Conditional Extreme Value Analysis for Random Vectors using Polar Representations

Miriam Isabel Seifert

Dissertation

Technische Universität München

Lehrstuhl für Mathematische Statistik

2015

Technische Universität München
Fakultät für Mathematik
Lehrstuhl für Mathematische Statistik

Conditional Extreme Value Analysis for Random Vectors using Polar Representations

Miriam Isabel Seifert

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Noam Berger Steiger
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Claudia Klüppelberg
	2.	Prof. Dr. Paul Embrechts, ETH Zürich / Schweiz
	3.	Univ.-Prof. Dr. Alexander Lindner, Universität Ulm

Die Dissertation wurde am 20. April 2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 11. Juni 2015 angenommen.

Zusammenfassung

Die vorliegende Dissertation befasst sich mit der Untersuchung des asymptotischen Verhaltens von bivariaten Zufallsvektoren unter Verwendung polarer Darstellungen $(X, Y) = R \cdot (u(T), v(T))$. Unsere Resultate erweitern bestehende Grenzwertsätze für elliptische und verwandte Verteilungen in verschiedene Richtungen.

Wir wählen das populäre CEV (*conditional extreme value*) Modell, d.h. wir leiten die bedingte Grenzverteilung von Y her, gegeben dass die Vektorkomponente X extrem groß wird. Dazu verwenden wir durchgängig einen geometrischen Ansatz, welcher ein tiefgehendes Verständnis der Resultate sowie der zugrundeliegenden Ideen ermöglicht. Außerdem analysieren wir die Asymmetrie zwischen den Variablen X und Y im CEV-Modell und zeigen, dass die wesentlichen Voraussetzungen an die bedingende Variable X zu stellen sind.

Wir schwächen die Annahmen an die Kurven $(u(t), v(t))$ für die Existenz von Grenzwertaussagen so weit ab, dass nur noch natürliche Einschränkungen vorhanden sind. Dabei zeigen wir, dass in einigen interessanten Fällen nichtdegenerierte Grenzwertresultate nur mit stochastischer Normierung (*random norming*) erhalten werden können.

In der Literatur wird für elliptische und verallgemeinerte Verteilungen grundsätzlich angenommen, dass R und T stochastisch unabhängig sind. Wir untersuchen inwieweit sich diese Annahme abschwächen lässt. Dazu führen wir ein neues Maß für die Abhängigkeitsstruktur ein und leiten elegante Kriterien für die Gültigkeit von Grenzwertaussagen her. Unsere Ergebnisse zeigen somit auch die Stabilität der bestehenden Grenzwertsätze für einen gewissen Grad an Abhängigkeit, welche insbesondere für Anwendungen von Bedeutung ist.

Summary

In this thesis we investigate the asymptotic behavior of bivariate random vectors using polar representations $(X, Y) = R \cdot (u(T), v(T))$. Our results extend the established limit statements for the elliptical and related distributions in several directions.

We work with the popular conditional extreme value (CEV) model, i.e. we derive the limiting conditional distribution of Y given that the vector component X becomes extreme. For this purpose, we use a geometric approach which permits a deeper understanding of the results and the ideas behind them. We also analyze the asymmetry between X and Y in the CEV model and show that the essential assumptions are met on the conditioning variable X.

We weaken the assumptions on the level curves $(u(t), v(t))$ for the existence of limit statements such that only natural restrictions remain. Thereby, we show that in some interesting cases one can get a non-degenerate limit result only with *random norming*.

In the literature for elliptical and generalized distributions, R and T are assumed to be stochastically independent. We investigate how far this assumption can be weakened. For this purpose, we introduce a novel measure for dependence structure and deduce convenient criteria for validity of limit theorems. In addition, our results verify stability of the established limit results for a certain degree of dependence, which is of importance particularly for applications.

Acknowledgements

At this point I would like to take the opportunity to thank all those who contributed to the success of my thesis.

First of all, I would like to thank my supervisor Prof. Claudia Klüppelberg. I am very indebted to her for the all-round support, wise advice and helpful recommendations.

Next, I want to give my thanks to Prof. Paul Embrechts and Prof. Alexander Lindner for accepting the invitation to be reviewers of my thesis and for coming to Munich to my defense. Further, I am grateful to Prof. Holger Drees for turning my interest to the research area of my thesis. I also would like to thank all my current and past colleagues in the "stochastic group" at the Helmut-Schmidt-University Hamburg for a pleasant working atmosphere.

Finally, I want to express my gratitude to my parents and to a special person in my life, W., for all their loving support they have been giving me at all times.

