XXII ESCUELA VENEZOLANA DE MATEMÁTICAS

Some applications of regular variation in probability and statistics

Philippe Soulier

MÉRIDA, VENEZUELA, 9 AL 15 DE SEPTIEMBRE DE 2009

XXII ESCUELA VENEZOLANA DE MATEMÁTICAS

Some applications of regular variation in probability and statistics

Philippe Soulier

Université de Paris Ouest-Nanterre

email philippe.soulier@u-paris10.fr

MÉRIDA, VENEZUELA, 9 AL 15 DE SEPTIEMBRE DE 2009

XXII ESCUELA VENEZOLANA DE MATEMÁTICAS

La Escuela Venezolana de Matemáticas es una actividad de los postgrados en matemáticas de las siguientes instituciones: Centro de Estudios Avanzados del Instituto Venezolano de Investigaciones Científicas, Facultad de Ciencias de la Universidad Central de Venezuela, Facultad de Ciencias de la Universidad de Los Andes, Universidad Simón Bolívar, Universidad Centroccidental Lisandro Alvarado y Universidad de Oriente, y se realiza bajo el auspicio de la Asociación Matemática Venezolana.

La XXI ESCUELA VENEZOLANA DE MATEMÁTICAS recibió financiamiento de la Academia de Ciencias Físicas, Matemáticas y Naturales de Venezuela, el Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT), el Instituto Venezolano de Investigaciones Científicas (Departamento de Matemáticas y Ediciones IVIC), la Universidad de Los Andes (CEP, CDCHT, Facultad de Ciencias y Departamento de Matemáticas) y el Rectorado de la Unversidad Centroccidental Lisandro Alvarado.

2000 Mathematics Subject Classification: 60E05, 60F05, (60G18, 60G70).

Some applications of regular variation in probability and statistics
Philippe Soulier

Diseño y edición: Escuela Venezolana de Matemáticas
Preprensa e impresión: Editorial Texto
Depósito legal If66020095102444
ISBN 978-980-261-111-9
Caracas, Venezuela
2009

Preface

The purpose of these notes is to study some applications of univariate regular variation in probability and statistics. One early field of application of regular variation is extreme value theory, which has in turn contributed to developments in the theory of regular variation. Both theories and their links are very well described in classical textbooks such as Feller (1971), Sidney (1987), Bingham et al. (1989) and the recent monograph on extreme value theory De Haan and Ferreira (2006).

More recently, the study of long memory (or long range dependent) stochastic processes has given rise to an important probabilistic and statistical literature where regular variation is used both for modelization of some random phenomena and as a tool for studying the properties of the models. Long range dependence has no formal definition: a stationary stochastic process is said to be long range dependent if some statistic of interest behaves in a very different way from the same statistic computed on an i.i.d. sequence. In many applications, there is enough evidence of a non standard behaviour of many commonly used statistics to look for models providing a theoretical justification. Long memory processes are now widely used in hydrology, finance, teletrafic modeling. The most commonly used definition of long memory for a stationary process is the regular variation of the autocovariance function, but it can be defined in many other ways, in particular when the autocovariance function does not exist. In all definitions, regular variation of some functional parameter is involved.

Regular variation is not only a formal similarity between extreme value theory and long memory processes. There are also structural links between certain long memory processes and heavy tailed random variables. The long memory property may be the consequence of the extreme duration of some random effect, and long memory can in turn transfer to heavy tails of the asymptotic distribution of some statistic of a long memory process. These structural links have been widely investigated, most notably in teletrafic modeling and financial time series.

The formal similarity of extreme value theory and long memory processes can also be seen from the statistical point of view. The nature of the statistical problem of estimating an index of regular variation is in some ways independent of the underlying stochastic structure. In some cases, the methodology developped for estimating the index of regular variation in extreme value theory can be transfered to obtain estimates of the index of regular which characterizes a long memory process.

Due to time constraints, these notes are entirely devoted to the univariate case. The applications of multivariate regular variation to extreme value theory are well developped and can be found in the monograph De Haan and Ferreira (2006). There are also some recent applications of multivariate regular variation to self-similar random fields which extend the univariate structural links mentionned above. The theory of regular variation in metric spaces has some applications in extreme value theory and is developping.

These notes are organized as follows. The first chapter recalls the basics of regular variation from an analytical point of view. The main source is the classical monograph Bingham et al. (1989). Chapter two recalls the main properties of random variables with regularly varying tails. The results of the first chapter are used to prove that certain transformations of independent regularly varying random variables are again regularly varying. The third chapter is concerned with the role of regular variation in limit theorems. The fourth chapter introduces the concept of long range dependence. Several types of models are described and the consequences of long memory are illustrated by non standard limit theorems. The last chapter presents two statistical problems related to regular variation: the estimation of the tail index and of the long memory parameter. The point of view is minimax semiparametric estimation.

To conclude this preface, I thank the Escuela Venezolana de Matemáticas for giving me the opportunity to teach this course and for publishing these lecture notes. I thank Dr. Carenne Ludeña for suggesting my name and for inviting me to Venezuela.

Contents

Chapter 1

Univariate regular variation

This chapter presents the main definition and results that will be used later. Its contents are well-known and mainly taken from the classical monography Bingham et al. (1989). To quote the preface of this book, the contents of this chapter are what "the mathematician of the street ought to know about regular variation". For the purpose of these notes, they have been streamlined to what the "probabilist or statistician of the street" should know. They include the representation theorem, Karamata theorem about integrals, and some elementary Abelian and Tauberian theorems about Laplace and Fourier transforms.

1.1 Karamata Theory

Let f be a function defined on an unbounded interval I with value in $\mathbb{R}_+^*$, such that the limit

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} \tag{1.1}$$

exists for all $x \in I$. What can then be said of the function f? Nothing without additional assumptions. Fortunately, the assumption of measurability is sufficient to obtain a characterization of the functions that satisfy (1.1).

Theorem 1.1. *Let f be a positive measurable function defined on an unbounded interval I of $(0, \infty)$ that for all $x \in I$, the limit (1.1) exists.*

Then there exists $\alpha \in \mathbb{R}$ *such that*

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^\alpha \;, \tag{1.2}$$

for all $x \in (0,\infty)$. *Moreover, the convergence is uniform on compact subsets of* $(0,\infty)$.

If $\alpha \neq 0$ in (1.2), the function f is said to be regularly varying at infinity with index α, denoted $f \in RV_\infty(\alpha)$. If $\alpha = 0$, f is said to be slowly varying at infinity, denoted $f \in RV_\infty(0)$ or $f \in SV_\infty$. In the sequel, the function f will always be assumed to be measurable.

Corollary 1.2. *Let* $\alpha \neq 0$ *and* $f \in RV_\infty(\alpha)$. *There exists a slowly varying function* ℓ *such that* $f(x) = x^\alpha \ell(x)$.

The properties of regularly varying functions are thus deduced from those of slowly varying functions.

Example 1.3. Powers of logarithm and iterated logarithms are slowly varying at infinity. For $\gamma, \delta \geq 0$ such that $\gamma + \delta < 1$ and $c > 0$, the function $x \mapsto \exp\{c \log^\gamma(x) \cos^\delta(\log(x))\}$ is slowly varying at infinity. If ℓ is slowly varying, then so is $[\ell]$.

In view of applications, it is convenient to characterize regular variation by convergence along subsequences. We omit the proof of the following results, cf. Bingham et al. (1989, Theorem 1.9.2, 1.10.3).

Theorem 1.4. *Let* f *and* g *be positive continuous functions. Let* $0 < a < b < \infty$ *and let* $\{a_n\}$ *and* $\{b_n\}$ *be sequences such that*

$$\limsup_{n\to\infty} a_n = \infty \;, \quad \limsup_{n\to\infty} a_n/a_{n+1} = 1 \;,$$
$$\lim_{n\to\infty} b_n f(a_n x) = g(x)$$

for all $x \in (a,b)$. *Then* f *is regularly varying.*

Continuity may be too strong an assumption. It can be replaced by monotonicity, which will be sufficient in applications.

Theorem 1.5. *Let f be a positive and ultimately monotone function. Let g be a function and $\{a_n\}$ and $\{x_n\}$ be sequences such that*

$$\limsup_{n\to\infty} x_n = \infty \ , \quad \limsup_{n\to\infty} a_n/a_{n+1} = 1 \ ,$$
$$\lim_{n\to\infty} a_n f(x_n t) = g(t)$$

for t in a dense subset of $(0,\infty)$. Then f is regularly varying.

Regular and slow variation can be defined at zero as well. A function f defined on $(0,\infty)$ is said to be regularly varying at zero with index $\alpha \in \mathbb{R}$ if the function $x \to f(1/x)$ is regularly varying at infinity with index $-\alpha$.

1.1.1 The representation Theorem

Theorem 1.6. *Let ℓ be a function defined on $(0,\infty)$, slowly varying at infinity. There exist x_0 and functions c and η defined on $[x_0,\infty)$ such that*

$$\lim_{x\to\infty} c(x) = c^* \in (0,\infty) \ , \tag{1.3}$$
$$\lim_{x\to\infty} \eta(x) = 0 \tag{1.4}$$

and for all $x \geq x_0$,

$$\ell(x) = c(x) \exp \int_{x_0}^{x} \frac{\eta(s)}{s} \, \mathrm{d}s \ . \tag{1.5}$$

The functions c can η be chosen in such a way that η is infinitely differentiable.

Remark 1.7. If ℓ is slowly varying and $\tilde{\ell} \sim \ell$, then $\tilde{\ell}$ is slowly varying. Any slowly varying function is equivalent to an infinitely differentiable slowly varying function.

Remark 1.8. A slowly varying function that admits the representation (1.5) where the function c is constant is called *normalized*. It can be proved that a slowly varying function is normalized slowly varying if and only if it belongs to the Zygmund class, defined as those functions ℓ such that for all $\alpha > 0$, $x^{-\alpha}\ell(x)$ is ultimately non increasing and $x^{\alpha}\ell(x)$ is ultimately non decreasing.

Proof of Theorem 1.6. Let $\ell \in SV(\infty)$ and define $m(x) = \log \ell(\mathrm{e}^x)$. Then

$$\lim_{x\to\infty} m(x+y) - m(x) = 0 \; . \tag{1.6}$$

We shall first prove that the convergence is uniform on compact sets. Without loss of generality, we prove the uniform convergence for $y \in [0,1]$.

Let $\epsilon > 0$ and for $x \geq 0$, define $E_x = \{z \in \mathbb{R} \mid |z-x| \leq 1 \;\; |m(z) - m(x)| \geq \epsilon/2\}$, $T_x = E_x - x$. Condition (1.6) implies that a given t in $[-1,1]$ cannot belong to T_x x large enoug, i.e. $\lim_{x\to\infty} \mathbb{1}_{T_x}(t) = 0$ for all t. By bounded convergence, we thus have

$$\lim_{x\to\infty} \mathrm{Leb}(T_x) = 0 \; .$$

Note that the measurability assumption is used here. Similarly, for x large enough, $\mathrm{Leb}(E_x) < 1/2$. For all $y \in [0,1]$, it also holds that $\mathrm{Leb}(E_{x+y}) < 1/2$ if $x+y$ is large enough. Thus, there exists x_0 such that if $x \geq x_0$, then $\mathrm{Leb}(E_x) < 1/2$ and $\mathrm{Leb}(E_{x+y}) < 1/2$ for all $y \in [0,1]$. Thus $E_x \cup E_{x+y} \subset [x-1, x+y+1]$ and $\mathrm{Leb}(E_x \cup E_{x+y}) < 1$, which implies that there exists some $z \in [x, x+1]$ such that $z \notin E_x \cup E_{x+y}$, and

$$|m(z) - m(x)| \leq \epsilon/2 \; , \;\; |m(z) - m(x+y)| \leq \epsilon/2 \; ,$$

whence

$$|m(x+y) - m(x)| \leq \epsilon \; ,$$

for all $y \in [0,1]$, if $x \geq x_0$. We have proved that the convergence is uniform.

It is now easy to see that the function m is bounded hence integrable on any compact interval $[a,b]$ for a large enough.

Let a be such that m is locally bounded on $[a,\infty)$ and let g be defined for $x \geq 0$ by

$$g(x) = \int_x^{x+1} m(u)\,\mathrm{d}u - m(x) = \int_0^1 \{m(x+u) - m(x)\}\,\mathrm{d}u \; .$$

The uniform convergence implies that $\lim_{x\to\infty} g(x) = 0$. Note now that

$$\int_x^{x+1} m(u)\,\mathrm{d}u = \int_a^{a+1} m(u)\,\mathrm{d}u + \int_a^x \{m(u+1) - m(u)\}\,\mathrm{d}u\ .$$

Denote

$$\gamma = \int_a^{a+1} m(u)\,\mathrm{d}u\ , \quad \zeta(x) = m(x+1) - m(x)\ .$$

Then, $\lim_{x\to\infty} \zeta(x) = 0$ and

$$m(x) = \int_x^{x+1} \{m(u+1) - m(u)\}\,\mathrm{d}u - g(x) = \gamma + \int_a^x \zeta(u)\,\mathrm{d}u - g(x)\ .$$

Altogether, we have obtained

$$\begin{aligned}\ell(x) = \exp\{m(\log x)\} &= \exp\{\gamma - g(\log(x)\} \exp\left\{\int_a^{\log(x)} \zeta(u)\,\mathrm{d}u\right\} \\ &= \exp\{\gamma - g(\log(x)\} \exp\left\{\int_a^x \frac{\zeta(\log(s))}{s}\,\mathrm{d}s\right\} \\ &= c(x) \exp\left\{\int_a^x \frac{\zeta(\log(s))}{s}\,\mathrm{d}s\right\}\ ,\end{aligned}$$

where the functions c and η are defined by $c(x) = \gamma - g(x)$ and $\eta(s) = \zeta(\log(s)$ and satisfy $\lim_{x\to\infty} c(x) = \mathrm{e}^\gamma$ et $\lim_{x\to\infty} \eta(x) = 0$ as required.

There now remain to prove that the function η can be chosen arbitrarily smooth, by a suitable modification of the function c. Using the previous notation, denote $F_1(x) = \int_a^x \{m(u+1) - m(u)\}\,\mathrm{d}u$. Then,

$$\begin{aligned}&F(x+y) - F(x) \\ &\qquad = \int_x^{x+y} \{m(u+1) - m(u)\}\,\mathrm{d}u \\ &\qquad = \int_0^y \{m(x+u+1) - m(x+u)\}\,\mathrm{d}u \\ &\qquad = \int_0^y \{[m(x+u+1) - m(x)] - [m(x+u) - m(x)]\}\,\mathrm{d}u\ .\end{aligned}$$

By the local uniform convergence to zero of the function $m(x+\cdot)-m(x)$, we have

$$\lim_{x\to\infty} F(x+y) - F(x) = 0$$

for all y. The previous result can thus be applied: there exists functions δ_1 and ζ_1 such that $\lim_{x\to\infty} \delta_1(s)$ exists in $\mathbb{R}$, $\lim_{s\to\infty} \zeta_1(s) = 0$ and

$$F(x) = \delta_1(x) + \int_a^x \zeta_1(s)\mathrm{d}s \; .$$

The function ζ_1 can be chosen as previously, i.e.

$$\zeta_1(s) = F(s+1) - F(s) \; ,$$

hence ζ_1 is continuous. The result follows by iteration of this method. □

Let ℓ be a slowly varying function, such that (1.5) holds with c a constant function and η differentiable. The function η can then be expressed in terms of the function ℓ as follows: for $x \geq x_0$

$$\eta(x) = \frac{x\ell'(x)}{\ell(x)} \; . \tag{1.7}$$

Conversely, any continuously differentiable function ℓ such that

$$\lim_{x\to\infty} \frac{x\ell'(x)}{\ell(x)} = 0 \; , \tag{1.8}$$

is slowly varying at infinity. This extends to a characterization of differentiable regularly varying functions.

Corollary 1.9. *A continuously differentiable function is regularly varying at infinty with index $\alpha \in \mathbb{R}$ if and only if*

$$\lim_{x\to\infty} \frac{x\ell'(x)}{\ell(x)} = \alpha \; .$$

If $f \in RV_\infty(\alpha)$ and $\rho \neq 0$, uniform convergence holds on larger subsets of $[0,\infty]$ than compact subsets of $(0,\infty)$.

Proposition 1.10. *If $f \in RV_\infty(\alpha)$, then, for any $a > 0$, convergence in (1.2) is uniform on*
- *$[a,\infty)$ if $\alpha < 0$;*

- *$(0,a]$ if $\alpha > 0$ and f is locally bounded on $(0,a]$.*

Exercise 1.1.1. Prove that the functions of Example 1.3 are slowly varying.

Exercise 1.1.2. Prove that the function f defined on $\mathbb{R}_+^*$ by $f(x) = \exp\{\cos(\log(x))\}$ is not slowly varying.

Exercise 1.1.3. Let $f \in RV_\infty(\alpha)$ with $\alpha > 0$. Prove that

$$\lim_{x\to\infty} \frac{\inf_{u\geq x} f(u)}{f(x)} = 1 ,$$

which means that a regularly varying function with positive index is equivalent to an increasing function.

1.1.2 Bounds

The representation Theorem yields important and useful bounds for slowly and regularly varying functions, known as Potter's bounds. We start with a result for slowly varying functions.

Proposition 1.11. *Let $\ell \in SV(\infty)$. Then, for all $\epsilon > 0$,*

$$\lim_{x\to\infty} x^\epsilon \ell(x) = \infty , \quad \lim_{x\to\infty} x^{-\epsilon} \ell(x) = 0 ,$$
$$\lim_{x\to\infty} \frac{\log \ell(x)}{\log(x)} = 0 .$$

This last limit follows from the representation Theorem and the following Lemma, which is a continuous version of Cesaro's lemma.

Lemma 1.12. *Let f be defined on $[a,\infty)$ such that $\lim_{x\to\infty} f(x)$ exists. Let g be non negative and $\int_a^\infty g(x)\,\mathrm{d}x = \infty$. Then*

$$\lim_{x\to\infty} \frac{\int_a^x f(x)g(x)\,\mathrm{d}x}{\int_a^x g(x)\,\mathrm{d}x} = \lim_{x\to\infty} f(x) .$$

We can now state the celebrated Potter's bounds.

Theorem 1.13 (Potter's bounds)**.** *Let $f \in RV_\infty(\alpha)$. For any $\epsilon > 0$ and $C > 1$, there exists x_0 such that for all $y \geq x \geq x_0$,*

$$C^{-1}(y/x)^{\alpha-\epsilon} \leq \frac{f(y)}{f(x)} \leq C(y/x)^{\alpha+\epsilon} .$$

Proof. By Theorem 1.6, we can write $f(x) = x^\alpha c(x)\ell(x)$, with

$$\ell(x) = \exp \int_{x_0}^{x} \frac{\eta(s)}{s} \, \mathrm{d}s \; ,$$

$\lim_{x\to\infty} \eta(x) = 0$ and $\lim_{x\to\infty} c(x) = c \in (0,\infty)$. For any $\epsilon > 0$, fix some x_0 such that $|c(x) - c| \leq c\epsilon$ and $|\eta(x)| \leq \epsilon$ for $x \geq x_0$. Thus, if $y \geq x \geq x_0$, we have

$$\begin{aligned} \frac{f(y)}{f(x)} &= \frac{c(y)}{c(x)} (y/x)^\alpha \exp \int_x^y \frac{\eta(s)}{s} \, \mathrm{d}s \\ &\leq \frac{1+\epsilon}{1-\epsilon} (y/x)^\alpha \exp \epsilon \int_x^y \frac{\mathrm{d}s}{s} \leq \frac{1+\epsilon}{1-\epsilon} (y/x)^{\alpha+\epsilon} \; . \end{aligned}$$

The lower bound is obtained similarly. □

Lemma 1.14 (Jumps of a regularly varying function)**.** *Let $f \in RV_\infty(\alpha)$ and assume moreover that f is non decreasing on $[x_0,\infty)$. For any $x \geq x_0$, define*

$$\Delta f(x) = \lim_{y\to x, y>x} f(x) - \lim_{y\to x, y<x} f(x) \; .$$

Then $\lim_{x\to\infty} \Delta f(x)/f(x) = 0$.

Proof. Since f is increasing, for any $s < 1 < t$ and large enough x, we have

$$0 \leq \limsup_{x\to\infty} \frac{\Delta f(x)}{f(x)} \leq \limsup_{x\to\infty} \frac{f(tx) - f(sx)}{f(x)} = t^\alpha - s^\alpha \; ,$$

and this can be made arbitrarily small by choosing s and t close enough to 1. □

1.1.3 Integration

One of the main use regular variation is to obtain equivalent of integrals. The next result is often referred to as Karamata's Theorem.

Theorem 1.15. *Let ℓ be slowly varying at infinity and let $\beta > 0$. Then*

$$\lim_{x\to\infty} \frac{1}{x^\beta \ell(x)} \int_a^x \ell(t) t^{\beta-1} \, \mathrm{d}t = \lim_{x\to\infty} \frac{1}{x^{-\beta}\ell(x)} \int_x^\infty \ell(t) t^{-\beta-1} \, \mathrm{d}t = \frac{1}{\beta} \; . \tag{1.9}$$

The function $x \to \int_a^x t^{-1}\ell(t)\,\mathrm{d}t$ *is slowly varying at infinity and*

$$\lim_{x\to\infty} \frac{1}{\ell(x)} \int_a^x \ell(t)t^{-1}\,\mathrm{d}t = \infty\ . \tag{1.10}$$

Example 1.16. The functions log and $1/\log$ are slowly varying at infinity and

$$\int_1^x \frac{\log(t)}{t}\,dt = \frac{1}{2}\log^2(x) \gg \log x\ ,$$
$$\int_{\mathrm{e}}^x \frac{1}{t\log(t)}\,dt = \log\log(x) \gg 1/\log(x)\ .$$

Proof of Theorem 1.15. Theorem 1.6 implies that there exist functions c and η such that $\ell(x) = c(x)\exp\{\int_a^x s^{-1}\eta(s)\,\mathrm{d}s\}$, with $\lim_{x\to\infty} c(x) = c^* \in (0,\infty)$ and $\lim_{x\to\infty}\eta(x) = 0$. Let $\epsilon > 0$ and b be such that $|\eta(s)| \leq \epsilon$ and $1-\epsilon \leq c(x)/c^* \leq 1+\epsilon$ for $s \geq b$. Then, for $t \geq x \geq b$, it holds that

$$\frac{\ell(t)}{\ell(x)} = \frac{c(t)}{c(x)}\exp\left\{\int_x^t \frac{\eta(s)}{s}\,\mathrm{d}s\right\} \leq \frac{1+\epsilon}{1-\epsilon}\exp\left\{\epsilon\int_x^t \frac{\mathrm{d}s}{s}\right\} = \frac{1+\epsilon}{1-\epsilon}\left(\frac{t}{x}\right)^{\epsilon}\ .$$

A lower bound is obtained similarly: for all $t \geq x \geq b$,

$$\frac{\ell(t)}{\ell(x)} \geq \frac{1-\epsilon}{1+\epsilon}\left(\frac{t}{x}\right)^{-\epsilon}\ .$$

Next, we write

$$\frac{1}{x^\beta\ell(x)}\int_a^x t^{\beta-1}\ell(t)\,\mathrm{d}t = \frac{1}{x^\beta\ell(x)}\int_a^b t^{\beta-1}\ell(t)\,\mathrm{d}t + \int_b^x \frac{t^\beta\ell(t)}{x^\beta\ell(x)}\,\frac{\mathrm{d}t}{t}\ .$$

By Proposition 1.11, $\lim_{x\to\infty} x^\beta\ell(x) = \infty$ since $\beta > 0$ and thus the first term in the right-hand side converges to 0. For ϵ small enough, the second term is bounded above and below as follows:

$$\int_b^x \frac{t^\beta\ell(t)}{x^\beta\ell(x)}\,\frac{\mathrm{d}t}{t} \leq \frac{1+\epsilon}{1-\epsilon}\int_b^x \frac{t^{\beta+\epsilon-1}}{x^{\beta-\epsilon}}\,\mathrm{d}t \sim \frac{1}{\beta+\epsilon}\frac{1+\epsilon}{1-\epsilon}\ ,$$
$$\int_b^x \frac{t^\beta\ell(t)}{x^\beta\ell(x)}\,\frac{\mathrm{d}t}{t} \geq \frac{1-\epsilon}{1+\epsilon}\int_b^x \frac{t^{\beta-\epsilon-1}}{x^{\beta-\epsilon}}\,\mathrm{d}t \sim \frac{1}{\beta-\epsilon}\frac{1-\epsilon}{1+\epsilon}\ .$$

Finally, for $\epsilon > 0$ small enough, we obtain

$$\liminf_{x\to\infty} \frac{1}{x^\beta \ell(x)} \int_a^x t^{\beta-1}\ell(t)\,\mathrm{d}t \geq \frac{1}{\beta} - \epsilon \;,$$
$$\limsup_{x\to\infty} \frac{1}{x^\beta \ell(x)} \int_a^x t^{\beta-1}\ell(t)\,\mathrm{d}t \leq \frac{1}{\beta} + \epsilon \;.$$

This proves the first part of (1.9). The other limit is proved similarly. Let us now prove (1.10). Let $\epsilon > 0$. For x such that $a/x \leq \epsilon$, the change of variables $t = sx$ yields

$$\frac{1}{\ell(x)} \int_a^x \frac{\ell(t)}{t}\,\mathrm{d}t \geq \int_\epsilon^1 \frac{\ell(sx)}{\ell(x)} \frac{\mathrm{d}s}{s} \to \log(1/\epsilon)$$

by uniform convergence. This proves (1.10). Denote $L(t) = \int_a^x \ell(t) t^{-1}\mathrm{d}t$. We now prove that L is slowly varying at infinity. Fix some $t > 0$. Then,

$$\frac{L(tx)}{L(x)} = \frac{\int_a^{tx} \ell(s) s^{-1}\,\mathrm{d}s}{\int_a^x \ell(s) s^{-1}\,\mathrm{d}s} = 1 + \frac{\int_1^t \ell(sx) s^{-1}\,\mathrm{d}s}{L(x)} \;.$$

By uniform convergence, we have $\int_1^t \ell(sx) s^{-1}\,\mathrm{d}s \sim \ell(x)\log(t)$, thus

$$\frac{L(tx)}{L(x)} \sim 1 + \log(t) \frac{\ell(x)}{L(x)} \to 1 \;.$$

□

The class of slowly varying function m which can be expressed as $\int_{x_0}^x t^{-1}\ell(t)\,\mathrm{d}t$ where ℓ is slowly varying is called the class Π. This class is very important in extreme value theory since the distribution functions that are in the domain of attraction of the Gumbel law are inverse of increasing functions of the class Π.

If the function f has bounded variation and is regularly varying at infinity, then the former result can be extended to the Stieltjes integral.

Theorem 1.17. *If $f \in RV_\infty(\alpha)$ and f has locally bounded variation on $[x_0, \infty)$, then for any β such that $\alpha + \beta > 0$, it holds that*

$$\int_{x_0}^x t^\beta \,\mathrm{d}f(t) \sim \frac{\alpha}{\alpha+\beta} x f(x) \;.$$

Proof. By integration by parts and Theorem 1.15, we have

$$\int_{x_0}^x t^\beta \, \mathrm{d}f(t) = x^\delta f(x) - x_0^\delta f(x_0) - \delta \int_{x_0}^x t^{\delta-1} f(t) \, \mathrm{d}t$$
$$\sim x^\delta f(x) - \frac{\delta}{\alpha+\delta} x^\delta f(x) = \frac{\alpha}{\alpha+\delta} x^\delta f(x) \; .$$

□

Exercise 1.1.4. Let ψ be a locally integrable function and assume that both limits

$$M(\psi) = \lim_{x\to\infty} \eta x^{-\eta} \int_1^x s^{\eta-1} \psi(s) \, \mathrm{d}s \; , \; m(\psi) \lim_{x\to\infty} \eta x^{-\eta} \int_1^x s^{\eta-1} \psi(1/s) \, \mathrm{d}s$$

exist. Then the (improper) integral

$$\int_0^\infty \frac{\psi(xt) - \psi(t)}{t} \, \mathrm{d}t$$

is convergent and equal to $(M - m)\log(x)$.

Karamata's Theorem 1.15 and Potter's bounds can be combined in many ways to yield convergence of integrals over non compact intervals. We give one example.

Proposition 1.18. *Let g be a function defined on $[0, \infty)$, locally bounded and regularly varying at infinity with index α. Let h be a measurable function on $[0, \infty)$ such that*

$$|h(t)| \le c(t \wedge 1)^\beta + C(t \vee 1)^\gamma \tag{1.11}$$

with $\alpha + \beta > -1$ and $\gamma + \alpha < -1$. If $\beta \le -1$ (which is possible only if $\alpha > 0$), assume moreover that $|g(x)| \le cx^\delta$ in a neighborhood of zero, for some δ such that $\delta + \beta > -1$. Then

$$\lim_{x\to\infty} \int_0^\infty \frac{g(tx)}{g(x)} h(t) \, \mathrm{d}t = \int_0^\infty t^\alpha h(t) \, \mathrm{d}t \; . \tag{1.12}$$

Proof. The idea is to use the uniform convergence on some compact set $[1/A, A]$ and Potter's bounds outside $[1/A, A]$. Fix some $\epsilon > 0$ such that $1 + \alpha + \beta - \epsilon > 0$ and $\alpha + \gamma + \epsilon < -1$. By Theorem 1.13, there exists X

such that $g(tx)/g(x) \leq (1+\epsilon)t^{\alpha+\epsilon}$ if $x \geq X$ and $tx \geq X$. Decompose the integral over $[0,\infty)$ into four subintervals $[0, X/x]$, $[X/x, 1/A]$, $[1/A, A]$, $[A,\infty)$. By uniform convergence and (1.11), we have

$$\lim_{A\to\infty}\lim_{x\to\infty}\int_{1/A}^{A}\frac{g(tx)}{g(x)}h(t)\,\mathrm{d}t = \lim_{A\to\infty}\int_{1/A}^{A} t^{\alpha}h(t)\,\mathrm{d}t = \int_{0}^{\infty} t^{\alpha}h(t)\,\mathrm{d}t\ . \tag{1.13}$$

By Theorem 1.13 and (1.11), we obtain the bounds

$$\begin{aligned}\int_{X/x}^{\eta}\frac{|g(tx)|}{|g(x)|}|h(t)|\,\mathrm{d}t + \int_{A}^{\infty}\frac{|g(tx)|}{|g(x)|}|h(t)|\,\mathrm{d}t \\ \leq C\int_{0}^{1/A} t^{\alpha-\epsilon+\beta}\,\mathrm{d}t + C\int_{A}^{\infty} t^{\alpha+\epsilon+\gamma}\,\mathrm{d}t \\ \leq C(A^{-\alpha-\beta+\epsilon-1} + A^{1+\alpha+\gamma+\epsilon}) \to 0\end{aligned}$$

as $A \to \infty$. Thus $\lim_{A\to\infty}\limsup_{x\to\infty}\int_A^\infty g(tx)/g(x)|h(t)|\,\mathrm{d}t = 0$. To bound the remaining term, let δ be as in the assumption of the theorem if $\beta \leq -1$ or $\delta = 0$ if $\beta > -1$. Since g is locally bounded and for $x \geq X$, $|g(x)| \geq cx^{\alpha-\epsilon}$, we bound the remaining term by a constant times

$$x^{-\alpha+\delta+\epsilon}\int_{0}^{X/x} t^{\delta+\beta}\,\mathrm{d}t = \frac{1}{1+\beta+\delta}X^{1+\beta+\delta}x^{-\alpha-\beta-1+\epsilon} \to 0$$

as $x \to \infty$. Finally, we have obtained that

$$\lim_{x\to\infty}\int_{t\notin[1/A,A]}\frac{|g(tx)|}{|g(x)|}|h(t)|\,\mathrm{d}t = 0\ ,$$

which, together with (1.13), concludes the proof of (1.12). □

Theorem 1.15 has a converse.

