$ as $g_{[0,n)} + g_{[n,n+k)}$.

We then obtain the following asymptotic equivalence.

**Lemma 5.5.2.** Uniformly in $x$ and $c$ in any compact subset of the positive half-line,

$$
\Psi_{[xU(t)]}^{-1}(t) \sim \frac{t}{g^*}(1 + x)
$$

as $t$ tends to infinity.

**Proof.** The lemma follows from Lemma 5.5.1.
Our next lemma is stated so that it can be easily referred to. It involves a real number $t_0$ defined in Lemma 5.5.1 when $k^*$ exists, and otherwise, one can take $t_0$ to be 1.

**Lemma 5.5.3.** For any $n$ at least $n_0$ and any $t$ at least $t_0$,

$$\Psi_{n}^{-1}(t) \geq t/g^*$$

and, as $n$ tends to infinity,

$$\inf_{t \geq t_0} \Psi_{n}^{-1}(t) \gtrapprox \frac{g(1 - 1/n)}{g^*}.$$ 

**Proof.** The proof is the same as that of Lemmas 5.3.3 and 5.3.4. ■

**6. Proof of the results of section 2.** The proof follows by an application of Theorem 3.2.1, after completion of all the steps described in section 3.1. While most arguments depend on the asymptotic behavior of $(g_n)_{n \geq 0}$ at infinity, the calculation of the conditional expectation, encoded in the functions $\psi_{i,n}$ and $\psi_n$, can be done once and for all. Indeed, without any loss of generality, we assume that $EX_i = -1$. Then

$$\psi_{i,n}(x) = E(S_i|x_n = x) = \begin{cases} -g_{[0,i)} & \text{if } i < n, \\ (x + 1)g_{i-n} - g_{(0,i]} & \text{if } i \geq n. \end{cases} \quad (6.1)$$

It follows that

$$\psi_{n}(x) = \max_{i \geq 0} \psi_{i,n}(x) = \max_{0 \leq i < n} (-g_{[0,i)}) \vee \max_{i \geq n} ((x + 1)g_{i-n} - g_{(0,i]}).$$

Therefore, if $x$ is such that $\psi_n(x)$ is both positive and greater than $\max_{i \geq 0} -g_{(0,i)}$, then

$$\psi_{n}(x) = 0 \vee \max_{i \geq n} (x + 1)g_{i-n} - g_{(0,i]}.$$ 

Writing $t_1$ for $0 \vee \max_{i \geq 0} -g_{[0,i)}$, this shows that, with the notation of section 5.2, for any $n$ positive and any $x$ in the preimage under $\psi_n$ of $(t_1, \infty)$,

$$\psi_{n}(x) = \Psi_{n}(x + 1).$$

60
Therefore, for \( t \) greater than \( t_1 \),

\[
\psi_n^{-1}(t) = \Psi_n^{-1}(t) - 1.
\]

Define the sequence of independent and identically distributed centered random variables \((Z_i)_{i \geq 1}\) by \( Z_i = X_i - \mu \), \( i \geq 1 \). A useful remark for completing step 6 is that

\[
S_i - \psi_{i,n}(X_n) = \sum_{0 \leq j < i \atop i-j \neq n} g_j Z_{i-j}.
\]

Finally we will use also the following weak law of large numbers.

**Lemma 6.1.** Let \((S_n)\) be a \((g,F)\)-process with negative mean innovations. Then,

\[
\lim_{n \to \infty} \frac{S_n}{g(1 - 1/n)} = \frac{\mu}{\Gamma(\gamma + 1)}
\]

in probability; in other words, \( S_n/ES_n \) converges to 1 in probability.

**Proof.** Lemma 5.1.1.i shows that only a finite number of \( g_n \) may be nonpositive. Therefore, here, we can assume without any loss of generality that all the coefficients \( g_n \) are positive. Lemma 5.1.1 and the uniform convergence theorem for regularly varying functions (Bingham, Goldie and Teugels, 1989, Theorem 1.2.1) imply that uniformly in \( \alpha \) in any compact subset of \((0, 1]\),

\[
\lim_{n \to \infty} g_{[\alpha n]}/g_{(0,n)} = 0.
\]

Therefore, Theorem 3 in Jamison, Orey and Pruitt (1965) implies

\[
\lim_{n \to \infty} S_n/g_{(0,n)} = \mu
\]

in probability. The result then follows from Lemma 5.1.1.ii. □

**Notation.** In the remainder of this paper we write \( s_n \) for the expected value of \( S_n \), that is, \( s_n = \mu g_{(0,n)} \).

**6.1. Proof of Theorem 2.1.1 – upper bound.** We complete all the steps described in section 3. Recall that without any loss of
generality, we assume that $\mu$ is $-1$. Also, replacing $t$ by $t/(-\mu g^*)$, and $g$ by $g/g^*$, we assume without loss of generality that $g^* = 1$.

**Step 1.** Since $\psi_n^{-1} = \Psi_n^{-1} - 1$ on $(t_1, \infty)$, relation (S1) follows from Lemma 5.3.2 with $\chi$ being Id and $\rho(x) = 1 + x^\gamma/\Gamma(1 + \gamma)$.

**Step 2.** Recall the asymptotic equivalence given in (3.1.1). Lemma 5.3.3 and regular variation of $F$ and $U$ imply that as $t$ tends to infinity,

$$
\sum_{n \leq \epsilon U(t)} F \circ \psi_n^{-1} (t) \lesssim \sum_{n \leq \epsilon U(t)} F(t) \sim \epsilon(UF)(t),
$$

while Lemma 5.3.4 implies

$$
\sum_{n \geq \epsilon U(t)/\epsilon} F \circ \psi_n^{-1} (t) \lesssim \sum_{n \geq \epsilon U(t)/\epsilon} F\left(\frac{g(1 - 1/n)}{\Gamma(1 + \gamma)}\right).
$$

Using the regular variation of $F$ and $g$, the approximation of the sum in this upper bound by a Riemann integral and Karamata’s theorem, we obtain an asymptotic upper bound equivalent to

$$
\Gamma(1 + \gamma)^\alpha \int_{U(t)/\epsilon}^{\infty} F \circ g(1 - 1/u) \, du \sim \Gamma(1 + \gamma)^\alpha \frac{e^{\alpha \gamma - 1}}{\alpha \gamma - 1} (UF)(t).
$$

Since $\alpha \gamma$ is larger than 1, this completes the proof of (S2).

**Step 3.** We prove that the maximum of the process is unlikely to occur at a time of smaller order than $U(t)$.

**Lemma 6.1.1.** The following limit holds,

$$
\lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{P\{ \exists n : n \leq \epsilon U(t), S_n > t \}}{(UF)(t)} = 0.
$$

**Proof.** Let $n$ be at least $n_0$. Recall that $s_n$ is the expectation of $S_n$. Since $t - s_n$ is at least $t$, if $S_n$ exceeds $t$ then $\sum_{0 \leq j < n} g_j Z_{n-j}$ exceeds $t$ as well. By the standard estimate for weighted convolution of distribution functions with regularly varying tails and Bonferroni’s inequality, for any fixed $k$ and any positive $\delta$,

$$
\lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{P\{ \exists n : n \leq \epsilon U(t), \sum_{0 \leq j < k} g_j Z_{n-j} > \delta t \}}{(UF)(t)} = 0.
$$

62
Therefore, taking $\delta$ to be less than $1/2$, it suffices to prove that for some $k$,

$$\lim_{\epsilon \to 0} \limsup_{t \to \infty} P\{ \exists n : k < n \leq \epsilon U(t), \sum_{k \leq j < n} g_j Z_{n-j} > \delta t \} \frac{1}{(UF)(t)} = 0.$$  

Since

$$P\{ \exists j : j \leq \epsilon U(t), Z_j > t \} \lesssim \epsilon (UF)(t),$$

it suffices to prove that

$$\lim_{\epsilon \to 0} \limsup_{t \to \infty} \epsilon U(t) \max_{k \leq n \leq \epsilon U(t)} P\{ \sum_{k \leq j < n} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq t \} > \delta t \} \frac{1}{(UF)(t)} = 0. \quad (6.1.1)$$

