

# MULTIVARIATE REGULAR VARIATION ON CONES: APPLICATION TO EXTREME VALUES, HIDDEN REGULAR VARIATION AND CONDITIONED LIMIT LAWS

SIDNEY I. RESNICK

ABSTRACT. We attempt to bring some modest unity to three subareas of heavy tail analysis and extreme value theory:

- limit laws for componentwise maxima of iid random variables;
- hidden regular variation and asymptotic independence;
- conditioned limit laws when one component of a random vector is extreme.

The common theme is multivariate regular variation on a cone and the three cases cited come from specifying the cones  $[0, \infty]^d \setminus \{\mathbf{0}\}$ ;  $(0, \infty]^d$ ; and  $[0, \infty] \times (0, \infty]$ .

## CONTENTS

1. Introduction	2
1.1. Some background.	2
1.1.1. Vague convergence.	2
2. Multivariate regular variation	3
2.1. The polar coordinate transform.	3
2.2. The one (or more) point uncompactification.	4
2.3. Multivariate regular variation of tail probabilities.	4
2.4. Some remarks on cones.	5
2.5. Form of the limit measure $\nu$ .	7
2.5.1. Expressing $\nu$ in terms of the angular measure $S$ .	7
3. Extreme value theory	8
3.1. Normalizing constants $a(\cdot), b(\cdot)$ .	9
3.2. Some conclusions.	10
4. Hidden regular variation	10
4.1. A construction of multivariate regularly varying distributions.	10
4.2. The <i>découpage de Lévy</i> .	11
4.3. Peaks over threshold and the POT method.	11
4.4. Asymptotic independence.	11
4.5. Hidden regular variation; definition and properties.	12
4.5.1. Consequences and properties.	13
4.5.2. Hidden regular variation and marginal distributions.	15
4.5.3. Examples of hidden regular variation.	15
4.5.4. The coefficient of tail dependence $\eta$ of Ledford and Tawn	18
5. Conditioned limit theorems	18
5.1. General formulation.	19
5.1.1. Standardizing the $Y$ -variable.	19
5.1.2. Standardizing the $X$ -variable.	19

---

*Key words and phrases.* Regular variation, risk, domain of attraction, copula, Pareto.

Sidney Resnick's research was partially supported by ARO Contract W911NF-07-1-0078 and NSA Grant H98230-06-1-0069 at Cornell University. Most of the synthesis and writing of this paper took place during May–June 2007 when Sidney Resnick was visiting the Lehrstuhl für Mathematische Statistik, Technical University München and grateful acknowledgement is made for hospitality and support.

5.1.3 The limit measure is a product.	21
6. Appendix	22
6.1. Vector notation.	22
6.2. Symbol and concept list.	23
References	23

## 1. INTRODUCTION

Several problems in extreme value theory and heavy tail analysis have as their core idea standard multivariate regular variation on a cone. Somewhat different theories and applications emerge by choosing different cones, but the common thread is that all of the problems have a reduction to regular variation of multivariate distributions on a particular cone. We consider the following three cones, labelled  $\mathbb{C}$ , and the associated three theories and application areas.

Cone	Application
$\mathbb{C} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$	multivariate extreme value theory
$\mathbb{C} = (\mathbf{0}, \infty]$	hidden regular variation, coefficient of tail dependence;
$\mathbb{C} = [0, \infty] \times (0, \infty]$	Conditioned limit theorems when one component is extreme.

TABLE 1. Three theories stemming from standard multivariate regular variation on three different cones.

**1.1. Some background.** We assume familiarity with the univariate theory of regularly varying functions. Fine surveys are available in Bingham et al. (1987), de Haan and Ferreira (2006), de Haan (1970), Feller (1971), Geluk and de Haan (1987), Resnick (2006, 1987), Seneta (1976). Vector notation and commonly used symbols are listed in an appendix at the end and can be consulted as needed. We briefly review the relationship of vague convergence and one dimensional regular variation of distribution tails. When a function  $U : \mathbb{R}_+ \mapsto \mathbb{R}_+$  is regularly varying with index  $\rho \in \mathbb{R}$ , we write  $U \in RV_\rho$ .

**1.1.1. Vague convergence.** Suppose  $(\mathbb{E}, \mathcal{E})$  is a nice space (that is, locally compact with countable base). We set

$$M_+(\mathbb{E}) = \{\text{all Radon measures on } \mathbb{E}\}.$$

So  $\mu \in M_+(\mathbb{E})$ , means that  $\mu$  is a measure on  $\mathcal{E}$  and  $\mu(K) < \infty$  for all compact sets  $K \in \mathcal{E}$ . Denote

$$C_K^+(\mathbb{E}) : \{f : f : \mathbb{E} \rightarrow \mathbb{R}_+, f \text{ continuous with compact support}\}.$$

If  $\mu_n, \mu \in M_+(\mathbb{E})$  then  $\mu_n \xrightarrow{v} \mu$  if for any  $f \in C_K^+(\mathbb{E})$ ,

$$\mu_n(f) := \int_{\mathbb{E}} f d\mu_n \rightarrow \int_{\mathbb{E}} f d\mu =: \mu(f).$$

See Daley and Vere-Jones (2003), Kallenberg (1983), Neveu (1977), Resnick (2006, 1987).

The following is from Resnick (1987) and Resnick (2006, page 62).

**Theorem 1.** *Regular variation of distribution function tails in dimension  $d = 1$ : Suppose  $X \geq 0$  is a random variable with distribution function  $F$  and set  $\bar{F}(x) = 1 - F(x) = P[X > x]$ . The following are equivalent:*

- (i)  $\bar{F} \in RV_{-\alpha}$ ,  $\alpha > 0$ .

(ii) There exists  $b_n \rightarrow \infty$  such that

$$n\bar{F}(b_n x) \rightarrow x^{-\alpha}, \quad x > 0.$$

(iii) There exists  $b_n \rightarrow \infty$  such that

$$(1.1) \quad \mu_n(\cdot) = nP\left[\frac{X}{b_n} \in \cdot\right] \xrightarrow{v} \nu_\alpha(\cdot),$$

in  $M_+(0, \infty]$ , where  $\nu_\alpha \in M_+(0, \infty]$  satisfies  $\nu_\alpha(x, \infty] = x^{-\alpha}$ , for  $x > 0$ .

**Remark 1.** (a) We can always set

$$(1.2) \quad b(t) = \left(\frac{1}{1-F}\right)^\leftarrow(t) = F^\leftarrow\left(1 - \frac{1}{t}\right),$$

which we call the *quantile function* and then define  $b_n = b(n)$ .

(b) Note the use of the unusual state space  $(0, \infty]$ , which is a *one point uncompactification* of the compact set  $[0, \infty]$ . See Resnick (2006, page 170). This topology makes neighborhoods of  $\infty$  relatively compact. This is required because vague convergence only controls behavior of measures on relatively compact sets and  $(x, \infty]$  is a natural set when dealing with right tail problems.

## 2. MULTIVARIATE REGULAR VARIATION

Suppose  $\mathbb{C}$  is a cone in  $\mathbb{R}^d$ ,  $d \geq 1$  so that

$$\mathbf{x} \in \mathbb{C} \text{ implies } t\mathbf{x} \in \mathbb{C}, \quad \forall t > 0.$$

We assume that  $\mathbf{1} \in \mathbb{C}$ . A function  $h : \mathbb{C} \rightarrow \mathbb{R}_+$  is regularly varying at  $\infty$  with limit function  $\lambda(\mathbf{x}) > 0$ , if for all  $\mathbf{x} \in \mathbb{C}$ ,

$$(2.1) \quad \lim_{t \rightarrow \infty} \frac{h(t\mathbf{x})}{h(t\mathbf{1})} = \lambda(\mathbf{x}) > 0, \quad \mathbf{x} \in \mathbb{C}.$$

For more details, see de Haan and Ferreira (2006), Resnick (2006).

We now consider an immediate implication: Fix  $\mathbf{x} \in \mathbb{C}$  and define

$$U(t) = h(t\mathbf{x}), \quad t > 0.$$

Then  $U \in RV_\rho$  for some  $\rho \in \mathbb{R}$ , since

$$\lim_{t \rightarrow \infty} \frac{U(ts)}{U(t)} = \lim_{t \rightarrow \infty} \frac{h(ts\mathbf{x})}{h(t\mathbf{x})} = \lim_{t \rightarrow \infty} \frac{h(ts\mathbf{x})}{h(t\mathbf{1})} \frac{h(t\mathbf{1})}{h(t\mathbf{x})} = \frac{\lambda(s\mathbf{x})}{\lambda(\mathbf{x})} = s^{\rho(\mathbf{x})}.$$

Fact: The exponent does not depend on  $\mathbf{x}$  (see Resnick (2006, page 169) and therefore the limit function  $\lambda$  is homogeneous:

$$(2.2) \quad \lambda(s\mathbf{x}) = s^\rho \lambda(\mathbf{x}).$$

**2.1. The polar coordinate transform.** Let  $\|\cdot\|$  be a norm on  $\mathbb{R}^d$  and denote the unit sphere of  $\mathbb{R}^d$  in this norm by

$$\mathfrak{N} = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = 1\}.$$

Define

$$T : \mathbb{R}^d \setminus \{\mathbf{0}\} \mapsto (0, \infty) \times \mathfrak{N}$$

by

$$(2.3) \quad T(\mathbf{x}) = \left(\|\mathbf{x}\|, \frac{\mathbf{x}}{\|\mathbf{x}\|}\right) =: (r, \mathbf{a}).$$

The inverse map  $T^\leftarrow : (0, \infty) \times \mathfrak{N} \mapsto \mathbb{R}^d \setminus \{\mathbf{0}\}$  yields

$$T^\leftarrow(r, \mathbf{a}) = r\mathbf{a}.$$

Note we do not define  $T$  for  $\mathbf{x} = \mathbf{0}$  or if  $\mathbf{x}$  has infinite components.

**2.2. The one (or more) point uncompactification.** The *one point uncompactification* identifies the compact sets in a compact space modified by the removal of one point. It is based on the following simple result. (See Resnick (2006, page 170).)

**Proposition 1.** *Let  $(\mathbb{X}, \mathcal{T})$  be a topological space with open sets  $\mathcal{T}$ . From  $\mathbb{X}$ , chop away  $\mathbb{D}$  to get*

$$\mathbb{X}^\sharp = \mathbb{X} \setminus \mathbb{D} = \mathbb{X} \cap \mathbb{D}^C$$

with relative topology

$$(2.4) \quad \mathcal{T}^\sharp = \mathcal{T} \cap \mathbb{D}^C = \mathcal{T} \cap \mathbb{X}^\sharp.$$

Then the compact sets of  $\mathbb{X}^\sharp$  are

$$(2.5) \quad \mathcal{K}(\mathbb{X}^\sharp) = \{K \in \mathcal{K}(X) : K \cap \mathbb{D} = \emptyset\}.$$

So the compact sets in the new, modified space are the original compact sets prior to modification, provided they miss the deleted set  $\mathbb{D}$ .

**Example 1.** Here are some examples of modified spaces that have been useful.

- (1)  $(0, \infty] = [0, \infty] \setminus \{0\}$  has compacta which are closed in  $[0, \infty]$  and do not contain 0.
- (2)  $[0, \infty]^d \setminus \{\mathbf{0}\} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$  has compacta which are compact in  $[0, \infty]^d$  and do not contain  $\{\mathbf{0}\}$ . We informally refer to such sets as bounded away from  $\mathbf{0}$ .
- (3)  $(0, \infty]^d = [0, \infty]^d \setminus \cup_{i=1}^d \{te_i, t \geq 0\}$ , where

$$e_i = (0, \dots, 0, 1, 0, \dots, 0).$$

The compacta of this compact set modified by removal of the axes are closed subsets of  $[0, \infty]^d$  not touching the axes. We refer to such sets as being neighborhoods of  $\infty$ .