Theorem 1.19. *Let f be positive and locally integrable on some interval $[x_0,\infty)$.*

(i) If for some σ, ρ such that $\sigma+\rho > -1$,

$$\lim_{x\to\infty}\frac{x^{\sigma}+1}{\int_{x_0}^{x} t^{\sigma}f(t)\,\mathrm{d}t} = \sigma+\rho+1\ ,$$

then f is regularly varying at infinity with index ρ.

(ii) If for some σ, ρ such that $\sigma + \rho < -1$,

$$\lim_{x\to\infty} \frac{x^{\sigma+1}}{\int_x^\infty t^\sigma f(t)\,\mathrm{d}t} = -(\sigma+\rho+1)\ ,$$

then f is regularly varying at infinity with index ρ.

Proof. We only prove case (i). Case (ii) follows by similar arguments. Define $g(x) = x^{\sigma+1} f(x) / \int_{x_0}^x t^\sigma f(t)\,\mathrm{d}t$ and fix some $x_1 > x_0$. Then, for $x > x_1$,

$$\int_{x_1}^x \frac{g(t)}{t}\,\mathrm{d}t = \log\left\{C^{-1}\int_{x_0}^x t^\sigma f(t)\,\mathrm{d}t\right\}\ ,$$

where $C = \int_{x_0}^{x_1} t^\sigma f(t)\,\mathrm{d}t$. Thus

$$\begin{aligned} f(x) &= x^{-\sigma-1} g(x) \int_{x_0}^x t^\sigma f(t)\,\mathrm{d}t = C x^{-\sigma-1} g(x) \exp \int_{x_1}^x \frac{g(t)}{t}\,\mathrm{d}t \\ &= C x_1^{-\sigma-1} g(x) \exp \int_{x_1}^x \frac{g(t)-\sigma-1}{t}\,\mathrm{d}t\ . \end{aligned}$$

Since $\lim_{x\to\infty} g(x) - \sigma - 1 = \rho$, we conclude by applying the representation theorem that f is regularly varying with index ρ. □

1.1.4 Densities

The following result is very important and will be used frequently. It will be referred to as the monotone density theorem.

Theorem 1.20. *Let $F \in RV_\infty(\alpha)$ and f be locally integrable on $[1,\infty)$ such that*

$$F(x) = \int_1^x f(t)\,\mathrm{d}t\ .$$

If f is ultimately monotone, then

$$\lim_{x\to\infty} \frac{x f(x)}{F(x)} = \alpha\ . \tag{1.14}$$

If moreover $\alpha \neq 0$, then $f \in RV_\infty(\alpha - 1)$.

Proof. Since $F \in RV_\infty(\alpha)$, (1.14) with $\alpha \neq 0$ implies that $f \in RV_\infty(\alpha - 1)$. So we only have to prove (1.14). Assume first that f is ultimately non decreasing. Then, for large enough $a < b$, we have

$$\frac{xf(ax)}{F(x)} \leq \frac{F(bx) - F(ax)}{(b-a)F(x)} \leq \frac{xf(bx)}{F(x)} \; .$$

Thus, for $a = 1$ and $b = 1 + \epsilon$, we have

$$\limsup_{x\to\infty} \frac{xf(x)}{F(x)} \leq \frac{(1+\epsilon)^\alpha - 1}{\epsilon} \leq \alpha(1+\epsilon)^\alpha \; .$$

Thus $\limsup_{x\to\infty} xf(x)/F(x) \leq \alpha$. The lower bound is obtained similarly. □

Lemma 1.21. *Let f be a monotone function on $[a, \infty)$. For $r \in \mathbb{R}$, denote $F_r(x) = \int_a^x t^r f(t)\,\mathrm{d}t$. If F_r is regularly varying at infinity with index $\beta > 0$, then f is regularly varying at infinity with index $\beta - r - 1$.*

Proof. We give a proof when f is non increasing and $r \geq 0$. Let $a > 0$ and $b > 1$. For x large, we have

$$\frac{\int_{ax}^{abx} t^r f(t)\,\mathrm{d}t}{\int_{x/b}^{x} t^r f(t)\,\mathrm{d}t} \leq \frac{(abx - ax)(abx)^r f(ax)}{(x - x/b)(x/b)^r f(x)} = \frac{a^{r+1} b^{2r+1} f(ax)}{f(x)} \; .$$

The first term above converges to $(ab)^\beta$ as $x \to \infty$, thus

$$\liminf_{x\to\infty} \frac{f(ax)}{f(x)} \geq a^{\beta - r - 1} b^{\beta - 2r - 1} \; .$$

Letting b tend to 1 yields

$$\liminf_{x\to\infty} \frac{f(ax)}{f(x)} \geq a^{\beta - r - 1} \; .$$

An upper bound for the limsup is obtained similarly. □

1.1.5 Inversion

Theorem 1.22. *Let $\alpha > 0$. Let $f \in RV_\infty(\alpha)$ be ultimately monotone. Then there exists a function $g \in RV_\infty(1/\alpha)$ such that*

$$f \circ g(x) \sim g \circ f(x) \sim x \; . \qquad (1.15)$$

Moreover g can be chosen as the left-continuous or the right-continuous inverse of f.

Proof. By Theorem 1.6 and Corollary 1.9, the function f is equivalent at infinity to a continuously differentiable function h such that $\lim_{x\to\infty} xh'(x)/h(x) = \alpha$. Thus h' is ultimatley positive and h is invertible on some interval $[x_0,\infty)$. Let g be its inverse and $y_0 = h(x_0)$. Then for all $x \geq x_0$, $g \circ h(x) = x$ and for all $y \geq y_0$, $h \circ g(y) = y$. Moreover, g is regularly varying with index $1/\alpha$, since

$$\frac{yg'(y)}{g(y)} = \frac{h(g(y))}{g(y)h'(g(y))} \to \frac{1}{\alpha} \; .$$

Since $f \sim g$, we have $f \circ g(y) \sim y$. Conversely, since $f \sim h$ and g is regularly varying, $g \circ f(x) = g(h(x)f(x)/h(x)) \sim g(h(x))$ by uniform convergence. Thus $g \circ f(x) \sim x$. To prove that $f^{\leftarrow}$ satisfy (1.15), note that by Lemma 1.14, f is equivalent to its right-continous inverse, so we can assume without loss of generality that f is right continuous. Then its left-continuous inverse is defined by $f^{\leftarrow}(y) = \inf\{x, f(x) \geq y\}$ (see Section A.2), and thus

$$y \leq f \circ f^{\leftarrow}(y) \leq y + \Delta f(f^{\leftarrow}(y)) \; ,$$

where $\Delta f(x)$ denotes the jump of f at x (possibly zero). This yields

$$1 - \frac{\Delta \circ f^{\leftarrow}(y)}{f \circ f^{\leftarrow}(y)} \leq \frac{y}{f \circ f^{\leftarrow}(y)} \leq 1 \; .$$

By Lemma 1.14, the left-had side tends to one, so $f \circ f^{\leftarrow}(y) \sim y$, and thus $f^{\leftarrow} \sim g$. □

Corollary 1.23 (De Bruyn conjugate)**.** *Let $\ell \in SV_\infty$. There exists a slowly varying function $\ell^\sharp$ such that* $\lim_{x\to\infty} \ell(x)\ell^\sharp(x\ell(x)) = 1$.

Example 1.24. Let a, b be real numbers such that $ab > 0$. Let $f(x) = x^{ab}\ell^a(x^b)$. Then if f is ultimately monotone and g is its left or right continuous inverse, then

$$g(x) \sim x^{\frac{1}{ab}}(\ell^\sharp)^{\frac{1}{b}}(x^{\frac{1}{a}}) \; .$$

Example 1.25. If f is monotone and regularly varying with non zero index, then for any two sequences a_n and b_n, $a_n \sim f(b_n)$ if and only if $f^{\leftarrow}(a_n) \sim b_n$.

This is not true if f is not regularly varying. If $f(x) = \log(x)$, then it is possible to have $a_n \sim \log(b_n)$ while e^{a_n} is not equivalent to b_n. Take for instance $a_n = n$ and $b_n = \log(n) + \sqrt{\log(n)}$.

1.2 Tauberian Theorems

Theorems that relate the behaviour of a function and its transform are called Abelian or Tauberian. Tauberian results are usually harder to prove and need some additional assumption. We give in this section results for the Laplace and Fourier transforms. The results for Fourier transforms will be used in the context of stable distributions and long memory processes.

1.2.1 Laplace-Stieltjes transform

For any function f with bounded variation on $\mathbb{R}_+$, denote

$$\mathcal{L}f(s) = \int_0^\infty \mathrm{e}^{-s} \mathrm{d}f(s)$$

the Laplace-Stietjes transform of f.

Theorem 1.26. *Let f be a right continuous non decreasing function on $\mathbb{R}_+$ such that $\mathcal{L}F(s) < \infty$ for $s > 0$. For $\alpha \geq 0$, $f \in RV_\infty(\alpha)$ if and only if $\mathcal{L}f \in RV_0(-\alpha)$ and then*

$$\lim_{x\to\infty} \frac{\mathcal{L}f(1/x)}{f(x)} = \Gamma(1+\alpha) \; . \tag{1.16}$$

Proof. By integration by parts, we have for all $s, x > 0$,

$$\frac{\mathcal{L}f(s/x)}{f(x)} = \int_0^\infty s\mathrm{e}^{-st} \frac{f(tx)}{f(x)} \, \mathrm{d}t \; .$$

If $f \in RV_\infty(\alpha)$, as $x \to \infty$, the integrand converges to $s\mathrm{e}^{-st}t^\alpha$, and

$$s \int_0^\infty \mathrm{e}^{-st} t^\alpha \, \mathrm{d}t = s^{-\alpha}\Gamma(1+\alpha) \; ,$$

so we only need a bounded convergence argument to prove (1.16). Since f is non decreasing, if $t \leq 1$, then $f(tx)/f(x) \leq 1$. If $t > 1$, by Theorem 1.13, for large enough x, $f(tx)/f(x) \leq 2t^{\alpha+1}$. Thus, for all $t \geq 0$, and large enough x, we have

$$\frac{f(tx)}{f(x)} \leq 2(t \vee 1)^{\alpha+1} \; .$$

Since $\int_0^\infty e^{-st}(t \vee 1)^{\alpha+1}\,dt < \infty$ for all $s > 0$, we can apply the bounded convergence Theorem and (1.16) holds, which implies $\mathcal{L}f \in RV_0(-\alpha)$.

Conversely, if $\mathcal{L}f \in RV_0(-\alpha)$, then

$$\lim_{x\to\infty} \frac{\mathcal{L}f(s/x)}{\mathcal{L}f(1/x)} = s^{-\alpha} = \mathcal{L}g(s)$$

with $g(x) = x^\alpha/\Gamma(1+\alpha)$. On the other hand, the function $s \to \mathcal{L}f(s/x)/\mathcal{L}f(1/x)$ is the Laplace transform of the non decreasing function g_x defined by

$$g_x(y) = \frac{f(xy)}{\mathcal{L}f(1/x)} \ .$$

Thus, $\mathcal{L}g_x$ converges to $\mathcal{L}g$ as $x \to \infty$ and thus g_x converges pointwise to g by the continuity theorem for Laplace transform (See Theorem A.5). □

We derive from Theorem 1.26 a corollary which concerns probability distribution functions and will be of use in Chapter 2.

Corollary 1.27. *Let $\alpha \in (0,1)$ and let F be an increasing function on $[0,\infty)$ such that $F(0) = 0$, $\lim_{x\to\infty} F(x) = 1$. Then $1 - F \in RV_\infty(-\alpha)$ if and only if $1 - \mathcal{L}F \in RV_0(\alpha)$.*

Proof. Denote $\bar{F} = 1 - F$ and $U(x) = \int_0^x \bar{F}(t)\,dt$. By integration by parts, we obtain that $1 - \mathcal{L}F(t) = t\mathcal{L}U(t)$. Thus, by Theorem 1.26, $1 - \mathcal{L}F \in RV_0(\alpha)$ if and only if $U \in RV_\infty(1-\alpha)$. By the monotone density theorem, this is equivalent to $1 - F \in RV_\infty(-\alpha)$. □

1.2.2 Power series

Theorem 1.28. *Let $\{q_n\}$ be a sequence of non negative real numbers and assume that the series*

$$Q(z) = \sum_{j=0}^\infty q_j z^j$$

converges for all $z \in [0,1)$. Let $\alpha \ge 0$ and let the function ℓ be slowly varying at infinity. Then the following relations are equivalent

$$Q(z) \sim (1-z)^{-\alpha}\ell(1/(1-z)) \ , \quad (z \to 1^-) \ , \tag{1.17}$$

$$q_0 + \cdots + q_n \sim \frac{1}{\Gamma(1+\alpha)} n^\alpha \ell(n) \ . \tag{1.18}$$

If moreover the sequence q_j is ultimately monotone and $\alpha > 0$, then both relations are equivalent to

$$q_n \sim \frac{1}{\Gamma(\alpha)} n^{\alpha-1} \ell(n) . \tag{1.19}$$

Proof. Define $U(x) = \sum_{0 \le j \le x} q_j$. Then

$$Q(\mathrm{e}^{-t}) = \int_0^\infty \mathrm{e}^{-tx} \, \mathrm{d}U(x) .$$

Thus, by Theorem 1.28, (1.17) is equivalent to $U(x) \sim x^\alpha \ell(x)/\Gamma(1+\alpha)$. By uniform convergence, $U(n) \sim U(x)$ for all $x \in [n, n+1]$, thus (1.18) is equivalent to (1.17). If q_j is non decreasing, define $u(x) = q_{[x]}$. Then u is non decreasing, $u(x) \sim u(n)$ for all $x \in [n, n+1)$, and

$$U(x) \sim \int_0^x u(s) \, \mathrm{d}s .$$

Thus (1.19) follows by the monotone density theorem. □

1.2.3 Fourier series

In this section, we want to obtain equivalents at zero for trigonometric series whose coefficients are regularly varying:

$$\sum_{j=1}^\infty c_j \mathrm{e}^{\mathrm{i}jx}$$

with $c_j = j^{-\alpha} \ell(j)$, $\alpha > 0$ and and ℓ is slowly varying. The main idea is to consider first the case $0 < \alpha < 1$ and $\ell \equiv 1$ and then apply summation by parts. in order to do this, we need a restriction on the class of admissible slowly varying functions. This is the reason for introducing quasi-monotone functions in Definition 1.29. The main results in this section are Theorem 1.34 and Corollary 1.36.

Definition 1.29 (Quasi- and near-monotonicity)**.** *A function f is said to be quasi-monotone if $f \in BV_{\mathrm{loc}}([0,\infty))$ and for all $\delta > 0$,*

$$\int_0^x t^\delta |\mathrm{d}f(t)| = O(x^\delta f(x)) . \tag{1.20}$$

A function is said to be near-monotone if for all $\delta > 0$,

$$\int_0^x t^\delta |\mathrm{d}f(t)| = o(x^\delta f(x)) \; . \tag{1.21}$$

A monotone function is quasi-monotone. A monotone slowly varying function is near-monotone. In particular, a normalized slowly varying function is quasi-monotone.

Example 1.30. Let $\alpha > 0$ and $f \in RV_\infty(-\alpha)$ be non increasing. Denote $\ell(x) = x^\alpha f(x)$. Then ℓ is slowly varying and near-monotone. Indeed, $\mathrm{d}\ell(t) = \alpha x^{\alpha-1}\ell(t)\,\mathrm{d}t + x^\alpha\,\mathrm{d}f(t)$ and for all $\delta > 0$,

$$\int_0^x t^\delta |\mathrm{d}\ell(t)| \leq \alpha \int_0^x t^{\alpha+\delta-1}\ell(t)\,\mathrm{d}t - \int_0^x t^{\alpha+\delta}\,\mathrm{d}f(t) \; .$$

By Theorems 1.15 and 1.17, both terms on the right-hand side of the last displayed equation are equivalent to $\alpha/(\alpha+\delta)\,x^\delta f(x)$, thus (1.21) holds.

Lemma 1.31. *Let ℓ be quasi-monotone and $\alpha > 0$. Then*

$$\sum_{j=n+1}^{\infty} j^{-\alpha}|\ell(j) - \ell(j+1)| = O(n^{-\alpha}\ell(n)) \; .$$

Proof. Since ℓ has bounded variation and is quasi-monotone, we have

$$\sum_{j=n+1}^{\infty} j^{-\alpha}|\ell(j) - \ell(j+1)| \leq \int_n^\infty t^{-\alpha}|\mathrm{d}\ell(t)| \; ,$$

and we now prove that this last integral is $O(n^{-\alpha}\ell(n))$. Fix some $\eta > 0$ and define $u(t) = \int_0^t s^\eta\,|\mathrm{d}\ell(s)|$. By assumption, $u(t) = O(t^\eta\ell(t))$. Integration by parts yields

$$\begin{aligned}\int_x^\infty t^{-\alpha}|\mathrm{d}\ell(t)| &= \int_x^\infty t^{-\alpha-\eta}\mathrm{d}u(t) \\ &= x^{-\delta-\eta}u(x) + (\delta+\eta)\int_x^\infty t^{-\delta-\eta-1}u(t)\,\mathrm{d}t \\ &\leq Cx^{-\delta}\ell(x) + C\int_x^\infty t^{-\delta-1}\ell(t)\,\mathrm{d}t \leq C'x^{-\delta}\ell(x) \; ,\end{aligned}$$

where the integral was bounded by applying Theorem 1.15. □

Lemma 1.32. *For $\alpha \in (0,1]$,*

$$\sum_{j=1}^{\infty} j^{-\alpha} \sin(jx) \sim \frac{\pi}{2\Gamma(\alpha)\sin(\pi\alpha/2)}\, x^{\alpha-1} \;, \quad x \to 0^+ \;. \tag{1.22}$$

For $\alpha \in (0,1)$,

$$\sum_{j=1}^{\infty} j^{-\alpha} \cos(jx) \sim \frac{\pi}{2\Gamma(\alpha)\cos(\pi\alpha/2)}\, x^{\alpha-1} \;, \quad x \to 0^+ \;. \tag{1.23}$$

Proof. We must first prove that the series is convergent (but not absolutely convergent; this type of convergence is sometimes called semi-convergence or conditional convergence). To that purpose, we need the following well-known bound: for all $x \in (0,\pi]$,

$$\left|\sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}jx}\right| = \frac{|\sin(nx/2)|}{|\sin(x/2)|} \le \frac{1}{|\sin(x/2)|} \le \frac{\pi}{x} \;. \tag{1.24}$$

Denote $s_0(x) = 0$ and $s_n(x) = \sum_{j=1}^{n} \sin(jx)$. Applying summation by parts yields

$$\begin{aligned}
&\sum_{j=n}^{m+n} j^{-\alpha} \sin(jx) \\
&\quad = \sum_{j=n}^{n+m} j^{-\alpha} \{s_j(x) - s_{j-1}(x)\} \\
&\quad = \sum_{j=n}^{n+m} s_j(x)\{j^{-\alpha} - (j+1)^{-\alpha}\} + (n+m+1)^{-\alpha} s_{n+m}(x) \;.
\end{aligned}$$

Applying (1.24) yields

$$\left|\sum_{j=n}^{m+n} j^{-\alpha} \sin(jx)\right| \le 2\pi x^{-1} n^{-\alpha} \;.$$

Thus the series converges, and letting $m \to \infty$ above yields, for any $\alpha \in (0,1]$ and $x \in (0,\pi]$,

$$\left|\sum_{j=n}^{\infty} j^{-\alpha} \sin(jx)\right| \le \pi n^{-\alpha} x^{-1} \;. \tag{1.25}$$

Let c_j denote the sine coefficient of the Fourier expansion of the odd function f defined on $[-\pi,\pi]$ defined by $f(x)=x^{\alpha-1}$ if $x>0$; i.e.

$$c_j = 2\int_0^\pi x^{\alpha-1}\sin(jx)\,\mathrm{d}x = 2j^{-\alpha}\int_0^{j\pi} x^{\alpha-1}\sin(x)\,\mathrm{d}x\ .$$

Recall that for $0<\alpha<1$, it holds that

$$\int_0^\infty \frac{\mathrm{e}^{\mathrm{i}x}}{x^{1-\alpha}}\,\mathrm{d}x = \Gamma(\alpha)\mathrm{e}^{\mathrm{i}\pi\alpha/2}\ . \tag{1.26}$$

Moreover, since the function f is differentiable at all $x\in(0,\pi)$, its Fourier series is convergent and sums up to πf, i.e.

$$\sum_{j=1}^\infty c_j\sin(jx) = \pi x^{\alpha-1}\ .$$

In order to prove (1.22), it is sufficient to show that

$$\sum_{j=1}^\infty \{c_j - 2\Gamma(\alpha)\sin(\pi\alpha/2)j^{-\alpha}\}\sin(jx) = o(x^{\alpha-1})\ . \tag{1.27}$$

Note that

$$\begin{aligned}c_j - 2\Gamma(\alpha)\sin(\pi\alpha/2)j^{-\alpha} &= j^{-\alpha}\int_{j\pi}^\infty \frac{\sin(x)}{x^{1-\alpha}}\,\mathrm{d}x \\ &= (-1)^j j^{-\alpha}\int_0^\infty \frac{\sin(x)}{(j\pi+x)^{1-\alpha}}\,\mathrm{d}x\ .\end{aligned}$$

Denote $r_j = \int_0^\infty (j\pi+x)^{\alpha-1}\sin(x)\,\mathrm{d}x$. Then $\{r_j\}$ is a positive and decreasing sequence. Indeed, denote $a_k = \int_0^\pi ((j+k)\pi+x)^{1-\alpha}\sin(x)\,\mathrm{d}x$. Then

$$r_j - r_{j+1} = \sum_{k=0}^\infty (a_{2k} - 2a_{2k+1} - a_{2k+2}) \geq 0$$

since the function $x\to x^{\alpha-1}$ is convex. Thus the sequence $j^{-\alpha}r_j$ is also decreasing. For $x\in[0,\pi/2]$ $\sin((x+\pi)/2)\geq 1/\sqrt{2}$, thus (1.24) and summation by parts yield

$$\left|\sum_{j=n}^{m+n}(-1)^j j^{-\alpha}r_j\sin(jx)\right| \leq 2\sqrt{2}j^{-\alpha}r_n\ .$$

Thus the series in (1.27) is bounded in a neighborhood of zero and this proves (1.27). □

Remark 1.33. For $\alpha = 1$, it is well known that for $x \in (0, \pi]$,

$$\sum_{j=1}^{\infty} \frac{\mathrm{e}^{\mathrm{i}jx}}{j} = -\log(1 - \mathrm{e}^{\mathrm{i}x}) = -\log(2\sin(x/2)) + \mathrm{i}(\pi - x)/2 \; .$$

Moreover, the partial sums of the sine series are uniformly bounded

$$\sup_{k \geq 1} \sup_{x \in [0,\pi]} \left| \sum_{j=1}^{k} \frac{\sin(jx)}{j} \right| < \infty \; .$$

Theorem 1.34. *Let ℓ be quasi-monotone slowly varying. Then the following equivalences hold as $x \to 0^+$.*

$$\sum_{j=1}^{\infty} j^{-\alpha} \ell(j) \sin(jx) \sim \frac{\pi}{2\Gamma(\alpha)\sin(\pi\alpha/2)} \, x^{\alpha-1} \ell(1/x) \; , \quad 0 < \alpha \leq 1 \; , \tag{1.28}$$

$$\sum_{j=1}^{\infty} j^{-\alpha} \ell(j) \cos(jx) \sim \frac{\pi}{2\Gamma(\alpha)\cos(\pi\alpha/2)} \, x^{\alpha-1} \ell(1/x) \; , \quad 0 < \alpha < 1 \; . \tag{1.29}$$

If the series $\sum_{j=1}^{\infty} j^{-1} \ell(j)$ is divergent, then

$$\sum_{j=1}^{\infty} j^{-1} \ell(j) \cos(jx) \sim \sum_{j \leq 1/x} j^{-1} \ell(j) \sim \int_1^{1/x} t^{-1} \ell(t) \, \mathrm{d}t \; . \tag{1.30}$$

Remark 1.35. For $\alpha < 1$, (1.28) and (1.29) can be summarized as

$$\sum_{j=1}^{\infty} j^{-\alpha} \ell(j) \mathrm{e}^{\mathrm{i}jx} \sim \Gamma(1 - \alpha) \mathrm{e}^{\mathrm{i}\pi(1-\alpha)/2} \, x^{\alpha-1} \ell(1/x) \; . \tag{1.31}$$

Proof. The proof follows straightforwardly from Lemmas 1.31 and 1.32. Denote $S_n(x) = \sum_{j=n}^{\infty} j^{-\alpha} \sin(jx)$. Applying summation by parts, we have

$$\sum_{j=n}^{n+m} j^{-\alpha}\ell(j)\sin(jx) = \sum_{j=n}^{n+m}\{S_j(x) - S_{j+1}(x)\}\ell(j)$$
$$= \sum_{j=n}^{n+m} S_j(x)\{\ell(j) - \ell(j-1)\}$$
$$- S_{n+m+1}(x)\ell(n+m) + S_n(x)\ell(n-1) \,.$$

Applying the bound (1.25) and Lemma 1.31, we obtain

$$\Big|\sum_{j=n}^{n+m} j^{-\alpha}\ell(j)\sin(jx)\Big|$$
$$\leq \frac{C}{x}\left\{\sum_{j=n}^{\infty} j^{-\alpha}|\ell(j) - \ell(j-1)| + (n+m)^{-\alpha}\ell(n+m) + n^{-\alpha}\ell(n-1)\right\}$$
$$= O\left(\frac{n^{-\alpha}\ell(n)}{x}\right) \,.$$

Thus, the series in (1.28) is convergent. We now prove the equivalence by proving that

$$\sum_{j=1}^{\infty} j^{-\alpha}\{\ell(j) - \ell(1/x)\}\sin(jx) = o(x^{\alpha-1}\ell(1/x)) \,. \tag{1.32}$$

Fix $0 < a < b$. By the uniform convergence theorem,

$$\lim_{x\to 0^+}\sup_{a/x<k\leq b/x}\left|\frac{\ell(k) - \ell(1/x)}{\ell(1/x)}\right| \leq \lim_{x\to 0^+}\sup_{a\leq t\leq b}\left|\frac{\ell(t/x) - \ell(1/x)}{\ell(1/x)}\right| = 0 \,.$$

Denote $R_2(x) = \sum_{a/x<k\leq b/x} j^{-\alpha}\{\ell(j) - \ell(1/x)\}\sin(jx)$. We thus have

$$\frac{|R_2(x)|}{x^{\alpha-1}\ell(1/x)} \leq \sup_{a/x<k\leq b/x}\frac{|\ell(k) - \ell(1/x)|}{|\ell(1/x)|}x\sum_{a/x<k\leq b/x} k^{1-\alpha}$$
$$\leq \sup_{a/x<k\leq b/x}\frac{|\ell(k) - \ell(1/x)|}{|\ell(1/x)|}x^{1-\alpha}\int_{a/x}^{b/x} t^{-\alpha}\,\mathrm{d}t$$
$$= \frac{b^{2-\alpha} - a^{2-\alpha}}{2-\alpha}\sup_{a/x<k\leq b/x}\frac{|\ell(k) - \ell(1/x)|}{|\ell(1/x)|} \,.$$

Thus

$$\lim_{x\to 0^+} R_2(x)/(x^{\alpha-1}\ell(1/x)) = 0 \ . \tag{1.33}$$

Denote now $R_1 = \sum_{1\le k\le a/x} j^{-\alpha}\{\ell(j)-\ell(1/x)\}\sin(jx)$ and set $\ell(0)=0$. Applying summation by parts, (1.25) and Lemmas 1.31, we have

$$\begin{aligned}
&|R_1(x)| \\
&\quad = \left|\sum_{j=1}^{[a/x]} S_j(x)\{\ell(j)-\ell(j-1)\} - S_{[a/x]+1}(x)\{\ell([a/x])-\ell(1/x)\}\right| \\
&\quad \le Cx^{-1}\sum_{j=1}^{[a/x]} j^{-\alpha}|\ell(j)-\ell(j-1)| + Ca^{1-\alpha}x^{\alpha-1}|\ell([a/x])-\ell(1/x)| \\
&\quad \le C\sum_{j=1}^{[a/x]} j^{1-\alpha}|\ell(j)-\ell(j-1)| + Ca^{1-\alpha}x^{\alpha-1}|\ell([a/x])-\ell(1/x)| \\
&\quad \le C'[a/x]^{1-\alpha}|\ell([a/x])| + Ca^{1-\alpha}x^{\alpha-1}|\ell([a/x])-\ell(1/x)| \ .
\end{aligned}$$

Thus

$$\limsup_{x\to 0^+} \frac{|R_1(x)|}{x^{\alpha-1}|\ell(1/x)|} \le C'a^{1-\alpha} \ . \tag{1.34}$$

Denote finally $R_3 = \sum_{j>b/x} j^{-\alpha}\{\ell(j)-\ell(1/x)\}\sin(jx)$. Applying summation by parts, (1.25) and Lemma 1.31 yields

$$\begin{aligned}
&|R_3(x)| \\
&\quad = \left|\sum_{j>b/x} S_j(x)\{\ell(j)-\ell(j-1)\} + S_{[b/x]+1}(x)\{\ell([b/x])-\ell(1/x)\}\right| \\
&\quad \le Cx^{-1}\sum_{j>b/x} j^{-\alpha}|\ell(j)-\ell(j-1)| + Cb^{-\alpha}x^{\alpha-1}|\ell([b/x])-\ell(1/x)| \\
&\quad \le C'b^{-\alpha}x^{\alpha-1}\ell(b/x) + Cb^{-\alpha}x^{\alpha-1}|\ell([b/x])-\ell(1/x)| \ .
\end{aligned}$$

Thus,

$$\limsup_{x\to\infty} \frac{R_3(x)}{x^{\alpha-1}\ell(1/x)} \le C'b^{-\alpha} \ . \tag{1.35}$$

Gathering (1.33), (1.34) and (1.35) yields

$$\limsup_{x\to\infty} \frac{\left|\sum_{j=1}^{\infty} j^{-\alpha}\{\ell(j)-\ell(1/x)\}\sin(jx)\right|}{x^{\alpha-1}|\ell(1/x)|} = O(a^{1-\alpha}+b^{-\alpha}) \ .$$

Since a and b were chosen arbitrarily, this proves (1.32). We now prove (1.30). Denote $S_n = \sum_{k=1}^{n} n^{-1}\ell(n)$ and $I(x) = \int_1^x t^{-1}\ell(t)\,\mathrm{d}t$. By uniform convergence, $I(x) \sim I(n) \sim S_n$, uniformly for $x \in [n, n+1)$, and $\ell(x) = o(I(x))$ by Theorem 1.15. By summation by parts and using the quasi-monotonicity of ℓ, we show as previously that the series in (1.30) is convergent and that

$$\sum_{j=n}^{\infty} \frac{\cos(kx)}{k}\ell(x) = O\left(\frac{\ell(n)}{nx}\right) \ . \tag{1.36}$$

Now,

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k}\ell(k) = \sum_{k\le 1/x} \frac{\ell(k)}{k} + \sum_{k\le 1/x} \frac{\ell(k)}{k}(\cos(kx)-1) + \sum_{k>1/x} \frac{\ell(k)}{k}\cos(kx) \ .$$

Applying the bound (1.36) and (1.10), we get

$$\sum_{k>1/x} \frac{\ell(k)}{k}\cos(kx) = o(I(x)) \ .$$

Applying the bound $|\cos(kx)-1| \le \frac{1}{2}k^2x^2$ and Theorem 1.15 (1.9), we have

$$\left|\sum_{k\le 1/x} \frac{\ell(k)}{k}(\cos(kx)-1)\right| \le \frac{1}{2}x^2 \sum_{k\le 1/x} k|\ell(k)| = O(\ell(x)) = o(I(x)) \ .$$