Using that $-\log F \sim \alpha \log$ at infinity, and taking $\lambda$ of the form $c t^{-1} \log t$ in inequality (4.1.1) with $\delta t$ in place of $t - s_n$ there, the logarithm of the ratio

$$P\{ \sum_{k \leq j < n} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq t \} > \delta t \} / F(t),$$

for $n$ at most $\epsilon U(t)$, is ultimately at most

$$-c\delta \log t + \eta c t^{-1} \log t g_{[0,n)} E|Z|$$

$$- c t^{-1} \log t \sum_{0 \leq j < n} g_j E Z \mathbb{1}\{ \lambda g_j Z \leq \log(1 - \eta) \}$$

$$+ \epsilon U(t) \exp\left( c \log t \max_{k \leq j < U(t)} g_j \right) \Pi\left( \frac{t \log(1 + \eta)}{c \log t} \right) + 2\alpha \log t.$$

Referring to the second summand in this bound, $g_{[0,n)}$ is at most $g_{[0,\epsilon U(t))] \sim \epsilon^3 t / \Gamma(1 + \gamma)$. Thus, the second summand is ultimately at most $2\eta c^3 E|Z| \log t$. Since $\lambda$ tends to 0 as $t$ tends to infinity and so $E \mathbb{1}\{ \lambda Z \leq \log(1 - \eta) \} = o(1)$, the third summand is negligible compared to the second one. Therefore, ultimately, (6.1.2) is at most

$$(-c\delta + 2\eta c t^{-1} E|Z| + 2\alpha) \log t$$

$$+ \epsilon \exp\left( \log U(t) + c \max_{j \geq k} g_j \log t + \log \Pi\left( \frac{t \log(1 + \eta)}{c \log t} \right) \right). \quad (6.1.3)$$
We take $\eta$ small enough so that $-\delta + 2\eta \gamma E|Z|$ is negative. Then, we take $c$ large enough so that whenever $\epsilon$ is less than 1,

$$c(-\delta + 2\eta \gamma E|Z|) + 2\alpha < -2,$$

say. Since $\alpha \gamma > 1$ and $(g_n)_{n \geq 0}$ converges to 0, we can then fix $k$ large enough so that

$$\gamma^{-1} + c \max_{j \geq k} g_j - \alpha < 0.$$

For such $k$, we obtain that

$$\lim_{t \to \infty} \log U(t) + c \max_{j \geq k} g_j \log t + \log \prod_{j \geq k} \left( \frac{t \log(1 + \eta)}{c \log t} \right) = -\infty.$$

Hence, ultimately, (6.1.3) is at most $-\log t$. It follows that (6.1.1) holds as well as the conclusion of the lemma.

**Step 4.** We prove that the maximum of the process is unlikely to occur at a time of larger order than $U(t)$.

**Lemma 6.1.2.** The following limit holds,

$$\lim_{\epsilon \to 0} \lim_{t \to \infty} \sup_{\epsilon} \frac{P\{ \exists n : n \geq U(t)/\epsilon, S_n > t \}}{(U(t))(t)} = 0.$$

**Proof.** Referring to step 4 in section 3, it suffices to show that (3.1.4) holds. In the current context, $\chi \circ U^{-1} \sim g(1-1/\text{Id})$ at infinity. Thus, given two positive real number $\theta$ and $\eta$, set

$$a_n = \theta \left( g\left( 1 - \frac{1}{n} \right) + \frac{t}{\eta} \right).$$

If all $X_n$ are at most $a_n + \mu$, then

$$S_n - s_n = \sum_{0 \leq i < n} g_i Z_{n-i} 1\{ Z_{n-i} \leq a_{n-i} \}.$$

Hence, the probability involved in (3.1.4) is at most

$$\sum_{n \geq U(t)/\epsilon} \sum_{0 \leq i < n} g_i Z_{n-i} 1\{ Z_{n-i} \leq a_{n-i} \} > t - s_n.$$
The usual estimate for the tail of weighted convolutions of heavy-tail distribution functions shows that for any fixed $k$,

$$
\sum_{n \geq U(t)/\epsilon} P\left\{ \sum_{0 \leq j < k} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq a_{n-j} \} > \frac{t-s_n}{2} \right\} \lesssim k2^\alpha \sum_{n \geq U(t)/\epsilon} F(t-s_n)
$$

$$
\sim k2^\alpha \int_0^\infty \frac{F(t-s_n)}{\Gamma(1+\gamma)} d\mu
$$

which, by the proof of Lemma 4.2.1, is negligible compared to $(U_F)(t)$ as first $t$ tends to infinity and then $\epsilon$ tends to 0. Therefore, it suffices to prove that for some fixed $k$,

$$
\lim_{\epsilon \to 0} \lim_{t \to \infty} \sum_{n \geq U(t)/\epsilon} P\left\{ \sum_{0 \leq j < n} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq a_{n-j} \} > \frac{t-s_n}{2} \right\} (U_F)(t) = 0. \quad (6.1.4)
$$

Using inequality (4.1.1), the logarithm of each summand is at most asymptotically bounded by an equivalent of

$$
-\lambda \frac{t-s_n}{2} + \eta \lambda g_{[0,n)} E|Z| - \lambda \sum_{0 \leq j < n} g_j EZ \mathbb{1}\{ \lambda g_j Z \leq \log(1-\eta) \}
$$

$$
+ \sum_{k \leq j < n} e^{\lambda g_j a_{n-j}} g_j \mathbb{1}(1+\eta) - \left( \frac{1}{\gamma} - \alpha \right) \log t(1+o(1)) . \quad (6.1.4)
$$

We take $\lambda$ of the form $2c(t-s_n)^{-1} \log g(1-1/n)$ for a constant $c$ to be determined later.

Referring to the successive terms in (6.1.4), we have

$$
\lambda \frac{t-s_n}{2} = c \log g(1-1/n) \sim c \gamma \log n
$$

as $n$ tends to infinity. Furthermore, since $t-s_n \geq -s_n = g_{[0,n)}$,

$$
\eta \lambda g_{[0,n)} E|Z| \sim 2\eta c \frac{\log g(1-1/n)}{t-s_n} g_{[0,n)} E|Z| \lesssim 2\eta c \gamma \log n E|Z| .
$$
Next, since $\lambda$ tends to 0 as $t$ tends to infinity and uniformly in $n \geq U(t)/\epsilon$,

$$\lambda \sum_{0 \leq j < n} g_j E Z \mathbb{1}\{\lambda g_j Z \leq \log(1 - \eta)\} = \lambda o(g_{(0,n)}) = o(\log n).$$

Moreover, since $n \geq U(t)/\epsilon$ and therefore, $g(1 - 1/n) \geq \epsilon^{-\gamma}t$, we have $t \lesssim -\epsilon^\gamma \Gamma(1 + \gamma)s_n$, and since $\gamma$ is at most 1,

$$\lambda g_j a_{n-j} \lesssim \frac{2c \log g(1 - 1/n)}{t - s_n} \theta \left( g \left(1 - \frac{1}{n}\right) + \frac{t}{\eta}\right)$$

$$\lesssim 4c\gamma \theta \Gamma(1 + \gamma) \left(1 + \frac{\epsilon^\gamma}{\eta}\right) \log n.$$

Finally, for the same reason, $\log t \lesssim \log g(1 - 1/n) \sim \gamma \log n$. In particular, this implies

$$\log \bar{H} \left(\frac{\log(1 + \eta)}{\lambda g_j}\right) \lesssim \alpha \log \lambda \sim -\alpha \log(t - s_n) \lesssim -\alpha \log(-s_n)$$

$$\sim -\alpha \gamma \log n.$$

Therefore,

$$\sum_{k \leq j < n} e^{\lambda g_j a_{n-j}} \Pi \left(\frac{\log(1 + \eta)}{\lambda g_j}\right) \lesssim n^{1 - \alpha \gamma + 4c\gamma \theta \Gamma(1 + \gamma) + o(1)}.$$

It follows that (6.1.4) is asymptotically bounded by an equivalent of

$$\left(-c\gamma + 2\eta c\gamma E|Z| + o(1) + \alpha - \frac{1}{\gamma}\right) \log n + n^{1 - \alpha \gamma + 4c\gamma \theta \Gamma(1 + \gamma) + o(1)}.$$  

(6.1.5)

We take $\eta$ less than $1/6E|Z|$. We take $c$ large enough so that

$$\frac{-c\gamma}{3} + \alpha - \frac{1}{\gamma} < -3.$$  

Since $\alpha \gamma$ is larger than 1, we can take $\theta$ small enough so that

$$1 - \alpha \gamma + 8c\gamma \theta \Gamma(1 + \gamma) < 0.$$  

These choices lead to that (6.1.5) as well as (6.1.4) are asymptotically bounded by $-3 \log n$. Since

$$\sum_{n \geq U(t)/\epsilon} n^{-2} \sim \frac{\epsilon}{U(t)},$$

66
This proves the lemma.

**Step 5.** We now prove that for $M$ to exceed $t$ it is likely that we must have at least one random variable $X_n$ to be large. Recall that $B_j$ is the event \( \{ X_j \leq \theta t \} \). For any $t$ large enough, up to increasing $\theta$ slightly, we can replace $X_j$ by $Z_j$ in the definition of $B_j$, and so we set $B_i = \{ Z_i \leq \theta t \}$.