- (4)  $[0, \infty] \times (0, \infty] = [0, \infty]^2 \setminus \{t(1, 0), t > 0\}$ . The compact sets are closed subsets of  $[0, \infty]^2$  not touching the horizontal axis.
- (5)  $(0, \infty] \times \aleph_+$ , where  $\aleph_+ = \{\mathbf{x} \geq \mathbf{0} : \|\mathbf{x}\| = 1\}$ .

**2.3. Multivariate regular variation of tail probabilities.** We characterize regular variation of tail probabilities on the cone  $\mathbb{C} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$ . Suppose  $\mathbf{Z} \geq \mathbf{0}$  is a random vector on with distribution  $F$  concentrating on  $[\mathbf{0}, \infty)$ . The following are equivalent but the symbols  $(\nu, b(\cdot), b_n)$  may differ slightly in each instance.

- (1) There exists a Radon measure on  $\mathbb{C}$  called  $\nu$  such that

$$(2.6) \quad \lim_{t \rightarrow \infty} \frac{1 - F(t\mathbf{x})}{1 - F(t\mathbf{1})} = \lim_{t \rightarrow \infty} \frac{P[t^{-1}\mathbf{Z} \in [0, \mathbf{x}]^c]}{P[t^{-1}\mathbf{Z} \in [0, \mathbf{1}]^c]} = \nu([\mathbf{0}, \mathbf{x}]^c),$$

for all  $\mathbf{x} \in [\mathbf{0}, \infty) \setminus \{\mathbf{0}\}$  which are continuity points of the limit  $\nu([\mathbf{0}, \cdot]^c)$ .

- (2) There exist  $b(t) \rightarrow \infty$  and a Radon measure  $\nu$  on  $\mathbb{C}$  such that

$$(2.7) \quad P[\mathbf{Z}/b(t) \in \cdot] \xrightarrow{v} \nu(\cdot), \quad t \rightarrow \infty, \text{ in } M_+(\mathbb{C}).$$

- (3) There exist  $b_n \rightarrow \infty$ , and a Radon measure  $\nu$  on  $\mathbb{C}$  such that in  $M_+(\mathbb{C})$ ,

$$nP[\mathbf{Z}/b_n \in \cdot] \xrightarrow{v} \nu(\cdot), \quad n \rightarrow \infty.$$

- (4) Let  $\aleph_+ = \aleph \cap [0, \infty]^d$ . Then there exist a probability measure  $S(\cdot)$  on Borel subsets of  $\aleph_+$  called the angular measure, and a function  $b(t) \rightarrow \infty$  such that with

$$(R, \Theta) = \left( \|\mathbf{Z}\|, \frac{\mathbf{Z}}{\|\mathbf{Z}\|} \right)$$

we have

$$(2.8) \quad tP\left[\left(\frac{R}{b(t)}, \Theta\right) \in \cdot\right] \xrightarrow{v} c\nu_\alpha \times S$$

in  $M_+((0, \infty] \times \mathbb{N}_+)$  or, equivalently,

$$(2.9) \quad nP\left[\left(\frac{R}{b_n}, \Theta\right) \in \cdot\right] \xrightarrow{v} c\nu_\alpha \times S.$$

The standard proofs are in Resnick (2006, page 173).

We will need several extensions of the definition of multivariate regular variation given in (2.6), (2.7), or (2.8). We should consider other cones besides  $[\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$ . We should consider what happens if the components of  $\mathbf{Z}$  not normalized by the same scaling function  $b(t)$ . And finally, we should assess methods for analyzing dependence structure and how to statistically estimate such dependence.

**2.4. Some remarks on cones.** Let  $\bar{\mathbb{C}}$  be a closed compact cone in  $[-\infty, \infty]^d$  with  $\mathbf{0} \in \bar{\mathbb{C}}$ . Suppose

$$\mathbb{C} := \bar{\mathbb{C}} \setminus (\{\mathbf{0}\} \cup \mathbb{D})$$

is still a cone. Compact subsets of  $\mathbb{C}$  are the compact sets of  $\bar{\mathbb{C}}$  not touching  $\mathbf{0}$  or  $\mathbb{D}$ . The distribution of  $\mathbf{Z}$  is regularly varying on  $\mathbb{C}$  if there exists a scaling function  $b(t) \rightarrow \infty$ , and a limit Radon measure  $\nu \in M_+(\mathbb{C})$  such that

$$tP\left[\frac{\mathbf{Z}}{b(t)} \in \cdot\right] \rightarrow \nu(\cdot), \quad \text{in } M_+(\mathbb{C}).$$

**Example 2.** Consider the following examples.

	Cone	$\mathbb{D}$
1.	$\mathbb{C} = [-\infty, \infty]^d \setminus \{\mathbf{0}\}$	$\mathbb{D} = \emptyset$
2.	$\mathbb{C} = (0, \infty]^2$	$\mathbb{D} = (0, \infty] \times \{0\} \cup \{0\} \times (0, \infty]$
3.	$\mathbb{C} = [0, \infty] \times (0, \infty]$	$\mathbb{D} = (0, \infty] \times \{0\}$

TABLE 2. Three cones which are used for various applications.

Note that ‘‘Radon’’ means something different for each example since the notion of compact is different for each example. For instance:

- In the first cone,  $\nu\{\mathbf{x} : \|\mathbf{x}\| > 1\} < \infty$ .
- Relative to the second cone,  $\nu\{\mathbf{x} > \mathbf{0} : \|\mathbf{x}\| > 1\}$  is not necessarily finite.
- For the third cone in the table,  $\nu\{(x, y) : \|(x, y)\| > 1\}$  is also not necessarily finite.

**Example 3.** Consider  $d = 1$  and suppose  $\mathbb{C} = [-\infty, \infty] \setminus \{0\}$ . Suppose  $Z$  is a random variable with distribution  $F$  and suppose  $F$  has a regularly varying tail. There is no mass at  $\{\pm\infty\}$  and regular variation means

$$(2.10) \quad tP[Z/b(t) \in \cdot] \rightarrow \nu(\cdot) \quad \text{in } M_+([-\infty, \infty] \setminus \{0\}).$$

This implies that

$$(2.11) \quad tP[Z/b(t) > x] \rightarrow \nu(x, \infty], \quad x > 0, \quad t \rightarrow \infty,$$

and from the sequential form of regular variation, this yields

$$(2.12) \quad \nu(x, \infty] = c_+ x^{-\alpha}, \quad \alpha > 0, \quad c_+ \geq 0.$$

Also, in a similar manner,

$$(2.13) \quad tP[Z/b(t) \leq -x] \rightarrow \nu[-\infty, -x], \quad x > 0,$$

which implies

$$(2.14) \quad \nu[-\infty, -x] = c_- x^{-\alpha}.$$

Why is the  $\alpha$  the same for both tails? There is only an issue if  $c_+ > 0$ ,  $c_- > 0$  since otherwise  $\nu$  only concentrates on a half line emanating from the origin. So suppose  $c_+ > 0$ ,  $c_- > 0$ . Then we have as  $t \rightarrow \infty$ ,

$$t\bar{F}(b(t)x) \rightarrow x^{-\alpha}, \quad x > 0,$$

or equivalently after taking reciprocals

$$\frac{1}{t(1 - F(b(t)x))} = \frac{U(b(t)x)}{t} \rightarrow x^\alpha,$$

where  $U = 1/(1 - F)$  and we may take  $U^\leftarrow(t) = b(t)$ . But  $U \in RV_\alpha$  implies  $b \in RV_{1/\alpha}$ .

Likewise for the other tail, as  $t \rightarrow \infty$ ,

$$tF(b(t)(-x)) \rightarrow \nu[-\infty, -x], \quad x > 0,$$

leads to

$$\nu(-\infty, -x] = c_- x^{-\beta}$$

for some  $\beta > 0$  which would imply  $b(t) \in RV_{1/\beta}$  so we have  $\beta = \alpha$ .

We conclude the limit measure is of the form

$$(2.15) \quad \nu(dx) = c_+ \alpha x^{-\alpha-1} dx 1_{(0, \infty)}(x) + c_- |x|^{-\alpha-1} dx 1_{[-\infty, 0)}(x),$$

and also that

$$(a) \quad P[|X| > x] \in RV_{-\alpha},$$

$$(b) \quad \lim_{t \rightarrow \infty} \frac{P[X > t]}{P[|X| > t]} = \lim_{t \rightarrow \infty} \frac{tP[X/b(t) > 1]}{tP[|X|/b(t) > 1]} = \frac{\nu(1, \infty]}{\nu\{x : |x| > 1\}} = \frac{c_+}{c_+ + c_-} =: p,$$

and

$$(c) \quad \lim_{t \rightarrow \infty} \frac{P[X < -t]}{P[|X| > t]} = \frac{c_-}{c_+ + c_-} =: q$$

where  $p + q = 1$ .

If  $0 < \alpha < 2$ , these are the necessary and sufficient conditions for convergence of sums of iid random variables with common distribution  $F$  to a stable law.

**Example 4.** Suppose  $d = 2$  and let  $\mathbf{Z} = (Z^{(1)}, Z^{(2)})$  where  $Z^{(1)}, Z^{(2)}$  are iid, and

$$P[Z^{(i)} > x] = x^{-1}, \quad x \geq 1.$$

The vector  $\mathbf{Z}$  has two different regular variation properties on two different cones.

(a)  $\mathbf{Z}$  has a regularly varying distribution on  $[0, \infty]^2 \setminus \{\mathbf{0}\}$ . To see this, note that for  $i = 1, 2$ ,

$$tP[Z^{(i)} > tx] = t(tx)^{-1} = x^{-1}, \quad x > 0, \text{ for all } t \text{ big,}$$

so we have marginal regular variation. For  $\mathbf{x} = (x^{(1)}, x^{(2)})$ , we have

$$(2.16) \quad \begin{aligned} \lim_{t \rightarrow \infty} tP\{\mathbf{Z} \leq t\mathbf{x}\}^c &= \lim_{t \rightarrow \infty} tP[Z^{(2)} > tx^{(2)}] + \lim_{t \rightarrow \infty} tP[Z^{(1)} > tx^{(1)}] \\ &\quad - \lim_{t \rightarrow \infty} tP[Z^1 > tx^{(1)}]P[Z^{(2)} > tx^{(2)}] \\ &= (x^{(2)})^{-1} + (x^{(1)})^{-1} - 0 = \nu([0, \mathbf{x}]^c). \end{aligned}$$

This implies that  $\nu$  concentrates on the axes and has no mass on  $(0, \infty]^2$ .

(b)  $\mathbf{Z}$  has a distribution which is regularly varying on  $(0, \infty]^2$ : To see this, observe that

$$(2.17) \quad tP\left[\frac{\mathbf{Z}}{\sqrt{t}} \in (\mathbf{x}, \infty]\right] = \sqrt{t}P[Z^{(1)} > \sqrt{t}x^{(1)}]\sqrt{t}P[Z^{(2)} > \sqrt{t}x^{(2)}] \rightarrow (x^{(1)})^{-1}(x^{(2)})^{-1}.$$

So on  $[0, \infty]^2 \setminus \{\mathbf{0}\}$ ,  $\mathbf{Z}$  has a distribution which is regularly varying with limit measure  $\nu$  given in (2.16). On  $(0, \infty]^2$ ,  $\mathbf{Z}$  has a distribution which is regularly varying with limit measure  $\nu_*(\mathbf{x}, \infty] = (x^{(1)}x^{(2)})^{-1}$  and  $b_*(t) = \sqrt{t}$ .