□

Let us now consider the case $1 < \alpha < 2$. The series $\sum_{j=1}^{\infty} j^{-\alpha}\ell(j)\mathrm{e}^{\mathrm{i}jx}$ is summable for any slowly varying function ℓ, and an equivalent at zero is easily obtain by summation by parts and applying Theorem 1.34. No additional assumption on the slowly varying function ℓ is needed, because the series $\sum_{j=1}^{\infty} j^{-\alpha}\ell(j)$ is absolutely summable and, if $\ell \sim \ell^\sharp$ where $\ell^\sharp$ is in the Zygmund class or normalized slowly varying, hence quasi-monotone, then

$$\sum_{j=n}^{\infty} j^{-\alpha}\ell(j) \sim \sum_{j=n}^{\infty} j^{-\alpha}\ell^\sharp(j) \sim \frac{n^{1-\alpha}\ell^\sharp(n)}{\alpha-1} \sim \frac{n^{1-\alpha}\ell(n)}{\alpha-1} \;, \tag{1.37}$$

$$\sum_{j=1}^{n} j^{1-\alpha}\ell(j) \sim \sum_{j=1}^{n} j^{1-\alpha}\ell^\sharp(j) \sim \frac{n^{2-\alpha}\ell^\sharp(n)}{2-\alpha} \sim \frac{n^{2-\alpha}\ell(n)}{2-\alpha} \;. \tag{1.38}$$

Recall that Rieman's ζ function is defined by

$$\zeta(\alpha) = \sum_{j=1}^{\infty} j^{-\alpha} \;.$$

Corollary 1.36. *Let ℓ be slowly varying and $1 < \alpha < 2$. Then*

$$\sum_{j=1}^{\infty} j^{-\alpha}\ell(j)\mathrm{e}^{\mathrm{i}jx} \sim \zeta(\alpha) - \Gamma(1-\alpha)\mathrm{e}^{\mathrm{i}\pi\alpha/2}x^{\alpha-1}\ell(1/x) \;. \tag{1.39}$$

Proof. Denote $S(x) = \sum_{j=1}^{\infty} j^{-\alpha}\ell(j)\mathrm{e}^{\mathrm{i}jx}$ and $T(x) = \sum_{j=1}^{\infty} j^{-\alpha}\mathrm{e}^{\mathrm{i}jx}$. We show by a method similar to that used to prove (1.32) that $S(x) - \zeta(\alpha) \sim \ell(1/x)\{T(x) - \zeta(\alpha)\}$ and by summation by parts and (1.22) that

$$T(x) \sim \zeta(\alpha) - \Gamma(1-\alpha)\mathrm{e}^{\mathrm{i}\pi\alpha/2}x^{\alpha-1} \;. \tag{1.40}$$

As peviously, let $0 < a < b$ and denote

$$\begin{aligned} S(x) &- \zeta(1/\alpha) - \ell(1/x)\{T(x) - \zeta(1/\alpha\} \\ &= \sum_{j=1}^{\infty} j^{-\alpha}\{\ell(j) - \ell(1/x)\}\{\mathrm{e}^{\mathrm{i}jx} - 1\} = R_1(x) + R_2(x) + R_3(x) \;, \end{aligned}$$

where R_1, R_2 and R_3 are the sums for the indices $j \le a/x$, $a/x < j \le b/x$ and $b/x \le j$, respectively. By the uniform convergence theorem, we obtain

$$\lim_{x\to 0^+} \frac{R_2(x)}{x^{\alpha-1}\ell(1/x)} = 0 \;.$$

To bound R_3, we use (1.37), wich is obviously true with $|\ell|$ instead of ℓ. We get, for some constant C,

$$
\begin{aligned}
|R_3(x)| &\le \sum_{j \ge b/x} j^{-\alpha} \{|\ell(j)| + |\ell(1/x)|\} \\
&\le C(b/x)^{1-\alpha} \{|\ell(b/x)| + |\ell(1/x)|\} \; .
\end{aligned}
$$

Thus

$$
\limsup_{x \to 0^+} \frac{|R_3(x)|}{x^{\alpha-1} |\ell(1/x)|} \le 2Cb^{1-\alpha} \; .
$$

Finally, applying (1.38) and the trivial bound $|\mathrm{e}^{\mathrm{i}jx} - 1| \le jx$, we obtain, for some constant C,

$$
\begin{aligned}
|R_1(x)| &\le x \sum_{1 \le j < a/x} j^{1-\alpha} \{|\ell(j)| + |\ell(1/x)|\} \\
&\le Ca^{2-\alpha} x^{\alpha-1} \{|\ell(a/x)| + |\ell(1/x)|\} \; .
\end{aligned}
$$

Thus

$$
\limsup_{x \to 0^+} \frac{|R_1(x)|}{x^{\alpha-1} |\ell(1/x)|} \le 2Ca^{2-\alpha} \; .
$$

Altogether, we obtain

$$
\limsup_{x \to 0^+} \frac{|S(x) - \zeta(\alpha) - \{T(x) - \zeta(\alpha)\} \ell(1/x)|}{x^{\alpha-1} |\ell(1/x)|} \le C'(a^{2-\alpha} + b^{1-\alpha}) \; ,
$$

and since a and b are arbitrary, we obtain

$$
\lim_{x \to 0^+} \frac{|S(x) - \zeta(\alpha) - \{T(x) - \zeta(\alpha)\} \ell(1/x)|}{x^{\alpha-1} |\ell(1/x)|} = 0 \; .
$$

We must now prove (1.40). Denote $u_j = \sum_{k=j}^{\infty} k^{-\alpha}$. Then $u_j \sim j^{1-\alpha}/(\alpha - 1)$, and applying summation by parts and (1.31), we obtain

$$
\begin{aligned}
T(x) - \zeta(\alpha) &= \sum_{j=1}^{\infty} u_j \mathrm{e}^{\mathrm{i}jx} \{1 - \mathrm{e}^{-\mathrm{i}x}\} \\
&\sim \mathrm{i}x \, \frac{\Gamma(2-\alpha)}{\alpha - 1} \, \mathrm{e}^{\mathrm{i}\pi(\alpha-1)/2} \, x^{\alpha-2} = -\Gamma(1-\alpha) \mathrm{e}^{\mathrm{i}\pi\alpha/2} x^{\alpha-1} \; .
\end{aligned}
$$

This concludes the proof of (1.39). □

Theorem 1.34 has some partial converses.

Theorem 1.37. *Let $0 < \alpha < 1$. Let ℓ be slowly varying at infinity and quasi-monotone. Let f be an integrable function on $[0,\pi]$ such that*

$$f(x) = \ell(1/x)\, x^{\alpha-1} \;, \; (x \to 0^+) \;.$$

Then

$$\int_0^\pi f(x) \mathrm{e}^{\mathrm{i}kx} \,\mathrm{d}x \sim \Gamma(\alpha) \mathrm{e}^{\mathrm{i}\pi\alpha/2}\, \ell(n) n^{-\alpha} \;.$$

The previous theorem has assumptions on f. Instead, we can make additional assumptions on the Fourier coefficients.

Theorem 1.38. *Let $0 < \alpha < 1$. Let f be an integrable function on $[0,\pi]$. Define $c_k = \int_0^\pi f(x)\cos(kx)\,\mathrm{d}x$ and assume that $c_k \geq 0$ for k large. Then*

$$f(x) \sim \frac{1}{\Gamma(\alpha)\cos(\pi\alpha/2)} x^{\alpha-1}\ell(1/x) \;, \; (x \to 0^+)$$

if and only if

$$c_1 + \cdots + c_n \sim \frac{1}{1-\alpha}\ell(n) n^{1-\alpha} \;. \tag{1.41}$$

If moreover the sequence c_k is non increasing, then

$$c_n \sim \ell(n) n^{-\alpha} \;.$$

Sketch of proof. For $r \in (0,1)$

$$1 + 2\sum_{k=1}^\infty r^k \cos(kx) = \frac{1-r^2}{1-2r\cos x + r^2} \;.$$

This series is uniformly convergent for $x \in [0,\pi]$, hence by Fubini Theorem,

$$1 + 2\sum_{k=1}^\infty r^k c_k = \int_0^\pi \frac{1-r^2}{1-2r\cos x + r^2} f(x) \,\mathrm{d}x \;.$$

The method of proof is to obtain an equivalent for the integral on the right-hand side, and then to apply Theorem 1.28 to obtain (1.41). □

1.2.4 Fourier transform

By similar techniques, results for Fourier transforms are obtained.

Theorem 1.39. *Let ℓ be quasi-monotone slowly varying. Then, for $0 < \alpha < 1$,*

$$\int_0^\infty \frac{\mathrm{e}^{\mathrm{i}tx}}{x^\alpha} \ell(x)\,\mathrm{d}x \sim \Gamma(1-\alpha)\,\mathrm{e}^{\mathrm{i}\pi(1-\alpha)/2}\,t^{\alpha-1}\ell(1/t)\;, \tag{1.42}$$

$$\int_0^\infty \frac{\mathrm{e}^{\mathrm{i}tx} - \mathbb{1}_{\{tx\le 1\}}}{x} \ell(x)\,\mathrm{d}x \sim \ell(1/t) \int_0^\infty \frac{\mathrm{e}^{\mathrm{i}x} - \mathbb{1}_{\{x\le 1\}}}{x}\,\mathrm{d}x\;. \tag{1.43}$$

The result for $1 < \alpha < 2$ is obtained by integration by parts, and there is no need to assume that ℓ is quasi-monotone.

Corollary 1.40. *Let ℓ be slowly varying and $1 < \alpha < 2$. Then*

$$\int_0^\infty \frac{\mathrm{e}^{\mathrm{i}tx} - 1}{x^\alpha} \ell(x)\,\mathrm{d}x \sim \Gamma(1-\alpha)\,\mathrm{e}^{\mathrm{i}\pi(1-\alpha)/2}\,t^{\alpha-1}\ell(1/t)\;. \tag{1.44}$$

A partial converse of Theorem 1.39 can be obtained under a monotonicity assumption.

Theorem 1.41. *Let $\alpha \in (0,1)$. Let f be bounded and non increasing on $[0,\infty)$ with $\lim_{x\to\infty} f(x) = 0$. Denote $\hat{f}(t) = \int_0^\infty f(x)\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}x$. If either $\mathrm{Re}(\hat{f})$ or $\mathrm{Im}(\hat{f})$ is regularly varying at zero with index $\alpha - 1$, then $f \in RV_\infty(-\alpha)$ and (1.42) holds (with $\ell(x) = x^\alpha f(x)$).*

Proof. Since f is non increasing and $\sup_{t\ge 0} |\int_0^t \mathrm{e}^{\mathrm{i}x}\,\mathrm{d}x| \le 2$, integration by parts yields that $\hat{f}$ is well defined for all $t \ge 0$ and the bound

$$t|\hat{f}(t)| \le 2f(0)\;. \tag{1.45}$$

Assume for instance that $\mathrm{Re}(\hat{f})$ is regularly varying and denote $h = \mathrm{Re}(\hat{f})$. The bound (1.45) justifies the use of a Parseval-type Formula (see Theorem A.6):

$$\begin{aligned}\int_0^x f(t)\,\mathrm{d}t &= \int_0^\infty \mathbb{1}_{[0,x]}(t) f(t)\,\mathrm{d}t \\ &= \frac{2}{\pi}\int_0^\infty \frac{\sin(tx)}{t} h(t)\,\mathrm{d}t = \frac{2}{\pi}\int_0^\infty \frac{\sin(t)}{t} h(t/x)\,\mathrm{d}t\;.\end{aligned}$$

We can apply Proposition 1.18 (with $g(x) = h(1/x)$) to obtain

$$\lim_{x\to\infty} \frac{\int_0^x f(t)\,\mathrm{d}t}{h(1/x)} = \lim_{x\to\infty} \frac{2}{\pi}\int_0^\infty \frac{\sin(t)}{t}\frac{h(t/x)}{h(1/x)}\,\mathrm{d}t = \frac{2}{\pi}\int_0^\infty \frac{\sin(t)}{t^{2-\alpha}}\,\mathrm{d}t \; .$$

Thus the function $x \to \int_0^x f(t)\,\mathrm{d}t$ is regularly varying at infinity with index $1-\alpha$, and since f is monotone, we can apply Theorem 1.20 to conclude f is regularly varying at infinity with index $-\alpha$. □

Remark 1.42. Monotonicity is used twice in the proof. First to prove the convergence of the integral defining $\hat{f}$ and to obtain the bound (1.45) and finally to apply the monotone density Theorem.

Fourier-Stieltjes transform

By integration by parts, we can use the results of the present section to obtain the following corollaries which will be useful in Section 3.3.

Corollary 1.43. *Let F be a non decreasing function such that $F(0) = 0$, $\lim_{x\to\infty} F(x) = 1$ and $1 - F$ is regularly varying at infinity with index $-\alpha$ with $\alpha \in (0, 2]$. Then, as $t \to 0^+$, the following relations hold.*

(i) *If $0 < \alpha < 1$,*

$$\int_0^\infty \{\mathrm{e}^{\mathrm{i}tx} - 1\} F(\mathrm{d}x) \sim -\Gamma(1-\alpha)\mathrm{e}^{-\mathrm{i}\pi\alpha/2}\,\bar{F}(1/t) \; .$$

(ii) *If $1 < \alpha < 2$,*

$$\int_0^\infty \{\mathrm{e}^{\mathrm{i}tx} - 1 - itx\} F(\mathrm{d}x) \sim -\Gamma(1-\alpha)\mathrm{e}^{-\mathrm{i}\pi\alpha/2}\,\bar{F}(1/t) \; .$$

(iii) *If $\alpha = 1$,*

$$\int_0^\infty \{\mathrm{e}^{\mathrm{i}tx} - 1 - itx\mathbb{1}_{\{x\le 1\}}\} F(\mathrm{d}x) \sim -\bar{F}(1/t)\left\{\frac{\pi}{2} - \mathrm{i}\mu + \mathrm{i}\log(t)\right\} \; .$$

with $\mu = \int_0^\infty x^{-2}\{\sin(x) - x\mathbb{1}_{\{x\le 1\}}\}\mathrm{d}x$.

(iv) *If $\alpha = 2$, then*

$$\int_0^\infty \{\mathrm{e}^{\mathrm{i}tx} - 1 - \mathrm{i}tx\} F(\mathrm{d}x) \sim -t^2 \int_0^{1/t} x\bar{F}(x)\,\mathrm{d}x \; .$$

Proof. We only have to prove (iv). Denote $h(t) = t^2 \int_0^{1/t} x\bar{F}(x)\,\mathrm{d}x$. Note that since $x\bar{F}(x)$ is regularly varying with index -1, Karamata Theorem 1.15 yields that

$$\bar{F}(1/t) = o\left(h(t)\right) . \tag{1.46}$$

By integration by parts,

$$\begin{aligned} &\int_0^\infty \{1-\cos(tx)\}F(\mathrm{d}x) \\ &\qquad = t\int_0^\infty \sin(tx)\bar{F}(x)\,\mathrm{d}x \\ &\qquad = t\int_0^{1/t} \sin(tx)\bar{F}(x)\,\mathrm{d}x + \int_1^\infty \sin(x)\bar{F}(x/t)\,\mathrm{d}x . \end{aligned}$$

By Proposition 1.18,

$$\lim_{t\to 0}\int_1^\infty \sin(x)\frac{\bar{F}(x/t)}{\bar{F}(1/t)}\,\mathrm{d}x = \int_1^\infty \frac{\sin(x)}{x^2}\,\mathrm{d}x ,$$

thus $\int_1^\infty \sin(x)\bar{F}(x/t)\,\mathrm{d}x = O(\bar{F}(1/t))$. Next, by Karamata Theorem and (1.46),

$$\begin{aligned} &\int_0^{1/t} |\sin(tx) - tx|\bar{F}(x)\,\mathrm{d}x \\ &\qquad \le Ct^3\int_0^{1/t} x^3\bar{F}(x)\,\mathrm{d}x = O(t^{-1}\bar{F}(1/t)) = o(h(t)) . \end{aligned}$$

By similar techniques, we obtain that $\int_0^\infty \sin(tx)F(\mathrm{d}x) = t\int_0^\infty xF(\mathrm{d}x) + O(\bar{F}(1/t))$. This concludes the proof of (iv). □

Converses

Let ϕ be the characteristic function associated to the distribution function F on $\mathbb{R}$ and denote $\phi = U + \mathrm{i}V$, i.e. let U and V be defined by

$$U(t) = \int_{-\infty}^\infty \cos(tx)\,F(\mathrm{d}x) , \quad V(t) = \int_{-\infty}^\infty \sin(tx)\,F(\mathrm{d}x) .$$

Introduce the tail sum and tail difference fucntions H and K defined by

$$H(x) = 1 - F(x) + F(-x) \;, K(x) = 1 - F(x) - F(-x) \;.$$

Integration by parts yields

$$U(t) = 1 - t \int_0^\infty H(x) \sin(tx) \, \mathrm{d}x \;,$$
$$V(t) = t \int_0^\infty K(x) \cos(tx) \, \mathrm{d}x \;.$$

The following inversion formulae hold

$$H(x) = \frac{2}{\pi} \int_0^\infty \frac{1 - U(t)}{t} \sin(xt) \, \mathrm{d}t \;.$$
$$K(x) = \frac{2}{\pi} \int_0^\infty \frac{V(t)}{t} \cos(xt) \, \mathrm{d}t \;.$$

Thus H and K can be studied separately.

Theorem 1.44 (Pitman, 1968)**.** *If $1 - U(t)$ is regularly varying at zero with index $\alpha \in (0, 2)$, then H is regularly varying at infinity with index $-\alpha$ and*

$$H(x) \sim \frac{2\Gamma(\alpha) \sin(\pi\alpha/2)}{\pi} \{1 - U(1/x)\} \;. \tag{1.47}$$

Proof. Let H_2 be defined by

$$H_2(x) = \int_0^x \int_0^t u H(h) \, \mathrm{d}u \, \mathrm{d}t \;.$$

Applying the inversion formula for H and Fubini Theorem, we get

$$\begin{aligned} H_2(x) &= \frac{2}{\pi} \int_0^\infty \{1 - U(t)\} \frac{2 - 2\cos(tx) - tx \sin(tx)}{t^4} \, \mathrm{d}t \\ &= \frac{2x^3}{\pi} \int_0^\infty \{1 - U(t/x)\} h(t) \, \mathrm{d}t \;, \end{aligned}$$

where $h(t) = t^{-4}\{2 - 2\cos(t) - t\sin(t)\}$. The function h is continuous and bounded on $[0, \infty)$ and $h(t) = O(t^{-4})$ at infinity, so we can apply Proposition 1.18 (with $\beta = 0$ and $\gamma = -4$) to obtain

$$\lim_{x \to \infty} \frac{H_2(x)}{x^3 \{1 - U(1/x)\}} = \frac{2}{\pi} \int_0^\infty h(t) t^\alpha \, \mathrm{d}t \;.$$

The function H_2 is thus regularly varying at infinty with index $3 - \alpha$. By the monotone density theorem, the function $x \to \int_0^x uH(u)\,\mathrm{d}u$ is regularly varying with index $2 - \alpha$. By Lemma 1.21, this implies that H is regularly varying at infinity with index $-\alpha$. and we obtain the Tauberian result (1.47) by applying the Abelian Corollary 1.43. □

By similar techniques, we can obtain a Tauberian result in the case $\alpha = 2$.

Theorem 1.45. *If the function* $1 - U$ *is regularly varying at 0 with index* 2*, then*

$$\int_0^x tH(t)\,\mathrm{d}t \sim x^2\{1 - U(1/x)\} .$$

Remark 1.46. In view of Chapter 3, note that if X is a random variable with distribution function F, then

$$2\int_0^x tH(t)\,\mathrm{d}t = \mathbb{E}[|X|^2 \mathbb{1}_{\{|X| \leq x\}} , \tag{1.48}$$

and if $1 - U$ is regularly varying with index 2 at 0 or equivalently if H is regularly varying at infinity with index -2, then both sides of (1.48) are slowly varying functions.

1.3 Second order regular variation

In many applications, in particular statistical applications, it is necessary to know the rate of convergence (1.2).

Definition 1.47. *A slowly varying function* ℓ *is said to be second order slowly varying with auxiliary function* b *if there exists a positive function* A *such that*

$$\lim_{x\to\infty} \frac{1}{b(x)} \log \frac{\ell(tx)}{\ell(x)} = A(t) . \tag{1.49}$$

Proposition 1.48. *There exist* $\rho \leq 0$ *and* $c > 0$ *such that* $A(t) = c\int_1^t u^{\rho-1}\,\mathrm{d}u$*, and* b *is regularly varying with index* ρ*.*

Proof. Define $h(x) = \log \ell(x)$. Condition (1.49) can be expressed as

$$\lim_{x\to\infty} \frac{h(tx) - h(x)}{b(x)} = A(t) \; .$$

Let $s, t > 0$. Then

$$\frac{h(stx) - h(x)}{b(x)} = \frac{b(tx)}{b(x)} \frac{h(stx) - h(tx)}{b(tx)} + \frac{h(tx) - h(x)}{b(x)} \; ,$$

whence

$$\lim_{x\to\infty} \frac{b(tx)}{b(x)} = \frac{A(st) - A(t)}{A(s)} \; .$$

Thus b is regularly varying with index $\rho \leq 0$ since $\lim_{x\to\infty} b(x) = 0$. This implies

$$\frac{A(st) - A(t)}{A(s)} = t^\rho \; .$$

If $\rho = 0$, then $A(st) = A(s) + A(t)$, hence $A(t) = c\log(t)$. If $\rho < 0$, then

$$A(st) = t^\rho A(s) + A(t) \; ,$$

and by interchanging the roles of s and t, it follows that

$$t^\rho A(s) + A(t) = s^\rho A(t) + A(s) \; ,$$

i.e. $A(s)(1 - t^\rho) = A(t)(1 - s^\rho)$, which means that the function $(1 - s^\rho)^{-1}A(s)$ is constant; in other words, there exists $c > 0$ such that $A(s) = c(1 - s^\rho)$. □

Example 1.49. If the slowly varying function ℓ is in the Zygmund class, then it can be expressed as $\ell(x) = c \exp \int_{x_0}^\infty s^{-1}\eta(s)\,\mathrm{d}s$ with $c \in (0,\infty)$ and $\lim_{s\to\infty} \eta(s) = 0$. The function η is not necessarily regularly varying, but if it is, then its index of regular variation, say ρ, is necessarily less than or equal to zero. In that case, we have

$$\frac{1}{\eta(x)} \log\left(\frac{\ell(tx)}{\ell(x)}\right) = \int_1^t \frac{\eta(sx)}{\eta(x)} \frac{\mathrm{d}s}{s} \to \int_1^t s^{\rho-1}\,\mathrm{d}s \; ,$$

as $s \to \infty$ by the uniform convergence theorem. Thus, ℓ is second order regularly varying with auxiliary function η.

Chapter 2

Random variables with regularly varying tails

A random variable X with distribution function F is said to have a regularly varying upper or right tail (lower or left tail) if its survival function $\bar{F}$ (its left tail $x \to F(-x)$) is regularly varying at infinity. If X is non negative, we will say indifferently that X is regularly varying or $\bar{F}$ is regularly varying at infinity.

In this chapter, we state some fundamental properties of regularly varying random variables, and study finite and random sums, and products of independent regularly varying random variables. In the last section, we consider solution of stochastic recurrence equations. We will apply the results of the previous chapter, notably, we use the Abelian and Tauberian theorems to characterize regular variation.

2.1 Moments and quantile function

Moments and truncated moments

For a distribution function F concentrated on $[0, \infty)$ and $\beta \geq 0$, denote

$$M_\beta(x) = \int_0^x y^\beta F(\mathrm{d}y) \; , \quad R_\beta(x) = \int_x^\infty y^\beta \, F(\mathrm{d}y) \; .$$

The following result is an immediate consequence of the representation Theorem 1.6 and of Karamata Theorem 1.15.

Proposition 2.1 (Truncated Moments)**.** *Let X be a random variable with distribution function F. If $\bar{F}$ is regularly varying at infinity with index $-\alpha$, $\alpha \geq 0$, then $\mathbb{E}[X^\beta] < \infty$ if $\beta < \alpha$ and $\mathbb{E}[X^\beta] = \infty$ if $\beta > \alpha$, and $\mathbb{E}[X^\alpha]$ may be finite or infinite. Moreover,*

$$M_\beta(x) = \mathbb{E}[X^\beta \mathbb{1}_{\{X \leq x\}}] \sim \frac{1}{\beta - \alpha} x^\beta \bar{F}(x) \; , \; \beta > \alpha \; ,$$
$$R_\beta(x) = \mathbb{E}[X^\beta \mathbb{1}_{\{X > x\}}] \sim \frac{1}{\beta - \alpha} x^\beta \bar{F}(x) \; , \; \alpha > \beta \; ,$$

Laplace transform

Let F be a distribution function on $[0, \infty)$ and define $m_k = \int_0^\infty x^k F(\mathrm{d}x)$. Recall that the Laplace transform $\mathcal{F}$ of F is defined by

$$\mathcal{L}F(s) = \int_0^\infty \mathrm{e}^{-s} F(\mathrm{d}s) \; ,$$

or equivalently, if X is a non negative random variable with distribution function F,

$$\mathcal{L}F(s) = \exp[\mathrm{e}^{-sX}] \; .$$

The Laplace transform is k times differentiable at zero if and only if $\mathbb{E}[X^k] < \infty$. If $\mathbb{E}[X^n] < \infty$ for some $n \geq 1$, define then

$$g_n(s) = (-1)^{n+1} \left\{ \mathcal{L}(s) - \sum_{k=0}^{n} \frac{m_k}{k!} (-s)^k \right\} \; .$$

The results of Section 1.2.1 can be adapted and extended. The behaviour of g_n at zero can be related to the behaviour of F at infinity. Cf. (Bingham et al., 1989, Theorem 8.1.6).

Theorem 2.2. *Let ℓ be a slowly varying function at infinity and $\alpha > 0$. If $\alpha \notin \mathbb{N}$, the following are equivalent:*

$$g_n(s) = s^\alpha \ell(1/s) \; , \tag{2.1}$$
$$1 - F(x) = \frac{(-1)^n}{\Gamma(1-\alpha)} x^{-\alpha} \ell(x) \; . \tag{2.2}$$

There is no such equivalence when α is an integer. For $\alpha = 1$, the following result holds.

Theorem 2.3. *If ℓ is slowly varying at infinity, the following are equivalent:*

$$\mathcal{L}F(s) = 1 - s\ell(1/s) \;, \tag{2.3}$$

$$\int_0^x \{1 - F(t)\}\,\mathrm{d}t \sim \ell(x) \;. \tag{2.4}$$

Proof. The Laplace transform of the function $x \to \int_0^x \{1 - F(t)\}\,\mathrm{d}t$ is $(1 - \mathcal{L}(s))/s$, and we conclude by applying Theorem 1.26. □

Remark 2.4. We cannot conclude that F is regularly varying with index -1 in the previous result, since this is precisely the case where the monotone density theorem does not apply. The function $\int_0^x \{1-F(t)\}\,\mathrm{d}t$ belongs to the Π-class. Cf. Section 1.1.3.

Quantile function

The quantile function of a distribution function F is its left-continuous inverse $F^{\leftarrow}$, defined on $[0, 1]$ by

$$F^{\leftarrow}(t) = \inf\{x \;,\; F(x) \geq t\} \;.$$

It is well known that if U is uniformly distributed on $[0, 1]$, then the distribution function of $F^{\leftarrow}(U)$ is F. If F is continuous and X is a random variable with distribution function F, then $F(X)$ is uniformly distributed on $[0, 1]$.

By the inversion theorem, $\bar{F}$ is regularly varying at infinity with index $-\alpha < 0$, if and only if the quantile function is regularly varying at 0 with index $-\alpha$.

It is also customary to consider the function Q defined on $[1, \infty)$ by $Q(t) = F^{\leftarrow}(1 - 1/t)$, which is regularly varying at infinity with index $1/\alpha$. If $\bar{F}(x) = x^{-\alpha}\ell(x)$, then

$$Q(t) = t^{1/\alpha}\ell^{\#}(t^{1/\alpha}) \;,$$

where $\ell^{\#}$ is the De Bruyn conjugate of $\ell^{-1/\alpha}$.

2.2 Sums

2.2.1 Finite sums

Sums of independent regularly varying random variables are regularly varying.

Theorem 2.5. *Let* $X_1, \dots, X_k$ *be i.i.d. random variables with distribution function* F *such that* $1 - F \in RV_\infty(-\alpha)$. *Then*

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_1 + \cdots + X_k > x)}{\mathbb{P}(X_1 > x)} = k \; . \tag{2.5}$$

Remark 2.6. This property characterizes the class of subexponential distribution functions, which is wider than the class of distribution functions with regularly varying tails. It can be shown that a necessary condition for (2.5) is that $\mathbb{E}[\mathrm{e}^{sX_1}] = \infty$ for all $s > 0$.

Proof. The proof is by induction. Let X and Y be independent random variables with distribution functions F and G, respectively. Assume that $1 - F \in RV_\infty(-\alpha)$ and that there exists c such that

$$\lim_{x\to\infty} \frac{1 - G(x-y)}{1 - F(x)} = c \geq 0 \; , \tag{2.6}$$

uniformly with respect to y in compact subsets of $[0, \infty)$. This obviously holds, with $c = 1$, if $G = F$ by the uniform convergence theorem. Then

$$\begin{aligned} &\frac{\mathbb{P}(X + Y > x)}{\mathbb{P}(X > x)} \\ &= 1 + \int_0^{x(1-\epsilon)} \frac{\mathbb{P}(Y > x - y)}{\mathbb{P}(X > x)} F(\mathrm{d}y) + \int_{x(1-\epsilon)}^{x} \frac{\mathbb{P}(Y > x - y)}{\mathbb{P}(X > x)} F(\mathrm{d}y) \; . \end{aligned}$$

We thus get the lower bound

$$\begin{aligned} \liminf_{x\to\infty} \frac{\mathbb{P}(X + Y > x)}{\mathbb{P}(X > x)} &\geq 1 + \liminf_{x\to\infty} \int_0^{x(1-\epsilon)} \frac{\mathbb{P}(Y > x - y)}{\mathbb{P}(X > x)} F(\mathrm{d}y) \\ &= 1 + c \lim_{x\to\infty} F(x(1-\epsilon)) = 1 + c \; , \end{aligned}$$

and the upper bound

$$\limsup_{x\to\infty} \frac{\mathbb{P}(X+Y>x)}{\mathbb{P}(X>x)} \leq 1 + c \lim_{x\to\infty} F(x(1-\epsilon)) + \lim_{x\to\infty} \frac{\mathbb{P}(X>x(1-\epsilon)) - \mathbb{P}(X>x)}{\mathbb{P}(X>x)} = c + (1-\epsilon)^{-\alpha} \; .$$

Since ϵ is arbitrary, this proves that

$$\lim_{x\to\infty} \frac{\mathbb{P}(X+Y>x)}{\mathbb{P}(X>x)} = 1 + c \; , \tag{2.7}$$

We can now prove our result by induction. Taking $F = G$ yields (2.5) for $k = 2$. Assuming (2.5) true for some $k \geq 2$. Then, taking $G = F^{*k}$ yields (2.6) with $c = k$. This implies (2.5) for $k+1$. This concludes the induction argument and the proof. □

Remark 2.7. By taking $c = 0$ in (2.6) and (2.7), we have also proved the seemingly obvious result that if X and Y are two independent nonnegative random variables, if X has a regularly varying right tail and the tail of Y is lighter than the tail of X, then the tail of the sum is equivalent to the tail of X. This is not necessarily true if X is not regularly varying. Let for instance X be a $\Gamma(n,1)$ random variable, i.e. the sum of n i.i.d. standard exponential, and let Y be independent of X and have a standard exponential distribution. Then $X+Y$ has a $\Gamma(n+1,1)$ distribution, which has a heavier tail than that of X.

A converse of Theorem 2.5 can be proved by means of the Tauberian Corollary 1.27. The proof is adapted from the proof of Faÿ et al. (2006, Proposition 4.8).

Theorem 2.8. *Let $X_1,\dots,X_n$ be i.i.d. nonnegative random variables. If $\mathbb{P}(X_1+\cdots+X_n>x)$ is regularly varying with index $-\alpha$, $\alpha \geq 0$, then $\mathbb{P}(X_1>x)$ is also regularly varying with index $-\alpha$ and (2.5) holds.*

Proof. Assume first that $\alpha \in (0,1)$. Let F denote the distribution function of X_1 and $\mathcal{L}$ its Laplace transform. By assumption and by Theorem 1.27, $1-\mathcal{L}^n$ is regularly varying at zero with index α, i.e.

$$1 - \mathcal{L}^n(t) \sim t^\alpha \ell(t) \; , \; t \to 0^+ \; ,$$

where ℓ is slowly varying at zero. Writing $\mathcal{L}^n = \{1 - (1 - \mathcal{L})\}^n$, this yields

$$1 - \mathcal{L}(t) \sim n^{-1} t^{\alpha} \ell(t) \ .$$

Thus $\bar{F}$ is regularly varying at infinity with index $-\alpha$.