**Lemma 6.1.3.** For any positive $\epsilon$, there exists positive $\theta$ such that
\[
P \left( \exists n \in I : S_n > t; \bigcap_{0 \leq i \leq n} B_i \right) = o(U^T)(t)
\]
as $t$ tends to infinity.

**Proof.** Recall that $N$ denotes $\lfloor U(t)/\epsilon \rfloor$. On $\bigcap_{1 \leq i \leq n} B_i$ we have
\[
\sum_{0 \leq i < n} g_i Z_{n-i} = \sum_{0 \leq i < n} g_i Z_{n-i} \mathbb{1} \{ Z_{n-i} \leq \theta t \}.
\]
We apply inequality (4.1.1) and use that $n$ is at most $N$ to obtain that the logarithm of $P \{ S_n > t; \bigcap_{1 \leq i \leq n} B_i \}$ is at most
\[
-\lambda t + \eta \lambda g_{[0,N]} E|Z| - \lambda \sum_{0 \leq j < n} g_j EZ \mathbb{1} \{ \lambda g_j Z \leq \log(1 - \eta) \}
\]
\[
+ \sum_{0 \leq j < n} e^{\lambda g_j g_1} P \left( \frac{\log(1 + \eta)}{\lambda g_j} \right) \text{.} \quad (6.1.6)
\]
We choose $\lambda$ of the form $c t^{-1} \log t$ where $c$ will be specified later. With this choice, we examine all the terms in (6.1.6). We have $\lambda t = c \log t$. Lemma 5.1.1 implies
\[
\lambda g_{[0,N]} \sim c \frac{\log t g(1 - \epsilon/U(t))}{\Gamma(1 + \gamma)} \sim \frac{c e^{-\gamma}}{\Gamma(1 + \gamma)} \log t \text{.}
\]
Since $\lambda$ tends to 0, we also have, referring to the third summand in (6.1.6),
\[
-\lambda \sum_{0 \leq j < n} g_j EZ \mathbb{1} \{ \lambda g_j Z \leq \log(1 - \eta) \} \leq \lambda g_{[0,n]} o(1)
\]
as $t$ tends to infinity. Finally, the bound
\[
\lambda g_j \theta t \leq c \theta \log t
\]
shows that
\[ \sum_{0 \leq j < n} e^{\lambda g_j \theta t} T \left( \frac{\log(1 + \eta)}{\lambda g_j} \right) \leq \frac{U(t)}{\epsilon} t^\theta T \left( \frac{t}{\log t} \right) e^\alpha \log^{-\alpha}(1 + \eta). \] (6.1.7)

Since \( \alpha \gamma > 1 \), Potter’s bounds imply that if
\[ \frac{1}{\gamma} + c \theta - \alpha < 0 \] (6.1.8)
then (6.1.7) tends to 0 as \( t \) tends to infinity. Since this is the case by choosing \( c \) and \( \theta \) such that \( c \theta \) is sufficiently small, (6.1.6) is bounded by an asymptotic equivalent of
\[ \left(-c + \frac{\eta \epsilon \gamma E|Z|}{\Gamma(1 + \gamma)}\right) \log t. \]
Let \( p \) be any positive number. We take \( \eta \) such that \( \eta \epsilon \gamma / \Gamma(1 + \gamma) < 1/2 \) say. Then, we take \( c \) large enough so that \(-c/2 < -p - 2/\gamma\). Then, we choose \( \theta \) small enough so that (6.1.8) holds. This shows that
\[ \max_{n \in I} P \{ S_n > t; \bigcap_{1 \leq i \leq n} B_i \} \leq t^{-p - 2/\gamma} \]
ultimately in \( t \). Therefore,
\[ P \{ \exists n \in I : S_n > t; \bigcap_{1 \leq i \leq n} B_i \} \leq \epsilon^{-1} U(t) t^{-p - 2/\gamma}, \]
which, by Potter’s bounds is \( o(t^{-p}) \) as \( t \) tends to infinity. Taking \( p \) greater than \( \alpha \) proves the lemma.

**Step 6.** We prove the law of large number which allows one to approximate all the \( S_i \), \( 0 \leq i \leq U(t) / \epsilon \), given that \( X_n \) is large. Recall that the sets \( B_n \) and \( C_n \) involved in step 6 depend on the parameter \( \theta \), while \( D_n \) depends on a parameter \( \delta \) and the set \( I \) depends on a parameter \( \epsilon \). In the following lemma, \( D_n \) refers in fact to the one sided event
\[ D_n = \bigcap_{1 \leq i \leq N} \{ S_i - \psi_{i,n}(X_n) \leq \delta t \}. \]

**Lemma 6.1.4.** Let \( \delta \) be a positive real number. For any \( \epsilon \) positive small enough, there exists a positive \( \theta \) such that
\[ P \left( \bigcup_{n \in I} B_n^c \cap C_n \cap D_n^c \right) = o(r(t)) \]

68
as $t$ tends to infinity.

**Proof.** The result follows from the same estimate as in the previous lemma, taking $p$ to be large enough in that proof upon using that $S_i - \psi_i, n(X_n) = \sum_{0 \leq j < i} g_j Z_{i-j}$. ■

Having completed steps 1–5 and the one-sided version of step 6, Theorem 3.2.1 yields the upper bound pertaining to Theorem 2.1.1.

### 6.2. Proof of Theorem 2.1.1 – lower bound.

To prove a lower bound matching the upper bound, we could use a tail balance condition and a two-sided version of the events $B_n$ to check the two-sided version of (S6) — which would then follow from the proof of Lemma 6.1.4— and then check that (S7) holds. To verify (S7) is particularly easy because (6.1) and Lemma 5.3.2 show that $i(n, x) = n + k^*$. However, in order not to impose a tail balance condition, we give a proof inspired by Zachary’s (2004) probabilistic proof of Veraverbeke’s theorem. Zachary’s proof, suitably modified, is remarkably robust to the choice of the process.

For this proof, we keep assuming, without any loss of generality that the mean $\mu$ of $X_i$ is $-1$ and that $g^*$ is $1$. Recall that $k^*$ is the smallest integer $k$ such that $g_k$ is maximal, that is equal to $g^*$. Recall we are in the case where the sequence $(g_i)_{i \geq 0}$ is nonnegative and tends to $0$ at infinity, so that the sequence attains its maximum value, assumed to be positive, and, by our convention, $g_{k^*} = 1$. Let

$$\hat{S}_n = \sum_{i \geq 0; i \neq k^*} g_i X_n - i.$$

We write $s_n$ for the expectation of $S_n$ and $\hat{s}_n$ for that of $\hat{S}_n$. Lemma 5.1.1.ii shows that $s_n \sim \hat{s}_n$ as $n$ tends to infinity. Moreover, Lemma 6.1 implies that for any positive $\epsilon$ and any $n$ larger than some $n_2$,

$$P\{\hat{S}_n > (1 + \epsilon)s_n\} \geq 1 - \epsilon.$$

If the event

$$\{\hat{S}_n > (1 + \epsilon)s_n; X_{n-k^*} > t - (1 + \epsilon)s_n\}$$

occurs, then $S_n$ is greater than $t$, and so is $M$. Consequently, applying Bonferroni’s inequality, the probability that $M$ is greater than $t$ is at least

$$\sum_{n \geq n_2} P\{\hat{S}_n > (1 + \epsilon)s_n; X_{n-k^*} > t - (1 + \epsilon)s_n\}$$
\[ - \sum_{m,n \geq n_2 \atop m \neq n} P \{ X_{n-k^*} > t - (1+\epsilon)s_n \mid X_{m-k^*} > t - (1+\epsilon)s_m \}. \quad (6.2.1) \]