**2.5. Form of the limit measure  $\nu$ .** Suppose that  $\nu$  is a Radon measure on Borel subsets of the cone  $\mathbb{C}$ , that  $b(t) \rightarrow \infty$  is a scaling function and  $F$  is a probability distribution such that

$$tF(b(t)\cdot) \xrightarrow{v} \nu(\cdot), \quad \text{in } M_+(\mathbb{C}).$$

For the cones we consider, we have the following properties:

- (i) The scaling function  $b(\cdot)$  satisfies  $b \in RV_{1/\alpha}$  for some  $\alpha > 0$ .
- (ii) The limit measure  $\nu(\cdot)$  has a homogeneity property

$$\nu(t\cdot) = t^{-\alpha}\nu(\cdot), \quad \text{on } \mathbb{C}.$$

- (iii) The scaling property (ii) of the limit measure  $\nu$  is equivalent to a product form when using polar coordinates. To see this, suppose  $\Lambda \in \mathcal{B}(\mathfrak{N} \cap \mathbb{C})$ . Then

$$\begin{aligned} \nu\{\mathbf{x} : \|\mathbf{x}\| > r, \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \Lambda\} &= \nu\{\mathbf{x} : \|r^{-1}\mathbf{x}\| > 1, \frac{r^{-1}\mathbf{x}}{\|r^{-1}\mathbf{x}\|} \in \Lambda\} \\ &= \nu\{r\mathbf{y} : \|\mathbf{y}\| > 1, \frac{\mathbf{y}}{\|\mathbf{y}\|} \in \Lambda\} \\ (2.18) \qquad \qquad \qquad &= r^{-\alpha}\nu\{\mathbf{y} : \|\mathbf{y}\| > 1, \frac{\mathbf{y}}{\|\mathbf{y}\|} \in \Lambda\}. \end{aligned}$$

Define a measure  $S$  on  $\mathcal{B}(\mathfrak{N} \cap \mathbb{C})$  by

$$S(\Lambda) := \nu\{\mathbf{y} : \|\mathbf{y}\| > 1, \frac{\mathbf{y}}{\|\mathbf{y}\|} \in \Lambda\}$$

and we conclude

$$(2.19) \qquad \qquad \qquad \nu\{\mathbf{x} : \|\mathbf{x}\| > r, \frac{\mathbf{x}}{\|\mathbf{x}\|} \in \Lambda\} = r^{-\alpha}S(\Lambda).$$

Call  $S$  the *angular or spectral measure*.

**Remark 2.** Note the following features.

- (1) If  $\mathfrak{N} \cap \mathbb{C}$  is compact,  $S$  is finite in which case we assume it is a probability measure.
- (2) If  $\mathbb{C} = [0, \infty]^d \setminus \{\mathbf{0}\}$ , then  $\mathfrak{N} \cap \mathbb{C} = \{\mathbf{x} \geq \mathbf{0} : \|\mathbf{x}\| = 1\}$  is compact.
- (3) If  $\mathbb{C} = (0, \infty]^2$ , then  $\mathfrak{N} \cap \mathbb{C} = \{\mathbf{x} > \mathbf{0} : \|\mathbf{x}\| = 1\}$  is not compact and  $S$  is not necessarily finite.
- (4) If  $\mathbb{C} = [0, \infty] \times (0, \infty]$ ,

$$\mathfrak{N} \cap \mathbb{C} = \{(x, y) : 0 \leq x \leq \infty, \infty \geq y > 0, \|(x, y)\| = 1\}$$

is not compact.

**2.5.1. Expressing  $\nu$  in terms of the angular measure  $S$ .** Suppose  $B \in \mathcal{B}(\mathbb{C})$ . Then integrating by means of polar coordinates we obtain

$$\nu(B) = \iint_{\{(r, \mathbf{a}) : r \cdot \mathbf{a} \in B, r > 0, \|\mathbf{a}\| = 1\}} \alpha r^{-\alpha-1} dr S(d\mathbf{a}).$$

For example if  $\mathbb{C} = [0, \infty]^d \setminus \{\mathbf{0}\}$ ,

$$(2.20) \qquad \qquad \qquad \nu([\mathbf{0}, \mathbf{x}]^c) = \iint_{\{r\mathbf{a} \leq \mathbf{x}\}^c} \alpha r^{-\alpha-1} dr S(d\mathbf{a}).$$

Now

$$\begin{aligned} \{r\mathbf{a} \leq \mathbf{x}\}^c &= \{(r, \mathbf{a}) : r > \frac{x^{(i)}}{a^{(i)}}, \text{ for some } i\} \\ &= \{(r, \mathbf{a}) : r > \bigwedge_{i=1}^d \frac{x^{(i)}}{a^{(i)}}\}. \end{aligned}$$

So integrate first in (2.20) with respect to  $r$  via Fubini's theorem and then

$$(2.21) \quad \begin{aligned} \nu([\mathbf{0}, \mathbf{x}]^c) &= \int_{\mathbb{N} \cap \mathbb{C}} \left( \bigwedge_{i=1}^d \frac{x^{(i)}}{a^{(i)}} \right)^{-\alpha} S(d\mathbf{a}) \\ &= \int_{\mathbb{N} \cap \mathbb{C}} \bigvee_{i=1}^d \left( \frac{a^{(i)}}{x^{(i)}} \right)^{\alpha} S(d\mathbf{a}). \end{aligned}$$

### 3. EXTREME VALUE THEORY

Regular variation on the cone  $[\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$  leads to the classical theory of multivariate extreme value distributions as limits of componentwise maxima of an iid sample. See de Haan and Resnick (1977), Resnick (1987, Chapter 5) and de Haan and Ferreira (2006).

Suppose  $\mathbf{X}_*$  has distribution  $F_*$  which is regularly varying on the cone  $[\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$ . We suppose the regular variation is in *standard form*, which means we may set  $b(t) = t$  and there exists a Radon measure  $\nu_*$  such that

$$(3.1) \quad tP\left[\frac{\mathbf{X}_*}{t} \in \cdot\right] \xrightarrow{v} \nu_*(\cdot), \quad \text{in } M_+([\mathbf{0}, \infty] \setminus \{\mathbf{0}\}).$$

From this we get the multivariate extreme value distributions in the following way, using the theory of extended regularly varying functions (Bingham et al., 1987, Geluk and de Haan, 1987). The story goes like this: Suppose  $a(\cdot) > 0$ ,  $b(\cdot) \in \mathbb{R}$  are measurable functions on  $(0, \infty)$  satisfying,

$$(3.2) \quad \frac{b(tx) - b(t)}{a(t)} \rightarrow \psi(x), \quad x > 0, t \rightarrow \infty,$$

where  $\psi \neq 0$ , and  $\psi$  is not constant. Then  $\psi$  must be of the form

$$(3.3) \quad \psi(x) = \begin{cases} k \left( \frac{x^\rho - 1}{\rho} \right), & \rho \in \mathbb{R}, \rho \neq 0, x > 0 \\ k \log x & \rho = 0, x > 0 \end{cases}$$

for some  $k \neq 0$ . Furthermore,

- (a) If  $\rho > 0$ , then  $b(\cdot) \in RV_\rho$ , and  $b(t) \sim \frac{1}{\rho} a(t)$ , and (3.2) really says that

$$\frac{b(tx)}{b(t)} \rightarrow x^\rho.$$

- (b) If  $\rho = 0$ , then (3.2) defines  $b(\cdot)$  as a  $\Pi$ -varying function. ( $\Pi$ -varying functions are a subclass of the slowly varying functions; see Bingham et al. (1987), de Haan and Ferreira (2006), de Haan (1970), Geluk and de Haan (1987), Resnick (1987).) Also,  $a(\cdot)$  is slowly varying and  $b(t)/a(t) \rightarrow \infty$ .

- (c) If  $\rho < 0$ , then  $b(\infty) := \lim_{t \rightarrow \infty} b(t)$  exists finite and

$$(3.4) \quad b(\infty) - b(t) \in RV_\rho, \quad b(\infty) - b(t) \sim \frac{1}{|\rho|} a(t).$$

The inverted form of extended regular variation: If the function  $b$  is nondecreasing with inverse  $b^\leftarrow$ , we can invert (3.2). Supposing for convenience that  $k = 1$ , the inversion of (3.2) yields

$$(3.5) \quad \frac{b^\leftarrow(b(t) + ya(t))}{t} \rightarrow \psi^\leftarrow(y) = (1 + \rho y)^{1/\rho}, \quad (1 + \rho y > 0).$$

Now assume (3.1) and suppose  $(b^{(i)}, a^{(i)})$ ,  $i = 1 \dots d$  satisfy (3.2) with limits  $\psi_i$  and parameters  $\rho_i$  and that each  $b^{(i)}$  is non-decreasing. Then

$$\begin{aligned} tP\left\{ \left[ \frac{b^{(i)}(X_*^{(i)}) - b^{(i)}(t)}{a^{(i)}(t)} \leq x^{(i)}, i = 1 \dots d \right]^c \right\} \\ = tP\left[ \frac{X_*^{(i)}}{t} \leq b^{(i)\leftarrow}(a^{(i)}(t)x^{(i)} + b^{(i)}(t))/t, i = 1 \dots d \right]^c \end{aligned}$$

$$(3.6) \quad \rightarrow \nu_* (\{\mathbf{y} : y^{(i)} \leq \psi_i^{\leftarrow}(x^{(i)}), i = 1 \dots d\}^c).$$

Where does the relation (3.6) lead? Set  $X^{(i)} = b^{(i)}(X^{(i)})$  and suppose  $b^{(i)}$  non-decreasing. Then (3.6) can be rephrased as

$$tP\{[X^{(i)} \leq a^{(i)}(t)x^{(i)} + b^{(i)}(t), i = 1, \dots, d]^c\} \rightarrow \nu_* (\{\mathbf{y} : y^{(i)} \leq \psi_i^{\leftarrow}(x^{(i)}); i = 1, \dots, d\}^c)$$

or, replacing  $t$  with  $n$

$$(3.7) \quad nP\{[X^{(i)} \leq a^{(i)}(n)x^{(i)} + b^{(i)}(n); i = 1, \dots, d]^c\} \rightarrow \nu_* (\{\mathbf{y} : y^{(i)} \leq \psi_i^{\leftarrow}(x^{(i)}); i = 1, \dots, d\}^c).$$

If  $np_n \rightarrow \ell \in (0, \infty)$  so that  $p_n \rightarrow 0$ , then  $n(-\log(1 - p_n)) \rightarrow \ell$ , so

$$n(-\log P[X^{(i)} \leq a^{(i)}(n)x^{(i)} + b^{(i)}(n); i = 1, \dots, d]) \rightarrow \nu_* (\{\mathbf{y} : y^{(i)} \leq \psi_i^{\leftarrow}(x^{(i)}); i = 1, \dots, d\}^c)$$

or exponentiating

$$(3.8) \quad (P[X^{(i)} \leq a^{(i)}(n)x^{(i)} + b^{(i)}(n), i = 1, \dots, d])^n \rightarrow \exp\{-\nu_* (\{\mathbf{y} : y^{(i)} \leq \psi_i^{\leftarrow}(x^{(i)}), i = 1, \dots, d\}^c)\}.$$

Since

$$P\left[\bigvee_{j=1}^n \mathbf{X}_j \leq \mathbf{x}\right] = (P[\mathbf{X}_1 \leq \mathbf{x}])^n,$$

when  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are iid random vectors and “ $\leq$ ” is interpreted componentwise, we conclude, as  $n \rightarrow \infty$ ,

$$(3.9) \quad P\left[\bigvee_{j=1}^n \left(\frac{X_j^{(i)} - b^{(i)}(n)}{a^{(i)}(n)}\right) \leq x^{(i)}; i = 1, \dots, d\right] \rightarrow \exp\{-\nu_* (\{\mathbf{y} : y^{(i)} \leq \psi_i^{\leftarrow}(x^{(i)}); i = 1, \dots, d\}^c)\}.$$

The limit is the form of the general max-stable distributions which are approximations for distributions of componentwise maxima of iid random vectors whose common distribution is in a maximal domain of attraction.