If $\alpha \geq 1$, we prove that $\mathbb{P}(X_1^2 + \cdots + X_n^2 > x)$ is regularly varying with index $\alpha/2$. Denote $S = X_1 + \cdots + X_n$ and $Y = X_1^2 + \cdots + X_n^2$. By assumption, S^2 has a regularly varying tail with index $\alpha/2$. Fix some $\epsilon > 0$ arbitrarily small. Then

$$\mathbb{P}(Y > x) \leq \mathbb{P}(S^2 > x) \leq \mathbb{P}(Y > (1-\epsilon)x) + \mathbb{P}(Z > \epsilon x)$$

with $Z = \sum_{i \neq j} X_i X_j$. For any $\beta \in (\alpha/2, \alpha)$, $\mathbb{E}[Z^\beta] < \infty$, thus Z has a lighter tail than S^2. This implies that $\mathbb{P}(Y > y) \sim \mathbb{P}(S^2 > x)$. This proves our claim.

If $\alpha \in [1, 2)$, the first part of the proof implies that X_1^2 has a regularly varying tail with index $\alpha/2$, hence $\bar{F}$ is regularly varying with index α. The proof is continued for $\alpha \geq 2$ by induction. □

Relative behaviour of sums and maxima

The previous result has an interpretation in terms of the relative behaviour of the maximum and the sum of i.i.d. random variables. Since $\mathbb{P}(\max_{1 \leq i \leq k} X_i > x) \sim k\bar{F}(x)$ as $x \to \infty$, the limit (2.5) can be expressed as

$$\lim_{x \to \infty} \frac{\mathbb{P}(\sum_{i=1}^k X_i > x)}{\mathbb{P}(\max_{1 \leq i \leq k} X_i > x)} = 1 \ .$$

Regular variation can be characterized by the relative behaviour of the sums and maxima. We cite as an illustration this result. See (Bingham et al., 1989, Theorem 8.15.1).

Theorem 2.9. *Let $\{X_n\}$ be a sequence of i.i.d. non negative random variables with distribution function F. Denote $S_n = X_1 + \cdots + X_n$ and $M_n = \max_{1 \leq i \leq n} X_i$. Then*

(i) $M_n/S_n \to 0$ in probability if and only if $\int_0^y y \, F(\mathrm{d}y)$ is slowly varying;

(ii) $M_n/S_n \to 1$ *in probability if and only if* $1-F$ *is slowly varying at infinity;*

(iii) M_n/S_n *has a non degenerate limiting distribution if and only if* $1-F$ *is regularly varying with index* $\alpha \in (0,1)$.

Exercise 2.2.1. Let F be an absolutely continuous distribution function with density f and failure rate $r = f/(1-F)$. Prove that if r is decreasing with limit 0 at infinity and

$$\int_0^\infty \mathrm{e}^{xr(x)} f(x)\,\mathrm{d}x < \infty ,$$

then F is subexponential, i.e. (2.5) holds.

2.2.2 Weighted sums

The results of the previous section can be extended to finite or infinite weighted sums of i.i.d. regularly varying random variables.

Theorem 2.10. *Let* $\{Z_j\}$ *be a sequence of i.i.d. random variable with regularly varying tails that satisfy the tail balance condition*

$$\lim_{x\to\infty} \frac{\mathbb{P}(Z_1 > x)}{\mathbb{P}(|Z| > x)} = p \in [0,1] . \tag{2.8}$$

Let $\{\psi_j\}$ *be a sequence of deterministic weights such that one of the following conditions hold:*

$$\alpha > 2 ,\ \mathbb{E}[Z_1] = 0 ,\ \sum_{j=1}^\infty \psi_j^2 < \infty , \tag{2.9}$$

$$\alpha \in (1,2] ,\ \mathbb{E}[Z_1] = 0 ,\ \sum_{j=1}^\infty |\psi_j|^{\alpha-\epsilon} < \infty ,\ \textit{for some } \epsilon \in (0,\alpha) , \tag{2.10}$$

$$\alpha \in (0,1] ,\ \sum_{j=1}^\infty |\psi_j|^{\alpha-\epsilon} < \infty ,\ \textit{for some } \epsilon \in (0,\alpha) . \tag{2.11}$$

Then

$$\mathbb{P}\left(\sum_{j=1}^\infty \psi_j Z_j > x\right) \sim \mathbb{P}(|Z_1| > x) \sum_{j=1}^\infty \{p(\psi_j)_+^\alpha + (1-p)(\psi_j)_-^\alpha\} , \tag{2.12}$$

with $x_+ = \max(x,0)$ *and* $x_- = \max(-x,0)$ *for any real number* x.

Proof. The proof consists in proving (2.12) for a finite sum, and then using a truncation argument. For a finite sums of i.i.d. random variables satisfying the tail balance condition (2.8), the result is a simple extension of Theorem 2.5. Denote $c_m = \sum_{j=1}^m \{p(\psi_j)_+^\alpha + (1-p)(\psi_j)_-^\alpha\}$. Denote $X = \sum_{j=1}^\infty \psi_j Z_j$ and $X_m = \sum_{j=1}^m \psi_j Z_j$. For each m, we have

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_m > x)}{\mathbb{P}(|Z_1| > x)} = c_m .$$

Since $c_m \to c = \sum_{j=1}^\infty \{p(\psi_j)_+^\alpha + (1-p)(\psi_j)_-^\alpha\}$, it is then sufficient to prove that

$$\lim_{m\to\infty} \limsup_{x\to\infty} \frac{\mathbb{P}(X - X_m > x)}{\mathbb{P}(|Z_1| > x)} = 0 . \tag{2.13}$$

Indeed, if (2.13) holds, then

$$\begin{aligned}
&\limsup_{x\to\infty} \frac{\mathbb{P}(X > x)}{\mathbb{P}(|Z_1| > x)} \\
&\qquad \le \limsup_{x\to\infty} \frac{\mathbb{P}(X_m > (1-\epsilon)x)}{\mathbb{P}(|Z_1| > x)} + \limsup_{x\to\infty} \frac{\mathbb{P}(X - X_m > \epsilon x)}{\mathbb{P}(|Z_1| > x)} \\
&\qquad = c_m(1-\epsilon)^{-\alpha} + \epsilon^{-\alpha} \limsup_{x\to\infty} \frac{\mathbb{P}(X - X_m > \epsilon x)}{\mathbb{P}(|Z_1| > \epsilon x)} \\
&\qquad \to c \quad \text{as } m \to \infty\ ,\ \epsilon \to 0 .
\end{aligned}$$

Since X_m and $X - X_m$ are independent, we also have

$$\begin{aligned}
\liminf_{x\to\infty} \frac{\mathbb{P}(X > x)}{\mathbb{P}(|Z_1| > x)} &\ge \liminf_{x\to\infty} \frac{\mathbb{P}(X_m > (1+\epsilon)x)\, \mathbb{P}(X - X_m > -\epsilon x)}{\mathbb{P}(|Z_1| > x)} \\
&= c_m(1+\epsilon)^{-\alpha} \to c \quad \text{as } m \to \infty\ ,\ \epsilon \to 0 .
\end{aligned}$$

We now prove (2.13) in the case $\alpha < 1$. For the other cases, see Mikosch and Samorodnitsky (2000, Lemma A3).

$$\begin{aligned}
&\mathbb{P}(X - X_m > x) \\
&\quad = \mathbb{P}(X - X_m > x\ ,\ \vee_{j=m+1}^\infty |\psi_j||Z_j| > x) \\
&\quad + \mathbb{P}(X - X_m > x\ ,\ \vee_{j=m+1}^\infty |\psi_j||Z_j| \le x) \\
&\quad \le \sum_{j=m+1}^\infty \mathbb{P}(|\psi_j||Z_j| > x) + \mathbb{P}\left(\sum_{j=m+1}^\infty |\psi_j||Z_j| \mathbb{1}_{\{|\psi_j||Z_j| \le x\}} > x \right) .
\end{aligned}$$

For large enough j, $|\psi_j| \le 1$, hence we can apply Theorem 1.13 and obtain, for some constant C and large enough x,

$$\sum_{j=m+1}^{\infty} \frac{\mathbb{P}(|\psi_j||Z_j| > x)}{\mathbb{P}(|Z_1| > x)} \le C \sum_{j=m+1}^{\infty} |\psi_j|^{\alpha-\epsilon}$$

and this tends to 0 as $m \to \infty$ by assumption (2.11). Next by Markov inequality and Karamata Theorem, we have

$$\mathbb{P}\left(\sum_{j=m+1}^{\infty} |\psi_j||Z_j| \mathbb{1}_{\{|\psi_j||Z_j| \le x\}} > x \right) \\ \le \frac{1}{x} \sum_{j=m+1}^{\infty} |\psi_j| \, \mathbb{E}[|\psi_j||Z_j| \mathbb{1}_{\{|Z_j| \le x\}}] \le C \sum_{j=m+1}^{\infty} \mathbb{P}(|\psi_j||Z_j| > x) \; ,$$

and we conclude as previously. □

2.2.3 Random sums

Let N be an integered-valued random variable, $\{X_i\}$ be a sequence of i.i.d. non-negative random variables with regularly varying tail. Define

$$S = \sum_{j=1}^{N} X_j \; .$$

The tails of random sums have been widely studied. Two cases are of interest: the random number of terms is independent of the summands, or it is a stopping time with respect to the filtration of the summands. We only consider the former case, for which the results are quite exhaustive. Based on the Tauberian result for Laplace transforms, we can provide an easy proof in the case where the tail of each summand is regularly varying with index $\alpha \in (0, 1)$ and the tail of the number of terms is lighter than the tail of the summand.

Theorem 2.11. *If* $\mathbb{P}(X_1 > x)$ *is regularly varying at infinity with index* $\alpha \in (0, 1)$ *and* $\mathbb{E}[N] < \infty$ *then*

$$\mathbb{P}(S > x) \sim \mathbb{E}[N] \, \mathbb{P}(X_1 > x) \; .$$

Proof. Let ψ denote the Laplace transform of X, i.e. $\psi(s) = \mathrm{e}^{-sX_1}$. Then the Laplace transform ψ_S of S is given by

$$\psi_S(t) = \mathbb{E}[\mathrm{e}^{-tS}] = \mathbb{E}[\psi^N(t)] = \sum_{k=0}^{\infty} \psi^k(t)\mathbb{P}(N = k)$$
$$= 1 + \sum_{k=1}^{\infty} \{\psi^k(t) - 1\}\mathbb{P}(N = k) \ .$$

For $\epsilon > 0$, there exists some t_0 such that $|1 - \psi(t)| < \epsilon$ if $t \leq t_0$. Then, for all $k \geq 1$,

$$k(1-\epsilon)^{k-1}\{1 - \psi(u)\} \leq 1 - \psi^k(u) \leq k\{1 - \psi(u)\} \ .$$

Summing these relations, we obtain

$$\{1 - \psi(u)\} \sum_{k=1}^{\infty} k\mathbb{P}(N = k)(1-\epsilon)^{k-1} \leq 1 - \psi_S(t) \leq \{1 - \psi(u)\}\mathbb{E}[N] \ .$$

Let the function H be defined by on $[0, 1]$ by $H(z) = \mathbb{E}[z^N]$. Since $\mathbb{E}[N] < \infty$, H is continuously differentiable on $[0, 1]$ and $H'(1) = \mathbb{E}[N]$. Thus, for any $\eta > 0$, ϵ van be chosen small enough so that $H'(1 - \epsilon) \geq \mathbb{E}[N](1 - \eta)$. This yields

$$1 - \eta \leq \frac{1 - \psi_S(t)}{\mathbb{E}[N]\{1 - \psi(t)\}} \leq 1 \ .$$

Thus $1 - \psi_S(t) \sim \mathbb{E}[N]\{1 - \psi(t)\}$. Since the right tail of of X_1 is regularly varying with index $\alpha \in (0, 1)$, we have, for some fucntion ℓ slowly varying at infinity, that $\mathbb{P}(X_1 > x) \sim x^{-\alpha}\ell(x)$, which implies by Corollary 1.27 that $1 - \psi(t) \sim t^\alpha \ell(1/t)$. Thus we have just proved that $1 - \psi_S(t) \sim \mathbb{E}[N]t^\alpha \ell(1/t)$ and this in turn implies that $\mathbb{P}(S > x) \sim \mathbb{E}[N]x^{-\alpha}\ell(x) = \mathbb{E}[N]\mathbb{P}(X_1 > x)$. □

The Tauberian theorems allow to extend easily these results to all non integer indices of regular variation, provided N has enough finite moments. If $\mathbb{E}[z^N] < \infty$ for some $z > 1$, then a simple proof valid for all α can be obtained by means of the following Lemma (Asmussen, 2000, Lemma IX.1.8)

Lemma 2.12. *If $1 - F$ is regularly varying, then for all $\epsilon > 0$, there exists a constant C such that $\overline{F^{*n}}(x) \leq C(1+\epsilon)^n \bar{F}(x)$.*

Theorem 2.13. *If $\mathbb{P}(X_1 > x)$ is regularly varying at infinity and if there exists $z > 1$ such that $\mathbb{E}[z^N] < \infty$, then*

$$\mathbb{P}(S > x) \sim \mathbb{E}[N]\,\mathbb{P}(X_1 > x) .$$

Proof. Let $\epsilon \in (0, z-1)$ where $\mathbb{E}[z^N] < \infty$. By Lemma 2.12, there exists C such that $\overline{F^{*n}}(x) \leq C(1+\epsilon)^n \bar{F}(x)$. Thus,

$$\frac{\mathbb{P}(S > x)}{\mathbb{P}(X > x)} = \sum_{n=1}^{\infty} \mathbb{P}(N = n) \frac{\overline{F^{*n}}(x)}{\bar{F}(x)} .$$

Now, for each n, $\mathbb{P}(N = n)\overline{F^{*n}}(x)/\bar{F}(x)$ converges to $n\mathbb{P}(N = n)$ and is bounded by $C(1+\epsilon)^n \mathbb{P}(N = n)$ which is a summable series by assumption. Hence bounded convergence applies and the proof is concluded. □

The full story is told in Faÿ et al. (2006, Proposition 4.1).

Theorem 2.14. *If $\mathbb{P}(X_1 > x)$ is regularly varying at infinity with index $\alpha > 0$, $\mathbb{E}[N] < \infty$ and $\mathbb{P}(N > x) = o(\mathbb{P}(X_1 > x))$ then*

$$\mathbb{P}(S > x) \sim \mathbb{E}[N]\,\mathbb{P}(X_1 > x) .$$

This result has a partial converse, which generalizes Theorem 2.8 to random sums. It can be found in Faÿ et al. (2006, Proposition 4.8).

Theorem 2.15. *Let $\{X_i\}$ be a sequence of i.i.d. nonnegative random variables, K be a an integer valued random variable, independent of $\{X_i\}$ and define $S = \sum_{k=1}^{K} X_k$. If $\mathbb{P}(S > x)$ is regularly varying with index $-\alpha$, $\alpha > 0$ and if $\mathbb{E}[K^{1\vee\alpha'}] < \infty$ for some $\alpha' > \alpha$, then $\mathbb{P}(X > x)$ is regularly varying with index $-\alpha$ and $\mathbb{P}(X > x) \sim \mathbb{P}(S > x)/\mathbb{E}[K]$.*

Ruin probability

This result has an application in insurance mathematics. Let $\{X_k\}$ be a sequence of i.i.d random variables representing the claim sizes and assume that the claim arrivals form a homogeneous Poisson process N

with intensity λ. Let u be the initial capital of the insurance company, and let c be the premium. The risk process is defined by

$$S(t) = u + ct + \sum_{i=1}^{N} X_i \ .$$

The ruin probability is the probability that S be negative at some time t:

$$\psi(u) = \mathbb{P}(\exists t > 0 \, S(t) < 0) \ .$$

If $\mathbb{E}[X_i] < \infty$, then a necessary and sufficient condition for $\psi(u) < 1$ is

$$c > \lambda \mathbb{E}[X_1] \ . \tag{2.14}$$

This condition means that on average, the company earns more than it loses at each unit period of time. If (2.14) holds, the *safety loading* is defined as

$$\rho = \frac{c}{\lambda \mathbb{E}[X_1]} - 1 \ .$$

Let F denote the distribution of X_1, and still assuming that $\mathbb{E}[X_1] < \infty$, let H be the distribution function defined by

$$H(u) = \frac{1}{\mathbb{E}[X_1]} \int_0^u \{1 - F(s)\} \, \mathrm{d}s \ .$$

If $\bar{F}$ is regularly varying at infinity with index $\alpha > 1$, then by Karamata's Theorem 1.15, H is regularly with index $\alpha - 1$ and

$$\bar{H}(u) \sim \frac{u \bar{F}(u)}{\mathbb{E}[X_1](\alpha - 1)}$$

The celebrated Pollaczek-Khinčin formula gives an expression of the ruin probability as a compound geometric distribution.

$$1 - \psi(u) = \frac{\rho}{1 + \rho} \sum_{n=0}^{\infty} (1 + \rho)^{-n} H^{*n}(u) \ .$$

Thus the ruin probability can be expressed as follows. Let $\{Y_i\}$ be an i.i.d. sequence of random variables with distribution H and let K

be an integer-valued random variable with geometric distribution with parameter $\rho/(1+\rho)$, which means here

$$\mathbb{P}(K = k) = \rho(1+\rho)^{-k-1} \;, \quad k \geq 0 \;.$$

It is then easy to see that

$$\psi(u) = \mathbb{P}(\sum_{i=1}^{K} Y_i > u) \;.$$

Since $\mathbb{E}[z^K] = \rho/(1+\rho-z)$ for all $z < 1+\rho$, the assumptions of Theorem 2.13 hold, and we obtain

$$\psi(u) \sim \mathbb{E}[K]\, \bar{H}(u) \sim \frac{u\bar{F}(u)}{\rho\mathbb{E}[X_1](\alpha-1)} \;.$$

Exercise 2.2.2. Prove Theorem 2.14 using the following *large devation* inequality of Nagaev (1979) ($\alpha > 2$) and Cline and Hsing (1998) ($\alpha \leq 2$). For any $\delta > 0$,

$$\lim_{n\to\infty} \sup_{y \geq \delta n} \left| \frac{\mathbb{P}(S_n - cn > y)}{k\mathbb{P}(X_1 - y)} - 1 \right| = 0 \;,$$

where $\{X_k\}$ is a sequence of i.i.d non negative random variables with regularly varying tails with index $\alpha > 0$ and $c = \mathbb{E}[X_1]$ when finite or $c \in \mathbb{R}$ is arbitrary if $\mathbb{E}[X_1] = \infty$.

Tail of infinitely divisible distributions

The Lévy-Khinčin Theorem states that an infinitely divisible distribution has a characteristic function that never vanishes and is given by

$$\log \phi(t) = \mathrm{i}bt - \frac{\sigma^2}{2}t^2 + \int \{\mathrm{e}^{\mathrm{i}tx} - 1 - \mathrm{i}tx\mathbb{1}_{\{|x|\leq 1\}}\}\, \nu(\mathrm{d}x) \;, \qquad (2.15)$$

where $b, \sigma \in \mathbb{R}$ and ν is a measure on $\mathbb{R}$ such that

$$\nu(\{x \in \mathbb{R}\,,\ |x| \geq \eta\}) < \infty \quad \text{for all } \eta > 0 \;,$$
$$\int_{\{|x|\leq 1\}} x^2\, \nu(\mathrm{d}x) < \infty \;.$$

Theorem 2.16. *The infinitely divisible distribution F with Lévy measure ν has regularly varying right tails (resp. left tails) if and only if the function $x \to \nu(x,\infty)$ (resp. $x \to \nu(-\infty,-x)$) is regularly varying, in which case they have the same index of regular variation.*

Proof. Let X be a random variable with an infinitely divisible distribution F with characteristic function given by (2.15). We can assume without loss of generality that $b = 0$. Then X can be expressed as $X = X_0 + X_- + X_+ + Z$, where X_1, X_2 and Z are independent, Z is a centered Gaussian random variable with variance σ^2, and the characteristic functions of X_0, X_- and X_+ are ϕ_0, ϕ_- and ϕ_+, respectively given by

$$\begin{aligned}
\log \phi_0(t) &= \int_{|x|\le 1} \{\mathrm{e}^{\mathrm{i}tx} - 1 - \mathrm{i}tx\}\, \nu(\mathrm{d}x) , \\
\log \phi_-(t) &= \int_{x<-1} \{\mathrm{e}^{\mathrm{i}tx} - 1\}\, \nu(\mathrm{d}x) , \\
\log \phi_+(t) &= \int_{x>1} \{\mathrm{e}^{\mathrm{i}tx} - 1\}\, \nu(\mathrm{d}x) .
\end{aligned}$$

Define $\theta_+ = \nu(\{x > 1\})$; then $\theta < \infty$ by assumption. Let F_+ be the distribution function corresponding to the probability measure $\theta_+^{-1}\nu$ on $(1,\infty)$. Then ϕ_+ is the characteristic function of the compound Poisson distribution with intensity θ and jump distribution F_+, i.e. X_+ can be represented as

$$X_+ = \sum_{k=1}^{N} Y_i ,$$

where N has a Poisson distribution with parameter θ and the Y_i are i.i.d. with distribution F_+ and independent of N. Similarly, X_- is a compound Poisson distribution with intensity $\theta_- = \nu(-\infty,-1)$ and jump distribution $F_- = \theta_-^{-1}\nu$ restricted to $(-\infty,-1)$. Since the Poisson distribution has exponentail tails, if $\nu(x,\infty)$ is regularly varying at infinity, Theorem 2.14 yields that $1-F$ is regularly varying with the same index as $\nu(x,\infty)$. The same argument holds for the left tail.

Since X_+ and X_- are independent, the left and right tails of $X_+ - X_-$ are precisely the tails of X_- and X_+. The random variable Z is gaussian so its tails are lighter than those of X_- and X_+. Finally, since $\log \phi_0$ is infinitely differentiable at zero, all moments of X_0 are finite. Thus, if $\nu(x,\infty)$ is regularly varying at infinity, then so is X, with the same index.

Conversely, if $1-F$ is regularly varying at infinity, then so is X_+. By Theorem 2.15, this implies that $\nu(x,\infty)$ is regularly varying with the same index as $1-F$. □

2.3 Products

Let X be a nonnegative random variable with a regularly varying tail of index $-\alpha$, and let Y be independent of X. A question arising in many applications is to know wether the product XY has a regularly varying tail. Two cases are of interest: if Y has a lighter tail than X, and if X of Y have the same distribution.

The first case is dealt with by the celebrated Breiman (1965)'s Lemma.

Theorem 2.17. *Let X and Y be independent non negative random variables, such that X is regularly varying at infinity with index $\alpha > 0$ and there exists $\epsilon > 0$ such that $\mathbb{E}[Y^{\alpha+\epsilon}] < \infty$. Then XY is regularly varying with index α and*

$$\lim_{x\to\infty} \frac{\mathbb{P}(XY > x)}{\mathbb{P}(X > x)} = \mathbb{E}[Y^\alpha] . \tag{2.16}$$

Proof. If $y \le 1$, then obviously $\mathbb{P}(yX > x) \le \mathbb{P}(X > x)$. If $y \ge 1$, then for any $\epsilon > 0$, Markov's inequality yields

$$\begin{aligned}
\frac{\mathbb{P}(yX > x)}{\mathbb{P}(X > x)} &= \frac{\mathbb{P}(yX > x\,,\ X > x)}{\mathbb{P}(X > x)} + \frac{\mathbb{P}(yX > x\,,\ X \le x)}{\mathbb{P}(X > x)} \\
&\le 1 + \frac{\mathbb{P}(X\mathbb{1}_{\{X\le x\}} > (x/y))}{\mathbb{P}(X > x)} \le 1 + \frac{\mathbb{E}[X^{\alpha+\epsilon}\mathbb{1}_{\{X\le x\}}]}{(x/y)^{\alpha+\epsilon}\mathbb{P}(X > x)} \\
&\le 1 + C\frac{x^{\alpha+\epsilon}\mathbb{P}(X > x)}{(x/y)^{\alpha+\epsilon}\mathbb{P}(X > x)} = 1 + Cy^{\alpha+\epsilon} .
\end{aligned}$$

where the last bound is obtained by applying Proposition 2.1. Thus we have proved that there exists a constant C such that for any $x, y > 0$,

$$\frac{\mathbb{P}(yX > x)}{\mathbb{P}(X > x)} \le C(1 \vee y)^{\alpha+\epsilon} . \tag{2.17}$$

Define the function $G_x(y) = \mathbb{P}(yX > x)/\mathbb{P}(X > x)$. By definition of regular variation, the sequence of functions G_x converges pointwise to the function $y \to y^\alpha$ as $x \to \infty$. If $\mathbb{E}[Y^{\alpha+\epsilon}] < \infty$, the bound (2.17) allows to apply the bounded convergence theorem, which yields (2.16). □

Remark 2.18. We have proved slightly more than Breiman's Lemma, namely that the sequence of functions G_x converges in $L^p(\nu)$ for any $p < 1 + \epsilon/\alpha$ and where ν is the distribution of Y.

If X and Y have the same distribution, then Breiman's Lemma does not apply. If for instance $\mathbb{P}(X > x) = \mathbb{P}(Y > y) = x^{-\alpha}\mathbb{1}_{\{x \geq 1\}}$, then it is easily seen that $\mathbb{P}(XY > x) \sim x^{-\alpha}\log(x)$. More generally, we have the following result. Cf. Jessen and Mikosch (2006).

Proposition 2.19. *Let X and Y be independent nonnegative random variables and have the same distribution F such that $1 - F$ is regularly varying at infinity. Then XY is regularly varying and moreover:*

(i) if $\mathbb{E}[X^\alpha] = \infty$, then $\lim_{x\to\infty} \mathbb{P}(XY > x)/\mathbb{P}(X > x) = \infty$;

(ii) if $\mathbb{E}[X^\alpha] < \infty$ and if the limit $\lim_{x\to\infty} \mathbb{P}(XY > x)/\mathbb{P}(X > x)$ exists, then it is equal to $2\mathbb{E}[X^\alpha]$.

These results have been extended in many ways, including for subexponential distributions, cf. Embrechts and Goldie (1980), Cline and Samorodnitsky (1994); and recently for dependent X and Y, where the situation is much more involved, cf. Maulik and Resnick (2004).

2.4 Stochastic recurrence equations

Random variable that are solutions of certain kind of stochastic recurrence equations typically have regularly varying tails. The seminal paper of Kesten (1973) established a general method to prove existence and regular variation of solutions of stochastic recurrence equation. In the one-dimensional case, Goldie (1991) gave a method to prove regular variation and to have an exact equivalent of the tail of stochastic recurrence equations. This results have had many applications, most notably to the study of the famous ARCH and GARCH processes. Cf.Bougerol and Picard (1992), Basrak et al. (2002). Since we restrict to the one dimensional case, we will only state the main result of Goldie (1991).

Let Ψ be a random real valued application defined on $\mathbb{R}$, i.e. a measurable family of applications $t \to \Psi(\omega, t)$ and ω will be omitted in the notation as usual. Consider the stochastic recurrence equation

$$X_t = \Psi(X_{t-1}) \ . \tag{2.18}$$

Theorem 2.20 (Goldie (1991), Corollary 2.4)**.** *Assume that there exists a stationary solution $\{X_t\}$ to equation (2.18) with marginal distribution F, and let R be a random variable with distribution F, independent of Ψ. Assume that there exists a nonnegative random variable M, independent of R and a real number $\kappa > 0$ such that*

$$\mathbb{E}[M^\kappa] = 1 \ , \quad \mathbb{E}[M^\kappa \log_+(M)] < \infty \ . \tag{2.19}$$

Then $m := \mathbb{E}[M^\kappa \log(M)] \in (0, \infty)$ and if

$$\mathbb{E}[|\Psi(R)_+^\kappa - (MR)_+^\kappa|] < \infty \ , \tag{2.20}$$

then

$$\lim_{t\to\infty} t^\kappa(1 - F(t)) = \frac{1}{\kappa m}\mathbb{E}[\Psi(R)_+^\kappa - (MR)_+^\kappa] \ . \tag{2.21}$$

and $1 - F$ is regularly varying at infinity with index κ if the right hand side is nonzero. If

$$\mathbb{E}[|\Psi(R)_-^\kappa - (MR)_-^\kappa|] < \infty \ , \tag{2.22}$$

then

$$\lim_{t\to\infty} t^\kappa F(-t) = \frac{1}{\kappa m}\mathbb{E}[\Psi(R)_-^\kappa - (MR)_-^\kappa] \ . \tag{2.23}$$

and $t \to F(-t)$ is regularly varying at infinity with index κ if the right hand side is nonzero.

The existence of the stationary solution is assumed in this result. It can usually be proved by applying some result of Kesten's. The result is meaningful only if the constants wich appear in (2.21) and (2.23) are nonzero. This happens in particular if $\mathbb{E}[|R|^\kappa < \infty]$, so regular variation of the tails of R is proved only if $\mathbb{E}[|R|^\kappa] = \infty$. .

The random variable M that appear in the statement of this theorem may seem mysterious, but it is in each examples what is its relation to the random map Ψ.

Application to the GARCH(1,1) process

Let a, α, β be positive real numbers and $\{\epsilon_t\}$ be an i.i.d. sequence with mean zero and unit variance. A GARCH(1,1) process is the stationary solution of the equation

$$X_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = a + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 \; .$$

It is straightforward to prove that there exists a stationary solution as soon as $\alpha + \beta < 1$, which can be expressed as $X_t = \sigma_t \epsilon_t$ and

$$\sigma_t^2 = a + a \sum_{j=1}^{\infty} M_{t-1} \cdots M_{t-j} \; ,$$

with $M_j = \alpha \epsilon_j^2 + \beta$. We can apply Theorem 2.20 by defining $\Psi(R) = a + MR$ and $M = \alpha \epsilon_1^2 + \beta$. The stationary distribution of σ_t^2 is the solution of the equation $R \stackrel{\text{law}}{=} a + MR$. Assume that there exists $\kappa > 0$ such that $\mathbb{E}[(\alpha\epsilon_1^2 + \beta)^\kappa] = 1$ and $\mathbb{E}[(\alpha\epsilon_1^2 + \beta)^\kappa \log_+(\alpha\epsilon_1^2 + \beta)] < \infty$. Since a, M and R are non negative, $|(a + MR)^\kappa - (MR)^\kappa|$ is bounded, hence (2.20) holds for any value of κ. Since $\mathbb{E}[\epsilon_1^2] = 1$, the value of κ such that $\mathbb{E}[M^\kappa] = 1$ is necessarily greater than 1. Thus,

$$\mathbb{E}[R^\kappa] = \mathbb{E}\left[\left(\sum_{j=1}^{\infty} M_1 \cdots M_j\right)^\kappa\right] \geq \sum_{j=1}^{\infty} (\mathbb{E}[M_1^\kappa])^j = \infty \; .$$

Thus we can conlude that the stationary solution σ_t^2 of the GARCH(1,1) process is regularly varying with index κ.

Chapter 3

Applications to limit theorems

In this chapter, we show how regular variation is linked to the main limit theorems of probability theory. Regular variation characterizes domains of attraction, normalizations and limit distributions. One of the main tool to relate regular variation to limit theorems is the convergence to type theorem. We use it to characterize the domain of attraction of the maximum. We then characterize the domain of attraction of the stable laws. Here again we will rely on the results of Chapter 1. In Section 3.4 we introduce the notion of self-similarity and give examples of self-similar processes. Lamperti's theorem reveals that regular variation is present in most limit theorems, at least through the normalizing sequences. We conlude this chapter by the link between regular variation and point processes, and briefly illustrate their use in limit theorems in extreme value theory.