Since \( \hat{S}_n \) and \( X_{n-k^*} \) are independent, the first sum in (6.2.1) is at least
\[
(1 - \epsilon) \sum_{n \geq n_2} F(t - (1 + \epsilon)s_n).
\]

Since \( s_n = -g_{[0,n)} \), Lemma 5.1.1 and regular variation of both \( F \) and \( g \) imply
\[
\sum_{n \geq n_2} F(t - (1 + \epsilon)s_n) \sim \int_1^\infty F(t + (1 + \epsilon)g(1 - 1/u)/\Gamma(1 + \gamma)) \, du
\]
as \( t \) tends to infinity. Under the claim to be proved that the second sum in (6.2.1) is asymptotically negligible with respect to the first sum, since \( \epsilon \) is arbitrary, upon using Lemma 4.2.1, we obtain
\[
\lim inf_{t \to \infty} \frac{P \{ M > t \}}{\int_1^\infty F(t + g(1 - 1/u)/\Gamma(1 + \gamma)) \, du} \geq 1.
\]

The second sum in (6.2.1) is less than
\[
\sum_{m,n \geq n_2} P \{ X_{n-k^*} > t - (1+\epsilon)s_n \mid X_{m-k^*} > t - (1+\epsilon)s_m \} \sum_{n \geq n_2} F(t - (1 + \epsilon)s_n).
\]

By the previous arguments, this last quantity is of smaller order than the first sum in (6.2.1). This proves (6.2.2), which, using Lemma 4.2.1, is the lower bound pertaining to Theorem 2.1.1.

6.3. Proof of Theorem 2.1.2 – upper bound. Again, we complete all the steps described in section 3. To prove this upper bound, we assume without any loss of generality that \( \mu = -1 \).

Step 1. The asymptotic equivalence (S1) follows from the equality \( \psi_n^{-1} = \psi_n^{-1} - 1 \) and Lemma 5.4.3. We now take \( \chi = U \) and \( \rho = \xi_{\Gamma(1+\gamma)} \). Thus, following (3.1.2), here
\[
r(t) = (\text{Id}F) \circ U(t) \int_0^\infty \xi_{\Gamma(1+\gamma)}^{-\alpha}(v) \, dv.
\]
Step 2. Let $\epsilon$ be a positive real number. Lemma 5.4.5 implies that
\[
\sum_{n_1 \leq n \leq \epsilon U(t)} F \circ \Psi_n^{-1}(t) \leq \sum_{n_1 \leq n \leq \epsilon U(t)} F(U(t)\xi_{\Gamma(1+\gamma)}(0)) \\
\lesssim \epsilon^{-\alpha} \xi_{\Gamma(1+\gamma)}(0)(\text{Id} F) \circ U(t).
\]
Furthermore, regular variation of $F$ and Lemma 5.4.4 show that
\[
\sum_{n \geq U(t)/\epsilon} F \circ \Psi_n^{-1}(t) \lesssim \sum_{n \geq U(t)/\epsilon} F(n\xi_0(1)).
\]
This last sum can be approximated by an integral, which, by Karamata's theorem is asymptotically equivalent to
\[
\frac{U(t)}{(\alpha - 1)\epsilon} F(U(t)\xi_0(1)/\epsilon) \sim \frac{\epsilon^{-\alpha}}{\alpha - 1} \xi_0(1)^{-\alpha} (\text{Id} F) \circ U(t)
\]
as $t$ tends to infinity. In view of (3.1.1), this completes step 2.

Step 3. Our next lemma shows that the maximum is unlikely to occur at a time of order smaller than $U(t)$. Its proof is inspired by that of Lemma 2.4 in Mikosch and Samorodnitsky (2000).

Lemma 6.3.1. The following limit holds
\[
\lim_{\epsilon \to 0} \limsup_{t \to \infty} P\{ \exists n : n \leq \epsilon U(t), S_n \geq t \} = 0.
\]

Proof. Let $Z$ be a random variable having the same distribution as $Z_1$ say. For $n$ large enough,
\[
S_n - s_n = \sum_{0 \leq j < n} g_j Z_{n-j} \leq 2g_n \sum_{0 \leq i < n} |Z_{n-i}|
\]
Therefore, with $N = \lfloor \epsilon U(t) \rfloor$ and $t$ large enough,
\[
P\{ \exists n : n \leq \epsilon U(t) : S_n > t \} \leq P\left\{ 2g_N \sum_{0 \leq i < N} |Z_{N-i}| > t \right\} \\
= P\left\{ \sum_{0 \leq i < N} |Z_{n-i} - E|Z| > \frac{t}{2g_N} - NE|Z| \right\}.
\]
To apply the large deviation result of Nagaev (1969 a,b) and Cline and Hsing (1991) stated as Lemma A.1 in Mikosch and Samorodnitsky (2000), we check that for some positive \( \delta \) and any \( t \) large enough
\[
\frac{t}{2g_N} - NE|Z| > \delta N. \tag{6.3.2}
\]
Since
\[
g_N \sim \frac{g\left(1 - \frac{1}{eU(t)}\right)}{eU(t)\Gamma(\gamma)} \sim \frac{\epsilon^{\gamma-1}t}{\Gamma(\gamma)U(t)},
\]
the left hand side of (6.3.2) is asymptotically equivalent to
\[
\left(\frac{1}{2}\epsilon^{1-\gamma}\Gamma(\gamma) - \epsilon E|Z|\right)U(t),
\]
while the right hand side is asymptotically equivalent to \( \delta \epsilon U(t) \).
Since \( \gamma \) is at least 1, we see that (6.3.2) holds if \( \epsilon \) is small enough and \( t \) is large enough. Hence, applying Lemma A.1 in Mikosch and Samorodnitsky (2000), (6.3.1) is at most
\[
2NP\{ |Z| - E|Z| \geq (t/2g_N) - NE|Z| \} \\
\sim 2N\overline{F}_*((t/2g_N) - NE|Z|) \\
\sim 2\epsilon\left(\frac{1}{2}\epsilon^{1-\gamma}\Gamma(\gamma) - \epsilon E|Z|\right)^{-\alpha}(\text{Id}\overline{F}_*) \circ U(t).
\]
The result follows since \( \gamma \) is at least 1 and \( \overline{F}_* \prec \overline{F} \).

**Step 4.** We now prove that the maximum of the process is unlikely to occur at a time of order larger than \( U(t) \).

**Lemma 6.3.2.** The following holds,
\[
\lim_{\epsilon \to 0} \limsup_{t \to \infty} \frac{\mathbb{P}\{ \exists n : n \geq U(t)/\epsilon, S_n > t \}}{(\text{Id}\overline{F}) \circ U(t)} = 0.
\]

**Proof.** It suffices to prove (3.1.4). Thus, we need to evaluate
\[
P\left\{ \exists n : n \geq U(t)/\epsilon, S_n > t \right\} \\
\cap \left\{ \forall n \geq 1, X_n \leq \theta(n + U(t)/\epsilon) \right\}. \tag{6.3.3}
\]

72
Define
\[ a_n = \theta(n + U(t)/\epsilon) - \mu, \]
so that \( X_n \leq \theta(n + U(t)/\epsilon) \) is equivalent to \( Z_n \leq a_n \). If \( Z_n \) is at most \( a_n \) for all \( n \), then
\[ S_n - s_n = \sum_{0 \leq i < n} g_i Z_{n-i} \mathbb{1} \{ Z_{n-i} \leq a_{n-i} \}. \]

Hence (6.3.3) is at most
\[ \sum_{n \geq U(t)/\epsilon} P \left\{ \sum_{0 \leq i < n} g_i Z_{n-i} \mathbb{1} \{ Z_{n-i} \leq a_{n-i} \} > t - s_n \right\}. \]

We use (4.1.1) to bound each probability involved in this sum, that is to bound
\[ \log P \left\{ \sum_{0 \leq i < n} g_i Z_{n-i} \mathbb{1} \{ Z_{n-i} \leq a_{n-i} \} > t - s_n \right\}. \] (6.3.4)