**3.1. Normalizing constants  $a(\cdot), b(\cdot)$ .** Where do the functions  $a(\cdot), b(\cdot)$  satisfying (3.2) come from? Suppose  $d = 1$  and let  $X_1, \dots, X_n$  be iid random variables with common distribution function  $F(x)$  satisfying a one dimensional version of (3.9); that is, suppose there exist functions  $a(t) > 0, b(t) \in \mathbb{R}$  and a distribution  $G$  such that

$$(3.10) \quad P\left[\bigvee_{j=1}^n \frac{X_j - b(n)}{a(n)} \leq x\right] \rightarrow G(x).$$

Then we say  $F$  is in the *maximal domain of attraction of  $G$*  and write  $F \in MDA(G)$ . The limit relation (3.10) leads to

$$P\left[\bigvee_{j=1}^n X_j \leq a(n)x + b(n)\right] = F^n(a(n)x + b(n)) \rightarrow G(x),$$

so taking logarithms we get

$$n(-\log F(a(n)x + b(n))) \rightarrow -\log G(x).$$

This leads to

$$(3.11) \quad t(-\log F(a(t)x + b(t))) \rightarrow -\log G(x)$$

and

$$t(1 - F(a(t)x + b(t))) \rightarrow -\log G(x)$$

and then taking reciprocals,

$$(3.12) \quad \frac{1}{t} \left( \frac{1}{1 - F(a(t)x + b(t))} \right) \rightarrow \frac{1}{-\log G(x)}.$$

Invert to get

$$(3.13) \quad \frac{\left(\frac{1}{1-F}\right)^{\leftarrow}(yt) - b(t)}{a(t)} \rightarrow \left(\frac{1}{-\log G}\right)^{\leftarrow}(y), \quad y > 0.$$

So we may set  $b(t) = \left(\frac{1}{1-F}\right)^{\leftarrow}(t)$  and  $a(t) = b(te) - b(t)$ .

**3.2. Some conclusions.** Assume again that  $d > 1$ . We summarize the procedure outlined in this section.

- (a) Start with a random vector  $\mathbf{X}_*$  satisfying the standard form of regular variation (3.1) on  $[0, \infty] \setminus \{0\}$ .
- (b) Assume  $b^{(i)} : (0, \infty) \mapsto (0, \infty)$  is non-decreasing and assume there exist  $a^{(i)}(\cdot) > 0$  such that for  $i = 1, \dots, d$  each  $b^{(i)}, a^{(i)}$  pair satisfies (3.2).
- (c) Transform  $\mathbf{X}_*$  to  $\mathbf{X} = (X^{(1)}, \dots, X^{(d)})$  by

$$X^{(i)} = b^{(i)}(X_*^{(i)}); \quad i = 1, \dots, d.$$

- (d) Then  $\mathbf{X}$  satisfies (3.6) and if  $(\mathbf{X}_1, \dots, \mathbf{X}_n)$  are iid and  $\mathbf{X}_1 \stackrel{d}{=} \mathbf{X}$  we get (3.9).
- (e) We can choose

$$(3.14) \quad b^{(i)} = \left(\frac{1}{1 - F_{(i)}}\right)^{\leftarrow}$$

where  $F_{(i)}$  is a distribution in the maximal domain of attraction of a one dimensional extreme value distribution.

#### 4. HIDDEN REGULAR VARIATION

Hidden regular variation is regular variation corresponding to the cone  $\{\mathbf{x} \geq \mathbf{0} : x^{(i)} \wedge x^{(j)} > 0, \text{ for some } 1 \leq i, j \leq d\}$ . Before reviewing this, there are several transitional topics that should be understood.

**4.1. A construction of multivariate regularly varying distributions.** In this section  $\mathbb{C} = [0, \infty] \setminus \{0\}$ . Suppose the random variable  $R$  satisfies  $R \geq 1$  and

$$P[R > r] = r^{-\alpha}, \quad r > 1,$$

and  $\Theta$  is a random variable concentrating on  $\mathfrak{N}_+ = \mathfrak{N} \cap \mathbb{R}_+^d$  with distributions  $S(d\theta)$ . We assume  $R$  and  $\Theta$  are independent, written  $R \perp \Theta$ . Then for  $r > 0$ ,  $\Lambda \in \mathcal{B}(\mathfrak{N}_+)$ , we have

$$(4.1) \quad \begin{aligned} tP\left[\frac{R}{t^{1/\alpha}} > r, \Theta \in \Lambda\right] &= tP\left[\frac{R}{t^{1/\alpha}} > r\right]P[\Theta \in \Lambda] \\ &\rightarrow r^{-\alpha}S(\Lambda), \quad \text{as } t \rightarrow \infty. \end{aligned}$$

Note the same conclusion would hold if we only assumed  $P[R > r] \in RV_{-\alpha}$  and replaced  $t^{1/\alpha}$  by

$$b(t) := \left(\frac{1}{1 - P[R > \cdot]}\right)^{\leftarrow}(t).$$

Let  $T$  be the polar coordinate transform:  $T\mathbf{x} = (\|\mathbf{x}\|, \mathbf{x}/\|\mathbf{x}\|) = (r, \mathbf{a})$  and define

$$(4.2) \quad \mathbf{Z} = T^{\leftarrow}(R, \Theta).$$

Then  $\mathbf{Z}$  has a regularly varying distribution

$$(4.3) \quad tP\left[\frac{\mathbf{Z}}{b(t)} \in \cdot\right] \rightarrow \nu(\cdot)$$

where

$$\nu \circ T^{\leftarrow} = \alpha r^{-\alpha-1} dr S(d\theta).$$

**4.2. The *découpage de Lévy*.** The following result of Lévy helps to understand "peaks over threshold." See Resnick (1987) for more detail on the proof.

**Theorem 2** (Lévy). *Let  $\{X_n, n > 1\}$  be iid random elements of a nice space  $(\mathbb{E}, \mathcal{E})$ . Fix  $B \in \mathcal{E}$  with  $P[X_1 \in B] > 0$  and define  $\tau_0^\pm = 0$  and*

$$\tau_i^+ = \inf\{j > \tau_{i-1}^+ : X_j \in B\}, \quad \tau_i^- = \inf\{j > \tau_{i-1}^- : X_j \in B^c\}.$$

*Let  $K_n = \sup\{i : \tau_i^+ \leq n\}$  be the (renewal) counting function for visits to  $B$ . Then*

$$\{X_{\tau_j^+}\}, \{X_{\tau_j^-}\}, \{K_n\}$$

*are independent with*

$$P[X_{\tau_1^+} \in A] = P[X_1 \in A | X_1 \in B], \quad A \subset B$$

*and*

$$P[X_{\tau_1^-} \in A] = P[X_1 \in A | X_1 \in B^c], \quad A \subset B^c.$$

*Further  $\{K_n\}$  is a renewal counting function,  $E(K_n) = nP[X_1 \in B]$  and both  $\{X_{\tau_j^+}\}, \{X_{\tau_j^-}\}$  are iid.*

**4.3. Peaks over threshold and the POT method.** Peaks over threshold is an important concept. A good way to understand it is by means of the *découpage de Lévy*.

Suppose  $\{\mathbf{Z}_j, j \geq 1\}$  are iid  $\mathbb{R}_+^d$ -valued random vectors and set  $B$  in the *découpage* to be

$$B = \{\mathbf{x} \in \mathbb{R}_+^d : \|\mathbf{x}\| > t\}.$$

Then for  $s > 1$ ,  $\Lambda \in \mathcal{B}(\mathbb{N}_+)$  and  $ts > 1$ ,  $\{\mathbf{Z}_{\tau_j^+}\}$  are iid and

$$(4.4) \quad \begin{aligned} P[R_{\tau_1^+} > ts, \Theta_{\tau_1^+} \in \Lambda] &= P[R_1 > ts, \Theta_1 \in \Lambda | R_1 > t] \\ &= \frac{P[R_1 > ts, \Theta_1 \in \Lambda]}{P[R_1 > t]} \approx s^{-\alpha} S(\Lambda), \end{aligned}$$

where " $\approx$ " means this is only true exactly as  $t \rightarrow \infty$ . The "peaks over threshold" or POT philosophy says *pretend* there is actual equality for some  $t$ . Then the method suggests,

(a) Consider the subsample of exceedances  $\{\mathbf{Z}_{\tau_j^+}, j \geq 1\}$  with distribution in polar coordinates:

$$P[T(\mathbf{Z}_{\tau_1^+}) \in drd\theta] = \alpha r^{-\alpha-1} dr S(d\theta), \quad r \geq 1, \theta \in \mathbb{N}_+.$$

(b) Think of exceedances as coming from the construction of Section 4.1.

This method is popular for inference purposes but there is no obvious way to quantify error when replacing " $\approx$ " with " $=$ ". See Coles (2001) for a convincing and clear discussion of advantages; there is also further material in Resnick (2006).

**4.4. Asymptotic independence.** Suppose  $\mathbf{Z}$  is an  $\mathbb{R}_+^d$ -valued random vector with a distribution which is regularly varying on  $[\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$  so that there exists  $b(t) \rightarrow \infty$  such that

$$tP[\mathbf{Z}/b(t) \in \cdot] \rightarrow \nu(\cdot), \quad \text{in } M_+([\mathbf{0}, \infty] \setminus \{\mathbf{0}\}).$$

Then  $\mathbf{Z}$  (or its distribution) is asymptotically independent if

$$(4.5) \quad \nu\{\mathbf{x} \in [\mathbf{0}, \infty] \setminus \{\mathbf{0}\} : x^{(i)} \wedge x^{(j)} > 0, \text{ for some } 1 \leq i < j \leq d\} = 0$$

or equivalently for  $\mathbf{x} \geq \mathbf{0}$ ,  $\mathbf{x} \neq \mathbf{0}$ ,

$$(4.6) \quad \nu([\mathbf{0}, \mathbf{x}]^c) = \sum_{i=1}^d c_i (x^{(i)})^{-\alpha}.$$

Therefore,  $\nu$  concentrates on the lines through  $\mathbf{0}$  of the form  $\{t\mathbf{e}_i, t > 0\}$ ,  $i = 1, \dots, d$  where  $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$  with "1" in the  $i$ -th spot.

Recall the example  $\mathbf{Z} = (Z^{(1)}, \dots, Z^{(d)})$  where  $Z^{(i)}$ ,  $1 \leq i \leq d$  are iid and  $P[Z^{(i)} > x] = x^{-\alpha}$ ,  $x \geq 1$  leads to asymptotic independence, which is not surprising since independence ought to imply asymptotic independence.

There are two main reasons for the name *asymptotic independence* although this modifier has lots of other meanings in probability and statistics so we make no claim to universal appropriateness.

- (1) Suppose  $d = 2$  and  $\mathbf{Z}$  possesses asymptotic independence. Then

$$\begin{aligned} \lim_{t \rightarrow \infty} P[Z^{(2)} > t | Z^{(1)} > t] &= \lim_{n \rightarrow \infty} \frac{P[Z^{(1)} > b(n), Z^{(2)} > b(n)]}{P[Z^{(1)} > b(n)]} \\ &= \lim_{n \rightarrow \infty} cnP[Z^{(1)} > b(n), Z^{(2)} > b(n)] \\ &= c\nu(\mathbf{1}, \infty) = 0 \end{aligned}$$

since  $\nu$  concentrates on  $\{(0, x) : x > 0\} \cup \{(x, 0) : x > 0\}$ . So given one component is large, the other tends not to be large; there is negligible probability they are both large.

- (2) The regular variation condition for the vector  $\mathbf{Z}$  is equivalent to supposing  $\{\mathbf{Z}_j, j \geq 1\}$  iid with  $\mathbf{Z}_j = \mathbf{Z}, j \geq 1$ , and

$$P\left[\bigvee_{j=1}^n \mathbf{Z}_j/b(n) \leq \mathbf{x}\right] \rightarrow e^{-\nu([\mathbf{0}, \mathbf{x}]^c)} = P[\mathbf{Y} \leq \mathbf{x}],$$

for a limiting extreme value random vector  $\mathbf{Y}$  with Frechet marginals. Asymptotic independence is equivalent to  $Y^{(1)}, \dots, Y^{(d)}$ , the components of  $\mathbf{Y}$ , being independent random variables (Resnick, 1987, Chapter 5).

The name asymptotic independence is used frequently for many different concepts, so beware. Also note that it can sometimes be a surprising concept.