3.1 The convergence to type theorem

Theorem 3.1 (Convergence to type)**.** *Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of probability distribution functions. Let $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$, $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ be sequences of real numbers such that $a_n \geq 0$ and $\alpha_n > 0$ for all n. Let U and V be two probability distribution functions, each taking* *at least three values, such that*

$$\lim_{n\to\infty} F_n(a_n x + b_n) = U(x), \quad \lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = V(x) ,$$

at all continuity point x of U and V. Then, there exist $A > 0$ and $B \in \mathbb{R}$ such that

$$\lim_{n\to\infty} a_n/\alpha_n = A, \quad \lim_{n\to\infty} (b_n - \beta_n)/\alpha_n = B , \tag{3.1}$$

$$U(x) = V(Ax + B) . \tag{3.2}$$

The sequences $\{a_n\}$ et $\{\alpha_n\}$ are said to have same type or to be equivalent up to type. This result can be expressed in terms of random variables: If $(X_n - b_n)/a_n$ converges weakly to X and $(X_n - \beta_n)/\alpha_n$ converges weakly to Y then (3.1) holds and $Y \stackrel{d}{=} AX + B$

Proof. It is easily seen that (3.1) implies (3.2). Indeed, (3.1) implies that $a_n > 0$ for n large enough and

$$\frac{X_n - \beta_n}{\alpha_n} = \frac{a_n}{\alpha_n} \frac{X_n - b_n}{a_n} + \frac{b_n - \beta_n}{\alpha_n} .$$

By assumption, both sequences $(X_n - \beta_n)/\alpha_n$ and $(X_n - b_n)/a_n$ converge weakly, $a_n/\alpha_n \to A$ and $(b_n - \beta_n)/\alpha_n \to B$, hence (3.2) holds at all continuity points. Let us now prove (3.1).

Let $F_n^{\leftarrow}$, $U^{\leftarrow}$ and $V^{\leftarrow}$ denote the left-continuous inverses of the functions F_n, U and V, respectively. By Proposition A.4, $\{F_n^{\leftarrow} - b_n\}/a_n$ converges to $U^{\leftarrow}$ at all continuity point of $U^{\leftarrow}$ and $\{F_n^{\leftarrow} - \beta_n\}/\alpha_n$ converges to $V^{\leftarrow}$ at all continuity point of $V^{\leftarrow}$. Both distribution functions U and V are non degenerate, so there exists at least two common continuity points of $U^{\leftarrow}$ and $V^{\leftarrow}$, say y_1 and y_2, such that $U^{\leftarrow}(y_2) > U^{\leftarrow}(y_2)$ and $V^{\leftarrow}(y_2) > V^{\leftarrow}(y_2)$. Then, for $i = 1, 2$,

$$\lim_{n\to\infty} \frac{F_n^{\leftarrow}(y_i) - b_n}{a_n} = U^{\leftarrow}(y_i) ,$$

$$\lim_{n\to\infty} \frac{F_n^{\leftarrow}(y_i) - \beta_n}{\alpha_n} = V^{\leftarrow}(y_i) ,$$

whence

$$\lim_{n\to\infty} \frac{\alpha_n V^{\leftarrow}(y_i) + \beta_n}{a_n U^{\leftarrow}(y_i) + b_n} = 1 ,$$

for $i = 1, 2$. Solving these equations for α_n and β_n, we get

$$\lim_{n\to\infty} \frac{\alpha_n}{a_n} = \frac{U^{\leftarrow}(y_2) - U^{\leftarrow}(y_1)}{V^{\leftarrow}(y_2) - V^{\leftarrow}(y_1)} ,$$
$$\lim_{n\to\infty} \frac{\beta_n - b_n}{a_n} = \frac{U^{\leftarrow}(y_1)V^{\leftarrow}(y_2) - U^{\leftarrow}(y_2)V^{\leftarrow}(y_1)}{V^{\leftarrow}(y_2) - V^{\leftarrow}(y_1)} .$$

This proves (3.1) with

$$A = \frac{U^{\leftarrow}(y_2) - U^{\leftarrow}(y_1)}{V^{\leftarrow}(y_2) - V^{\leftarrow}(y_1)} , \quad B = \frac{U^{\leftarrow}(y_1)V^{\leftarrow}(y_2) - U^{\leftarrow}(y_2)V^{\leftarrow}(y_1)}{V^{\leftarrow}(y_2) - V^{\leftarrow}(y_1)} .$$

□

3.2 Convergence to extreme value distributions

Definition 3.2 (Extreme value distributions). *The extreme value distributions are the Gumbel, Fréchet and Weibull distribution, defined by*

- *Gumbel:* $G_0(x) = \mathrm{e}^{-\mathrm{e}^{-x}}$, $x \in \mathbb{R}$;
- *Fréchet:* $G_\gamma(x) = \mathrm{e}^{-x^{-1/\gamma}}$, $x > 0$, $\gamma > 0$;
- *Weibull:* $G_\gamma(x) = \mathrm{e}^{-|x|^{-1/\gamma}}$, $x < 0$, $\gamma < 0$.

Theorem 3.3. *Let F and G be non degenerate distribution functions, $\{a_n\}$ and $\{b_n\}$ be sequences of real numbers such that $\lim_{n\to\infty} F^n(a_n x + b_n) = G(x)$ at all continuity points of G. Then G is one of the extreme value distributions, up to a change in location and scale.*

Proof. For any $t \geq 0$ and continuity point x of G, it holds that

$$\lim_{n\to\infty} F^n(a_{[nt]}x + b_{[nt]}) = G^{1/t}(x) .$$

Theorem 3.1 implies that there exist A_t and B_t such that

$$\frac{a_{[nt]}}{a_n} = A_t$$
$$G(A_t x + B_t) = G^{1/t}(x) .$$

This implies that the function $t \to a_{[t]}$ is regularly varying at infinity with index $\gamma \in \mathbb{R}$, and $A_t = t^\gamma$. For $s, t > 0$, we obtain similarly

$$G^{1/st}(x) = G(A_{st}x + B_{st}) = G^{1/t}(A_s x + B_s) = G(A_t(A_s x + B_s) + B_t) \ ,$$

whence

$$B_{st} = t^\gamma B_s + B_t \ .$$

We have seen in the proof of Proposition 1.48 that this implies

$$B_t = c \int_1^t s^{\gamma-1} \, \mathrm{d}s \ ,$$

for some constant c.

We must now identify the limit distribution G in the three cases $\gamma = 0$, $\gamma > 0$ and $\gamma < 0$.

If $\gamma = 0$, then

$$G^{1/t}(x) = G(x + c\log(t)) \ . \tag{3.3}$$

Since G is assumed to be non degenerate, this implies that $c \neq 0$. This also implies that G has infinite right and left endpoints. Indeed, assume otherwise that there exists $x_0 < \infty$ such that $G(x_0) = 1$. The relation (3.3) would then imply

$$G(x_0 + c\log(t)) = 1 \ ,$$

for all $t > 0$, which contradicts the non degeneracy of G. Similarly, $G(x) > 0$ for all $x > -\infty$. Denote $u = -c\log(t)$ and $G(0) = \mathrm{e}^{-\mathrm{e}^{-p}}$, then

$$G(u) = \mathrm{e}^{-\mathrm{e}^{-u/c-p}} \ .$$

Thus G is the Gumbel law, up to a change in location and scale.

If $\gamma > 0$, then

$$G^{1/t}(x) = G(t^\gamma x + c(1 - t^\gamma)) = G(t^\gamma(x - c) + c) \ .$$

Denote $H(x) = G(x + c)$. The function H satisfies $H^{1/t}(x) = H(t^\gamma x)$. Thus $H^t(0) = H(0)$ for all $t > 0$, which implies $H(0) = 0$ or $H(0) = 1$. If $H(0) = 1$, then there exists $x < 0$ such that $0 < H(x) < 1$. This in turn implies that the function $t \to H^t(x)$ is non increasing, and that the function $H(t^{-\gamma}x)$ is non decreasing. This is a contradiction, thus $H(0) = 0$. It can be shown similarly that $0 < H(1) < 1$, and defining $H(1) = \mathrm{e}^{-p}$, $p > 0$ and $u = t^{-\gamma}$, we obtain

$$H(u) = \mathrm{e}^{-pu^{-1/\gamma}} .$$

This concludes the proof in the case $\gamma > 0$. The case $\gamma < 0$ is similar and omitted. □

The possible limit laws for renormalized maxima of i.i.d. samples are thus known. The remaining question is to find sufficient (and necessary if possible) condition for convergence to these distributions. We introduce a definition.

Definition 3.4. *A probability distribution function F is in the domain of attraction of an extreme value distribution G_γ, denoted $F \in MDA(G_\gamma)$, if there exist sequences of real numbers $\{a_n\}$ and $\{b_n\}$ such that*

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\gamma(x) . \tag{3.4}$$

It turns out that the proof of Theorem 3.3 contains as a by-product a necessary and sufficient condition for a distribution function F to be the domain of attraction of the Fréchet law G_γ ($\gamma > 0$).

Theorem 3.5. *A distribution function F is in the domain of attraction of the Fréchet law G_γ ($\gamma > 0$) if and only if it has a regularly varying upper tail with index $1/\gamma$.*

Proof. It has been established in the previous proof that if (3.4) holds, then the sequence a_n is regularly varying with index γ. This implies that $a_n \to \infty$ and $b_n \to \infty$ and thus $F(a_n x + b_n) \to 0$. Thus, (3.4) is equivalent to

$$n\{1 - F(a_n x + b_n)\} \to -\log G_\gamma(x) . \tag{3.5}$$

The function $1 - F(t)$ is positive and monotone, so we can apply Theorem 1.5 and conclude that $1 - F$ is regularly varying. □

The case $\gamma < 0$ can be deduced from the case $\gamma > 0$ by the transformation $X \to 1/(x^* - X)$ where x^* is the right-end point of the distribution of X, which is necessarily finite for distributions in the domain of attraction of the Weibull law. The case $\gamma = 0$ is more involved, and will not be discussed here.

3.3 Convergence to stable laws

In this section we characterize the stable laws and their domain of attraction. We want to find necessary and sufficient conditions on distribution functions F and to characterize the non degenerate distributions functions G such that there exist sequences $\{a_n\}$ and $\{b_n\}$ for which the limit

$$\lim_{n\to\infty} F^{*n}(a_n x + nb_n) = G(x) \tag{3.6}$$

holds at all continuity points of G. The function F^{*n} is the n-th convolution power of F, that is, the distribution of the sum of n i.i.d. random variables with distribution F. In other words, we want to find conditions on the common distibution of i.i.d. random variables X_n, $n \geq 1$ such that

$$a_n^{-1}(X_1 + \cdots + X_n - nb_n) \tag{3.7}$$

converges weakly to G. If ψ and ϕ denote the characteristic functions of F and G, respectively, then (3.6) and (3.7) are equivalent to

$$\lim_{n\to\infty} \mathrm{e}^{-\mathrm{i}tnb_n/a_n}\psi^n(t/a_n) = \phi(t)\ . \tag{3.8}$$

We now use the same arguments as in the proof of Theorem 3.3 to prove that $\phi(t) \neq 0$ for all t and that the sequence a_n is regularly varying. Taking moduli on both sides of (3.8), we have

$$\lim_{n\to\infty} |\psi^n(t/a_n)| = |\phi(t)|\ .$$

Thus, for any $r > 0$,

$$\lim_{n\to\infty} |\psi^n(t/a_{[nr]})| = |\phi(t)|^{1/r}\ .$$

Theorem 3.1 yields that the sequence a_n is regularly varying with index $\gamma \in \mathbb{R}$ and

$$|\phi(t)|^r = |\phi(r^\gamma t)|\ .$$

Since ϕ is non degenerate, there exists t_0 such that $|\phi(t_0)| < 1$, thus $\lim_{r\to\infty} |\phi(t_0)|^r = 0$, and this implies that $\gamma > 0$. Now, if there exists

t_0 such that $\phi(t_0) = 0$, then $\phi(t) = 0$ for all t, which is a contradiction. Thus $0 < |\phi(t)| < 1$ for all $t \neq 0$. Since convergence of characteristic functions is locally uniform, we obtain that $|\psi(t)| < 1$ for all $t \neq 0$ in a neighborhood of zero, and thus we can take logs in both of (3.8) to obtain:

$$\lim_{n\to\infty} n\{\log\psi(t/a_n) - \mathrm{i}tb_n/a_n\} = \log\phi(t) .$$

This is equivalent to

$$\lim_{n\to\infty} n\{\psi(t/a_n) - 1 - \mathrm{i}tb_n/a_n\} = \log\phi(t) .$$

Taking real parts, we obtain

$$\lim_{n\to\infty} n\{\mathrm{Re}(\psi(t/a_n)) - 1\} = \mathrm{Re}(\log\phi(t)) .$$

Since $\log(\phi)$ is continuous and $\phi(0) = 1$, $\mathrm{Re}(\log\phi(t))$ is strictly negative for $t \neq 0$ small enough. Moreover a_n is regularly varying with positive index, so we can apply Theorem 1.5 to obtain that $\mathrm{Re}(\psi) - 1$ is regularly varying at 0 and $\mathrm{Re}(\log\phi(t)) = -c|t|^{1/\gamma} = -c|t|^\alpha$ with $\alpha = 1/\gamma$ and some constant $c > 0$. Note that if ϕ is the characteristic function of a random variable X, then $2\mathrm{Re}(\log\phi)$ is the characteristic function of the random variable $X - X'$ where X' is independent of X and has the same distribution. Thus $\mathrm{e}^{-c|t|^\alpha}$ must be a characteristic function, which implies $\alpha \leq 2$ (see Remark 3.7 below). Thus, by Theorem 1.44 the function $x \to 1 - F(x) + F(-x)$ is regularly varying with index α if $\alpha \in (0, 2)$ and if $\alpha = 2$, the function

$$x \to \int_{-x}^{x} t^2 F(\mathrm{d}t)$$

is slowly varying at infinity. If $\alpha < 2$, we must still prove that both tails of F are regularly varying with index $-\alpha$ or that one tail is regularly varying with index α and the other tail is lighter. We follow Feller (1971, Section XVII.2). Let F_n be the Radon measure defined by

$$F_n(\mathrm{d}x) = nF(a_n\mathrm{d}x) .$$

Define ψ_n and ψ by

$$\psi_n(t) = n\{\phi(t/a_n) - 1 - \mathrm{i}tb_n/a_n\} , \quad \psi(t) = \log\phi(t) .$$

With this notation, we have

$$\psi_n(t) = \int_{-\infty}^{\infty} \{\mathrm{e}^{\mathrm{i}tx} - 1\} F_n(\mathrm{d}x) - \mathrm{i}tnb_n/a_n \; .$$

Since ψ_n converges to ψ, we also have $\mathrm{Im}(\psi_n(1))$ converges to $\mathrm{Im}(\psi(1))$, i.e.

$$\int_{-\infty}^{\infty} \sin(x) F_n(\mathrm{d}x) - nb_n/a_n \to \psi(1) \; .$$

Hence we can choose $b_n = n^{-1} a_n \int_{-\infty}^{\infty} \sin(x) F_n(\mathrm{d}x)$, and the convergence (3.6) will still hold. We make this choice, and this implies that ψ_n can now be expressed as

$$\psi_n(t) = \int_{-\infty}^{\infty} \{\mathrm{e}^{\mathrm{i}tx} - 1 - \mathrm{i}t\sin(x)\} F_n(\mathrm{d}x) \; ,$$

and $\psi_n(1)$ and $\psi(1)$ are real numbers. Fix some $h > 0$ and define k and ψ_n^* by

$$k(x) = 1 - \frac{\sin(hx)}{hx} \; , \quad \psi_n^*(t) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}tx} k(x) F_n(\mathrm{d}x) \; .$$

The function k is positive, $k(0) = 0$ and and $\lim_{x\to\infty} k(x) = 1$. With this notation, we have

$$\psi_n^*(t) = \psi_n(t) - \frac{1}{2h} \int_{-h}^{h} \psi_n(t+s)\,\mathrm{d}s + \mathrm{i}\frac{h}{2}\beta_n$$

with $\beta_n = \int_{-\infty}^{\infty} \sin(x) F_n(\mathrm{d}x)$. Note that β_n is a real number. Since convergence of characteristic functions is uniform on compact sets, we get that $\psi_n^* - \mathrm{i}\beta_n h/2$ converges to the function ψ^* defined by

$$\psi^*(t) = \psi(t) - \frac{1}{2h} \int_{-h}^{h} \psi(t+s)\,\mathrm{d}s \; .$$

In particular, $\psi_n^*(0) - \mathrm{i}\beta_n h/2$ converges to $\psi^*(0)$. But $\psi_n^*(0)$ is a sequence of real numbers, so this implies that $h\beta_n/2$ converges to $-\mathrm{Im}(\psi^*(0))$ and

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} k(x) F_n(\mathrm{d}x) = \mathrm{Re}(\psi^*(0)) = -\frac{1}{2h} \int_{-h}^{h} \mathrm{Re}(\psi(s))\,\mathrm{d}s \; .$$

Recall that $\mathrm{Re}(\psi(t)) = -c|t|^\alpha$ and $\mathrm{Re}(\psi^*(0)) > 0$. Denote $\mu_n = \int_{-\infty}^{\infty} k(x) F_n(\mathrm{d}x)$ and let G_n be the probability distribution function defined by

$$G_n(x) = \mu_n^{-1} \int_{-\infty}^{x} k(y) F_n(\mathrm{d}y) ,$$

i.e. the distribution function whose characteristic function is $\mu_n^{-1}\psi_n^*$. The sequence $\mu_n^{-1}\psi_n^*$ converges to the function $\check{\psi}$ defined by

$$\check{\psi}(t) = \frac{\psi^*(t) - \mathrm{Im}(\psi^*(0))}{\mathrm{Re}(\psi^*(0))}$$

which is continuous since ψ is continuous and $\check{\psi}(0) = 1$. Thus $\check{\psi}$ is the distribution function of a probability distribution, say G, and G_n converges weakly to G. Since k is bounded away from zero on all intervals $[x, \infty)$ with $x > 0$, we have, for all but a denumerable number of $x > 0$,

$$\lim_{n\to\infty} n\bar{F}(a_n x) = \lim_{n\to\infty} \int_x^\infty k^{-1}(t)\, G_n(\mathrm{d}t) = \int_x^\infty k^{-1}(t)\, G(\mathrm{d}t) ,$$
$$\lim_{n\to\infty} nF(-a_n x) = \lim_{n\to\infty} \int_{-\infty}^{-x} k^{-1}(t)\, G_n(\mathrm{d}t) = \int_{-\infty}^{-x} k^{-1}(t)\, G(\mathrm{d}t) .$$

Applying Theorem 1.5, we obtain that both the right and left tails of F are either regularly varying with index $-\alpha$ (since we already know that the sequence a_n is regularly varying with index $1/\alpha$), or $n\bar{F}(a_n x) \to 0$ or $nF(-a_n x) \to 0$ for all x. Both limits cannot be simultaneously vanishing, since we already know that the sum tail is regularly varying with index α. We have thus proved the following Theorem.

Theorem 3.6. *Let F be a probability distribution function such that (3.6) holds for some sequences a_n and b_n. Then either $1-F$ is regularly varying at infinity with index $-\alpha$, $\alpha \in (0, 2)$ and*

$$\lim_{x\to\infty} \frac{1 - F(x)}{1 - F(x) + F(-x)} = p \in [0, 1] , \tag{3.9}$$

or the function $x \to \int_{-x}^{x} x^2 F(\mathrm{d}x)$ is slowly varying.

Remark 3.7. If $\int_{-\infty}^{\infty} x^2 F(\mathrm{d}x) < \infty$, then the limiting distribution is of course the standard Gaussian distribution, because the characteristic function of F is twice differentiable, and the central limit theorem is then obtained by a second order Taylor expansion. If $\alpha > 2$ and the function $t \to \mathrm{e}^{-c|t|^\alpha}$ were a characteristic function, the corresponding distribution would have finite second moment, which would be equal to zero. This is a contradiction.

We now prove that the conditions of the previous results are sufficient for (3.6). In passing, we characterize the non Gaussian stable laws by their characteristic functions.

Theorem 3.8. *Let F be a distribution function such that (3.9) holds and the function $x \to 1 - F(x) + F(-x)$ is regularly varying at infinity with index $\alpha \in (0, 2)$. Let $\{a_n\}$ be a sequence of real numbers such that*

$$\lim_{n\to\infty} n\{1 - F(a_n) + F(-a_n)\} = 1 \; .$$

Let $\{X_n\}$ be a sequence of i.i.d. random variables with distribution F. Then there exists a sequence of real numbers $\{b_n\}$ such that

$$\frac{1}{a_n} \sum_{i=1}^{n} (X_i - b_n)$$

converges weakly to a stable law with characteristic function ϕ such that

$$\log \phi(t) = \begin{cases} -|t|^\alpha \Gamma(1-\alpha) \cos(\pi\alpha/2)\{1 - \mathrm{i}\,\mathrm{sgn}(t) \tan(\pi\alpha/2)\} \; , & \alpha \neq 1 \; , \\ -|t| \frac{\pi}{2} \{1 - \mathrm{i}\, \frac{2}{\pi} \,\mathrm{sgn}(t) \log(|t|)\} \; , & \alpha = 1 \; , \end{cases}$$

with $\beta = p - q$ and μ as in Corollary 1.43.

Remark 3.9. Condition (3.9) is called the tail balance condition. It was already used in Section 2.2.2. The coefficient β is called the skewness parameter. The distribution is symetric if $\beta = 0$.

Remark 3.10. The centering constants b_n can be chosen as follows:

- $b_n \equiv 0$ if $\alpha < 1$;
- $b_n = \mathbb{E}[X_1]$ if $\alpha > 1$;
- $b_n = \mathbb{E}[X_1 \mathbb{1}_{\{|X_1| \leq a_n\}}] + \beta a_n \mu / n$ if $\alpha = 1$ with μ as in Corollary 1.43.

Remark 3.11. Actually, the convergence can be strengthened to convergence of the sequence of processes $t \to a_n^{-1} \sum_{k=1}^{[nt]} (X_k - b_n)$ to a stable Lévy process in the space $\mathcal{D}[0,\infty)$ endowed with Skorohod's J_1 topology. See for instance (Resnick, 2007, Corollary 7.1)

Proof of Theorem 3.8. Let ν_n denote the measure defined on $(-\infty,0) \cup (0,\infty)$ by

$$\nu_n([x,\infty)) = n\{1 - F(a_n x)\} \ (x > 0) \ ,$$
$$\nu_n((-\infty,y]) = nF(a_n y) \ (y < 0) \ .$$

Then

$$n\{\psi(t/a_n) - 1\} - \mathrm{i}tnb_n/a_n = \int (\mathrm{e}^{\mathrm{i}tx} - 1)\, \nu_n(\mathrm{d}x) - \mathrm{i}tnb_n/a_n \ . \quad (3.10)$$

The proof now consists in applying Corollary 1.43 to (3.10) with a suitable choice of b_n. Indeed, with b_n as in Remark 3.10, we have

$$n\{\psi(t/a_n) - 1\} - \mathrm{i}tnb_n/a_n$$
$$= \begin{cases} \int (\mathrm{e}^{\mathrm{i}tx} - 1)\, \nu_n(\mathrm{d}x) & \alpha < 1 \ , \\ \int (\mathrm{e}^{\mathrm{i}tx} - 1 - \mathrm{i}tx\mathbb{1}_{\|x| \le 1\}})\, \nu_n(\mathrm{d}x) - \mathrm{i}t\beta\mu & \alpha = 1 \ , \\ \int (\mathrm{e}^{\mathrm{i}tx} - 1 - itx)\, \nu_n(\mathrm{d}x) & 1 < \alpha < 2 \ . \end{cases}$$

□

3.4 Self-similarity

A stochastic process $\{X_t \, , \ t \in \mathcal{T}\}$ is a collection of random variables indexed by a set $\mathcal{T}$. The distribution of a stochastic process is the collection of its finite dimensional marginal distributions, i.e. the distribution of all vectors $\{X_j \, , \ j \in J\}$ for all finite subset J of $\mathcal{T}$. If two stochastic processes X and Y have the same distribution, i.e. the same finite dimensional marginal distribution, we write $X \stackrel{\text{fi.di}}{=} Y$. The sequence of stochastic processes $\{X^{(u)}\}$, ($u \in \mathbb{N}$ or $u \in \mathbb{R}$) is said to converge in the sense of finite dimensional distributions to a process Y as $u \to \infty$, which we denote $X^{(u)} \stackrel{\text{fi.di}}{\to} Y$ if all the finite dimensional marginal distribution of $X^{(u)}$ converge weakly as $u \to \infty$ to the corresponding marginal distributions of Y. A stochastic process is said to be stochastically right-continuous if for all t, X_s converges weakly to X_t as s decreases to t.

Definition 3.12. *A stochastic process $\{X_t, t \in (0,\infty)\}$ is called self-similar if there exists functions a and b such that for each $s > 0$,*

$$\{X_{st}, t \in (0,\infty)\} \stackrel{\text{fi.di.}}{=} \{a(s)X_t + b(s), t \in (0,\infty)\}$$

If equality in law holds only for the one-dimensional marginal distributions, then the process is called marginally self-similar.

Theorem 3.13. *Let $\{X_t, t \in (0,\infty)$ be a stochastically right-continuous self-similar process such that X_1 is non-degenerate. Then there exists $\rho, b \in \mathbb{R}$ such that $\{X_{st}, t > 0\} \stackrel{\text{fi.di.}}{=} \{s^\rho X_t + b(s^\rho - 1)/\rho, t > 0\}$, (with the usual convention for $\rho = 0$).*

Proof. For $s, t > 0$, X_{st} has the same distribution as $a(st)X_1 + b(st)$ and $a(s)a(t)X_1 + a(s)b(t) + b(s)$. Since the distribution of X_1 is non degenerate, this implies that $a(st) = a(s)a(t)$ and $b(st) = a(s)b(t) + b(s)$. Since X is stochastically right-continuous, X_s converges weakly to X_t as s decreases to t. Since for all $s > 0$, X_s has the same distribution as $a(s)X_1 + b(s)$ and X_t has the same distribution as $a(t)X_1 + b(t)$, we obtain that a and b are right-continuous, hence measurable. Thus we can apply Corollary A.2 to obtain the form of a. The representation for b is obtained by the same arguments as in the proof of Theorem 3.3. □

It is well-known that the Brownian motion is $1/2$ self-similar and that Lévy α-stable processes are $1/\alpha$ self-similar. We now give other finite variance and infinite variance examples.

Example 3.14 (Fractional Brownian motion). The fractional Brownian motion B_H is the only Gaussian process wich is both self-similar with index H and with stationary increments. It is characterized by its autocovariance function.

$$\mathrm{cov}(B_H(s), B_H(t)) = \frac{1}{2}\{t^{2H} - |t-s|^{2H} + s^{2H}\} . \tag{3.11}$$

It can be represented as a Wiener stochastic integral with respect to a standard Brownian motion B.

$$B_H(t) = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{\infty} K_H(s,t) \, \mathrm{d}B(s) \; ,$$

with $K_H(s,t) = (t-s)_+^{H-1/2} - (-s)_+^{H-1/2}$. The process B_H is self-similar with index H which is called the Hurst index.

Example 3.15 (Hermite processes). Hermite processes will appear (in the next chapter) as limits of non linear functionals of Gaussian long memory processes. They are defined as multiple integrals with respect to the Brownian motion. See e.g. Embrechts and Maejima (2002) for a precise definition these multiple integrals. For $H > 1/2$, the m-th order Hermite process $Z_{m,H}$ is defined by

$$Z_{m,H}(t) = \int_{\mathbb{R}^m} \left\{ \int_0^t \prod_{j=1}^m (s-y_j)_+^{-(\frac{1}{2}+\frac{(1-H)}{m})} \, \mathrm{d}s \right\} \mathrm{d}B(y_1) \ldots \mathrm{d}B(y_m) \; .$$

The process $Z_{m,H}$ is self-similar with self-similarity index H.

Example 3.16 (Fractional stable processes). Let Λ_α be an α-stable Lévy process with $\alpha \in (1,2)$, i.e. a process with independent and stationary increments such that

$$\log \mathbb{E}[\mathrm{e}^{\mathrm{i}z\Lambda_\alpha(t)}] = t\sigma^\alpha |z|^\alpha \{1 - \mathrm{i}\beta \tan(\pi\alpha/2)\} \; ,$$

for some $\sigma > 0$ and $\beta \in [-1,1]$. For a deterministic function f such that $\int_{-\infty}^{\infty} |f(t)|^\alpha \mathrm{d}t$ the stochastic integral $\int_{-\infty}^{\infty} f(t) \mathrm{d}\Lambda_\alpha(t)$ can be defined. Cf. Samorodnitsky and Taqqu (1994). A self-similar stable process can be defined by analogy with the fractional Brownian motion as a stochastic integral with respect to Λ. Define the process $L_{H,\alpha}$ by

$$L_{H,\alpha}(t) = \int_{-\infty}^{\infty} K_{H,\alpha}(s,t) \mathrm{d}\Lambda_\alpha(s) \tag{3.12}$$

with $K_{H,\alpha}(s,t) = (t-s)_+^{H-1/\alpha} - (-s)_+^{H-1/\alpha}$ for $H \neq 1/\alpha$ and $K_{1/\alpha,\alpha}(x,t)$ $= \log(|t-s| - \log|s|)$. This process is well defined for $H \in (0,1)$ since $\int_{-\infty}^{\infty} |K_{H,\alpha}(s,t)|^\alpha \, \mathrm{d}s < \infty$. and is sometimes called linear fractional stable motion (LFSM).

The following result, which is essentially due to Lamperti (1962), shows that self-similar processes arise naturally in limit theorems.

Theorem 3.17 (Lamperti)**.** *Let* $\{X_t, t \in (0,\infty)\}$ *and* $\{Y_t, t \in (0,\infty)\}$ *be stochastic processes and* a *and* b *be functions such that*

$$\{a(s)X_{st} + b(s), t > 0\} \overset{\text{fi.di}}{\to} \{Y_t, t > 0\} \quad (s \to \infty) .$$

Then Y *is self-similar with index* ρ*. If the function* a *is measurable, then it is regularly varying with index* $-\rho$*.*

Proof. For any $t > 0$, we have both $a(st)X_{st} + b(st)$ converges weakly to Y_1 and $a(s)X_{st}+b(s)$ converges weakly to Y_t. By the convergence to type Theorem, this implies that there exists α_t and β_t such that $a(st)/a(t) \to \alpha_t$, $\{b(st)-b(s)\}/a(s) \to \beta_t$ and Y_t has the same distribution as $\alpha_t Y_1+\beta_t$. The argument holds for finite dimensional distributions so $\{Y_t\}$ is self-similar with index $\rho \in \mathbb{R}$ and a is regularly varying with index $-\rho$. □

Finally, we give a version for sequences of processes.

Theorem 3.18. *Let* $\{X_n\}$ *be a sequence of random variables,* $\{Y_t, t > 0\}$ *be a process, and* $\{a_n\}$*,* $\{b_n\}$ *be sequences of real numbers such that* $a_n X_{[nt]}+b_n$ *converges weakly to* Y_t *for all* $t > 0$*, and* Y_1 *is non degenerate. Then* Y *is self-similar with index* ρ *and* a_n *is regularly varing with index* $-\rho$*.*

These results illustrate the importance of regular variation and self-similar processes in limit results for processes: under mild restrictions, the normalization must be regularly varying and the limiting processes must be self-similar.

3.5 Point processes

Point processes are an extremely useful tool in extreme value theory. We briefly present their use to derive limit theorems in this section. The main reference is the classical monography Sidney (1987).

Definition 3.19 (Point process on $\mathbb{R}$, Laplace functional)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space and let* $\mathcal{B}$ *is the Borel sigma-field of* $\mathbb{R}$*. A point process on* $\mathbb{R}$ *is an application* N *on* $\Omega \times \mathcal{B}$ *such that* *(i) For allmost all $\omega \in \Omega$, $N(\omega, \cdot)$ is a σ-finite measure on $\mathbb{R}$.*

(ii) For each $A \in B$, the application $\omega \to N(\omega, A)$ is an integer valued random variable.