In (4.1.1), we take
\[ \lambda = c \frac{\log g(1 - 1/n)}{t - s_n} \]
for some positive number \( c \) to be determined later. On the range \( n \geq U(t)/\epsilon \), our chosen \( \lambda \) tends to 0 as \( t \) tends to infinity, uniformly in \( n \). Moreover, as \( t \), and hence \( n \), tends to infinity,
\[ \lambda \leq c \frac{\log g(1 - 1/n)}{-s_n} \sim c \Gamma(1 + \gamma) \frac{\log g(1 - 1/n)}{g(1 - 1/n)}. \]

Using the Karamata representation (Bingham, Goldie and Teugels, 1989, Theorem 1.3.1), \( \log g(1 - 1/n) \sim \gamma \log n \) as \( n \) tends to infinity. Therefore on the range \( n \geq U(t)/\epsilon \), since we assume that the sequence \( (g_n)_{n \geq 0} \) is asymptotically equivalent to a monotone sequence,
\[ \lambda \max_{0 \leq i \leq n} g_i \sim \lambda g_n \leq c \gamma \frac{\log n}{n}. \]

In particular, uniformly on that range of \( n \),
\[ \max_{0 \leq i < n} -EZ \mathbb{1} \{ \lambda g_i Z \leq \log(1 - \epsilon) \} = o(1). \]
as $t$ tends to infinity. Furthermore, referring to the last sum involved in (4.1.1),
\[
\max_{0 \leq i < n} \lambda g_i a_{n-i} \leq \lambda \max_{0 \leq i < n} g_i a_n \lesssim 2c\gamma^2 \theta \log n.
\]

Therefore, since by Potter’s bounds $\mathcal{H}(n/ \log n) = O((\log n)/n)^{\alpha - \epsilon}$, an application of (4.1.1) show that (6.3.4) is at most
\[
-c \log g(1 - 1/n) + \epsilon c E|Z| \log g(1 - 1/n) + \log g(1 - 1/n) o(1) + n^{3\theta \gamma^2 + 1 - \alpha} \log^{\alpha - \epsilon} n.
\]

Taking $\epsilon$ small enough, we first choose $c$ so that, say,
\[
-c + \epsilon c E|Z| \leq -\alpha - 3,
\]
and then $\theta$ small enough so that
\[
3\theta \gamma^2 + 1 - \alpha < 0.
\]

Then, as $t$ tends to infinity and uniformly in $n \geq U(t)/\epsilon$,
\[
P\left\{ \sum_{0 \leq i < n} g_i Z_{n-i} \mathbb{1}\{ Z_{n-i} \leq a_{n-i} \} > t - s_n \right\} = O(n^{-\alpha - 2}).
\]

Since $\text{Id}^{-\alpha - 1} = o(\text{Id} \mathcal{P})$, it follows that (6.3.3) is $o(\text{Id} \mathcal{P}) \circ U(t)$ as $t$ tends to infinity, and this concludes the proof of the lemma.

**Step 5.** We can now prove that if the process exceeds $t$ at a time between $\epsilon U(t)$ and $U(t)/\epsilon$, then at least one of the $X_i$ has to exceed $U(t)$.

**Lemma 6.3.3.** For any positive $\epsilon$ and $p$, there exists $\theta$ such that
\[
P\left( \bigcup_{n \in I} \left( \left\{ S_n > t \right\} \cap \bigcap_{0 < j \leq n} \left\{ Z_j \leq \theta U(t) \right\} \right) \right) = o(t^{-p}).
\]

**Proof.** Recall that $n_0$, defined before Lemma 5.2.1, is an integer such that whenever $n$ is at least $n_0$, both $g_n$ and $g_{[0,n)}$ are nonnegative. Let $t$ be sufficiently large so that $\epsilon U(t)$ is at least $n_0$. If all $Z_j$,
0 < j \leq n, are at most \theta U(t), then the event \( S_n > t \) occurs if and only if
\[
\sum_{0 \leq j < n} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq \theta U(t) \} > t - s_n.
\]
Since \( t - s_n \) exceeds \( t \), the logarithm of the probability of that event is bounded as in the next lemma by (6.3.5) hereafter, with \( \delta = 1 \) say. We conclude as in the proof of the next lemma by taking \( \lambda = ct^{-1} \log t \) with \( c \) large enough and using Bonferroni’s inequality.

**Step 6.** As for the proof of Theorem 2.1.1, we complete only the one-sided version of step 6, namely the version where \( D_n \) is defined as
\[
D_n = \bigcap_{1 \leq i \leq N} \{ S_i - \psi_{i,n}(X_n) \leq \delta t \}.
\]
As we argued when proving Lemma 6.1, Theorem 3 in Jamison, Orey and Pruitt (1965) yields a weak law of large numbers on the weighted sum \( \sum_{0 \leq j < n} g_j Z_{n-j} / g_{[0,n)} \) as \( n \) tends to infinity. The next lemma shows that this weak law of large numbers holds with some uniformity with respect to the weights.

**Lemma 6.3.4.** Let \( \delta \) be a positive real number. For any \( \epsilon \) positive small enough there exists a positive \( \theta \) such that
\[
P\left( \bigcup_{n \in i} B_n^c \cap C_n \cap D_n^c \right) = o(r(t))
\]
as \( t \) tends to infinity.

**Proof.** Let \( N \) be \( \lfloor U(t) / \epsilon \rfloor \). On \( C_n \),
\[
\sum_{0 \leq j < i \atop i-j \neq n} g_j Z_{i-j} = \sum_{0 \leq j < i \atop i-j \neq n} g_j Z_{i-j} \mathbb{1}\{ Z_{i-j} \leq \theta U(t) \}.
\]
In the following, we use the one-sided form of \( D_n \), namely
\[
D_n = \bigcap_{i \leq N} \{ S_i - \psi_{i,n}(X_n) \leq \delta t \}.
\]
Applying inequality (4.1.1) and using that \( n \) is at most \( N \), the logarithm of the probability that \( B_n^c \cap C_n \cap D_n^c \) occurs is at most
\[
-\lambda \delta t + \eta \lambda g_{[0,N)} E|Z| - \lambda \sum_{0 \leq j < N} g_j EZ \mathbb{1}\{ \lambda g_j Z \leq \log(1 - \eta) \} + \sum_{0 \leq j < N} e^{\lambda g_j \theta U(t)} \mathbb{1}\{ \lambda g_j \theta U(t) \} \frac{\log(1 + \eta)}{\lambda g_j} \). \tag{6.3.5}
\]
Note that because we used $N$ in this bound, it holds for all $n$ such that $\epsilon U(t) \leq n \leq U(t)/\epsilon$.

We choose $\lambda = ct^{-1}\log t$ where $c$ will be specified later. With this choice, we examine all terms in (6.3.5). We have $\lambda \delta t = c \delta \log t$. Moreover, Lemma 5.1.1 implies

$$
\lambda g_{\{0,N\}} \sim c \frac{\log t}{t} \frac{g(1-\epsilon/U(t))}{\Gamma(1+\gamma)} \sim \frac{c\epsilon^{-\gamma}}{\Gamma(1+\gamma)} \log t.
$$

It also implies

$$
\lambda g_N \sim c \frac{\log t}{t} \frac{g(1-\epsilon/U(t))}{\Gamma(\gamma)\epsilon t} \sim \frac{c\epsilon^{-\gamma+1}}{\Gamma(\gamma)\epsilon} \log t.
$$

Therefore, since $(\log t)/U(t)$ tends to 0 at infinity, the bound

$$
-Z1\{\lambda g_jZ \leq \log(1-\eta)\} \leq -Z1\{2\lambda g_NZ \leq \log(1-\eta)\}
$$

yields

$$
-\lambda \sum_{0 \leq j < N} g_j E Z 1\{\lambda g_jZ \leq \log(1-\eta)\} = o(\lambda g_{\{0,N\}}).
$$