**Example 5.** Consider the following two surprising examples and a third which is totally unsurprising.

- (1) Suppose  $U$  is  $U(0, 1)$  distributed and define

$$\mathbf{Z} = \left(\frac{1}{U}, \frac{1}{1-U}\right).$$

Then  $\mathbf{Z}$  possesses asymptotic independence as can be verified by direct calculation. This example was pointed out in de Haan and de Ronde (1998).

- (2) Let  $(N_1, N_2)$  be bivariate normal with normal distribution

$$N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

So  $\text{Corr}(N_1, N_2) = \rho$ . Provided  $\rho < 1$ ,

$$\mathbf{Z} = \left(\frac{1}{\Phi(N_1)}, \frac{1}{\Phi(N_2)}\right)$$

is asymptotically independent. See Sibuya (1960) or Resnick (1987, Chapter 5).

- (3) Suppose  $\mathbf{Z} = (Z^{(1)}, Z^{(2)})$  where  $Z^{(1)} \perp Z^{(2)}$  and for  $i = 1, 2$ ,

$$P[Z^{(i)} > x] = x^{-1}, \quad x > 1.$$

This was already discussed.

**4.5. Hidden regular variation; definition and properties.** Suppose we have two cones

$$\mathbb{C} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}, \quad \mathbb{C}^0 = \mathbb{C} \setminus \bigcup_{i=1}^d \{te_i, t > 0\}.$$

We may also express  $\mathbb{C}^0$  as

$$\mathbb{C}^0 = \{\mathbf{x} \in \mathbb{C} : \text{for some } 1 \leq i < j \leq d, x^{(i)} \wedge x^{(j)} > 0\}.$$

Suppose  $\mathbf{Z} \geq \mathbf{0}$  has distribution  $F$ . Then if  $F$  is regularly varying on  $\mathbb{C}$  we say  $F$  possesses hidden regular variation if it is also regularly varying on  $\mathbb{C}^0$ . Thus there exists a Radon measure  $\nu$  on  $\mathbb{C}$ , and also there exists a Radon measure  $\nu^0$  on  $\mathbb{C}^0$  such that for scaling functions  $b(t) \rightarrow \infty$  and  $b^0(t) \rightarrow \infty$  we have

$$(4.7) \quad tP[\mathbf{Z}/b(t) \in \cdot] \xrightarrow{v} \nu(\cdot) \quad \text{in } M_+(\mathbb{C})$$

$$(4.8) \quad tP[\mathbf{Z}/b^0(t) \in \cdot] \xrightarrow{v} \nu^0(\cdot) \quad \text{in } M_+(\mathbb{C}^0)$$

and, to ensure this is not a trivial concept, we assume

$$(4.9) \quad \frac{b(t)}{b^0(t)} \rightarrow \infty.$$

4.5.1. *Consequences and properties.* Here are some properties resulting from the definition. For fuller treatment, see Heffernan and Resnick (2005), Maulik and Resnick (2005), Resnick (2006, 2002). Hidden regular variation elaborates ideas of Ledford and Tawn (1996, 1997).

(a) There exists  $\alpha > 0$  such that for  $t > 0$ ,

$$\nu(t \cdot) = t^{-\alpha} \nu(\cdot) \quad \text{on } \mathbb{C}.$$

(b) There exists  $\alpha^0 \geq \alpha$  such that for  $t > 0$ ,

$$\nu^0(t \cdot) = t^{-\alpha^0} \nu^0(\cdot) \quad \text{on } \mathbb{C}^0.$$

(c) Consequently, we have

$$b \in RV_{1/\alpha}, \quad b^0 \in RV_{1/\alpha^0}.$$

(d)  $\aleph^0 := \aleph \cap \mathbb{C}^0$  is not compact.

(e) Suppose, as usual, that  $T(\mathbf{x}) = (\|\mathbf{x}\|, \frac{\mathbf{x}}{\|\mathbf{x}\|})$  is the polar coordinate transform. Then for  $c > 0$ , we have

$$(4.10) \quad \nu \circ T^{-1} = c\nu_\alpha \times S$$

where  $\nu_\alpha(x, \infty] = x^{-\alpha}$ , for  $x > 0$  and  $S$  is a probability measure on  $\aleph_+ = \aleph \cap \mathbb{C}$ . Also for  $c^0 > 0$ ,

$$(4.11) \quad \nu^0 \circ T^{-1} = c^0 \nu_{\alpha^0} \times S^0$$

where  $\nu_{\alpha^0}(x, \infty] = x^{-\alpha^0}$  for  $x > 0$ , and  $S^0$  is Radon on  $\aleph^0$  but not necessarily finite.

(f) Regions of the form  $(\mathbf{x}, \infty]$  are relatively compact in  $\mathbb{C}^0$  and

$$(4.12) \quad \begin{aligned} tP\left[\bigwedge_{i=1}^d \frac{Z^{(i)}}{b^0(t)} > x\right] &= tP\left[\frac{\mathbf{Z}}{b^0(t)} \in (x\mathbf{1}, \infty]\right] \\ &\rightarrow \nu^0(x\mathbf{1}, \infty] = x^{-\alpha^0} \nu^0(\mathbf{1}, \infty]. \end{aligned}$$

With  $x = 1$ , the limit is  $\nu^0(\mathbf{1}, \infty]$ .

(g) Hidden regular variation implies asymptotic independence since if  $\mathbf{z} > \mathbf{0}$

$$(4.13) \quad \begin{aligned} tP\left[\frac{\mathbf{Z}}{b(t)} > \mathbf{z}\right] &\leq tP\left[\frac{\mathbf{Z}}{b(t)} > \bigwedge_{i=1}^d z^{(i)}\mathbf{1}\right] \\ &= tP\left[\frac{\mathbf{Z}}{b^0(t)} > \frac{b(t)}{b^0(t)} \bigwedge_{i=1}^d z^{(i)}\mathbf{1}\right] \rightarrow 0, \end{aligned}$$

since  $b(t)/b^0(t) \rightarrow \infty$  and  $tP[\mathbf{Z}/b^0(t) \in \cdot] \rightarrow \nu^0(\cdot)$ . So in contrast to (4.12) we get with  $b(t)$  as the scaling function, that as  $t \rightarrow \infty$ ,

$$(4.14) \quad tP\left[\bigwedge_{i=1}^d \frac{Z^{(i)}}{b(t)} > x\right] \rightarrow 0.$$

For  $d = 2$ , (4.14) gives

$$0 = \lim_{t \rightarrow \infty} P[Z^{(2)} > t | Z^{(1)} > t]$$

and if you are a religious copularian, this is

$$\lambda := \lim_{u \rightarrow 1} P[F_1(Z^{(2)}) > u | F_1(Z^{(1)}) > u] = 0$$

where  $F_1$  is the marginal distribution.

(h) Choices for the scaling function  $b(t)$  appearing in (4.7) include

$$b(t) = \left( \frac{1}{1 - F_1} \right)^\leftarrow(t)$$

where  $F_1$  is the distribution of  $Z^{(1)}$  or alternatively,

$$b(t) = \left( \frac{1}{1 - P[\bigvee_{i=1}^d Z^{(i)} > \cdot]} \right)^\leftarrow(t).$$

Choices for  $b^0(t)$  include

$$b^0(t) = \left( \frac{1}{1 - P[\bigwedge_{i=1}^d Z^{(i)} > \cdot]} \right)^\leftarrow(t).$$

This results from the next property.

(i) Regular variation on  $\mathbb{C}$  implies  $\bigvee_{i=1}^d Z^{(i)}$  has a regularly varying distribution tail

$$P[\bigvee_{i=1}^d Z^{(i)} > x] \in RV_{-\alpha}$$

while regular variation on  $\mathbb{C}^0$  implies  $\bigwedge_{i=1}^d Z^{(i)}$  has a regularly varying distribution tail,

$$P[\bigwedge_{i=1}^d Z^{(i)} > x] \in RV_{-\alpha^0}.$$

More precisely, we have the following result. See Resnick (2006, 2002).

**Theorem 3.** Suppose  $b(\cdot) \in RV_{1/\alpha}$ ,  $b^0(\cdot) \in RV_{1/\alpha^0}$ ,  $0 < \alpha \leq \alpha^0$  and  $\frac{b(t)}{b^0(t)} \rightarrow \infty$ . Suppose also

$$(4.15) \quad tP\left[\frac{Z^{(i)}}{b(t)} > x\right] \rightarrow x^{-\alpha}, \quad x > 0, i = 1, \dots, d.$$

Then  $\mathbf{Z}$  possesses hidden regular variation iff

(a) Max-linear combinations have regularly varying tail probabilities,

$$(4.16) \quad tP\left[\bigvee_{i=1}^d \frac{s^{(i)} Z^{(i)}}{b(t)} > x\right] \rightarrow c(\mathbf{s})x^{-\alpha}, \quad x > 0, t \rightarrow \infty.$$

for  $\mathbf{s} \in [\mathbf{0}, \infty) \setminus \{\mathbf{0}\}$  and some  $c(\mathbf{s}) > 0$ ;

and

(b) Min-linear combinations have regularly varying tail probabilities,

$$(4.17) \quad tP\left[\bigwedge_{i=1}^d \frac{a^{(i)} Z^{(i)}}{b^0(t)} > x\right] \rightarrow d(\mathbf{a})x^{-\alpha^0}, \quad x > 0, t \rightarrow \infty,$$

for some  $d(\mathbf{a}) > 0$  where  $\mathbf{a} \in (\mathbb{C}^0)^{-1} = (\mathbf{0}, \infty] \setminus \bigcup_{i=1}^d \{t\mathbf{e}_i^{-1}, t > 0\}$  and

$$\mathbf{e}_i^{-1} = (\infty, \dots, \infty, 1, \infty, \dots, \infty).$$

- (j) The last item suggests a diagnostic for detection: Assume  $\mathbf{Z}_1, \dots, \mathbf{Z}_n$  is a random sample and compute the componentwise minima,

$$\left( \bigwedge_{i=1}^n Z_i^{(1)}, \dots, \bigwedge_{i=1}^n Z_i^{(d)} \right).$$

Estimate the tail index for the minima. If the problem has been standardized so that  $\alpha = 1$  we seek evidence that  $\hat{\alpha}^0 > 1$  which would be consistent with hidden regular variation.

4.5.2. *Hidden regular variation and marginal distributions.* Regular variation on  $\mathbb{C}^0$  plus correct behavior of the marginal distributions implies hidden regular variation.

**Proposition 2.** *Suppose  $\mathbf{Z} \geq \mathbf{0}$  satisfies*

$$Z^{(1)} \stackrel{d}{=} Z^{(2)} \stackrel{d}{=} \dots \stackrel{d}{=} Z^{(d)}$$

and  $\mathbf{Z}$  is regularly varying on  $\mathbb{C}^0$  so that there exist  $b^0(\cdot) \in RV_{1/\alpha^0}$  and  $\nu^0 \in M_+(\mathbb{C}^0)$  such that

$$tP\left[\frac{\mathbf{Z}}{b^0(t)} \in \cdot\right] \xrightarrow{\nu^0} \nu^0(\cdot), \quad \text{in } M_+(\mathbb{C}^0).$$

Suppose additionally that one dimensional marginals have regularly varying tails; that is, there exists  $b(\cdot) \in RV_{1/\alpha}$  with  $\alpha \leq \alpha^0$  and  $b(t)/b^0(t) \rightarrow \infty$ , and

$$tP\left[\frac{Z^{(i)}}{b(t)} > x\right] \rightarrow x^{-\alpha}, \quad x > 0, \quad i = 1, \dots, d.$$

Then  $\mathbf{Z}$  is regularly varying on  $\mathbb{C} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$ .