The Laplace functional of the point process N is the application $\mathcal{L}N$ defined on compactly supported continuous function by

$$\mathcal{L}N(f) = \mathbb{E}[\mathrm{e}^{N(f)}] .$$

Definition 3.20 (Poisson point process on the real line)**.** *A Poisson point process with mean measure μ on $\mathbb{R}$ is a random measure N defined on the Borel sets of $\mathbb{R}$ such that for each $k \geq 1$ and any pairwise disjoint Borel sets $A_1, \ldots, A_n$, $N(A_1), \ldots, N(A_k)$ are independent Poisson random variables with respective mean $\mu(A_1), \ldots, \mu(A_k)$.*

The Poisson point process is also characterized by its Laplace functional. If N is a Poisson point process with mean measure μ and f is a continuous compactly supported function on $\mathbb{R}$, then

$$\mathbb{E}[\mathrm{e}^{N(f)}] = \exp \int \{\mathrm{e}^{f(x)} - 1\} \, \mu(\mathrm{d}x) .$$

Point processes with value in $\mathbb{R}$ are random elements in the space of point measures on $\mathbb{R}$, endowed with the topology of vague convergence, which is metrizable. Hence weak convergence can be defined as usual. We only recall the following fundamental result: weak convergence of a sequence of point processes is equivalent to the pointwise convergence of the sequence of Laplace functionals. This means that the sequence of point processes N_n converge weakly to N if and only if for any compactly supported continuous function f,

$$\lim_{n\to\infty} \mathcal{L}N_n(f) = \mathcal{L}N(f) .$$

The result that links regular variation of random variables and point processes is due to (Resnick, 1986, Proposition 3.1, Corollary 3.2, Proposition 5.3). It states that regular variation of a probability distribution function is equivalent to the weak convergence of a sequence of point process to a Poisson point process or to the weak convergence of a sequence of random measures to a deterministic measure. For brevity, we state here a simplified version.

Theorem 3.21. *Let* $\{X_j\}$ *be sequence of i.i.d. random variables with distribution function* F*. The following are equivalent.*

(i) $1 - F$ *is regularly varying at infinity with index* $-\alpha < 0$*.*

(ii) There exist a Radon measure μ *on* $(0, \infty]$*, not concentrated at infinity, and a sequence of real numbers* $\{a_n\}$ *such that* $\lim_{n\to\infty} a_n = \infty$ *and the sequence of point processes* $\sum_{j=1}^n \delta_{X_j/a_n}$ *converges weakly to the Poisson point process with mean measure* μ*.*

(iii) There exists a Radon measure μ *on* $(0, \infty]$*, not concentrated at infinity, and a sequence of real numbers* $\{a_n\}$ *such that* $\lim_{n\to\infty} a_n = \infty$*, such that, for all sequences* k_n *satisfying* $k_n \to \infty$ *and* $n/k_n \to \infty$*, the sequence of random measures* $\frac{1}{k_n}\sum_{j=1}^n \delta_{X_j/a_{[n/k_n]}}$ *converges weakly to* μ*.*

Proof. Denote $N_n = \sum_{j=1}^n \delta_{X_j/a_n}$. We first compute the Laplace functional of N_n. Let f have compact support on $\mathbb{R}_+$. Let μ_n be the measure defined by $\mu_n(A) = n\mathbb{P}(X/a_n \in A)$. By independence and equidistribution of the X_is, we have

$$\log \mathcal{L}N_n(f) = n \log \mathbb{E}[\mathrm{e}^{f(X_1/a_n)}] = n \log\{1 + \frac{1}{n}\int (\mathrm{e}^f - 1)\,\mathrm{d}\mu_n\} .$$

By Lemma A.10, $1-F$ is regularly varying at infinity if and only if there exists a measure μ such that, for all continous function g with compact support in $(0, \infty)$,

$$\lim_{n\to\infty}\int_0^\infty g(x)\,\mu_n(\mathrm{d}x) = \int_0^\infty g(x)\,\mu(\mathrm{d}x)$$

and the measure μ is then necessarily the homogeneous measure μ_α on $(0, \infty)$ characterized by $\mu_\alpha(tA) = t^{-\alpha}\mu(A)$. Note that $n \log\{1 + \frac{1}{n}\int(\mathrm{e}^f - 1)\,\mathrm{d}\mu_n\}$ converges if and only if $\int(\mathrm{e}^f - 1)\,\mathrm{d}\mu_n$ has a limit. If f has compact support, then so has $\mathrm{e}^f - 1$, hence $\int(\mathrm{e}^f - 1)\,\mathrm{d}\mu_n$ converges to $\int(\mathrm{e}^f - 1)\,\mathrm{d}\mu$ for some measure μ if and only if $1 - F$ is regularly varying. This proves the equivalence between (i) and (ii). Now, (ii) implies (iii) by the weak law of large numbers. Assume that (iii) holds. Denote $\nu_n = \frac{1}{k_n}\sum_{j=1}^n \delta_{X_j/a_{[n/k_n]}}$. Let f be continuous and compactly supported on $(0, \infty)$. The Laplace functional of ν_n can be expressed as

$$\log \mathcal{L}\nu_n(f) = n \log\left\{1 + \frac{1}{n}\int f_n\,\mathrm{d}\tilde{\nu}_n\right\}$$

with $f_n = k_n(\mathrm{e}^{f/k_n} - 1)$ and $\tilde{\nu}_n(A) = \frac{n}{k_n}\mathbb{P}(X/a_{[n/k_n]} \in A)$. Since f is compactly supported, f_n converges uniformly to f. Thus if $\log \mathcal{L}\nu_n(f)$ converges, then $\int f \, \mathrm{d}\tilde{\nu}_n$ converges to the same limit. □

This result can be strengthened and extended in many ways. In the first place, it is valid not only for point process in $\mathbb{R}$ but also for any "nice" metric space (in particular finite dimensional euclidean spaces). Cf. (Sidney, 1987, Chapter 3). It can also be extended to weakly dependent points (Davis and Hsing, 1995). It is very useful to prove results in extreme value theory or convergence to stable laws. For the sake of illustration, we just give an elementary example.

Example 3.22 (Convergence of order statistics). Let $\{X_i\}$ be a sequence of i.i.d. random variables with distribution function F such that $1 - F$ is regularly varying with index $-\alpha < 0$. Let $X_{(n:1)}, \dots, X_{(n:n)}$ be the increasing order statistics and let N be the Poisson point process on $(0, \infty]$ with mean measure μ_α. Then, for a fixed $k \geq 1$, Theorem 3.21 straightforwardly yields that

$$\lim_{n\to\infty} \mathbb{P}(X_{(n:n-k+1)} \leq x) = \mathbb{P}(N((x, \infty)) < k) = \mathrm{e}^{-x^{-\alpha}} \sum_{j=0}^{k-1} \frac{x^{-j\alpha}}{j!} \ ,$$

For a sequence k (depending on n) such that $k \to \infty$ and $k/n \to 0$ (a so-called intermediate sequence), Theorem 3.21 (iii) yields the convergence of intermediate quantiles. It can be shown that $X_{(n:n-k+1)}/F^{\leftarrow}(1-k/n)$ converegees in probability to 1. Cf. (Resnick, 2007, Section 4.6).

Chapter 4

Long memory processes

Long memory or long range dependence are two equivalent concepts, but without a formal definition. These two denominations gather a wide range of properties of certain stochastic processes which separate them markedly from i.i.d. sequences in discrete time or processes with independent increments in continuous time. The best attempt to a definition is maybe given by Samorodnitsky (2006). A stochastic process may be said to have long memory if a phase transition phenomenon can be observed in the distribution (or asymptotic distribution) of a statistic of interest, according to the value of a parameter of the distribution of the process. More precisely, for certain value of this parameter, this statistic will have the same or a similar distribution as if the process were an i.i.d. sequence or had independent increments, and will drastically differ for the other value of the parameter. This parameter can be finite dimensional or real valued, as for instance the Hurst index of the fractional Brownian motion (cf. Example 3.14), or it can be a functional parameter, such as the autocovariance function of a Gaussian process.

This phenomenon has attracted a lot of attention and induced a huge literature during the last thirty years, due both to its theoretical interest and to potential applications in hydrology, teletrafic modeling and financial time series.

In this chapter, we will restrict the definition of long memory to a property of the autocovariance function of finite variance stationary processes, mainly in discrete time. This form of long memory is sometimes called second order long memory. For many other concepts of long memory, see Samorodnitsky (2006). We will give the main definitions in Section 4.1 and give examples in Section 4.2. Our main interest will be to relate the long memory property to the regular variation of some functional parameter. In Section 4.3 we will illustrate the phase transition phenomenon with examples of limit theorems for the partial sum processes.

4.1 Second order long memory

Let $\{X_k, k \in \mathbb{Z}\}$ be a weakly stationary process, i.e. such that

$$\mathbb{E}[X_k^2] < \infty \quad \text{for all } k \in \mathbb{Z} \, , \tag{4.1}$$

$$\text{the functions } \; k \to \mathbb{E}[X_k] \; , (k,n) \to \operatorname{cov}(X_k, X_{k+n}) \; \text{ do not depend on } k. \tag{4.2}$$

The function $\gamma : n \to \operatorname{cov}(X_0, X_n)$ is called the autocovariance function of the weakly stationary process $\{X_k\}$.

A (strict sense) stationary process is weakly stationary if (4.1) holds. Two weakly stationary processes may have the same autocovariance function, but different finite-dimensional distribution of any order. An i.i.d sequence with finite variance is strictly and weakly stationary. A weakly stationary sequence with zero mean, variance σ^2 and the same autocovariance function as an i.i.d. sequence is called a weak white noise.

Let γ be the autocovariance function of a weakly stationary process $\{X_k\}$. Denote $S_n = X_1 + \cdots + X_n$. Then

$$\operatorname{var}(S_n) = n\gamma(0) + 2\sum_{k=1}^{n} (n-k)\gamma(k) \; .$$

If

$$\sum_{k=1}^{\infty} |\gamma(k)| < \infty \; , \tag{4.3}$$

then

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{var}(S_n) = \gamma(0) + 2\sum_{k=1}^{\infty} \gamma(k) \; .$$

Most usual models of stationary processes, including of course i.i.d. sequences, satisfy (4.3). Such processes will be called *weakly dependent.* If (4.3) does not hold, then the process $\{X_k\}$ is called strongly dependent or is said to have long memory.

This definition of long memory is specific to weakly stationary processes. Other definitions are possible, and sometimes needed, e.g. for infinite variance processes. This definition is very weak, and has little distributional implications. In particular, it says nothing about the problem of weak convergence of the renormalized partial sum process

$$\mathrm{var}(S_n)^{-1/2} \sum_{k=1}^{[nt]} X_k \ . \tag{4.4}$$

If the sequence of processes in (4.4) converges (in the sense of finite dimensional distributions) to a non degenerate limit process, then by Lamperti's theorem, $\mathrm{var}(S_n)$ must be regularly varying. For historical reasons, the index of regular variation of $\mathrm{var}(S_n)$ is denoted $2H$ and H is called the Hurst index of the process $\{X_n\}$, which necessarily belongs to $[0,1]$.

Definition 4.1 (Hurst index)**.** *A weakly stationary* $\{X_k\}$ *has Hurst index* H *if* $\mathrm{var}(S_n)$ *is regularly varying with index* $2H$*. The process* $\{X_k\}$ *is called strongly dependent or is said to have long memory if* $H > 1/2$*; weakly dependent if* $H = 1/2$ *and anti-persitent if* $H < 1/2$*.*

Regular variation of $\mathrm{var}(S_n)$ does not imply regular variation of the autocovariance function. If there exist $H \in [0,1]$ and a slowly varying function L such that

$$\mathrm{var}(S_n) \sim L(n) n^{2H} \ , \tag{4.5}$$

then $\gamma(k)$ need not be regularly varying. Additional assumptions such as monotony are needed. Nevertheless, if the autocovariance function γ is assumed to be regularly varying, then (4.5) is implied by

$$\gamma(k) \sim 2H(2H-1)L(k)k^{2H-2} \ . \tag{4.6}$$

Definition 4.2 (Spectral density)**.** *A weakly stationary process* $\{X_k\}$ *with autocovariance function* γ *has a spectral density if there exists an* *even positive integrable function f on $[-\pi,\pi]$ such that*

$$\gamma(k) = \int_{-\pi}^{\pi} f(x)\, \mathrm{e}^{\mathrm{i}kx}\, \mathrm{d}x \; .$$

Stationary processes that admit a spectral density can be characterized. Cf. Brockwell and Davis (1991).

Theorem 4.3. *A stationary process $\{X_k\}$ admits a spectral density f if and only if there exist weak white noise $\{\epsilon_k\}$ and a sequence of real numbers $\{a_k\}$ such that $\sum_{k\in\mathbb{Z}} a_k^2 < \infty$ and*

$$X_k = \mu + \sum_{j\in\mathbb{Z}} a_j \epsilon_{k-j} \; . \tag{4.7}$$

The series in (4.7) is convergent in the mean square. If the series $\sum_{k\in\mathbb{Z}} |a_k|$ is convergent or if the white noise $\{\epsilon_k\}$ is an i.i.d. sequence then it is almost surely convergent. If (4.7) holds, then the spectral density f of the process is given by

$$f(x) = \frac{\sigma^2}{2\pi} \left| \sum_{\in\mathbb{Z}} a_j \mathrm{e}^{\mathrm{i}jx} \right| \; . \tag{4.8}$$

The autocovariance function of the process defined by (4.7) is given by

$$\gamma(n) = \sigma^2 \sum_{j\in\mathbb{Z}} a_j a_{j+n} \; ,$$

with $\sigma^2 = \mathrm{var}(\epsilon_1)$.

The Abelian and Tauberian Theorems of Chapter 1 link the behaviour at infinity of the coefficients of the linear represention, the behaviour at infinity of the autocovariance function, and the behaviour at zero of the spectral density.

- By Theorem 1.34 and 1.37, if ℓ is slowly varying at infinity and quasi-monotone and $f(x) = \ell(1/x)x^{\alpha-1}/(2\Gamma(\alpha)\cos(\pi\alpha/2))$, then $\gamma(k) \sim \ell(n)n^{-\alpha}$; conversely if $\gamma(k) = \ell(n)n^{-\alpha}$ then

$$f(x) \sim \ell(1/x)x^{\alpha-1}/(2\Gamma(\alpha)\cos(\pi\alpha/2)) \; .$$

- By Theorem 1.38, $f(x) = x^{\alpha-1}\ell(x)/(\Gamma(\alpha)\cos(\pi\alpha/2))$ with $0 < \alpha < 1$ and ℓ slowly varying at zero if and only if $\sum_{=1}^{n}\gamma(k) \sim n^{1-\alpha}\ell(1/n)/(1-\alpha)$. If the autocovariance is ultimately non increasing, then $\gamma(k) \sim \ell(n)n^{-\alpha}$. Similar results have been obtained by Inoue (1997) for time series whose spectral measure is not absolutely continuous.

- If the coefficents a_j are regularly varying at infinity with index $\gamma \in (1/2, 1)$, i.e.

$$a_j = \ell(j)j^{-\gamma} , \tag{4.9}$$

  then (4.8) and Theorem 1.38 imply that the spectral density is regularly varying at 0 with index $2\gamma - 2$. If moreover ℓ is quasi-monotone, then (4.6) holds with $H = 3/2 - \gamma$

We gather the previous results in the following table. If the coefficients of the linear representation and the autocovariance function are regularly varying at infinity and the spectral density is regularly varying at 0, then the following relations hold.

	Index of regular variation	Hurst index
linear filter	$-\gamma \in (1/2, 1)$ (at ∞)	$H = 3/2 - \gamma$
autocovariance	$-\beta \in (-1, 0)$ (at ∞)	$H = 1 - \beta/2$
spectral density	$-\alpha \in (-1, 0)$ (at 0)	$H = (1 + \alpha)/2$

4.2 Examples of long memory processes

4.2.1 Gaussian and linear processes

Since any weakly stationary process that admits a spectral density has a linear representation of the form (4.7) with respect to some weak white noise $\{\epsilon_k\}$, this representation (4.7) is not very informative unless more assumptions are made on the white noise. The denomination linear process usually means a stationary process $\{X_k\}$ that admits the representation (4.7) where $\{\epsilon_k\}$ is an i.i.d. sequence. If moreover $\{\epsilon_k\}$ is a sequence of i.i.d. standard Gaussian random variables, then $\{X_t\}$ is a Gaussian process and its distribution depends only on its spectral density.

Example 4.4 (Fractional differencing). For $d < 1/2$, $-d \notin \mathbb{Z}_+$, denote $\pi_j(d), j \geq$ the coefficients of the Taylor expansion of $(1-z)^{-d}$ for $|z| < 1$:

$$(1-z)^{-d} = 1 + \sum_{j=1}^{\infty} \frac{\Gamma(d+j)}{j!\,\Gamma(d)} z^j \ .$$

The coefficient $\pi_j(d)$ are absolutely summable for $d < 0$ and square summable for $a_j > 0$. For $d > 0$, $\pi_j(d) > 0$ for all d, and $\sum_{j=1}^{\infty} \pi_j(d) = \infty$. For $d < 0$, $d \notin \mathbb{Z}$, $\sum_{j=1}^{\infty} \pi_j(d) = 0$. The coefficients $\pi_j(d)$ decay polynomially fast at infinity:

$$\pi_j(d) \sim \Gamma(1-d) \sin(\pi d/2) j^{d-1} \ .$$

Let $\{a_j\}$ be a summable series and $\{\epsilon_t\}$ be a weak white noise. Let Y_t be the weakly stationary sequence defined by (4.7) with $\mu = 0$. Then one can define a weakly stationary process $\{X_t\}$ by

$$X_t = \sum_{j=0}^{\infty} \pi_j(d) Y_{t-j} \ .$$

Let f_Y be the spectral density of the process $\{Y_t\}$, given by (4.8). Then the spectral density f_X of the process $\{X_t\}$ is given by

$$f_X(x) = |1 - \mathrm{e}^{\mathrm{i}x}|^{-2d} f_Y(x) \ .$$

Since the coefficients $\{a_j\}$ are summable, f_Y is continuous. If $f_Y(0) > 0$, then

$$f_X(x) \sim x^{-2d} f_Y(0) \ , \quad (x \to 0) \ .$$

The autocovariance function γ_X of the process $\{X_k\}$ also decays slowly:

$$\gamma_X(k) \sim \Gamma(1-2d) \sin(\pi d) f_Y(0) k^{2d-1} \ .$$

The terminology fractional differencing comes from the formal representation

$$X = (I - B)^{-d} Y$$

where B is the backshift operator defined by $(BX)_t = X_{t-1}$, so that $(I-B)^{-d}$ is fractional differencing if $d < 0$ and fractional integration if $d > 0$.

A well-known example is the ARFIMA or FARIMA (where F stands for fractionally and I for integrated) process, where the process $\{Y_t\}$ is an ARMA process.

Example 4.5 (Fractional Brownian motion, fractional Gaussian noise). The fractional Browian motion $\{B_H(t), t \geq 0\}$ with Hurst index $H \in (0,1)$ was defined in Example 3.14. It is self-similar with index H and its covariance is given by

$$cov(B_H(s), B_H(t)) = \frac{1}{2}\{t^{2H} - |s-t|^{2H} + s^{2H}\} .$$

The fractional Gaussian noise $\{Z_H(t)\}$ is the only stationary Gaussian process with autocovariance function γ_H defined by

$$\mathrm{cov}(Z_H(0), Z_H(t)) = \frac{1}{2}\{|t+1|^{2H} - 2t^{2H} + |t-1|^{2H}\} . \tag{4.10}$$

The spectral density of the Fractional Gaussian noise (considered as a stationary process indexed by $\mathbb{Z}$) is given by an infinite series

$$f_H(x) = \frac{2\pi(1-\cos(x))}{H\Gamma(2H)\sin(\pi H)} \sum_{k\in\mathbb{Z}} |2k\pi + x|^{-2H-1} .$$

Thus f_H is regularly varying at zero with index $1-2H$

Subordination

Let X be a standard Gaussian random variable and let f be a function such that $\mathbb{E}[f^2(X)] < \infty$. Then f can be expanded as an infinite series

$$f = \sum_{j=0}^{\infty} \frac{c_m(f)}{m!} H_m , \tag{4.11}$$

where $H_m, m \geq 0$ are the Hermite polynomials which form an orthogonal family of the Hilbert space of functions which are square integrable under the standard Gaussian law. More precisely,

$$\mathbb{E}[H_m(X)H_{m'}(X)] = m!\mathbb{1}_{\{m=m'\}} ,$$

and the coefficients $c_m(f)$ are given by $c_m(f) = \mathbb{E}[f(X)H_m(X)]$. The Hermite rank of f is the smallest integer q such that $c_q \neq 0$.

An important property of the Hermite polynomials is the so-called Mehler formula. Let (X, Y) be a Gaussian vector with standard margins and correlation ρ. Then

$$\mathbb{E}[H_m(X)H_{m'}(Y)] = m!\mathbb{1}_{\{m=m'\}}\rho^m \; . \tag{4.12}$$

This formula and the expansion (4.11) yield the following covariance bound. Let f be a function such that $\mathbb{E}[f^2(X)] < \infty$ and let m be the Hermite rank of $f - \mathbb{E}[f(X)]$. Then

$$|\mathrm{cov}(f(X), f(Y))| \leq |\rho|^m \mathrm{var}(f(X)) \; . \tag{4.13}$$

To prove this inequality, we expand f as in (4.11) and use (4.12):

$$\mathrm{cov}(f(X), f(Y)) = \sum_{j=m}^{\infty} \frac{c_m^2(f)}{m!}\rho^j \leq \rho^m \sum_{j=m}^{\infty} \frac{c_m^2(f)}{m!} = \rho^m \mathrm{var}(f(X_1)) \; .$$

This also shows that the bound (4.12) is optimal as regards the correlation, since a lower bound also holds, if $\rho \geq 0$,

$$\mathrm{cov}(f(X), f(Y)) \geq \frac{c_m^2(f)}{m!}\rho^m$$

This implies that the dependence between functions of two jointly Gaussian random variables, measured by the correlation, is decreased by subordination, and this decrease can be exactly measured by the Hermite rank of the function. For a Gaussian process, this has the following consequence.

Proposition 4.6. *Let $\{X_j\}$ be a stationary Gaussian process with Hurst index H. Let f be a function with Hermite rank m. If $m(1 - H) < 1/2$, then the process $\{f(X_j)\}$ is long range dependent with Hurst index $1 - m(1 - H)$.*

Subordination of non Gaussian linear processes has also been widely investigated, but is a much more difficult matter.

4.2.2 Stochastic volatility processes

Long memory processes have been used to modelized some empirical properties of financial time series. The returns or log-returns of some financial time series have the property to be uncorrelated, while non linear functions (absolute value, squares, log-squares) are strongly correlated. Stochastic volatility processes have been introduced to model this feature. Let $\{Z_i\}$ be an i.i.d. centered sequence and $\{X_i\}$ be a stationary process. Let σ be a deterministic positive function. Define the process $\{Y_i\}$ by

$$Y_i = \sigma(X_i) Z_i \; . \tag{4.14}$$

If X_i is measurable with respect to the sigma field $\mathcal{F}^Z_{i-1}$ generated by $Z_j, j < i$, then $\sigma^2(X_i) = \mathrm{var}(Y_i \mid \mathcal{F}^Z_{i-1})$ is the conditional variance of Y_i given its past. This is the case of ARCH and GARCH processes which we do not consider here. Instead, we consider the case where the processes $\{X_i\}$ and $\{Z_i\}$ are independent. This model with $\sigma(x) = \mathrm{e}^{x/2}$ and $\{X_i\}$ is a Gaussian process is very popular and allows to model easily dependence and tails. Cf. Breidt et al. (1998), Harvey (1998). If the right tail of the Z_j is regularly varying with index $-\alpha$ and if $\mathbb{E}[\sigma^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$, then, by Breiman's lemma (Theorem 2.17),

$$\mathbb{P}(\sigma(X_1)Z_1 > y) \sim \mathbb{E}[\sigma^\alpha(X_1)]\mathbb{P}(Z_1 > y) \; .$$

The dependence of the process $\{Y_i\}$ lies in the process $\{X_i\}$ through subordination. The random variables Y_i are uncorrelated, but functions of the Y_i might be correlated. Assume that the process is Gaussian with Hurst index $H \in (1/2, 1)$ and that $\mathbb{E}[Z_1^2] < \infty$. We can then compute the covariance of the squared process. For $k > 1$, we have

$$\mathrm{cov}(Y_1^2, Y_k^2) = \mathbb{E}[Z_1^2]^2 \mathrm{cov}(\sigma^2(X_1), \sigma^2(X_k)) \; .$$

Thus if H is the Hurst index of the Gaussian process $\{X_j\}$ and m is the Hermite rank of σ^2, and if $m(1-H) < 1/2$, then Y_j^2 is long range dependent with Hurst index $1 - m(1-H)$.

4.2.3 Duration driven long memory processes

Let $\{t_j, \; j \in \mathbb{Z}\}$ be the points of a stationary point process, numbered for instance in such a way that $t_{-1} < 0 \le t_0$, and for $t \ge 0$. Let N be the counting measure of the points, i.e. $N(A) = \sum_{j\in\mathbb{Z}} \mathbb{1}_A(t_j)$. Stationarity of the point process means that the finite dimensional distributions of the random measure N are translation invariant. Let $N(t) = N(0,t] = \sum_{j\geq 0} \mathbb{1}_{\{t_j \leq t\}}$ be the number of points between time zero and t. Define then

$$X_t = \sum_{j\in\mathbb{Z}} W_j \mathbb{1}_{\{t_j \leq t < t_j + Y_j\}}, \quad t \geq 0. \tag{4.15}$$

In this model, the rewards $\{W_j\}$ are an i.i.d. sequence; they are generated at birth times $\{t_j\}$ and have durations $\{Y_j\}$. The observation at time t is the sum of all surviving present and past rewards. In model (4.15), we can take time to be continuous, $t \in \mathbb{R}$ or discrete, $t \in \mathbb{Z}$. We now describe several well known special cases of model (4.15).

(i) Renewal-reward process.
The durations are exactly the interarrival times of the renewal process: $Y_0 = t_0$, $Y_j = t_{j+1} - t_j$, and the rewards are independent of their birth times. Then there is exactly one surviving reward at time t:

$$X_t = W_{N(t)}. \tag{4.16}$$

This process was considered by Taqqu and Levy (1986) as an example of non linear long memory process and by Liu (2000) as a model for stochastic volatility.

(ii) ON-OFF model.
This process consists of alternating ON and OFF periods with independent durations. Let $\{Y_k\}_{\geq 1}$ and $\{Z_k\}_{k\geq 1}$ be two independent i.i.d. sequences of positive random variables with finite mean. Let t_0 be independent of these sequences and define $t_j = t_0 + \sum_{k=1}^{j}(Y_k + Z_k)$. The rewards W_j are deterministic and equal to 1. Their duration is Y_j. The Y_js are the ON periods and the Z_js are the OFF periods. The first interval t_0 can also be split into two successive ON and OFF periods Y_0 and Z_0. The process X can be expressed as

$$X_t = \mathbb{1}_{\{t_{N(t)} \leq t < t_{N(t)} + Y_{N(t)}\}}. \tag{4.17}$$

This process was considered by Taqqu et al. (1997) to modelize some features of internet trafic.

(iii) Error duration process.
This process was introduced by Parke (1999) to model some macroeconomic data. The birth times are deterministic, namely $t_j = j$, the durations $\{Y_j\}$ are i.i.d. with finite mean and

$$X_t = \sum_{j \leq t} W_j \mathbb{1}_{\{t < j + Y_j\}}. \tag{4.18}$$

This process can be considered in continous or discrete time, but the point process N is stationary only in discrete time.

(iv) Infinite Source Poisson model.
If the t_j are the points of a homogeneous Poisson process, the durations $\{Y_j\}$ are i.i.d. with finite mean and $W_j \equiv 1$, we obtain the infinite source Poisson model or M/G/∞ input model considered among others in Mikosch et al. (2002).

Maulik et al. (2002) have considered a variant of this process where the shocks (referred to as transmission rates in this context) are random, and possibly contemporaneously dependent with durations.

In the first two models, the durations satisfy $Y_j \leq t_{j+1} - t_j$, hence are not independent of the point process of arrivals (which is here a renewal process). Nevertheless Y_j is independent of the past points $\{t_k,\ k \leq j\}$. The process can be defined for all $t \geq 0$ without considering negative birth times and rewards. In the last two models, the rewards and durations are independent of the renewal process, and any past reward may contribute to the value of the process at time t. More general models are considered in Mikosch and Samorodnitsky (2007).

Stationarity and second order properties

• The renewal-reward process (4.16) is strictly stationary since the renewal process is stationary and the shocks are i.i.d. It is weakly stationary if the rewards have finite variance. Then $\mathbb{E}[X_t] = \mathbb{E}[W_1]$ and

$$\mathrm{cov}(X_0, X_t) = \mathbb{E}[W_1^2]\,\mathbb{P}(Y_0 > t) = \lambda\mathbb{E}[W_1^2]\,\mathbb{E}[(Y_1 - t)_+]\ , \tag{4.19}$$

where Y_0 is the first after zero whose distribution is the so-called delay distribution and $\lambda = \mathbb{E}[(t_1 - t_0)]^{-1}$ is intensity of the stationary renewal process. Cf. for instance Taqqu and Levy (1986) or Hsieh et al. (2007).

• The stationary version of the ON-OFF process was studied in Heath et al. (1998). The first On and OFF periods Y_0 and Z_0 can be defined in such a way that the process X is stationary. Let F_{on} and F_{off} be the distribution functions of the ON and OFF periods Y_1 and Z_1. (Heath et al., 1998, Theorem 4.3) show that if $1 - F_{\mathrm{on}}$ is regularly varying with index $\alpha \in (1,2)$ and $1 - F_{\mathrm{off}}(t) = o(F_{\mathrm{on}}(t))$ as $t \to \infty$, then

$$\mathrm{cov}(X_0, X_t) \sim c\mathbb{P}(Y_0 > t) = c\lambda\mathbb{E}[(Y_1 - t)_+] , \tag{4.20}$$

• Consider now the case when the durations are independent of the birth times. Let $\{(Y_j, W_j)\}$ be an i.i.d. sequence of random vectors, independent of the stationary point process of points $\{t_j\}$. Then the process $\{X_t\}$ is strictly stationary as soon as $\mathbb{E}[Y_1] < \infty$, and has finite variance if $\mathbb{E}[W_1^2 Y_1] < \infty$. In that case, $\mathbb{E}[X_t] = \lambda\mathbb{E}[W_1 Y_1]$ and

$$\begin{aligned}\mathrm{cov}(X_0, X_t) = \lambda\,\mathbb{E}[W_1^2\,(Y_1 - t)_+] - \lambda\mathbb{E}[W_1 W_2\,(Y_1 \wedge (Y_2 - t)_+]\} \\ + \mathrm{cov}(W_1\,N(-Y_1, 0], W_2\,N(t - Y_2, t]) ,\end{aligned} \tag{4.21}$$

where λ is the intensity of the stationary point process. The last term is not easy to study. An equivalent expression is given in Mikosch and Samorodnitsky (2007, Section 3.2) in terms of the covariance measure of the point process N, but in general it does not give much more insight. It vanishes in two particular cases:

- if N is a homogeneous Poisson point process;
- if W_1 is centered and independent of Y_1.