Finally, we have

$$
\lambda g_N \theta U(t) \sim \frac{c\epsilon^{-\gamma+1}}{\Gamma(\gamma)\theta} \log t,
$$

so that for $t$ large enough,

$$
\sum_{0 \leq j < N} e^{\lambda g_j \theta U(t)} \bar{H}\left(\frac{\log(1+\eta)}{\lambda g_j}\right) \lesssim N t^{2c\epsilon^{-\gamma+1}/\Gamma(\gamma)\theta} \bar{F}\left(\frac{\log(1+\eta)}{\lambda g_N}\right)
$$

$$
\sim U(t)t^{2c\epsilon^{-\gamma+1}/\Gamma(\gamma)\theta} \bar{F}\left(\frac{U(t)}{\log t}\right) O(1).
$$

We obtain that (6.3.5) is at most equivalent to

$$
c \log t\left(-\delta + \frac{\eta \epsilon^{-\gamma}}{\Gamma(1+\gamma)} E|Z| + o(1)\right) + U(t)t^{2c\epsilon^{-\gamma+1}/\Gamma(\gamma)\theta} \bar{F}\left(\frac{U(t)}{\log t}\right) O(1).
$$

Let $p$ be an arbitrary positive real number. We take $c$ large enough and $\eta$ small enough so that

$$
c\left(-\delta + \frac{\eta \epsilon^{-\gamma}}{\Gamma(1+\gamma)} E|Z|\right) \leq -p - 2\gamma - 3.
$$

76
We take $\theta$ small enough so that

$$U(t)t^{\mu - \gamma + 1/2}\gamma \mathcal{F}\left(U(t)/\log t\right) = o(1).$$

Such $\theta$ exists because Potter’s bound applied to both $\mathcal{F}$ and $U$ ensure that the function $U(t)\mathcal{F}(U(t)/\log t)$ tends to 0 at infinity at a rate at least some positive power of $1/t$. Therefore, we obtain, as $t$ tends to infinity,

$$\sup_{n \in I} P(B_n^c \cap C_n \cap D_n^c) = o(t^{-\gamma - 2}).$$

The result follows by an application of Bonferroni’s inequality, upon using Potter’s bound to bound $U(t)$ and taking $p$ to be greater than the negative of the index of regular variation of $r$, that is $(1 - \alpha)/\gamma$ here.

Having completed steps 1 through 5 indicated in section 3.1 as well as the one-sided version of step 6, the upper bound result follows by an application of Theorem 3.2.1.

**6.4. Proof of Theorem 2.1.2 – lower bound.** As with the proof of Theorem 2.1.1, proving the lower bound by an application of Theorem 3.2.1 upon completing step 7 requires a tail balance condition on the distribution function $F$ in order to prove the two-sided version of Step 6. This extra assumption is not needed with the following proof, again adapted from Zachary’s (2004) work.

We assume without loss of generality, that $\mu$ is $-1$. Let $\epsilon$ be a positive real number less than 1 and consider the corresponding set $I$. Let $\delta$ be a positive real number, and let $p$ be an integer depending on $n$, $t$ and $\delta$, such that $g_p \psi_1^{-1}((1 + \delta)t) + s_{n+p}$ is maximum and therefore asymptotically equivalent to $(1 + \delta)t$. For $n$ in $I$, Lemma 5.4.3 shows that $\psi_1^{-1}((1 + \delta)t)$ is of order $U(t)$. Arguments very similar to that of the proof of Lemma 5.4.3 show that $p/U(t)$ remains in a compact subset of the nonnegative half-line when $n$ stays in $I$. Then Lemma 5.1.1 implies that $s_{n+p}/t$ stays bounded over $n$ in $I$ and as $t$ tends to infinity. Therefore, we can find $\eta$ small enough so that $\min_{n \in I} \delta t + \eta s_{n+p}$ is positive for any $t$ large enough.

We consider the events

$$A_{n,t} = \{ S_{n+p} - g_pX_n \geq (1 + \eta)s_{n+p} \text{ and } X_n \geq \psi_1^{-1}((1 + \delta)t) \}.$$
If $A_{n,t}$ occurs, then

\[
S_{n+p} \geq g_p X_n + (1 + \eta)s_{n+p} \\
> g_p \psi_n^{-1}((1 + \delta)t) + (1 + \eta)s_{n+p} \\
\geq (1 + \delta)t + \eta s_{n+p} \\
\geq t,
\]

and therefore $M \geq t$. Consequently, provided $t$ is large enough, the inclusion $A_{n,t} \subset \{ M > t \}$ holds for every $n$. It follows that for $t$ large enough,

\[
P\{ M > t \} \geq P\left( \bigcup_{n \in I} A_{n,t} \right) \\
\geq \sum_{n \in I} P(A_{n,t}) - \sum_{n,m \in I, n \neq m} P(A_{n,t} \cap A_{m,t}) . \tag{6.4.1}
\]

Since $S_{n+p} - g_p X_n$ and $X_n$ are independent, for $n$ larger than $n_0$,

\[
P(A_{n,t}) = P\{ S_{n+p} - g_p X_n > (1 + \eta)s_{n+p} \} P\{ X_n \geq \psi_n^{-1}((1 + \delta)t) \}.
\]

Therefore, Lemma 6.1 implies that provided $n$ is large enough, $P(A_{n,t})$ is at least $(1 - \epsilon)\overline{F} \circ \psi_n^{-1}((1 + \delta)t)$. Hence, if we can prove that the second sum in (6.4.1) is negligible compare to the first one, then

\[
P\{ M > t \} \geq (1 - \epsilon) \sum_{n \in I} \overline{F} \circ \psi_n^{-1}((1 + \delta)t) .
\]

Then, since $\epsilon$ and $\delta$ are arbitrary, the arguments used to derive (3.1.2) shows that, in view of $\chi = U$ and $\rho = \xi_{\Gamma(1+\gamma)}$,

\[
\liminf_{t \to \infty} \frac{P\{ M > t \}}{(\Id F) \circ U(t)} \geq \int_{0}^{\infty} \xi_{\Gamma(1+\gamma)}^{-\alpha}(\nu) d\nu .
\]

But the double sum in (6.4.1) is at most the square of the first one and hence is of order $o(\Id F) \circ U(t)$.

\[\square\]

6.5. Proof of Theorem 2.1.3. Part of the proof is analogous to that of Theorem 2.1.1. Convergence of the sequence $(g_n)_{n \geq 0}$ to $g_\infty$ implies asymptotic equivalence $g(1 - x) \sim g_\infty/x$ as $x$ tends to 0, and therefore $U \sim \Id / g_\infty$ at infinity.
Step 1. Argue as in step 1 of Theorem 2.1.1, using Lemma 5.5.2 instead of Lemma 5.3.2, to show that we may take $\chi \sim \text{Id}$ at infinity, which yields $\psi_{1_{xU(t)}}^{-1}(t) \sim t(1 - \mu x)/g^*$ as $t$ tends to infinity.

Step 2. The arguments used in step 2 of the proof of Theorem 2.1.1 carry over, substituting Lemma 5.5.3 for Lemmas 5.3.3 and 5.3.4.

Step 3. We prove that the process is very unlikely to reach the level $t$ at a time of smaller order than $t$.

Lemma 6.5.1. The following limit hold,

$$\lim_{\epsilon \to 0} \lim_{t \to \infty} \frac{P\left\{ \exists n : 0 \leq n \leq \epsilon U(t), S_n > t \right\}}{tF(t)} = 0.$$  

Proof. Arguing as in the beginning of Lemma 6.1.1— see (6.1.1) — it suffices to prove that for any positive $\delta$, there exists some positive $\theta$ such that

$$\lim_{\epsilon \to 0} \lim_{t \to \infty} \epsilon t \max_{1 \leq n \leq \epsilon t} \frac{P\left\{ \sum_{0 \leq j < n} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq \theta t \} > \delta t \right\}}{tF(t)} = 0. \quad (6.5.1)$$

Let $c$ be a positive real number to be determined later. Using inequality (4.1.1) with $\lambda = c(\delta t - s_n)^{-1}\log t$ and $a_\xi = \theta t$, the logarithm of the ratio

$$P\left\{ \sum_{0 \leq j < n} g_j Z_{n-j} \mathbb{1}\{ Z_{n-j} \leq \theta t \} > \delta t \right\} / F(t)$$

is ultimately at most

$$-c(1 + \epsilon/\delta)^{-1}\log t + \eta cE|Z|\log t$$

$$- \frac{c \log t}{\delta t - s_n} \sum_{0 \leq j < n} g_j EZ \mathbb{1}\left\{ \frac{c \log t}{\delta t - s_n} g_j Z \leq \log(1 - \eta) \right\}$$

$$+ \sum_{0 \leq j < n} e^{\lambda g^* \theta t} \Pi\left( \frac{\log(1 + \eta)}{g^*} \frac{\delta t}{c \log t} \right) + (\alpha + \epsilon) \log t. \quad (6.5.2)$$