*Proof.* : For  $\mathbf{x} > \mathbf{0}$  we have by the inclusion-exclusion formula,

$$\begin{aligned} tP\left(\left[\frac{\mathbf{Z}}{b(t)} \leq \mathbf{x}\right]^c\right) &= t \sum_{i=1}^d P\left[\frac{Z^{(i)}}{b(t)} > x^{(i)}\right] - t \sum_{1 \leq i < j \leq d} P\left[\frac{Z^{(i)}}{b(t)} > x^{(i)}, \frac{Z^{(j)}}{b(t)} > x^{(j)}\right] \\ &\quad + t \sum_{1 \leq i < j < k \leq d} P\left[\frac{Z^{(i)}}{b(t)} > x^{(i)}, \frac{Z^{(j)}}{b(t)} > x^{(j)}, \frac{Z^{(k)}}{b(t)} > x^{(k)}\right] - \dots, \\ &\rightarrow \sum_{i=1}^d (x^{(i)})^{-\alpha} + 0, \quad (t \rightarrow \infty). \end{aligned}$$

The “0” comes from

$$tP\left[\frac{Z^{(1)}}{b(t)} > x^{(1)}, \frac{Z^{(2)}}{b(t)} > x^{(2)}\right] = tP\left[\frac{Z^{(1)}}{b^0(t)} > x^{(1)} \frac{b(t)}{b^0(t)}, \frac{Z^{(2)}}{b^0(t)} > x^{(2)} \frac{b(t)}{b^0(t)}\right] \rightarrow 0,$$

since  $b(t)/b^0(t) \rightarrow \infty$ . □

4.5.3. *Examples of hidden regular variation.* We now consider some examples to emphasize how hidden regular variation can arise in practice.

**Example 6.** (1) Recall Example 4. Suppose  $\mathbf{Z} = (Z^{(1)}, Z^{(2)})$  with  $Z^{(1)} \perp\!\!\!\perp Z^{(2)}$  and  $P[Z^{(i)} > x] = x^{-1}$  for  $x > 1$  and  $i = 1, 2$ . Then

$$\begin{aligned} tP\left[\frac{\mathbf{Z}}{t} \in \cdot\right] &\xrightarrow{\nu} \nu(\cdot), && \text{in } M_+(\mathbb{C}), \\ tP\left[\frac{\mathbf{Z}}{\sqrt{t}} \in \cdot\right] &\xrightarrow{\nu^0} \nu^0(\cdot), && \text{in } M_+(\mathbb{C}^0) \end{aligned}$$

where

$$(4.18) \quad \nu(\mathbf{x}, \infty] = 0, \quad \text{for } \mathbf{x} > \mathbf{0}, \text{ and}$$

$$(4.19) \quad \nu^0(\mathbf{x}, \infty] = \frac{1}{x^{(1)}x^{(2)}}, \quad \text{for } \mathbf{x} > \mathbf{0}.$$

(2) Suppose  $B, \mathbf{Y}, \mathbf{U}$  are independent with

(a)  $P[B = 0] = P[B = 1] = \frac{1}{2}$ .

(b)  $\mathbf{Y} = (Y^{(1)}, Y^{(2)})$  where  $Y^{(1)} \perp Y^{(2)}$  and  $P[Y^{(i)} > x] = x^{-1}$ , for  $x > 1$ . Set  $\alpha = 1$  and  $b(t) = t$  to describe the regular variation of the distribution of  $\mathbf{Y}$ .

(c)  $\mathbf{U}$  is multivariate regularly varying on  $\mathbb{C}$  with index  $\alpha^0$ ,  $1 < \alpha^0 < 2$ . Then there exists  $b^0 \in RV_{1/\alpha^0}$  and a Radon measure  $\nu^0 \in M_+(\mathbb{C})$  so that

$$tP[\mathbf{U}/b^0(t) \in \cdot] \xrightarrow{v} \nu^0(\cdot) \neq 0$$

in  $M_+(\mathbb{C})$ . Here we really do mean  $M_+(\mathbb{C})$  and not  $M_+(\mathbb{C}^0)$ .

Next, define

$$\mathbf{Z} = B\mathbf{Y} + (1 - B)\mathbf{U}.$$

For  $\mathbf{x} > \mathbf{0}$ , we have,

$$tP[\mathbf{Z}/b^0(t) > \mathbf{x}] = \frac{t}{2}P[\mathbf{Y} > b^0(t)\mathbf{x}] + \frac{t}{2}P[\mathbf{U} > b^0(t)\mathbf{x}] = I + II.$$

For  $I$  we have because of the independent Pareto random variables that as  $t \rightarrow \infty$ ,

$$\frac{t}{2} \frac{1}{b^0(t)^2 x^{(1)}x^{(2)}} \rightarrow 0$$

since  $(b^0(t))^2 \in RV_{2/\alpha^0}$  and  $t/(b^0(t))^2 \rightarrow 0$ .

For  $II$  we have

$$II \rightarrow \frac{1}{2}\nu^0(\mathbf{x}, \infty].$$

Also marginal distributions are  $RV_{-\alpha} = RV_{-1}$ , since

$$P[Z^{(1)} > x] = \frac{1}{2}P[Y^{(1)} > x] + \frac{1}{2}P[U^{(1)} > x] \in RV_{-1}$$

since the first term on the right is  $RV_{-1}$  and the second is  $RV_{-\alpha^0}$  with  $\alpha^0 > 1$ . This is enough to imply hidden regular variation by Proposition 2.

(3) Here is a way to construct a class of examples.

Let  $\Theta$  be a random variable with distribution  $G$  concentrating on  $(1, \infty)$ . Suppose  $B$  is an independent Bernoulli random variable,

$$P[B = 0] = \frac{1}{2} = P[B = 1].$$

Suppose  $R \perp \{B, \Theta\}$  and  $P[R > r] = r^{-1}$ ,  $r > 1$ . Define

$$\Theta = B(\Theta, 1) + (1 - B)(1, \Theta), \quad \mathbf{Z} = R\Theta.$$

So the distribution of  $\Theta$  concentrates on the lines emanating from  $(1, 1)$  in the horizontal and vertical directions.

We now give some properties of this construction.

(a) We have

$$Z^{(1)} = R\Theta^{(1)}, \quad Z^{(2)} = R\Theta^{(2)},$$

and

$$\Theta^{(1)} = B\Theta + (1 - B)1, \quad \Theta^{(2)} = B1 + (1 - B)\Theta.$$

(b) For the marginal distributions, we have,

$$P[Z^{(i)} > x] = P[R\Theta^{(i)} > x] = \int_1^\infty P[\Theta^{(i)} > x/r] \frac{dr}{r^2}$$

and setting  $s = \frac{x}{r}$  we obtain

$$\begin{aligned}
 &= \frac{1}{x} \int_0^x P[\Theta^{(i)} > s] ds = \frac{1}{2x} \left[ \int_0^1 2 ds + \int_1^x \bar{G}(s) ds \right] \\
 (4.20) \quad &= \frac{1}{2x} \left[ 2 + \int_1^x \bar{G}(s) ds \right].
 \end{aligned}$$

- (c) The random vector  $\mathbf{Z}$  is regularly varying on  $(0, \infty]^2$  with  $\alpha^0 = 1$ . Furthermore, let  $\nu^0$  be the limit measure. It has angular measure  $S^0$  and  $S^0$  is finite iff

$$\int_0^\infty \bar{G}(s) ds < \infty.$$

To see this, note for  $\nu^0$  we have for  $\mathbf{x} > \mathbf{0}$ ,

$$\nu^0(\mathbf{x}, \infty] = \lim_{t \rightarrow \infty} tP\left[\frac{\mathbf{Z}}{t} > \mathbf{x}\right].$$

Now

$$\begin{aligned}
 tP[\mathbf{Z}/t > \mathbf{x}] &= tP[R\Theta^{(1)} > tx^{(1)}, R\Theta^{(2)} > tx^{(2)}] \\
 &= tP[R\Theta^{(1)} > tx^{(1)}, R\Theta^{(2)} > tx^{(2)}, B = 1] \\
 &\quad + tP[R\Theta^{(1)} > tx^{(1)}, R\Theta^{(2)} > tx^{(2)}, B = 0] \\
 &= \frac{t}{2} P[R\Theta > tx^{(1)}, R > tx^{(2)}] + \frac{t}{2} P[R > tx^{(1)}, R\Theta > tx^{(2)}] \\
 &= \frac{t}{2} \int_{r > tx^{(2)}} \bar{G}\left(\frac{tx^{(1)}}{r} \vee 1\right) \frac{dr}{r^2} + \frac{t}{2} \int_{r > tx^{(2)}} \bar{G}\left(\frac{tx^{(2)}}{r} \vee 1\right) \frac{dr}{r^2}.
 \end{aligned}$$

Set  $s = t/r$  and we obtain

$$\begin{aligned}
 &= \frac{1}{2} \int_0^{1/x^{(2)}} \bar{G}(sx^{(1)} \vee 1) ds + \frac{1}{2} \int_0^{1/x^{(1)}} \bar{G}(sx^{(2)} \vee 1) ds \\
 &= \frac{1}{2x^{(1)}} \int_0^{x^{(1)}/x^{(2)}} \bar{G}(u \vee 1) du + \frac{1}{2x^{(2)}} \int_0^{x^{(2)}/x^{(1)}} \bar{G}(u \vee 1) du.
 \end{aligned}$$

Let, for instance,  $x^{(1)} \rightarrow 0$  and the limit is infinite iff  $\int_0^\infty \bar{G}(u) du = \infty$ .

- (d) If  $\bar{G} \in RV_{-\alpha}$ ,  $\alpha < 1$ , then  $\mathbf{Z}$  is regularly varying on  $\mathbb{C} = [0, \infty]^2 \setminus \{\mathbf{0}\}$  with index  $\alpha$ . The reason for this is that from item 3b and Karamata's theorem, we have

$$P[Z^{(i)} > x] \in RV_{-\alpha}$$

and the multivariate regular variation follows from Proposition 2.

**Remark 3.** Note the following.

- (1) For Example 1,  $\nu^0$  is infinite on  $\mathbb{C}^0 \cap \{\mathbf{x} : \|\mathbf{x}\| > 1\}$ . Hence  $S^0$  is infinite on  $\aleph^0$ . This follows from (4.19) by, say, setting  $x^{(1)} = 1$  and letting  $x^{(2)} \rightarrow 0$ .
- (2) For Example 2,  $\nu^0$  is finite on  $\mathbb{C}^0 \cap \{\mathbf{x} : \|\mathbf{x}\| > 1\}$ . Hence  $S^0$  is finite on  $\aleph^0 = \aleph \cap \mathbb{C}^0$ . This follows because we assumed  $\mathbf{U}$  is regularly varying on  $\mathbb{C}$  and not just on  $\mathbb{C}^0$ .
- (3) For Example 3,  $\nu^0$  may be finite or infinite on  $\mathbb{C}^0 \cap \{\mathbf{x} : \|\mathbf{x}\| > 1\}$ , depending on whether the first moment of  $G$  exists or not.

4.5.4. *The coefficient of tail dependence  $\eta$  of Ledford and Tawn.* Hidden regular variation elaborates some ideas of Ledford and Tawn (1996, 1997). They define  $\eta$  by assuming that  $\mathbf{Z} = (Z^{(1)}, Z^{(2)})$  is a two dimensional random vector with non-negative components such that  $Z^{(1)} \stackrel{d}{=} Z^{(2)}$  and

$$(4.21) \quad P[Z^{(1)} \wedge Z^{(2)} > t] \sim L(t)(P[Z^{(1)} > t])^{1/\eta},$$

where  $0 < \eta \leq 1$  and  $L \in RV_0$ .