In the latter case (4.19) holds, and in the former case, we obtain a formula which generalizes (4.19):

$$\mathrm{cov}(X_0, X_t) = \lambda\,\mathbb{E}[W_1^2\,(Y_1 - t)_+] . \tag{4.22}$$

If N is the counting measure of a stationary renewal process, and if the rewards are assumed non negative and with finite expectation, then it is possible to obtain a more precise expression for the covariance term in (4.21). If the interarrival distribution $\bar{F}_I$ of the renewal process is absolutely continuous with density f and finite expectation, then the renewal function is also absolutely continuous with density u such that $u \sim f + \lambda$ at infinity. See Daley and Vere-Jones (2003), in particular Examples 4.4(b) and 8.2(b). Then, applying Mikosch and Samorodnitsky (2007, Proposition 4.1), we obtain

$$\operatorname{cov}(W_1 N(-Y_1, 0], W_2 N(t - Y_2, t]) \\ = h(t) + \lambda \int_{-\infty}^{\infty} h(t-s)\{u(|s|) - \lambda\} \, \mathrm{d}s \;,$$

with $h(t) = (\mathbb{E}[W_1])^2 \int_0^\infty \bar{G}(s)\bar{G}(s+|t|) \, \mathrm{d}s$ and G is the distribution function defined by $G(t) = \mathbb{E}[W_1 \mathbb{1}_{\{Y \le t\}}]/\mathbb{E}[W_1]$. By bounded convergence, we get that $h(t) \sim (\mathbb{E}[W_1]\mathbb{E}[WY]) \, \bar{G}(t)$ as $(t \to \infty)$. If W_1 is independent of Y_1, then 4.23 implies that the function $\bar{G}$ is regularly varying. Otherwise, it must be assumed that $\bar{G}(t) = O(\ell(t)t^{-\alpha})$. Denote now $v(t) = u(t) - \lambda$, so that $v(t) \sim f(t)$ $(t \to \infty)$. Then

$$\int_{-\infty}^{\infty} h(t-s)\{u(|s|) - \lambda\} \, \mathrm{d}s \\ = \int_0^\infty h(s)v(s+t) \, \mathrm{d}s + \int_0^t h(t-s)v(s) \, \mathrm{d}s + \int_0^\infty h(t+s)v(s) \, \mathrm{d}s \;.$$

The last two terms above are easily seen to be $O(h(t))$, and the first one is $O(\bar{F}_I(t))$ since $v \sim f$ at infinity. Thus, is the tails of the interarrival distribution are thinner than the tails of the durations, then

$$\operatorname{cov}(W_1 N(-Y_1, 0], W_2 N(t - Y_2, t]) = O(h(t)) = o(\mathbb{E}[W_1^2 (Y - t)_+]) \;.$$

Thus in this case we see that (4.20) also holds.

In conclusion, we can say that these models exhibit long memory if if (4.20) holds and the durations have regularly varying tails with index $\alpha \in (1, 2)$ or,

$$\mathbb{E}[W_1^2 \mathbb{1}_{\{Y_1 > t\}}] = \ell(t)t^{-\alpha} \;. \tag{4.23}$$

Then X has long memory with Hurst index $H = (3 - \alpha)/2$ since by Karamata Theorem 1.15, (4.20) and (4.23) imply that

$$\operatorname{cov}(X_0, X_t) \sim \frac{\lambda}{\alpha - 1} \, \ell(t)t^{1-\alpha} \;. \tag{4.24}$$

If time is taken to be contiuous, this yields

$$\operatorname{var}\left(\int_0^t X_s \, \mathrm{d}s\right) \sim \frac{2\lambda}{(\alpha-1)(2-\alpha)(3-\alpha)} \, \ell(t) t^{3-\alpha} \, . \tag{4.25}$$

Thus, in all the cases where the relation between the tail of the phenomenon whose duration induces long memory and the Hurst index of the process can be made explicit, it reads

$$H = \frac{3-\alpha}{2} \, .$$

Examples of interest in teletraffic modeling where W_1 and Y_1 are not independent but (4.23) holds are provided in Maulik et al. (2002) and Faÿ et al. (2007).

4.2.4 Random parameters and aggregation

A "physical" explanation of long range dependence has been searched in applications where this phenomenon is rather well established, such as hydrology. Aggregation of random parameter short memory processes provide one such physical interpretation. We will only describe the simplest of these models, namely the random parameter AR(1) process, introduced by Granger (1980). Let $\{Z_k^{(j)}, k \in \mathbb{Z}, j \geq 1\}$ be a double array of i.i.d. centered random variables with variance σ^2 and let a be a random variable with value in $(-1, 1)$, independent of the sequence $\{Z_j\}$. Let $\{a_j\}$ be i.i.d. copies of a and define the sequence of processes

$$X_k^{(j)} = a_j X_{k-1}^{(j)} + Z_k^{(j)}$$

Conditionnally on a_j, the j-th process $\{X_k^{(j)}, k \in \mathbb{Z}\}$ is a stationary short memory AR(1) process with conditional autocovariance and spectral density

$$\gamma_j(n) = \frac{\sigma^2 a_j^k}{1 - a_j^2} \, ,$$

$$f_j(x) = \frac{\sigma^2}{2\pi} |1 - a_j \mathrm{e}^{-\mathrm{i}x}|^{-2} \, .$$

Unconditionnally, for each j, $X_k^{(j)}$ is a stationary sequence with autocovariance function and spectral density given by

$$\gamma(n) = \sigma^2 \mathbb{E}\left[\frac{a^k}{1-a^2}\right] , \tag{4.26}$$

$$f(x) = \frac{\sigma^2}{2\pi} \mathbb{E}\left[|1 - a\mathrm{e}^{-\mathrm{i}x}|^{-2}\right] . \tag{4.27}$$

Assume for simplicity that the distibution of a is concentrated on $[0,1]$ and admits a density g which is regularly varying at 1 with index $\beta \in (0,1)$, i.e. $g(1-u) \sim u^\beta \ell(u)$ with $\beta \in (0,1)$ and ℓ slowly varying. Then the spectral density is regularly varying at zero with index $\beta - 1$:

$$f(x) \sim \Gamma(2-\beta)\Gamma(\beta)\ell(x)x^{\beta-1} .$$

The aggregated process is defined by

$$X_{n,k} = n^{-1/2} \sum_{j=1}^{n} X_k^{(j)} .$$

Since the individual process are independent, the aggregated process converge (in the sense of finite dimensional distributions) to a Gaussian process with spectral density given by (4.27), hence exhibiting long memory with Hurst index H.

4.2.5 Infinite variance linear processes

Infinite variance time series with long memory can also be defined by analogy with finite variance linear processes. If $\{Z_j\}$ is an i.i.d. sequence and $\{a_j\}$ is a sequence of real numbers, then the series

$$\sum_{j=1}^{\infty} a_j Z_j$$

is almost surely convergent in the following cases:

(i) There exists $\alpha \in (0,1]$ such that $\mathbb{E}[|\epsilon|^p]$ is finite for all $p \in (0,\alpha)$ and infinite for $p > \alpha$ (e.g. if the tails of ϵ_1 are regularly varying with index $-\alpha$) and $\sum_{j=1}^{\infty} |a_j|^{\alpha'} < \infty$ for some $\alpha' \in (0,\alpha)$.

(ii) There exists $\alpha \in (1,2)$ such that $\mathbb{E}[|\epsilon|^p]$ is finite for all $p \in (0,\alpha)$, $\mathbb{E}[\epsilon_1] = 0$ and infinite for $p > \alpha$ (e.g. if the tails of ϵ_1 are regularly varying with index $-\alpha$) and $\sum_{j=1}^{\infty} |a_j|^{\alpha'} < \infty$ for some $\alpha' \in (0,\alpha)$.

If one of these conditions holds, a strictly stationary process $\{X_t\}$ can be defined by

$$X_t = \sum_{j=0}^{\infty} a_j Z_{t-j} \ . \tag{4.28}$$

In both cases $\mathbb{E}[|X_1|^p] < \infty$ for $p < \alpha$. If moreover the tails of ϵ_1 are regularly varying with index α and satisfy the tail balance condition, then the tails of X_1 are also regularly varying with index α by Theorem 2.10.

Long memory in this context will be characterized by the rate of decay of the coefficients a_j. The autocovariance makes no sense if $\alpha < 2$, but other measures of dependence have been defined for infinite variance processes and studied for moving averages by Kokoszka and Taqqu (1996). As in the finite variance case, long memory is related to the non summability of the coefficients a_j. If $\alpha < 1$, this is ruled out by condition (i). If $\alpha > 1$, Condition (ii) allows for non summability of the coefficients a_j. For instance, if a_j is regularly varying with index $H - 1 - 1/\alpha$, then Condition (ii) holds.

4.3 Limit theorems for long memory processes

One of the main characteristics of long memory processes is the weak convergence of the partial sum process to an unusual process, under an unusual normalization. In this section, we will state without proof some famous results that highlight the non standard behaviour of the processes defined in the previous section.

4.3.1 Gaussian and linear processes

Limit theorems for Gaussian long memory processes and functions of such processes have been investigated at least since the seminal paper Rosenblatt (1961), followed and generalized by Taqqu (1974/75), Dobrushin and Major (1979), Taqqu (1979), and finally Arcones (1994) for sequences of random vectors. These references have established the main tools to obtain limit theorems for subordinated Gaussian processes. If $\{X_t\}$ is a Gaussian process with Hurst index H, i.e.

$$\mathrm{cov}(X_0, X_n) = n^{2H-2}\ell(n) ,$$

and if f is a function such that $\mathbb{E}[f(X_1)^2] < \infty]$ and $f - \mathbb{E}[f(X_1)]$ has Hermite rank $m \geq 1$, then

(i) if $m(1-H) > 1/2$, the process $\{f(X_j)\}$ has short memory;

(ii) if $m(1-H) < 1/2$, the process $\{f(X_j)\}$ has long memory with Hurst index H' given by $H' = 1 - m(1-H)$.

For the partial sum processn, non Gaussian limits arise in the former case and Gaussian limits in the latter case. Define

$$S_{n,f}(t) = \sum_{j=1}^{[nt]} \{f(X_j) - \mathbb{E}[f(X_1)]\} . \tag{4.29}$$

In case (i), the covariance inequality (4.13) yields that the series

$$\sum_{j=1}^{\infty} |\mathrm{cov}(f(X_0), f(X_j))|$$

is summable, and thus

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{var}(S_{n,f}(1)) = \mathrm{var}(f(X_0)) + 2\sum_{j=1}^{\infty} \mathrm{cov}(f(X_0), f(X_j)) .$$

Denoting $\sigma^2(f)$ this limit, it can be shown that the finite dimensional distributions of $n^{-1/2}S_{n,f}$ converge weakly to those of $\sigma(f)W$, where W is the standard Brownian motion; cf. Breuer and Major (1983), (Arcones, 1994, Theorem 4).

In case (ii), applying Melher formula (4.12), we have

$$\mathrm{var}\left(\sum_{k=1}^{n} \sum_{j=m+1} \frac{c_j(f)}{j!} H_j(x)\right) \\ = O(n^{(2-2(m+1)(1-H))\vee 1}) = o(n^{2-2m(1-H)}) .$$

Thus, the asymptotic behaviour of $S_{n,f}$ is determined by

$$\frac{c_m(f)}{m!} \sum_{k=1}^{n} H_m(X_k) \; .$$

Denote $S_{n,m} = \sum_{k=1}^n H_m(X_k)$. It has been shown by Dobrushin and Major (1979) (see also Ho and Sun (1990), (Arcones, 1994, Section 3)) that

$$\ell^{-m/2}(n) n^{-1+q(1-H)} \sum_{k=1}^{[nt]} H_m(X_k) \Rightarrow Z_{m,H'}(t)$$

with $H' = 1-m(1-H)$, and the convergence is in $\mathcal{D}([0,\infty))$. Tightness is obtained without additional moment assumptions, since the covariance bound (4.13) yields

$$\mathrm{var}\left(\ell^{-m/2}(n) n^{-1+q(1-H)} \sum_{k=[ns]}^{[nt]} f(X_k) \right) \leq C|t-s|^{2H'} \; ,$$

and $2H' > 1$, so that (Billingsley, 1968, Theorem 15.6) applies. We conclude that $\ell^{-m/2}(n) n^{-1+q(1-H)} S_{n,f}$ converges in $\mathcal{D}([0,\infty))$ to the Hermite process $(c_m(f)/m!) Z_{m,H'}$.

Linear processes

The previous results can be partially extended to linear processes $\{X_t\}$ which admit the represention (4.7), in the case where $\{\epsilon_n\}$ is an i.i.d. non Gaussian sequence. If the coefficients $\{a_j\}$ of the linear representation satisfy the condition (4.9) with $\gamma \in (1/2, 1)$, then it is easily seen that the partial sum process converges to the fractional Brownian motion. The convergence of finite dimensional distribution follows straightforwardly from the central limit theorem for linear sequences of (Ibragimov and Linnik, 1971). Tightness is obtained without additional moment assumption by the same argument as above. Denote $\sigma_n^2 = \mathrm{var}(\sum_{k=1}^n X_k)$.

Proposition 4.7. *Let $\{X_t\}$ be a linear process which admit the represention (4.7) with respect to an i.i.d. sequence $\{\epsilon_n\}$ with mean zero and unit variance. If the coefficients $\{a_j\}$ of the linear representation* *satisfy the condition (4.9) with* $\gamma \in (1/2, 1)$*, then the partial sum process* $\sigma_n^{-1} \sum_{=1}^{[nt]} X_k$ *converges in* $\mathcal{D}([0, \infty))$ *to the fractional Brownian motion with Hurst index* $3/2 - \gamma$*.*

The results for nonlinear instantaneous transformations are more difficult to obtain and can be very different from the Gaussian case. If the noise ϵ_j has finite moments of all order, then the same dichotomy as in the Gaussian case arises, where the Hermite rank must be replaced by the Appell rank of the function f with respect to the distribution of ϵ_1. The partial sum process $S_{n,f}$ may either converge to the Brownian motion with renormalization $\sqrt{n}$ or to a Hermite process $Z_{m,H'}$, with a different normalization. See Giraitis (1985), Giraitis and Surgailis (1989), Ho and Hsing (1997).

If the noise ϵ_i has balanced regularly varying tails with index $\alpha \in (2, 4)$, (Surgailis, 2004, Theorem 2.1) showed that the partial sum process $S_{n,f}$ may converge to an $\alpha\gamma$-stable Lévy process, if $\alpha\gamma < 2$ (and some other restrictions), even if f is bounded, and if the Appel rank of f is at least 2. The long memory property of the original process affects the marginal distribution of the limiting process, but not its dependence structure.

Stochastic volatility processes

Limit theorems for stochastic volatility processes as defined in (4.14) with function $\sigma(x) = \mathrm{e}^{x/2}$ have been investigated by Surgailis and Viano (2002). The results are as expected, given moment conditions: the process being uncorrelated, its partial sums converge to the Browian motion, while partail sums of non linear functions of the process converge to the fractional Brownian motion.

4.3.2 Infinite variance moving averages

Consider the infinite variance moving average of Section 4.2.5. Kasahara and Maejima (1988, Theorem 5.1) proves the convergence of the partial sum process to a fractional stable process.

Theorem 4.8. *Assume that the process* $\{X_j\}$ *is defined by (4.28), where the sequence* $\{Z_j\}$ *is i.i.d. with marginal distribution* F *in the domain of attraction of a stable law with index* $\alpha \in (1, 2)$ *and centered. Assume* *that there exists a function ψ which is regularly varying at infinity with index $\gamma \in [0, 1 - 1/\alpha)$ such that*

$$\lim_{n\to\infty} \frac{1}{\psi(n)} \sum_{j=1}^{n} a_j = 1$$

and $c_j = O(\psi(j)/j)$. Then

$$\frac{1}{a(n)\psi(n)} \sum_{j=1}^{[nt]} X_j \stackrel{\text{fi.di}}{\to} L_{H,\alpha}(t)$$

where $1 - F(a_n) \sim 1/n$, $L_{H,\alpha}$ is defined in (3.12) and $H = \gamma + 1/\alpha$.

4.3.3 Duration driven long memory processes

The class of long memory processes described in Section 4.2.3 has been been introduced in particular for the modelization of teletrafic data. In this context, the superposition of many sources is considered, and the total workload of the network is the integral along time of the individual contributions, and the sum of all the trafic. More precisely, let $X^{(i)}$ be i.i.d. copies of one of the process defined by (4.15), and consider the workload process (in continuous time) for n sources, defined by

$$W_{n,T}(t) = \sum_{i=1}^{n} \int_0^{Tt} X_s^{(i)} \, \mathrm{d}s \; .$$

Under precise assumption on the underlying point process and duration sequence, a phase transition phenomenon arises when n and T tend to infinity. If the stationary distribution of the durations has a regularly varying right tail with index $\alpha \in (1, 2)$, the following dichotomy holds.

- If the number of sources n grows to infinity sufficiently fast, the limiting process is the fractional Brownian motion.
- If the number of sources grows slowly, the limit process is a α-stable Lévy process.

The former case is called the fast growth regime and the latter is the slow growth regime. In the slow growth regime, the transfer of regular variation from the tail of the duration to the memory of the original process is canceled, and the limiting process has heavy tail but short memory (independent increments). This heuristic dichotomy was first described by Taqqu and Levy (1986) in the case of the renewal reward process. It was then rigorously investigated by Mikosch et al. (2002) for the infinite source Poisson and the ON-Off processes. Mikosch and Samorodnitsky (2007) extended these results to a very general framework. We present here the result for the infinite source Poisson process. Since the superposition of n independent homogeneous Poisson processes with intensity λ is an homogeneous Poisson processes with intensity $n\lambda$, it is equivalent to consider an increasing number of independent sources or just one source with increasing intensity. This is the framework considered by Mikosch et al. (2002).

Theorem 4.9. *Assume that N is a homogeneous Poisson process with intensity λ. Let F be the distribution function of the durations $\{Y_j\}$ and assume that $1-F$ is regularly varying at infinity with index $\alpha \in (1,2)$. Denote $a(t) = F^{\leftarrow}(1-1/t)$. Let X_t be the infinite source Poisson process as defined in (4.15) and let $W_T(t) = \int_0^{Tt} X_s \, \mathrm{d}s$ be the total workload at time Tt. Assume that the intensity λ depends on T.*

(i) If $\lambda T\{1-F(T)\} \to 0$, then $a(\lambda T)^{-1}\{W_T(t) - \mathbb{E}[W_T(t)]\}$ converges weakly to a totally skewed to the right α-stable Lévy process.

(ii) If $\lambda T\{1-F(T)\} \to \infty$, then $[\lambda T^3\{1-F(T)\}]^{-1}\{W_T(t) - \mathbb{E}[W_T(t)]\}$ converges weakly to the fractional Browian motion.

This result doubly illustrates the phase transition phenomenon which characterizes long memory. The relative scaling of time and space (the number of sources) may yield two limits with completely different nature: one limiting process has infinite variance and independent increments and the other one has finite variance and dependent increments. Case (i) is referred to as slow growth and case (ii) as fast growth. There is another phase transition related to the tail index of the durations: if they have finite variance, the limiting process of the workload is the Brownian motion.

Chapter 5

Estimation

In the previous chapters several models whose main feature is the regular variation at zero or infinity of some function. The whole distribution of the model can be determined by this function, or only part of it. In an i.i.d. sequence of nonnegative random variables, the marginal distribution function charactizes the whole distribution of the sequence. This is no longer the case if the sequence is stationary but not i.i.d. The spectral density characterizes the distribution of a Gaussian stationary process, but not of a non Gaussian linear process as defined in Chapter 4. The statistical problem that we address in this chapter is the estimation of the index of regular variation of this function, or the estimation of the function.

The natural way to build an estimator of an index of regular variation is to exhibit a family of statistics $T_{n,k}$ that have a scaling property, for instance $\mathrm{var}(T_{n,k}) \sim a s_{n,k}^{-\alpha}$, where $s_{n,k}$ is a sequence of known constants, and to perform a logarithmic regression of $\log(T_{n,k})$ on $\log(s_{n,k})$. The heuristics of these methods are rather simple, but the rigorous derivation of the asymptotic properties of the estimators thus defined are usually rather hard to work out.

Since a regularly varying function can be expressed as $x^{-\alpha}\ell(x)$ where ℓ is a slowly varying function, the estimation of α is a called a semi-parametric problem, i.e. one where the parameter of interest is finite dimensional, and the nuisance parameter (which contains at least the slowly varying function ℓ) may possibly be infinite dimensional, unless a specific parametric model is assumed. This is usually not advisable, since misspecification of the model will make the estimators of α inconsistent. It is preferable to choose a data-driven finite dimensional model by model selection or adaptive estimation or any other method.

The point of view that we adopt here is minimax estimation, because it highlights the interest of second order regular variation. In the present context, this consists in studying the best possible rate of convergence for any estimator of the index α, given the a priori knowledge that ℓ is in some functional class $\mathcal{L}$. This is usually done in two steps: first a lower bound is obtained for some risk function, and then an estimator is exhibited which matches this lower bound, up to some multiplicative constant. This estimator is then called minimax rate optimal over the class $\mathcal{L}$. If the upper bound matches exactly the lower bound, then the estimator is called rate optimal up to the exact constant.

In the next two sections, we consider of estimating the tail index of an i.i.d. sequence and the memory parameter of a Gaussian long memory time series. We define convenient classes of slowly varying functions, for which we derive lower bounds. We then define estimators that are rate optimal and in some cases optimal up to the exact constant. Even though the stochastic structures of the models are very different, the statistical results are quite similar. The optimal rates of convergence of estimators depend only on the class of slowly varying functions and are the same. These rates are slower than the parametric rate of convergence $\sqrt{n}$ (n being the number of observations). They are of the form $n^{\beta/(2\beta+1}$, where β is a smoothness parameter which controls the bias of the estimators.

5.1 Tail index estimation

Let F be a distribution function on $[0,\infty)$ such that $1-F$ is regularly varying at infinity with index $-\alpha$, $\alpha>0$, and let $F^{\leftarrow}$ denote its left-continuous inverse. Then the function U defined on $[1,\infty)$ by $U(t) = F^{\leftarrow}(1-1/t)$ is regularly varying at infinity with index $1/\alpha$. Then the distribution F is in the domain of attraction of the GEV distribution G_γ with $\gamma = 1/\alpha$. The second order behaviour of the function U must be specified in order to obtain uniform rates of convergence.

Definition 5.1. *Let η^* be a non increasing function on $[1,+\infty)$, regularly varying at infinity with index $\rho \geq 0$ and such that* $\lim_{x\to\infty} \eta^*(x) =$

0. *Let* $SV_\infty(\eta^*)$ *be the class of measurable functions* L *defined on* $[1,\infty]$ *which can be expressed as*

$$L(x) = L(1)\exp\left\{\int_1^\infty \frac{\eta(s)}{s}\,ds\right\} \;, \tag{5.1}$$

for some measurable function η *such that* $|\eta| \le \eta^*$.

The lower bound for the estimation of the tail index was obtained by Hall and Welsh (1984) in the case $\rho < 0$ and by Drees (1998) in the case $\rho = 0$. We now quote the latter result.

Theorem 5.2. *Let* $0 < \gamma_0 < \gamma_1$ *be two positive real numbers. Let* η^* *be a non increasing function on* $[1,\infty)$*, regularly varying at* ∞ *with index* $\rho \le 0$ *and such that* $\lim_{x\to\infty}\eta^*(x) = 0$. *Let* t_n *be a sequence satisfying*

$$\lim_{n\to\infty} \eta^*(t_n)(n/t_n)^{1/2} = 1 \;. \tag{5.2}$$

Then, if $\rho < 0$,

$$\liminf_{n\to\infty} \inf_{\hat\gamma_n} \sup_{L\in SV_\infty(\eta^*)} \sup_{\gamma\in[\gamma_0,\gamma_1]} \mathbb{E}_{\gamma,L}[\eta^*(t_n)^{-1}|\hat\gamma_n - \gamma|] > 0 \;, \tag{5.3}$$

and if $\rho = 0$

$$\liminf_{n\to\infty} \inf_{\hat\gamma_n} \sup_{L\in SV(\eta^*)} \sup_{\gamma\in[\gamma_0,\gamma_1]} \mathbb{E}_{\gamma,L}[\eta^*(t_n)^{-1}|\hat\gamma_n - \gamma|] \ge 1 \;, \tag{5.4}$$

where $\mathbb{E}_{\gamma,L}$ *denotes the expectation with respect to the distribution of an i.i.d. sequence with marginal distribution* F *such that*

$$F^{\leftarrow}(1-1/u) = u^\gamma L(u)$$

and the infimum $\inf_{\hat\gamma_n}$ *is taken on all estimators of* γ *based on an i.i.d. sample of size* n *with distribution function* F.

Many estimators of the tail index or more generally of the extreme value index of the marginal distribution of an i.i.d. sequence have been introduced and investigated. When the extreme value index is positive, the Hill estimator is rate optimal. Cf. Hall (1982) and Hall and Welsh (1984) for the case $\rho < 0$ and Drees (1998), citing Csörgő et al. (1985) for the case $\rho = 0$.

We now define the Hill estimator. We will not prove the following results but refer to (Resnick, 2007, Sections 4 and 9) for the consistency and weak convergence.

Denote the increasing order statistics of a n-sample Let $X_1, \dots, X_n$ by $X_{(n:1)}, \dots, X_{(n:n)}$. Fix some threshold k and define the Hill estimator by

$$\hat{\gamma}_{n,k} = \frac{1}{k} \sum_{j=1}^{k} \log \frac{X_{(n:n-j+1)}}{X_{(n:n-k)}} \; . \tag{5.5}$$

If F is in the domain of attraction of an extreme value distribution with positive index γ, then $\hat{\gamma}_{n,k}$ converges in probability to γ for any choice of the intermediate sequence k, i.e. if k depends on n in such a way that

$$\lim_{k\to\infty} k = \lim_{n\to\infty} n/k = \infty \; . \tag{5.6}$$

Under the second order condition of Theorem 5.2, the optimal rate of convergence can be achieved if k is chosen suitably.

Theorem 5.3. *Under the assumptions of Theorem 5.2, let t_n be a sequence satisfying (5.2).*

(i) If $\rho < 0$ then $\eta^*(t_n)^{-1}|\hat{\gamma}_{n,[n/t_n]} - \gamma|$ is bounded in probability uniformly with respect to $\eta \in SV_\infty(\eta^*)$ and $\gamma \in [\gamma_0, \gamma_1]$.

(ii) If $\rho = 0$ and if t_n* is such that

$$\lim_{n\to\infty} t_n^*/t_n = 0 \; , \quad \lim_{n\to\infty} \eta^*(t_n^*)/\eta^*(t_n) = 1 \; , \tag{5.7}$$

then $\eta^*(t_n)^{-1}|\hat{\gamma}_{n,[n/t_n^*]} - \gamma|$ converges to 1 in probability, uniformly with respect to $\eta \in SV_\infty(\eta^*)$ and $\gamma \in [\gamma_0, \gamma_1]$.

Sketch of proof. We will only explain briefly how the uniformity with respect to the class $SV_\infty(\eta^*)$ is obtained. Let Q be defined by $Q(u) = F^{\leftarrow}(1 - 1/u)$. By assumption, $Q(u) = u^\gamma L(u)$. The random variables $X_1 \dots, X_n$ can then be expressed as $Q(1-U_1), \dots, Q(1-U_n)$, where $\{U_j\}$ is an i.i.d. sequence of uniform random variables. The Hill estimator can be expressed as

$$\hat{\gamma}_{n,k} = \frac{\gamma}{k} \sum_{j=1}^{k} \log \frac{U_{(n:k+1)}}{U_{(n:j)}} + \frac{1}{k} \sum_{j=1}^{k} \log \frac{L(1/U_{(n:j)})}{L(1/U_{(n:k+1)})}$$

The second term, say $b_{n,k}$, is a bias term. Since $L \in SV_\infty(\eta^*)$, there exists a function η such that $|\eta| \le \eta^*$ and (5.1) holds. Thus, if $x < y$,

$$\left|\log \frac{L(y)}{L(x)}\right| \le \int_x^y \frac{|\eta(s)|}{s} \mathrm{d}s \le \eta^*(x)\log(y/x) \ .$$

This yields a uniform bound for the bias term.

$$|b_{n,k}| \le \frac{\eta^*(1/U_{(n:k+1)})}{k} \sum_{j=1}^{k} \log \frac{U_{(n:k+1)}}{U_{(n:j)}} \ .$$

Denote $S_{n,k} = k^{-1}\sum_{j=1}^{k} \log(U_{(n:k+1)}/U_{(n:j)})$. It can be shown by standard technique such as the Renyi representation for order statistics (Cf. Rényi (1953), Ahsanullah and Nevzorov (2005)) or by point process techniques, cf. Resnick (2007) that $S_{n,k}$ converges weakly to 1 and that $k^{1/2}(S_{n,k}-1)$ converges weakly to the standard Gaussian law. It can also be shown that $U_{(n:k+1)}/(k/n)$ converges in probability to 1. Choose now $k = [n/t_n]$ if $\rho < 0$ and $k = [n/t_n^*]$ if $\rho = 0$. In both cases, we have that $\eta^*(1/U_{(n:k+1)})/\eta^*(t_n)$ converges in probability to 1 (by regular variation of η^* and by since by assumption $\eta^* * (t_n)/\eta^*(t_n^*) \to 1$) and

$$\eta^*(t_n)^{-1}(\hat\gamma_{n,k} - 1) = \frac{1}{k^{1/2}\eta^*(t_n)} k^{1/2}(S_{n,k} - 1) + \frac{\eta^*(1/U_{(n:k+1)})}{\eta^*(t_n)} S_{n,k} \ .$$

If $\rho < 0$, then by assumption, k is such that $k^{1/2}\eta^*(t_n) \to 1$, thus $\eta^*(t_n)^{-1}|\hat\gamma_{n,k} - 1|$ is bounded in probability, uniformly with respect to $\eta \in SV_\infty(\eta^*)$.

If $\rho = 0$, then (5.7) implies that k is such that $k^{1/2}\eta^*(t_n) \to 0$, thus $\eta^*(t_n)^{-1}|\hat\gamma_{n,k} - 1|$ converges in probability to 1, uniformly with respect to $\eta \in SV_\infty(\eta^*)$. □

Remark 5.4. In the case $\rho < 0$, a central limit theorem can be obtained with a suboptimal choice of the threshold k.

The proof of the rate optimality in the case $\rho = 0$ is based on the fact that it is possible to find sequences t_n and t_n^* such that (5.2) and (5.7) hold simultaneously, which we prove now.

Lemma 5.5. *Let the function η^* be non decreasing and slowly varying at infinity. There exist sequences t_n and t_n^* such that (5.2) and (5.7) hold.*

Proof. Let f be the function defined by $f(t) = t^{1/2}\eta^*(t)^{-1}$. By Theorem 1.22, there exists a function g such that $f(g(t)) \sim g(f(t)) \sim t$ as $t \to \infty$. Define then t_n by $t_n = g(n^{1/2})$. Then $f(t_n) \sim n^{1/2}$, i.e. (5.2) holds.

By the representation Theorem, there exist a function ζ such that $\lim_{s\to\infty} \zeta(s) = 0$ and a function c such that $\lim_{x\to\infty} c(x)$ exists i$n(0, \infty)$ such that

$$\frac{\eta^*(t_n)}{\eta^*(t_n^*)} = \frac{c(t_n)}{c(t_n^*)} \exp \int_{t_n^*}^{t_n} \frac{\zeta(s)}{s} \,\mathrm{d}s = \frac{c(t_n)}{c(t_n^*)} \exp \int_1^{t_n/t_n^*} \frac{\zeta(t_n^* s)}{s} \,\mathrm{d}s \,.$$

Since t_n^* tends to infinity, proving (5.7) is equivalent to proving that there exists a function $x(t)$ such that $\lim_{t\to\infty} x(t) = \infty$ and

$$\lim_{t\to\infty} \int_1^{x(t)} \frac{\zeta(ts)}{s} \,\mathrm{d}s = 0 \,. \tag{5.8}$$

Denote $h_t(x) = \int_1^{t_n/t_n^*} s^{-1}\zeta(t_n^* s)\,\mathrm{d}s$. Then for each x, $\lim_{t\to\infty} h_t(x) = 0$ and for each fixed t, $\lim_{x\to\infty} h_t(x) = \infty$. Define

$$x(t) = \sup\{y \,, h_t(y) \le 1/y\}.$$

It is then easily seen that (5.8) holds. □

The Hill estimator has been studied for weakly dependent stationary sequences. Cf. Resnick and Stărică (1997), Resnick and Stărică (1998), Drees (2000) and Rootzén (2009). The rate of convergence is not affected by the dependence considered in these papers, more or less related to some form of mixing. To the best of our knowledge, there are no results for the Hill (or any other) estimator applied to long range dependent stationary processes with heavy tailed stationary distribution.

5.2 Memory parameter estimation for a Gaussian stationary process

In this section we investigate the rate of convergence of estimators of the memory parameter of a stationary time series, defined as the index of regular variation at zero of the spectral density.