Referring to the third summand in this bound, since $n \leq \epsilon t$, it is at most

$$- \frac{c \log t}{\delta} g^* E|Z| \mathbb{1}\left\{ Z \leq \delta \log(1 - \eta) \frac{t}{g^* c} \right\} = o(\log t).$$
Let $\nu$ be a positive real number such that $1 - \alpha + 2\nu$ is negative. The fourth summand is ultimately at most

$$ne^{(cg^*\theta/\delta)}\log t \mathbb{H}\left(\frac{\log(1 + \eta)}{g^*} \frac{\delta t}{c \log t}\right)$$

$$\leq \epsilon t^{1+cg^*/\delta} \mathbb{H}\left(\frac{\log(1 + \eta)}{g^*} \frac{\delta t}{c \log t}\right)$$

$$\leq \epsilon t^{1+(cg^*/\delta)-\alpha+\nu},$$

where we used Potter’s bound to obtain the last inequality. Therefore, (6.5.2) is ultimately at most

$$(-c(1 + \epsilon/\delta)^{-1} + \eta c E|Z| + \alpha + \epsilon + o(1)) \log t + \epsilon t^{1+(cg^*/\delta)-\alpha+\nu}.$$ 

We take

$$\eta = 1/2E|Z|, \quad \epsilon = \delta/3, \quad c = 8\alpha$$

and $\theta$ small enough so that $cg^*/\delta \leq \nu$, which, given how $\nu$ was defined, guarantees that $1+(cg^*/\delta) - \alpha+\nu$ is negative. This proves (6.5.1) as well as the lemma.

**Step 4.** We need to prove that the process is very unlikely to reach the level $t$ at a time of larger order than $t$. This follows from Lemma 6.1.2 whose proof, and hence, conclusion, remains valid in the present context.

**Step 5.** Similarly to the previous step, Lemma 6.1.3 remains valid in the present context.

**Step 6.** Similarly to the previous step, Lemma 6.1.4 remains valid.

An application of Theorem 3.2.1 yields the upper bound. The proof of the lower bound of Theorem 2.1.1 carries over in the present setting, and this concludes the proof of Theorem 2.1.3.

**6.6. Proof of Theorem 2.2.1.** We only sketch the proof. Assume without loss of generality that $\mu = -1$ and define as before

$$\psi_{i,n}(x) = E(S_i \mid X_n = x) = \begin{cases} 
  s_i & \text{if } i < n, \\
  (x+1)g_{i-n} + s_i & \text{if } i \geq n.
\end{cases}$$
For $x$ positive, we define
\[ \psi^+_{i, n}(x) = \max_{k \geq 0} x g_k + s_{i+k} \]
and for $x$ negative define
\[ \psi^-_{i, n}(x) = \max_{k \geq 0} x g_k + s_{i+k} . \]

By the same arguments as in our heuristic, we expect to prove that to reach the level $t$, either $\psi^+_{+, n}(X_n) > t$ or $\psi^-_{-, n}(X_n) > t$ for some $n$. The actual proof can be done by redefining $B_n$ as the two-sided event $\{ |X_n| \leq \theta \chi(t) \}$ and using the tail balance condition. Thus we have
\[
P\{M > t\} \sim \sum_{n \geq n_1} F \circ \psi^{-1}_{+, n}(t) + \sum_{n \geq n_1} F \circ \psi^{-1}_{-, n}(t) .
\]

Similarly to what we proved previously, one has
\[ \psi^+_{+, n}(x) \sim g^* x + s_n \]
as $x$ tends to infinity and
\[ \psi^-_{-, n}(x) \sim g^* x + s_n \]
as $x$ tends to minus infinity. The result follows as in the proof of Theorem 2.1.1.

6.7. Proof of Theorem 2.3.1. The theorem is proved by applying Theorem 3.3.1. The tail balance condition (2.2.1) guarantees that (S6) holds with the two-sided event $D_n$; this can be seen by exactly the same arguments we used to prove Lemma 6.3.4, using two-sided versions of the events $B_n$. Thus, it remains to show that (3.3.1) holds.

We first assume without loss of generality that $\mu = -1$. Equality (6.1) shows that for any positive real number $\lambda$, $\tau$ and $y$,
\[
\begin{align*}
h_t(\lambda, \tau, y) &= t^{-1} \psi_{\lambda U(t)], [\tau U(t)]} (\chi(t)y) \\
&= -t^{-1} g_{[0, \lambda U(t)]} \\
&\quad + 1\{ \lambda U(t) \geq [\tau U(t)] \} t^{-1} g_{[\lambda U(t)] - [\tau U(t)]} (\chi(t)y + 1).
\end{align*}
\]
Therefore, setting
\[ h(\lambda, \tau, y) = -\frac{\lambda^\gamma}{\Gamma(1 + \gamma)} + 1\{\lambda \geq \tau\} \frac{(\lambda - \tau)^{\gamma - 1}}{\Gamma(\gamma)} y, \]
and using that we take \( \chi \) equal to \( U \) in this case, we have, using Lemma 5.1.1, the pointwise convergence
\[ \lim_{t \to \infty} h_t(\lambda, \tau, y) = h(\lambda, \tau, y). \]

Let \( \epsilon \) be a positive real number. We prove that \( h_t(\cdot, \tau, y) \) tends to \( h(\cdot, \tau, y) \) in \( D[0,1/\epsilon] \). For this, note that
\[ 1\{|\lambda U(t)| \geq |\tau U(t)|\} = 1\left\{\frac{\tau U(t)}{|\tau U(t)|} \lambda \geq \tau \right\}. \]

Set \( v_t(\lambda) = \lambda [\tau U(t)] / (\tau U(t)) \). Then Lemma 5.1.1 implies that, as \( t \) tends to infinity,
\[
\begin{align*}
&h_t(v_t(\lambda), \tau, y) - h(\lambda, \tau, y) = -t^{-1}g(0, v_t(\lambda) U(t)) + \frac{\lambda^\gamma}{\Gamma(1 + \gamma)} \\
&+ 1\{\lambda \geq \tau\} \left( t^{-1}g_{[\tau U(t)], \lambda} - [\tau U(t)] \gamma(\chi(t)y + 1) - \frac{(\lambda - \tau)^{\gamma - 1}}{\Gamma(\gamma)} y \right)
\end{align*}
\]
tends to 0 uniformly in \( \lambda \) in any fixed compact subset of the positive half-line. Moreover, since \( \gamma \) is positive and the \( g_i \) are ultimately positive, Lemma 5.1.1 also shows that \( t^{-1}g(0, \lambda U(t)) \) tends to \( \lambda^\gamma / \Gamma(1 + \gamma) \) uniformly on any interval of the form \([0,1/\epsilon]\). Therefore, taking \( \epsilon \) to be less than \( \tau \), this shows that \( h_t(v_t(\lambda), \tau, y) - h(\lambda, \tau, y) \) converges uniformly to 0 on \([0,1/\epsilon]\). Since \( v_t \) tends to the identity uniformly in \([0,1/\epsilon]\), it follows from the definition of the Skorohod topology (see Billingsley, 1968, definition of the distance \( d \) in section 14) that for every \( \tau \) and \( y \) the function \( h_t(\cdot, \tau, y) \) converges to \( h(\cdot, \tau, y) \) in \( D[0,1/\epsilon] \).