If  $\mathbf{Z}$  possesses hidden regular variation, then (4.21) is true: Suppose as usual  $b(\cdot)$  satisfies

$$tP[Z^{(1)} > b(t)] \sim 1.$$

Then (4.21) becomes

$$(4.22) \quad P[Z^{(1)} \wedge Z^{(2)} > b(t)] \sim L(b(t))t^{-1/\eta} = L'(t)t^{-1/\eta},$$

where  $L'(t) = L(b(t)) \in RV_0$ . We know from hidden regular variation (See Theorem 3) that

$$P[Z^{(1)} \wedge Z^{(2)} > t] \sim t^{-\alpha^0} \ell(t),$$

for some function  $\ell(t) \in RV_0$ . So

$$P[Z^{(1)} \wedge Z^{(2)} > b(t)] \sim b(t)^{-\alpha^0} \ell(b(t)).$$

Now for some slowly varying function  $\ell_*(t)$ ,  $b(t) \sim t^{1/\alpha} \ell_*(t)$  and  $\ell \circ b \in RV_0$ , so

$$(4.23) \quad P[Z^{(1)} \wedge Z^{(2)} > b(t)] \sim t^{-\alpha^0/\alpha} \ell_{**}(t),$$

for a slowly varying function  $\ell_{**} \in RV_0$ . From comparing (4.22) and (4.23), we conclude that

$$\eta = \frac{\alpha}{\alpha^0}.$$

## 5. CONDITIONED LIMIT THEOREMS

Suppose  $\mathbb{C}^{1/2} = [0, \infty] \times (0, \infty]$  and  $\mathbf{Z}$  is regularly varying on  $\mathbb{C}^{1/2}$  meaning

$$(5.1) \quad tP\left[\frac{\mathbf{Z}}{b(t)} \in \cdot\right] \rightarrow \nu(\cdot), \quad \text{in } M_+(\mathbb{C}^{1/2}).$$

This implies

$$tP\left[\frac{Z^{(1)}}{b(t)} \in [0, \infty], \frac{Z^{(2)}}{b(t)} > x\right] = tP\left[\frac{Z^{(2)}}{b(t)} > x\right] \rightarrow \nu([0, \infty] \times (x, \infty]) = cx^{-\alpha}.$$

Suppose, as a normalization, that  $c = 1$ , so that,

$$(5.2) \quad \nu([0, \infty] \times (1, \infty]) = 1.$$

Then from (5.1), for  $\Lambda \in \mathcal{B}(\mathbb{R}_+)$ ,

$$(5.3) \quad \begin{aligned} P\left[\frac{Z^{(1)}}{b(t)} \in \Lambda \mid \frac{Z^{(2)}}{b(t)} > 1\right] &= \frac{P\left[\frac{Z^{(1)}}{b(t)} \in \Lambda, \frac{Z^{(2)}}{b(t)} > 1\right]}{P\left[\frac{Z^{(2)}}{b(t)} > 1\right]} \\ &\sim tP\left[\frac{Z^{(1)}}{b(t)} \in \Lambda, \frac{Z^{(2)}}{b(t)} > 1\right] \rightarrow \nu(\Lambda \times (1, \infty]). \end{aligned}$$

So regular variation on  $\mathbb{C}^{1/2}$  implies a conditioned limit theorem: Given that the second component is large, the conditional distribution of the first variable is approximately determined by the limit measure in the definition of regular variation.

**5.1. General formulation.** The previous formulation was for regular variation where each component could be normalized by the same scaling function. Regular variation on  $\mathbb{C}^{1/2}$  naturally led to a conditioned limit theorem. How can we generalize this to make a more flexible class of statistical models? Here is an introduction; more detail and applications are found in Abdous et al. (2005), Heffernan and Resnick (2007), Heffernan and Tawn (2004).

We assume we have a two dimensional random vector  $\mathbf{Z} = (X, Y)$ . Suppose there exist  $a(t) > 0$ ,  $\alpha(t) > 0$ ,  $\beta(t) \in \mathbb{R}$ ,  $b(t) \in \mathbb{R}$  such that

$$(5.4) \quad tP\left[\left(\frac{X - \beta(t)}{\alpha(t)}, \frac{Y - b(t)}{a(t)}\right) \in \cdot\right] \xrightarrow{\nu} \mu(\cdot),$$

in  $M_+([-\infty, \infty] \times (-\infty, \infty])$  or equivalently

$$(5.5) \quad tP\left[\frac{X - \beta(t)}{\alpha(t)} \leq x, \frac{Y - b(t)}{a(t)} > y\right] \rightarrow \mu([-\infty, x] \times (y, \infty]), \quad -\infty \leq x \leq \infty, y > -\infty.$$

We require that the limit measure  $\mu$  satisfy the following constraints:

- (a) For each fixed  $y$ , the distribution function  $u[-\infty, x] \times (y, \infty]$  is not a degenerate distribution function in  $x$ .
- (b) For  $x \in [-\infty, \infty]$  and  $y > -\infty$ , we have

$$\mu([-\infty, x] \times (y, \infty]) < \infty.$$

**5.1.1. Standardizing the  $Y$ -variable.** Observe that (5.5) is in  $M_+([-\infty, \infty] \times (-\infty, \infty])$  and so we may insert the compact set  $[-\infty, \infty]$  for the  $X$ -variable, effectively marginalizing to  $Y$ . Therefore,

$$(5.6) \quad tP\left[\frac{Y - b(t)}{a(t)} > y\right] \rightarrow \mu([-\infty, \infty] \times (y, \infty]), \quad y \in \mathbb{R},$$

which entails that  $Y \in D(G_\gamma)$  for some  $\gamma \in \mathbb{R}$ ; that is, the distribution of  $Y$  is in a maximal domain of attraction (de Haan and Ferreira, 2006, Embrechts et al., 1997, Resnick, 1987). Note (5.6) is the defining property of univariate extreme value distributions (as in (3.11) in Section 3.1) so (5.6) entails (essentially by inverting (5.6)) that  $b(\cdot)$  can be chosen as a monotone function and for some  $\gamma \in \mathbb{R}$ , as  $t \rightarrow \infty$ ,

$$(5.7) \quad \frac{b(tx) - b(t)}{a(t)} \rightarrow c\left(\frac{x^\gamma - 1}{\gamma}\right), \quad x > 0, c \neq 0.$$

This allows us to standardize the  $Y$ -variable in the usual way. See Resnick (1987, Chapter 5), or de Haan and Ferreira (2006), de Haan and Resnick (1977). Therefore write  $Y^* = b^\leftarrow(Y)$  and

$$\begin{aligned} tP\left[\frac{X - \beta(t)}{\alpha(t)} \leq x, \frac{Y^*}{t} > y\right] &= tP\left[\frac{X - \beta(t)}{\alpha(t)} \leq x, \frac{Y - b(t)}{a(t)} > \frac{b(y) - b(t)}{a(t)}\right] \\ &\rightarrow \mu([-\infty, x] \times (c\left(\frac{y^\gamma - 1}{\gamma}\right), \infty]) =: \mu_{*/2}([-\infty, x] \times (y, \infty]), \quad x \in \mathbb{R}, y > 0. \end{aligned}$$

The subscript  $*/2$  on the limit reminds us that the limit is obtained from the original  $\mu$  by standardizing only the  $Y$  variable.

Our conclusion: We can always standardize the  $Y$ -variable by making the transformation  $Y \mapsto Y^* = b^\leftarrow(Y)$ ; then  $Y^*$  need only be normalized by  $t$ . The meaning of the phrase *standardization* is that normalization by  $t$  is adequate.

**5.1.2. Standardizing the  $X$ -variable.** What about the  $X$ -variable? Suppose we know that  $\beta$  is non-decreasing and write

$$(5.8) \quad tP\left[\frac{\beta^\leftarrow(X)}{t} \leq x, \frac{Y^*}{t} > y\right] = tP\left[\frac{X - \beta(t)}{\alpha(t)} \leq \frac{\beta(tx) - \beta(t)}{\alpha(t)}, \frac{Y^*}{t} > y\right].$$

Provided

$$(5.9) \quad \frac{\beta(tx) - \beta(t)}{\alpha(t)} \rightarrow \psi_2(x) \neq 0, \quad x > 0,$$

we get with  $X^* = \beta^{\leftarrow}(X)$  that for  $x > 0, y > 0$ ,

$$(5.10) \quad tP\left[\frac{X^*}{t} \leq x, \frac{y^*}{t} > y\right] \rightarrow \mu\left([-\infty, \psi_2(x)] \times \left(c\left(\frac{y^\gamma - 1}{\gamma}\right), \infty\right]\right) =: \mu_*([0, x] \times (y, \infty]),$$

and standardization is achieved.

**Proposition 3.** (Heffernan and Resnick, 2007) Suppose

$$(5.11) \quad tP\left[\frac{X - \beta(t)}{\alpha(t)} \leq x, \frac{y^*}{t} > y\right] \rightarrow \mu_{*/2}([-\infty, x] \times (y, \infty]), \quad x \in \mathbb{R}, y > 0,$$

is the assumption where the  $Y$ -variable is standardized. We assume

- (a) For each  $y > 0$ ,  $\mu_{*/2}([-\infty, x] \times (y, \infty])$  is not a degenerate distribution in  $x$ ; that is, it has more than one point of increase for the  $x$ -variable.
- (b) The distribution of  $Y^*$  is in the domain of attraction of  $G_1$ , the Frechet extreme value distribution with  $\gamma = 1$ . We write this as  $P[Y^* \leq t] \in D(G_1)$ . (This actually follows from (5.11).)
- (c) The measure  $\mu_{*/2}$  satisfies

$$(5.12) \quad \mu_{*/2}([-\infty, \infty] \times (1, \infty]) = 1.$$

Then there exist functions  $\psi_1(c), \psi_2(c), c > 0$  such that

$$(5.13) \quad \lim_{t \rightarrow \infty} \frac{\alpha(tc)}{\alpha(t)} = \psi_1(c)$$

and

$$(5.14) \quad \lim_{t \rightarrow \infty} \frac{\beta(tc) - \beta(t)}{\alpha(t)} = \psi_2(c),$$

and the convergences in [5.13], [5.14] are locally uniform.

*Proof.* Pick  $c > 0$  and assume  $(x, 1)$  and  $(x, c^{-1})$  are continuity points of  $\mu_{*/2}([-\infty, x] \times (y, \infty])$ . (There are only countably many points not continuity points.) Define

$$(5.15) \quad H(x) = \mu_{*/2}([-\infty, x] \times (y, \infty]), \quad x \in \mathbb{R},$$

and because of (5.12),  $H(x)$  is a probability distribution function. Similarly, we define the conditional probability distribution

$$(5.16) \quad H_t(\alpha(t)x + \beta(t)) = P\left[\frac{X - \beta(t)}{\alpha(t)} \leq x | Y^* > t\right].$$

Then on the one hand,

$$(5.17) \quad \begin{aligned} H_t(\alpha(t)x + \beta(t)) &= \frac{P\left[\frac{X - \beta(t)}{\alpha(t)} \leq x, Y^* > t\right]}{P[Y^* > t]} \\ &\sim tP\left[\frac{X - \beta(t)}{\alpha(t)} \leq x, Y^* > t\right] \rightarrow \mu_{*/2}([-\infty, x] \times (1, \infty]). \end{aligned}$$

On the other hand

$$(5.18) \quad \begin{aligned} H_t(\alpha(tc)x + \beta(tc)) &= P\left[\frac{X - \beta(tc)}{\alpha(tc)} \leq x | \frac{Y^*}{t} > 1\right] \sim tP\left[\frac{X - \beta(tc)}{\alpha(tc)} \leq x, \frac{Y^*}{t} > 1\right] \\ &= \frac{1}{c} \left( tcP\left[\frac{X - \beta(tc)}{\alpha(tc)} \leq x, \frac{Y^*}{tc} > \frac{1}{c}\right] \right) \\ &\rightarrow \frac{1}{c} \mu_{*/2}([-\infty, x] \times (c^{-1}, \infty]) =: H^{(c)}(x). \end{aligned}$$

Comparing (5.17) and (5.18) and applying the convergence to types theorem (Feller, 1971, Resnick, 1999) we conclude that (5.13), (5.14) must hold and that  $H$  and  $H^{(c)}$  are related by

$$(5.19) \quad H^{(c)}(x) = H(\psi_1(c)x + \psi_2(c)).$$

□

5.1.3. *The limit measure is a product.* When  $\psi_2 \not\equiv 0$ , this allows standardization. However, it can occur that  $\psi_2 \equiv 0$  as shown in the following example (Abdous et al., 2005, Heffernan and Resnick, 2007).