Definition 5.6. *Let η^* be a non decreasing function on $[0, \pi]$, regularly varying at zero with index $\rho \geq 0$ and such that $\lim_{x\to 0} \eta^*(x) = 0$. Let $SV(\eta^*)$ be the class of even measurable functions L defined on $[-\pi, \pi]$ which can be expressed for $x \geq 0$ as*

$$L(x) = L(\pi) \exp\left\{ -\int_x^\pi \frac{\eta(s)}{s} \, ds \right\} ,$$

for some measurable function η such that $|\eta| \leq \eta^*$.

The following result is proved in Soulier (2009) by adapting the proof of (Drees, 1998, Theorem 2.1).

Theorem 5.7. *Let η^* be a non decreasing function on $[0, \pi]$, regularly varying at 0 with index $\rho \geq 0$ and such that $\lim_{x\to 0} \eta^*(x) = 0$. Let t_n be a sequence satisfying*

$$\lim_{n\to\infty} \eta^*(t_n)(nt_n)^{1/2} = 1 . \tag{5.9}$$

Then, if $\rho > 0$,

$$\liminf_{n\to\infty} \inf_{\hat\alpha_n} \sup_{L \in SV(\eta^*)} \sup_{\alpha\in(-1,1)} \mathbb{E}_{\alpha,L}[\eta^*(t_n)^{-1} |\hat\alpha_n - \alpha|] > 0 , \tag{5.10}$$

and if $\rho = 0$

$$\liminf_{n\to\infty} \inf_{\hat\alpha_n} \sup_{L \in SV(\eta^*)} \sup_{\alpha\in(-1,1)} \mathbb{E}_{\alpha,L}[\eta^*(t_n)^{-1} |\hat\alpha_n - \alpha|] \geq 1 , \tag{5.11}$$

where $\mathbb{P}_{\alpha,L}$ denotes the distribution of any second order stationary process with spectral density $x^{-\alpha}L(x)$ and the infimum $\inf_{\hat\alpha_n}$ is taken on all estimators of α based on n observations of the process.

Proof. Let $\ell > 0$, t_n be a sequence that satisfies the assumption of Theorem 5.7, and define $\alpha_n = \eta^*(\ell t_n)$ and

$$\eta_n(s) = \begin{cases} 0 & \text{if } 0 \leq s \leq \ell t_n, \\ \alpha_n & \text{if } \ell t_n < s \leq \pi, \end{cases}$$

$$L_n(x) = \pi^{\alpha_n} \exp\left\{ -\int_x^\pi \eta_n(s) \, ds \right\} .$$

Since η^* is assumed non decreasing, it is clear that $L_n \in SV(\eta^*)$. Define now $f_n^-(x) = x^{-\alpha_n} L_n(x)$ and $f_n^+ = (f_n^-)^{-1}$. f_n^- can be written as

$$f_n^-(x) = \begin{cases} (\ell t_n/x)^{\alpha_n} & \text{if } 0 < x \le \ell t_n, \\ 1 & \text{if } \ell t_n < x \le \pi. \end{cases}$$

Straighforward computations yield

$$\int_0^\pi \{f_n^-(x) - f_n^+(x)\}^2 \,\mathrm{d}x = 8\ell t_n \alpha_n^2 (1 + O(\alpha_n^2)) = 8\ell n^{-1}(1 + o(1)). \tag{5.12}$$

The last equality holds by definition of the sequence t_n. Let $\mathbb{P}_n^-$ and $\mathbb{P}_n^+$ denote the distribution of a n-sample of a stationary Gaussian processes with spectral densities f_n^- et f_n^+ respectively, $\mathbb{E}_n^-$ and $\mathbb{E}_n^+$ the expectation with respect to these probabilities, $\frac{d\mathbb{P}_n^+}{d\mathbb{P}_n^-}$ the likelihood ratio and $A_n = \{\frac{d\mathbb{P}_n^+}{d\mathbb{P}_n^-} \ge \tau\}$ for some real $\tau \in (0,1)$. Then, for any estimator $\hat\alpha_n$, based on the observation $(X_1, \dots, X_n)$,

$$\begin{aligned} \sup_{\alpha,L} \mathbb{E}_{\alpha,L}[|\hat\alpha_n - \alpha|] &\ge \frac{1}{2}\left(\mathbb{E}_n^+[|\hat\alpha_n - \alpha_n|] + \mathbb{E}_n^-[|\hat\alpha_n + \alpha_n|]\right) \\ &\ge \frac{1}{2}\mathbb{E}_n^-\left[\mathbb{1}_{A_n}|\hat\alpha_n + \alpha_n| + \frac{d\mathbb{P}_n^+}{d\mathbb{P}_n^-}\mathbb{1}_{A_n}|\hat\alpha_n - \alpha_n|\right] \\ &\ge \frac{1}{2}\mathbb{E}_n^-\left[\{|\hat\alpha_n + \alpha_n| + \tau|\hat\alpha_n - \alpha_n|\}\mathbb{1}_{A_n}\right] \ge \tau\alpha_n \mathbb{P}_n^-(A_n). \end{aligned}$$

Denote $\epsilon = \log(1/\tau)$ and $\Lambda_n = \log(d\mathbb{P}_n^+/d\mathbb{P}_n^-)$. Then $\mathbb{P}_n^-(A_n) = 1 - \mathbb{P}_n^-(\Lambda_n \le -\epsilon)$. Applying (5.12) and (Giraitis et al., 1997, Lemma 2), we obtain that there exist constants C_1 and C_2 such that

$$\mathbb{E}_n^-[\Lambda_n] \le C_1\ell\ , \quad \mathbb{E}_n^-[(\Lambda_n - m_n)^2] \le C_2\ell\ .$$

This yields, for any $\eta > 0$ and small enough ℓ,

$$\mathbb{P}_n^-(A_n) \ge 1 - \epsilon^{-2}\mathbb{E}[\Lambda_n^2] \ge 1 - Cl\epsilon^{-2} \ge 1 - \eta\ .$$

Thus, for any $\eta, \tau \in (0,1)$, and sufficiently small ℓ, we have

$$\begin{aligned} \liminf_{n\to\infty} \inf_{L \in SV(\eta^*)} \inf_{\alpha \in (-1,1)} & \mathbb{E}_{\alpha,L}[\eta^*(t_n)^{-1}|\hat\alpha_n - \alpha|] \\ & \ge \tau(1-\eta)\lim_{n\to\infty}\frac{\eta^*(\ell t_n)}{\eta^*(t_n)} = \tau(1-\eta)\ell^\rho\ . \end{aligned}$$

This proves (5.10) and (5.11). □

Example 5.8. Define $\eta^*(s) = Cs^\beta$ for some $\beta > 0$ and $C > 0$. Then any function $L \in SV(\eta^*)$ satisfies $L(x) = L(0) + O(x^\beta)$, and we recover the case considered by Giraitis et al. (1997). The lower bound for the rate of convergence is $n^{\beta/(2\beta+1)}$.

Example 5.9. For $\rho > 0$, define $\eta^*(s) = \rho/\log(1/s)$, then

$$\exp\left\{\int_x^{1/e} \frac{\eta^*(s)}{s}\, ds\right\} = \exp\{\rho \log\log(1/x)\} = \log^\rho(1/x).$$

A suitable sequence t_n must satisfy $\rho^2/\log^2(t_n) \approx nt_n$. One can for instance choose $t_n = \log^2(n)/(n\rho^2)$, which yields $\eta^*(t_n) = \rho/\log(n)\{1 + o(1)\}$. Note that $\eta(s) = \rho/\log(s)$ belongs to $SV(\eta^*)$, and the corresponding slowly varying function is $\log^{-\rho}(1/x)$. Hence, the rate of convergence is not affected by the fact that the slowly varying function vanishes or is infinite at 0.

Example 5.10. Denote $\eta^*(x) = \{\log(1/x)\log\log(1/x)\}^{-1}$ and $L(x) = \log\log(1/x)$. Then $L \in SV(\eta^*)$ In that case, the optimal rate of convergence is $\log(n)\log\log(n)$. Even though the slowly varying function affecting the spectral density at zero diverges very weakly, the rate of convergence of any estimator of the memory parameter is dramatically slow.

These rates of convergence are achieved by two estimators of the memory parameter, the GPH estimator and the local Whittle or Gaussian semiparametric estimator. Both are based on the periodogram of a n-sample $X_1, \ldots, X_n$, evaluated at the Fourier frequencies $x_j = 2j\pi/n$, $j = 1, \ldots, n$, defined by

$$I_{X,j} = (2\pi n)^{-1} \left|\sum_{t=1}^n X_t \mathrm{e}^{-\mathrm{i}tx_j}\right|^2.$$

The frequency domain estimates of the memory parameter α are based on the following heuristic approximation: the renormalised periodogram ordinates $I_{X,j}/f(x_j)$, $1 \le j \le n/2$ are approximately i.i.d. standard exponential random variables. The GPH estimator is thus based on an ordinary least square regression of $\log(I_{X,k})$ on $\log(k)$ for $k = 1, \ldots, m$, where m is a bandwith parameter:

$$(\hat\alpha(m), \hat C) = \arg\min_{\alpha, C} \sum_{k=1}^m \{\log(I_{X,k}) - C + \alpha \log(k)\}^2.$$

The GPH estimator has an explicit expression as a weighted sum of log-periodogram ordinates:

$$\hat{\alpha}(m) = -s_m^{-2} \sum_{k=1}^{m} \nu_{m,k} \log(I_{X,k}),$$

with $\nu_{m,k} = \log(k) - m^{-1} \sum_{j=1}^{m} \log(j)$ and $s_m^2 = \sum_{k=1}^{m} \nu_{m,k}^2 = m\{1 + o(1)\}$. The first rigorous results were obtained by Robinson (1995) for Gaussian long memoyr processes. The rate optimality of the GPH estimator for Gaussian processes whose spectral density is second order regularly varying at zero was proved by Giraitis et al. (1997), and extended to more general spectral densities by Soulier (2009). The technique of proof relies on moment bounds for functions of Gaussian vectors, see e.g. Soulier (2001) for ad hoc bounds.

Theorem 5.11. *Let η^* be a non decreasing slowly varying function such that $\lim_{x\to 0} \eta^*(x) = 0$. Let $\mathbb{E}_{\alpha,L}$ denote the expectation with respect to the distribution of a* Gaussian *process with spectral density $x^{-\alpha}L(x)$. Let t_n be a sequence that satisfies (5.9) and let m be a non decreasing sequence of integers such that*

$$\lim_{n\to\infty} m^{1/2}\eta^*(t_n) = \infty \text{ and } \lim_{n\to\infty} \frac{\eta^*(t_n)}{\eta^*(m/n)} = 1 \; . \tag{5.13}$$

Assume also that the sequence m can be chosen in such a way that

$$\lim_{n\to\infty} \frac{\log(m) \int_{m/n}^{\pi} s^{-1}\eta^*(s)\,\mathrm{d}s}{m\eta^*(m/n)} = 0 \; . \tag{5.14}$$

Then, for any $\delta \in (0,1)$,

$$\limsup_{n\to\infty} \sup_{|\alpha|\le\delta} \sup_{L\in SV(\eta^*)} \eta^*(t_n)^{-2}\mathbb{E}_{\alpha,L}[(\hat{\alpha}(m)-\alpha)^2] \le 1. \tag{5.15}$$

Sketch of proof. Here again, we only briefly explain how the rate optimality is obtained. See Soulier (2009) for a detailed proof. Define $\mathcal{E}_k = \log\{x_k^{\alpha} I_k / L(x_k)\}$. The deviation of the GPH estimator can be split into a stochastic term and a bias term:

$$\hat{\alpha}(m) - \alpha = -s_m^{-2} \sum_{k=1}^{m} \nu_{m,k}\mathcal{E}_k - s_m^{-2} \sum_{k=1}^{m} \nu_{m,k} \log(L(x_k)). \tag{5.16}$$

Covariance bounds for functions of Gaussian vectors such as the covariance inequality of (Arcones, 1994, Lemma 1) yields the following bound:

$$\mathbb{E}\Big[\Big\{\sum_{k=1}^{m}\nu_{m,k}\log(\mathcal{E}_k)\Big\}^2\Big] \le C\,m\,, \tag{5.17}$$

for some constant C which depend only on δ and the function η^*. In order to deal with the bias term, it must be shown that

$$\Big|\sum_{k=1}^{m}\nu_{m,k}\log(L(x_k))\Big| \le m\eta^*(x_m)\{1+o(1)\}, \tag{5.18}$$

uniformly with respect to $|\eta| \le \eta^*$. This is where assumption (5.14) is needed. Then, choosing m as in (5.13) yields (5.15). □

We deduce that the GPH estimator achieves the optimal rate of convergence, up to the exact constant over the class $SV(\eta^*)$ when η^* is slowly varying. This happens because in this case, the bias term dominates the stochastic term if the bandwidth parameter m satisfies (5.13), as shown by the bounds (5.17) and (5.18). This result is of obvious theoretical interest, but is not completely devoid of practical importance, since when the rate of convergence of an estimator is logarithmic in the number of observations, constants do matter.

Corollary 5.12. *Let $\delta \in (0,1)$ and η^* be a non decreasing slowly varying function such that $\lim_{x\to 0}\eta^*(x) = 0$ and such that it is possible to choose a sequence m that satisfies (5.14). Then, for t_n as in (5.9),*

$$\liminf_{n\to\infty}\inf_{\hat\alpha_n}\sup_{L\in SV(\eta^*)}\sup_{\alpha\in(-\delta,\delta)}\mathbb{E}_{\alpha,L}[\eta^*(t_n)^{-1}|\hat\alpha_n-\alpha|] = 1\,. \tag{5.19}$$

Appendix A

Appendix

A.1 Cauchy's integral equation

Theorem A.1. *Let f be a measurable function defined on $\mathbb{R}$ such that for all $x, y \in \mathbb{R}$,*

$$f(x+y) = f(x) + f(y) \; . \tag{A.1}$$

Then there exists $A \in \mathbb{R}$ such that $f(x) = Ax$.

Proof. First, the additivity property (A.1) straightforwardly yields that $f(0) = 0$, $f(-x) = -f(x)$ and $f(r) = rf(1)$ for any rational number r.

We now prove that f is bounded in a neighborhood of the origin. Since f is measurable, the sets $\{x \in \mathbb{R} \mid f(x) \leq n\}$ are measurable and since their union is $\mathbb{R}$, at least one of them is of positive Lebesgue measure. By (Bingham et al., 1989, Corollary 1.1.3), the $A + A$ contains an interval $(y - \delta, y + \delta)$ for some $y \in \mathbb{R}$ and some $\delta > 0$. Thus if $t \in (-\delta, \delta)$, there exist $a, a' \in A$ such that $a + a' = y + t$. The additivity property implies that

$$f(t) = f(a) + f(a') - f(y) \leq 2n - k(y) \; .$$

Thus f is bounded above on $(-\delta, \delta)$. Simlarly, since $f(-t) = -f(t)$, f is bounded below on $(-\delta, \delta)$, say by M.

By additivity, if $|x| \leq \delta/n$, then $|f(x)| \leq M/n$. For any real number y, there exists a rational number r such that $|y - r| \leq \delta/n$. This implies

$$|f(x) - xf(1)| \leq |f(x-r)| + |r-x||f(1)| \leq M/n + \delta|f(1)|/n \; .$$

Since n is arbitrary, this proves that $f(x) = xf(1)$. □

Corollary A.2. *If g is measurable and $g(xy) = g(x)g(y)$ for all $x, y > 0$, then $g(x) = x^\rho$ for some $\rho \in \mathbb{R}$.*

A.2 Monotone functions

Definition A.3 (Left-continuous inverse). *Let $F : \mathbb{R} \to \mathbb{R}$ be a right-continuous non decreasing function. Its left-continuous inverse $F^{\leftarrow}$ is defined by*

$$F^{\leftarrow}(y) = \inf\{x,\ F(x) \geq y\}.$$

The following properties hold.

$$F(x) \geq y \Leftrightarrow x \geq F^{\leftarrow}(y)\ , \quad F(F^{\leftarrow}(y)) \geq y\ .$$

If F is a probability distribution function and U is a random variable uniformly distributed on $[0,1]$, then $F^{\leftarrow}(U)$ has distribution function F. If F is continuous at y, then $F \circ F^{\leftarrow}(y) = y$. If F is continuous, then $F(X)$ is uniformly distributed.

Proposition A.4 (Convergence of monotone functions).

(i) Let $\{H_n\}$ be a sequence of monotone functions on $\mathbb{R}$. If H_n converges pointwise to H and if H is continuous, then the convergence is uniform on compact sets. If $H_n(\infty) = H(\infty) = 1$, then the convergence is uniform on $\mathbb{R}$.

(ii) If H_n converges pointwise to H, then $H_n^{\leftarrow}$ converges to $H^{\leftarrow}$.

A.3 Continuity Theorem for Laplace transforms

Theorem A.5 (Feller (1971) Theorem XIII 2a). *Let U_n be a sequence of positive non decreasing functions on $[0, \infty)$. Let $\mathcal{L}U_n$ be the Laplace transform of U_n.*

(i) If U_n converges to U at any continuity point of U and if there exists a such that $\mathcal{L}U_n(a)$ is bounded, then $\mathcal{L}U_n(t)$ converges to $\mathcal{L}U(t)$ for all $t > a$.

(ii) *If $\mathcal{L}U_n$ converges to a fonction L pointwise on some interval $]a,\infty[$, then L is the Laplace transform of a function U and U_n converges to U at all continuity points.*

Proof. (i) Let $M = \sup_n LU_n(a)$ and let $t > 0$ be a continuity point of U. Then, by bounded convergence, for all $x \geq 0$,

$$\lim_{n\to\infty} \int_0^t \mathrm{e}^{-(a+x)s} \mathrm{d}U_n(s) = \int_0^t \mathrm{e}^{-(a+x)s} \mathrm{d}U(s) \; .$$

For any t such that $M\mathrm{e}^{-xt} \leq \epsilon$, this implies that, for n large enough,

$$\int_0^\infty \mathrm{e}^{-(a+x)s} \mathrm{d}U(s) - 2\epsilon \leq \int_0^t \mathrm{e}^{-(a+x)s} \mathrm{d}U_n(s) \leq \int_0^\infty \mathrm{e}^{-(a+x)s} \mathrm{d}U(s) + 2\epsilon \; .$$

Letting $t \to \infty$, this yields, for all $x > a$,

$$\mathcal{L}U(x) - 2\epsilon \leq \liminf_{n\to\infty} \mathcal{L}U_n(x) \leq \limsup_{n\to\infty} \mathcal{L}U_n(x) \leq \mathcal{L}U(x) + 2\epsilon \; .$$

(ii) Let $b > a$ and define $g_n(x) = \mathcal{L}U_n(b+x)/\mathcal{L}U_n(b)$. Then g_n is the Laplace transform of the probability distribution function $F_n(x) = \int_0^x \mathrm{e}^{-bt} U_n(\mathrm{d}t)/\mathcal{L}U_n(b)$. The sequence of functions g_n converge to the function g defined by $g(x) = L(b+x)/L(b)$. Let $F^\sharp$ be a limit of F_n along a converging subsequence. Since $g_n(0) = 1$ for all n, we can apply the first part of the theorem and obtain that g is the Laplace transform of $F^\sharp$. Thus F_n has only one limit along any subsequence, hence F_n converges to $F^\sharp$ at any continuity point of $F^\sharp$.

□

A.4 Fourier transform

Theorem A.6. *Let f be locally integrable and g non increasing and* $\lim_{x\to\infty} g(x) = 0$. *Assume*

$$\int_0^1 \frac{|f(t)|}{t} \, \mathrm{d}t \int_0^t g(u) \, \mathrm{d}u < \infty \; .$$

Then

$$\int_0^\infty f(x)g(x)\,\mathrm{d}x = \frac{2}{\pi}\int_0^\infty \mathrm{Re}(\hat{f}(x))\mathrm{Re}(\hat{g}(x))\,\mathrm{d}x = \frac{2}{\pi}\int_0^\infty \mathrm{Im}(\hat{f}(x))\mathrm{Im}(\hat{g}(x))\,\mathrm{d}x\ ,$$

where $\hat{f}(t) = \int_0^\infty \mathrm{e}^{\mathrm{i}tx} f(x)\,\mathrm{d}x$.

Proof. See Titchmarsh (1986, Theorem 38) □

A.5 Vague convergence

We gather here some results about vague convergence that are needed in Section 3.5. The classical reference is Sidney (1987).

Definition A.7 (Radon measure)**.** *A Radon measure on* $(0,\infty]$ *is a measure* ν *on the Borel sigma-field of* $(0,\infty]$ *such that* $\nu(K) < \infty$ *for all relatively compact set* K.

Definition A.8 (Vague convergence)**.** *A sequence of Radon measures* $\{\nu_n\}$ *on* $(0,\infty]$ *converges vaguely to the Radon measure* ν *iff for any function* f *with compact support in* $(0,\infty]$, $\lim_{n\to\infty} \nu_n(f) = \nu(f)$.

Lemma A.9. *A sequence of Radon measures* $\{\nu_n\}$ *converge vaguely to* ν *if* $\lim_{n\to\infty} \nu_n(A) = \nu(A)$ *for all relatively compact set* A *such that* $\nu(\partial A) = 0$.

Lemma A.10. *Let* F *be a distribution function. The survival function* $1-F$ *is regularly varying at infinity if and only if there exists a sequence of real numbers* $\{a_n\}$ *such that* $\lim_{n\to\infty} a_n = \infty$ *and the sequence of measures* μ_n *defined on* $(0,\infty)$ *by* $\mu_n(A) = n\int_0^\infty \mathbb{1}_{\{x/a_n\in A\}} F(\mathrm{d}x)$ *converges vaguely on* $(0,\infty)$ *to some Radon measure* μ. *If* μ *is not a point mass at* ∞, *then* $\mu = \mu_\alpha$ *for some* $\alpha > 0$, *where* μ_α *is defined by*

$$\mu_\alpha(A) = \int_0^\infty \mathbb{1}_A(x)\,\alpha x^{-\alpha-1}\mathrm{d}x\ .$$

Proof. The direct part ($1-F$ regularly varying implies vague convergence) is straightforward; the converse follows by applying Lemma 1.5. □

Bibliography

Mohammad Ahsanullah and Valery B. Nevzorov. *Order statistics: examples and exercises.* Nova Science Publishers Inc., Hauppauge, NY, 2005.

Miguel A. Arcones. Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Annals of Probability*, 15(4): 2243–2274, 1994.

Søren Asmussen. *Ruin probabilities*, volume 2 of *Advanced Series on Statistical Science & Applied Probability.* World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

Bojan Basrak, Richard A. Davis, and Thomas Mikosch. Regular variation of GARCH processes. *Stochastic Processes and their Applications*, 99(1):95–115, 2002.

Patrick Billingsley. *Convergence of probability measures.* New York, Wiley, 1968.

Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels. *Regular variation.* Cambridge University Press, Cambridge, 1989.

Philippe Bougerol and Nico Picard. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, 52 (1-2):115–127, 1992.

F. Jay Breidt, Nuno Crato, and Pedro de Lima. The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83(1-2):325–348, 1998.

Leo Breiman. On some limit theorems similar to the arc-sine law. *Theory of Probability and Applications*, 10:323–331, 1965.

Peter Breuer and Péter Major. Central limit theorems for nonlinear functionals of Gaussian fields. *Journal of Multivariate Analysis*, 13 (3):425–441, 1983.

Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. Springer Series in Statistics. New York, Springer-Verlag, 1991.

Daren B. H. Cline and Tailen Hsing. Large deviation probabilities for sums of random variables with heavy or subexponential tails. Statistics Dept., Texas A&M University, 1998.

Daren B. H. Cline and Gennady Samorodnitsky. Subexponentiality of the product of independent random variables. *Stochastic Process. Appl.*, 49(1):75–98, 1994.

Sándor Csörgő, Paul Deheuvels, and David Mason. Kernel estimates of the tail index of a distribution. *The Annals of Statistics*, 13(3): 1050–1077, 1985.

Daryl J. Daley and David Vere-Jones. *An introduction to the theory of point processes. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003.

Richard A. Davis and Tailen Hsing. Point process and partial sum convergence for weakly dependent random variables with infinite variance. *The Annals of Probability*, 23(2):879–917, 1995.

Laurens De Haan and Ana Ferreira. *Extreme value theory. An introduction.* Springer Series in Operations Research and Financial Engineering. New York, NY: Springer., 2006.

Roland L. Dobrushin and Peter Major. Non-central limit theorems for nonlinear functionals of Gaussian fields. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50(1):27–52, 1979.

Holger Drees. Optimal rates of convergence for estimates of the extreme value index. *The Annals of Statistics*, 26(1):434–448, 1998.

Holger Drees. Weighted approximations of tail processes for β-mixing random variables. *The Annals of Applied Probability*, 10(4):1274–1301, 2000.

Paul Embrechts and Charles M. Goldie. On closure and factorization properties of subexponential and related distributions. *Australian Mathematical Society. Journal. Series A*, 29(2):243–256, 1980.

Paul Embrechts and Makoto Maejima. *Selfsimilar processes.* Princeton University Press, Princeton, NJ, 2002.

Gilles Faÿ, Bárbara González-Arévalo, Thomas Mikosch, and Gennady Samorodnitsky. Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Systems. Theory and Applications*, 54(2):121–140, 2006.

Gilles Faÿ, François Roueff, and Philippe Soulier. Estimation of the memory parameter of the infinite-source Poisson process. *Bernoulli*, 13(2):473–491, 2007.

William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

Liudas Giraitis. The central limit theorem for functionals of a linear process. *Lithuanian Mathematical Journal*, 25(1):43–57, 1985.

Liudas Giraitis and Donatas Surgailis. A limit theorem for polynomials of a linear process with long-range dependence. *Lithanian Mathematical Journal*, 29(2):290–311, 1989.

Liudas Giraitis, Peter M. Robinson, and Alexander Samarov. Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long-range dependence. *Journal of Time Series Analysis*, 18(1):49–60, 1997.

Charles M. Goldie. Implicit renewal theory and tails of solutions of random equations. *The Annals of Applied Probability*, 1(1):126–166, 1991.

C. W. J. Granger. Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238, 1980.

Peter Hall. On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B. Methodological*, 44 (1):37–42, 1982.

Peter Hall and Alan H. Welsh. Best attainable rates of convergence for estimates of parameters of regular variation. *The Annals of Statistics*, 12(3):1079–1084, 1984.

Andrew C. Harvey. Long memory in stochastic volatility. In *In: Knight, J., Satchell, S. (Eds.), Forecasting volatility in financial markets.*, London, 1998. Butterworth-Heinemann.

David Heath, Sidney Resnick, and Gennady Samorodnitsky. Heavy tails and long range dependence in ON/OFF processes and associated fluid models. *Mathematics of Operations Research*, 23(1):145–165, 1998.

Hwai-Chung Ho and Tailen Hsing. Limit theorems for functionals of moving averages. *The Annals of Probability*, 25(4):1636–1669, 1997.

Hwai Chung Ho and Tze Chien Sun. Limiting distributions of nonlinear vector functions of stationary Gaussian processes. *The Annals of Probability*, 18(3):1159–1173, 1990.

Meng-Chen Hsieh, Clifford M. Hurvich, and Philippe Soulier. Asymptotics for duration-driven long range dependent processes. *J. Econometrics*, 141(2):913–949, 2007.

Idlar A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables.* Wolters-Noordhoff Publishing, Groningen, 1971.

Akihito Inoue. Abel-Tauber theorems for Fourier-Stieltjes coefficients. *Journal of Mathematical Analysis and Applications*, 211(2):460–480, 1997.

Anders Hedegaard Jessen and Thomas Mikosch. Regularly varying functions. *Institut Mathématique. Publications. Nouvelle Série*, 80(94): 171–192, 2006.

Yuji Kasahara and Makoto Maejima. Weighted sums of i.i.d. random variables attracted to integrals of stable processes. *Probability Theory and Related Fields*, 78(1):75–96, 1988.

Harry Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248, 1973.

Piotr S. Kokoszka and Murad S. Taqqu. Infinite variance stable moving averages with long memory. *Journal of Econometrics*, 73(1):79–99, 1996.

John Lamperti. Semi-stable stochastic processes. *Transactions of the American Mathematical Society*, 104:62–78, 1962.

Ming Liu. Modeling long memory in stock market volatility. *Journal of Econometrics*, 99:139–171, 2000.

Krishanu Maulik and Sidney Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67 (2005), 2004.

Krishanu Maulik, Sidney Resnick, and Holger Rootzén. Asymptotic independence and a network traffic model. *Journal of Applied Probability*, 39(4):671–699, 2002.

Thomas Mikosch and Gennady Samorodnitsky. The supremum of a negative drift random walk with dependent heavy-tailed steps. *The Annals of Applied Probability*, 10(3):1025–1064, 2000.

Thomas Mikosch and Gennady Samorodnitsky. Scaling limits for cumulative input processes. *Mathematics of Operations Research*, 32(4): 890–918, 2007.

Thomas Mikosch, Sidney Resnick, Holger Rootzén, and Alwin Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *The Annals of Applied Probability*, 12(1):23–68, 2002.

Sergey V. Nagaev. Large deviations of sums of independent random variables. *The Annals of Probability*, 7(5):745–789, 1979.

William R. Parke. What is fractional integration? *Review of Economics and Statistics*, pages 632–638, 1999.

Alfréd Rényi. On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungaricae*, 4:191–231, 1953.

Sidney Resnick and Cătălin Stărică. Asymptotic behavior of Hill's estimator for autoregressive data. *Communications in Statistics. Stochastic Models*, 13(4):703–721, 1997. Heavy tails and highly volatile phenomena.

Sidney Resnick and Catalin Stărică. Tail index estimation for dependent data. *The Annals of Applied Probability*, 8(4):1156–1183, 1998.

Sidney I. Resnick. Point processes, regular variation and weak convergence. *Advances in Applied Probability*, 18(1):66–138, 1986.

Sidney I. Resnick. *Heavy-Tail Phenomena.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

P. M. Robinson. Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23(3):1048–1072, 1995.

Holger Rootzén. Weak convergence of the tail empirical process for dependent sequences. *Stochastic Processes and Their Applications.*, 119(2):468–490, 2009.

Murray Rosenblatt. Independence and dependence. In *Proceedings of the 4th Berkeley Symposium in Mathematical Statistics and Probability, Vol. II*, pages 431–443. University of California Press, 1961.

Gennady Samorodnitsky. Long range dependence. *Foundations and Trends® in Stochastic Systems*, 1(3):163–257, 2006.

Gennady Samorodnitsky and Murad S. Taqqu. *Stable non-Gaussian random processes.* Stochastic Modeling. Chapman & Hall, New York, 1994. Stochastic models with infinite variance.

I. Resnick Sidney. *Extreme values, regular variation and point processes.* Applied Probability, Vol. 4,. New York, Springer-Verlag, 1987.

Philippe Soulier. Moment bounds and central limit theorem for functions of Gaussian vectors. *Statistics & Probability Letters*, 54(2):193–203, 2001.

Philippe Soulier. Best attainable rates of convergence for the estimation of the memory parameter. http://fr.arxiv.org/abs/0906.0180, 2009.

Donatas Surgailis. Stable limits of sums of bounded functions of long-memory moving averages with finite variance. *Bernoulli*, 10(2):327–355, 2004.

Donatas Surgailis and Marie-Claude Viano. Long memory properties and covariance structure of the EGARCH model. *ESAIM. Probability and Statistics*, 6:311–329, 2002.

Murad Taqqu, Walter Willinger, and Robert Sherman. Proof of a fundamental result in self-similar traffic modeling. *Computer Communication Review.*, 27, 1997.

Murad S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwande Gebiete*, 31:287–302, 1974/75.

Murad S. Taqqu. Convergence of integrated processes of arbitrary Hermite rank. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50(1):53–83, 1979.

Murad S. Taqqu and Joshua M. Levy. Using renewal processes to generate long range dependence and high variability. In E. Eberlein and M.S. Taqqu (eds), *Dependence in Probability and Statistics*. Boston, Birkhäuser, 1986.

Edward C. Titchmarsh. *Introduction to the theory of Fourier integrals*. Chelsea Publishing Co., New York, third edition, 1986.