We then apply Theorem 3.3.1 to obtain Theorem 2.3.1 when \( \mu = -1 \). For a general negative mean \( \mu \), let \( (X_i)_{i \geq 0} \) be as before a sequence of independent and identically distributed random variables with mean \(-1\) and let \( \tilde{X}_i = (-\mu)X_i, i \geq 0 \). We agree to cover by a tilde whatever quantity is calculated on the \( \tilde{X}_i \) and to leave uncovered quantities calculated on the \( X_i \). Then, with the notation of section 2,
\( \tilde{S}_n = (-\mu)S_n \),
\( \tilde{M} > t \) if and only if \( M > t/(-\mu) \),
\( \tilde{N}_t = N_t/(-\mu) \),
\( \tilde{J}_t = J_t/(-\mu) \),
\( \tilde{\tau}_t = \frac{\tilde{J}_t}{U(t)} = \frac{U(t/(-\mu))}{U(t)} \tau_t/(-\mu) \),
\( \tilde{Y}_t = \frac{\tilde{X}_t}{U(t)} = (-\mu) \frac{U(t/(-\mu))}{U(t)} Y_t/(-\mu) \),
\( \tilde{e}_t(\lambda) = -\mu S(\lambda U(t)) = \tilde{e}_{t/(-\mu)}\left( \frac{U(t)}{U(t/(-\mu))} \right) \).

It follows that the limiting random variables \((\tilde{e}, \tilde{\tau}, \tilde{Y})\) satisfy
\( \tilde{e} = e((-\mu)^{1/\gamma} \cdot) \), \( \tilde{\tau} = (-\mu)^{-1/\gamma} \tau \) and \( \tilde{Y} = (-\mu)^{1-1/\gamma} Y \),
and this completes the proof of Theorem 2.3.1.

Acknowledgements. Ph.B. thanks Florin Avram for showing him some works on combinatorics which was particularly enlightening and which, a year later, was very inspiring when working on this paper.

References

J. Akonom, Chr. Gouriéroux (1987). A functional central limit theorem for fractional processes, discussion paper 8801, CEPREMAP, Paris.
S. Asmussen (1987). Applied Probability and Queues, Wiley.
S. Asmussen (2000). Ruin Probabilities, World Scientific.
S. Asmussen, C. Klüppelberg (1996). Large deviation results for subexponential tails, with applications to insurance risk, Stoch. Proc. Appl., 64, 103–125.
S. Asmussen, H. Schmidli, V. Schmidt (1999). Tail probabilities for non-standard risk and queueing process with subexponential jumps, Adv. Appl. Probab., 31, 422–447.
F. Baccelli, S. Foss (2004). Moments and tails in monotone-separable stochastic networks, Ann. Probab., 14, 612–650.
F. Baccelli, S. Foss, M. Lelarge (2005). Tails in generalized Jackson networks with subexponential service-time distributions, J. Appl. Probab., 42, 513–530.
Ph. Barbe, M. Broniatowski (1998). Note on functional large deviation principle for fractional ARIMA processes, *Statistical Inference for Stochastic Processes*, 1, 17–27.

Ph. Barbe, A.-L. Fougères, Chr. Genest (2006). On the tail behavior of sums of dependent risks, *Astin Bulletin*, 36, 361–373.

J. Beran (1994). *Statistics for Long-Memory Processes*, Chapman and Hall.

P. Billingsley (1968). *Convergence of Probability Measures*, Wiley.

N.H. Bingham, C.M. Goldie, J.L. Teugels (1989). *Regular Variation*, 2nd ed. Cambridge University Press.

E. Bolthausen (1993). Stochastic processes with long range interactions of the paths, *Contemporary Mathematics*, 149, 297-319.

A.A. Borovkov (1971). *Stochastic Processes in Queueing Theory*, Springer.

A.A. Borovkov (2003). Large deviation probabilities for random walks in the absence of finite expectation jumps, *Probab. Theor. Relat. Fields*, 125, 421–446.

B.L.J. Braaksma, D. Stark (1997). A Darboux-type theorem for slowly varying functions, *J. Comb. Theor.*, A, 77, 51–66.

S.S. Choi, Th.M. Cover, I. Csizár (1987). Conditional limit theorem under Markov conditioning, *IEEE Trans. Inform. Theory*, IT-33, 788–801.

D.B.H. Cline, T. Hsing (1991). Large deviation probability for sums of random variables with heavy or subexponential tails, preprint, Texas A&M Univ.

A. Dembo, O. Zeitouni (1992). *Large Deviation Techniques and Applications*, Bartlett.

P. Embrechts, D.D. Lambrigger, M.V. Wütrich (2008). Multivariate extremes and the aggregation of dependent risks: example and counter-examples, preprint.

P. Embrechts, J. Nešlehová, M.V. Wütrich (2008). Additivity properties for Value-at-Risk under Archimedean dependence and heavy-tailedness, *Insurance, Mathematics and Economics*, to appear.

P. Embrechts, N. Veraverbeke (1982). Estimates for the probability of ruin with special emphasis on the possibility of large claims, *Insurance: Math. Economics*, 1, 55–77.

S. Foss, Z. Palmowski, S. Zachary (2005). The probability of exceeding a high boundary on a random time for a heavy-tailed random walk, *Ann. Appl. Probab.*, 15, 1936–1957.
C.W.J. Granger (1980). Long memory relationships and the aggregation of dynamic models, *J. Econometrics*, 14, 227–238.

B. Jamison, S. Orey, W. Pruitt (1965). Convergence of weighted averages of independent random variables, *Zeit. Wahrsch. Theor. Verw. Geb.*, 4, 40–44.

P.S. Kokoszka, M.S. Taqqu (1995). Fractional ARIMA with stable innovations, *Stoch. Proc. Appl.*, 60, 19–47.

D. Konstantinides, T. Mikosch (2005). Large deviations and ruin probabilities for solutions of stochastic recurrence equations with heavy-tailed innovations, *Ann. Probab.*, 33, 1992–2035.

D. Korshunov (1997). On distribution tail of the maximum of a random walk, *Stoch. Proc. Appl.*, 72, 97–103.

A.J. McNeail, R. Frey, P. Embrechts (2005). *Quantitative Risk Management: Concept, Techniques, Tools*, Princeton University Press.

T. Mikosch, G. Samorodnitsky (2000). The supremum of a negative drift random walk with dependent heavy-tail steps, *Ann. Appl. Probab.*, 10, 1025–1064.

A. Montanari, R. Rosso, M.S. Taqqu (1997). Frationally differenced ARIMA models applied to hydrologic time series: identification, estimation, and simulation, *Water Resources Research*, 33, 1035–1044.

A. Nagaev (1969a). Integral limit theorem for large deviations when Cramér’s condition is not fulfilled, I, II, *Theory Probab. Appl.*, 14, 51–64, 193–208.

A. Nagaev (1969b). Limit theorems for large deviations when Cramér’s conditions are violated, *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk*, 6, 17–22 (in Russian).

K.W. Ng, Q. Tang, J.-A. Yan, H. Yang (2004). Precise large deviations for sum of random variables with consistently varying tails, *J. Appl. Prob.*, 41, 93–107.

H. Nyrhnien (2005). Power estimates for ruin probabilities, *Adv. Appl. Probab.*, 37, 726–742.

K.B. Oldham, J. Spanier (2006). *The Fractional Calculus*, Dover.

A.G. Pakes (1975). On the tails of waiting-time distributions, *J. Appl. Probab.*, 12, 555–564.

P.C.B. Phillips (1987). Time series regression with a unit root, *Econometrica*, 55, 277–301.

D. Pollard (1984). *Convergence of Stochastic Processes*, Springer.
S.I. Resnick (2007). *Heavy-Tail Phenomena, Probabilistic and Statistical Modeling*, Springer.

G. Samorodnitsky, M.S. Taqqu (1992). Linear models with long-range dependence and finite or infinite variance, in *New Directions in Time Series Analysis, Part II*, D. Brillinger, P. Caines, J. Geweke, E. Parzen, M. Rosenblatt eds., Springer.

K. Tanaka (1999). The nonstationary fractional unit root, *Econometric Theory*, 15, 549–582.

S.R.S. Varadhan (1966). Asymptotic probabilities and differential equations, *Comm. Pure Appl. Math.*, 19, 261–286.

N. Veraverbeke (1977). Asymptotic behavior of Wiener-Hopf factors of a random walk. *Stoch. Proc. Appl.*, 5, 27–37.

Q. Wang, Y.-X. Lin, Ch. Gulati (2002). Asymptotics for general nonstationary fractionally integrated process without prehistoric influence, *J. Appl. Math. Decision Sci.*, 6, 255-269.

M. Woodroofe (1982). *Nonlinear Renewal Theory in Sequential Analysis*, SIAM.

S. Zachary (2004). A note on Veraverbeke’s theorem, *Queueing Systems*, 46, 9–14.