**Example 7.** Let  $N_1, N_2$  be iid  $N(0, 1)$  and define

$$(X, Y) = (\sqrt{1 - \rho^2}N_1 + \rho N_2, N_2)$$

so that  $(X, Y)$  has a bivariate normal distribution

$$N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

Define

$$\Phi(x) = \int_{-\infty}^x \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$$

for the normal cdf and then we may set (see, for instance, de Haan and Ferreira (2006), Resnick (1987))

$$(5.20) \quad a(t) = \frac{1}{\sqrt{2 \log t}},$$

$$(5.21) \quad b(t) = \left(\frac{1}{1 - \Phi}\right)^{\leftarrow}(t) = \sqrt{2 \log t} - \frac{\frac{1}{2}(\log \log t + \log 4\pi)}{\sqrt{2 \log t}} + o(a(t)).$$

This choice of  $a(\cdot)$  and  $b(\cdot)$  ensures

$$(5.22) \quad tP\left[\frac{N_1 - b(t)}{a(t)} > x\right] \rightarrow e^{-x}, \quad x \in \mathbb{R}.$$

Furthermore (Abdous et al., 2005, Heffernan and Resnick, 2007) for  $(x, y) \in \mathbb{R}^2$ ,

$$(5.23) \quad tP[X - \rho b(t) \leq x, \frac{Y - b(t)}{a(t)} > y] \rightarrow \Phi\left(\frac{x}{\sqrt{1 - \rho^2}}\right)e^{-y} = \mu([-\infty, x] \times (y, \infty]),$$

or after standardization

$$(5.24) \quad tP[X - \rho b(t) \leq x, \frac{Y^*}{t} > y] \rightarrow \Phi\left(\frac{x}{\sqrt{1 - \rho^2}}\right)y^{-1} = \mu_{*/2}([-\infty, x] \times (y, \infty]),$$

for  $x \in \mathbb{R}, y > 0$ . It requires some calculation to verify (5.23) or (5.24) but the calculations are elementary. From (5.24), we see that

$$\alpha(t) = 1, \quad \beta(t) = \rho b(t).$$

This implies

$$\frac{\beta(tc) - \beta(t)}{\alpha(t)} = \frac{\rho(b(tc) - b(t))}{1}.$$

Either from (5.22) or from the explicit forms in (5.20) and (5.21) we get that the function  $b \in \Pi$  with auxiliary function  $a(\cdot)$ , meaning (de Haan and Ferreira, 2006, de Haan, 1970, Resnick, 1987)

$$\frac{b(tc) - b(t)}{a(t)} \rightarrow \log c, \quad c > 0.$$

Thus we obtain

$$\frac{\beta(tc) - \beta(t)}{\alpha(t)} = \rho\left(\frac{b(tc) - b(t)}{a(t)}\right)a(t) \sim \rho \log c \cdot a(t) \rightarrow 0 \equiv \psi(c).$$

Also  $\alpha(t) \equiv 1$  implies  $\psi(c) \equiv 1$ .

The circumstance  $(\psi_1(c), \psi_2(c)) \equiv (1, 0)$  is associated with the limit measure being a product measure. We state the result (Heffernan and Resnick, 2007) next.

**Proposition 4.** *The limit measure  $\mu_{*/2}$  is a product measure*

$$(5.25) \quad \mu_{*/2} = H(\cdot) \times \nu_1 \quad \text{where } \nu_1(x, \infty] = x^{-1}, x > 0,$$

or equivalently

$$(5.26) \quad \mu_{*/2}([-\infty, x] \times (y, \infty]) = H(x)y^{-1}, \quad y > 1,$$

iff the functions  $\psi_1, \psi_2$  in (5.13), (5.14) satisfy

$$(5.27) \quad \psi_1(c) \equiv 1, \quad \psi_2(c) \equiv 0.$$

*Proof.* If  $\mu_{*/2}$  is a product then from (5.18)

$$H^{(c)} = H$$

since (5.18) implies  $\mu_{*/2}([-\infty, x] \times (1, \infty]) = H^{(c)}(x) = H(x)$ . Therefore,  $\psi_1 \equiv 1$ ,  $\psi_2 \equiv 0$ .

If  $\psi_1 \equiv 1$  and  $\psi_2 \equiv 0$ , then  $H^{(c)} = H$  and from (5.18)

$$\frac{1}{c}\mu_{*/2}([-\infty, x] \times (c^{-1}, \infty]) = H^{(c)}(x) = H(x)$$

or

$$\mu_{*/2}([-\infty, x] \times (c^{-1}, \infty]) = cH(x)$$

or

$$\mu_{*/2}([-\infty, x] \times (y, \infty]) = y^{-1}H(x).$$

So  $\mu_{*/2}$  is a product measure. □

## 6. APPENDIX

**6.1. Vector notation.** Vectors are denoted by bold letters, capitals for random vectors and lower case for non-random vectors. For example:  $\mathbf{x} = (x^{(1)}, \dots, x^{(d)}) \in \mathbb{R}^d$ . Operations between vectors should always be interpreted componentwise so that for two vectors  $\mathbf{x}$  and  $\mathbf{z}$

$$\begin{aligned} \mathbf{x} < \mathbf{z} &\text{ means } x^{(i)} < z^{(i)}, \quad i = 1, \dots, d, & \mathbf{x} \leq \mathbf{z} &\text{ means } x^{(i)} \leq z^{(i)}, \quad i = 1, \dots, d, \\ \mathbf{x} = \mathbf{z} &\text{ means } x^{(i)} = z^{(i)}, \quad i = 1, \dots, d, & \mathbf{z}\mathbf{x} &= (z^{(1)}x^{(1)}, \dots, z^{(d)}x^{(d)}), \\ \mathbf{x} \vee \mathbf{z} &= (x^{(1)} \vee z^{(1)}, \dots, x^{(d)} \vee z^{(d)}), & \frac{\mathbf{x}}{\mathbf{z}} &= \left( \frac{x^{(1)}}{z^{(1)}}, \dots, \frac{x^{(d)}}{z^{(d)}} \right), \end{aligned}$$

and so on. Also, if

$$\boldsymbol{\alpha} = (\alpha^{(1)}, \dots, \alpha^{(d)}) \geq \mathbf{0},$$

we write for  $\mathbf{x} \geq \mathbf{0}$

$$\mathbf{x}^{\boldsymbol{\alpha}} = ((x^{(1)})^{\alpha^{(1)}}, \dots, (x^{(d)})^{\alpha^{(d)}}).$$

Further, we define

$$\mathbf{0} = (0, \dots, 0), \quad \mathbf{1} = (1, \dots, 1), \quad \boldsymbol{\infty} = (\infty, \dots, \infty), \quad \mathbf{e}_i = (0, \dots, 1, \dots, 0),$$

where in  $\mathbf{e}_i$  the "1" occurs in the  $i$ th spot. For a real number  $c$ , write as usual  $c\mathbf{x} = (cx^{(1)}, \dots, cx^{(d)})$ . We denote the rectangles (or the higher dimensional intervals) by

$$[\mathbf{a}, \mathbf{b}] = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\}.$$

Higher dimensional rectangles with one or both endpoints open are defined analogously, for example,

$$(\mathbf{a}, \mathbf{b}] = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a} < \mathbf{x} \leq \mathbf{b}\}.$$

**6.2. Symbol and concept list.** Here is a glossary of miscellaneous symbols and terminology used throughout.

$RV_\rho$	The class of regularly varying functions on $[0, \infty)$ with index $\rho \in \mathbb{R}$ .
$D(\mathbb{E}, \mathbb{R}^d)$	The Skorokhod space of $\mathbb{R}^d$ -valued càdlàg functions on the metric space $\mathbb{E}$ equipped with the $J_1$ -topology.
$\epsilon_x$	The probability measure consisting of all mass at $x$ .
$f^\leftarrow$	The left continuous inverse of a non-decreasing function $f$ defined by $f^\leftarrow(x) = \inf\{y : f(y) \geq x\}$ .
$Leb$	Lebesgue measure.
$\mathbb{E}$	Often $[0, \infty]^d \setminus \{\mathbf{0}\}$ or $[-\infty, \infty] \setminus \{-\infty\}$ .
Radon	Adjective applied to a measure to indicate the measure is finite on relatively compact sets.
$\xrightarrow{v}$	Vague convergence of measures.
$M_+(\mathbb{E})$	The space of non-negative Radon measures on $\mathbb{E}$ .
$M_p(\mathbb{E})$	The space of Radon point measures on $E$ .
$\nu_\alpha$	A measure on $(0, \infty]$ given by $\nu_\alpha(x, \infty] = x^{-\alpha}$ , $\alpha > 0$ , $x > 0$ .
$\Rightarrow$	Convergence in distribution.
$b(t)$	Usually the quantile function of a distribution function $F(x)$ , defined by $b(t) = F^\leftarrow(1 - \frac{1}{t})$ but usage can vary somewhat by context.

## REFERENCES

- B. Abdous, Anne-Laure Fougères, and K. Ghoudi. Extreme behaviour for bivariate elliptical distributions. *The Canadian Journal of Statistics*, 33(2), 2005.
- N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, 1987.
- S.G. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. London: Springer. xiv, 210 p. , 2001.
- D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Vol. I: Elementary theory and methods. 2nd ed.* Probability and Its Applications. New York, NY: Springer. xxi, 469 p. , 2003.
- L. de Haan and J. de Ronde. Sea and wind: multivariate extremes at work. *Extremes*, 1(1):7–46, 1998.
- L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.
- L. de Haan and S. I. Resnick. Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(4):317–337, 1977.
- L. de Haan. *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematisch Centrum Amsterdam, 1970.
- P. Embrechts, C. Kluppelberg, and T. Mikosch. *Modelling Extreme Events for Insurance and Finance*. Springer-Verlag, Berlin, 1997.
- W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, New York, 2nd edition, 1971.
- J. L. Geluk and L. de Haan. *Regular Variation, Extensions and Tauberian Theorems*, volume 40 of *CWI Tract*. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam, 1987. ISBN 90-6196-324-9.
- J.E. Heffernan and S.I. Resnick. Hidden regular variation and the rank transform. *Adv. Appl. Probab.*, 37(2):393–414, 2005.
- J.E. Heffernan and S. Resnick. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.*, 17(2):537–571, 2007. ISSN 1050-5164. doi: 10.1214/105051606000000835.
- J.E. Heffernan and J.A. Tawn. A conditional approach for multivariate extreme values (with discussion). *JRSS B*, 66(3):497–546, 2004.
- O. Kallenberg. *Random Measures*. Akademie-Verlag, Berlin, third edition, 1983. ISBN 0-12-394960-2.

- A.W. Ledford and J.A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1): 169–187, 1996. ISSN 0006-3444.
- A.W. Ledford and J.A. Tawn. Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B*, 59(2): 475–499, 1997. ISSN 0035-9246.
- K. Maulik and S.I. Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67, 2005.
- J. Neveu. Processus ponctuels. In *École d'Été de Probabilités de Saint-Flour, VI—1976*, pages 249–445. Lecture Notes in Math., Vol. 598, Berlin, 1977. Springer-Verlag.
- S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2006. ISBN: 0-387-24272-4.
- S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York, 1987.
- S.I. Resnick. *A Probability Path*. Birkhäuser, Boston, 1999.
- S. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4): 303–336 (2003), 2002. ISSN 1386-1999.
- E. Seneta. *Regularly Varying Functions*. Springer-Verlag, New York, 1976. Lecture Notes in Mathematics, 508.
- M. Sibuya. Bivariate extreme statistics. *Ann. Inst. Stat. Math.*, 11:195–210, 1960.

PROF. SIDNEY RESNICK, SCHOOL OF OPERATIONS RESEARCH AND INFORMATION ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY 14853

*E-mail address:* `sir1@cornell.edu`