Contents

Chapter 1

Introduction

In this thesis we make investigations in the research area of bivariate extreme value theory. More precisely, we contribute to the limit behavior analysis of the conditional distribution of random vectors given that one vector component becomes extreme. For this purpose we use polar representations with a radial and an angular component, which are very appropriate for our studies because of the following main reasons. First, polar representations enable a natural extension for the popular and important class of elliptical distributions, and second, they permit a convenient description of random vectors for extreme value analysis as their asymptotic behavior is associated to the radial component. Up to now, however, the limit results are obtained only for a relatively narrow class of distributions with polar representations under rather technical and intransparent assumptions.

The major intention of this thesis is to weaken assumptions usually required on these polar representations in different directions, such that only intuitive and natural restrictions remain and the limit distribution of the random vector given an extreme component still can be deduced. Thus, our results facilitate a much better applicability of the obtained limit statements, which is of importance for practical purposes. Altogether, we enrich the class of distributions for bivariate random vectors in the context of modeling their extreme behavior. Throughout this thesis, we put a special emphasize on geometrical analysis of the assumptions, of the results, as well as of the ideas of proofs, which makes underlying characteristics visible and, hence, permits deeper insights strengthening the intuition behind the theoretical statements.

Next in this chapter, we present the structure of the thesis as well as its main contributions. At first in Section 1.1 we consider the multivariate extreme value theory and refer to major works in this research field. A particular attention is drawn to the conditional extreme value models, which are in the focus of our work. Then in Section 1.2 we specify the object of our analysis by providing some preliminary definitions and results. For this purpose we introduce the class of random vectors with a polar representation as a generalization of elliptical random vectors and review the recent literature regarding the corresponding conditional limit statements. In Section 1.3 we summarize the main results of the thesis with the emphasis on the geometric approach developed and exploited for our analysis. Finally, in Section 1.4 we provide a detailed review of the three research papers which constitute the main body of the dissertation.

1.1 The multivariate extreme value theory and the conditional extreme value models

Extreme events do not occur frequently, however, they are quite important as they may have devastating consequences. In a univariate case, we classify an event as *extreme* if the absolute value of the realization of the considered random variable is very large and exceeds a particularly designated high level (cf. McNeil et al. 2005, p. 275). The extreme value theory characterizes such events by probabilistic and statistical statements which are useful in numerous applications, e.g. in medicine, finance, meteorology, engineering etc. A lot of work has been done for studying the extreme behavior of univariate random variables, moreover, a particular interest is drawn to the expanding field of the multivariate extreme value theory. Next we briefly discuss some concepts and results of the extreme value theory which are of importance in this thesis.

In the univariate extreme value theory, the distribution of a random variable is studied under the assumption that it belongs to the max-domain of attraction of some extreme value distribution; this concept is explained below in detail, see (1.4). Analogously, the conventional approach for modeling multivariate extreme values assumes that the joint distribution is in a multivariate max-domain of attraction. This condition requires that each marginal distribution belongs to the max-domain of attraction of some univariate extreme value distribution. There exists many important contributions on the multivariate extreme value theory, among them de Haan and Resnick (1977), Resnick (1987, 2007), Tawn (1990), Coles and Tawn (1991), Beirlant et al. (2004), de Haan and Ferreira (2006). In this context the important concepts are for instance the coefficient of tail dependence developed by Ledford and Tawn (1996, 1997, 1998) or the hidden regular variation addressed among others by Resnick (2002), Maulik and Resnick (2004), Das et al. (2013) and Lindskog et al. (2014).

However, we are often interested in situations where not all components of a random vector but only some of them appear to be extreme. To cope with such situations, Heffernan and Tawn (2004) proposed an innovative and flexible model. They focused on a single component X_i of the random vector $(X_1, \ldots, X_n)$ being large and examined the limiting conditional distribution of the remaining components given $X_i = x$ for $x \to \infty$. This approach provides the opportunity for modeling multivariate data by assuming one component rather than all components of the vector to be extreme. This asymmetry in the treatment of components fits to several extreme situations in practice. For example, a flood is extreme if it exceeds a certain high level; and if a flood is extreme, we are also interested in its duration as it could threat the dyke stability.

Further, Heffernan and Resnick (2007) introduced a mathematical framework for a theory of conditional distributions given an extreme component and introduced the terminology of the conditional extreme value (CEV) model for a bivariate random vector (X, Y). Their theory bases on the assumption of the existence of a vague limit for the joint distribution of a suitably normalized random vector (X, Y), which is equivalent to assuming the existence of the limiting conditional distribution of the normalized variable Y given $X > x$, $x \to \infty$. This approach allows for connections between the CEV and the conventional multivariate extreme value theory as it implies that X is in the max-domain of attraction of some extreme value distribution. Furthermore, they developed the technique of *random norming*, which appears to be very useful in this thesis and is briefly explained below, see (1.10).

Das and Resnick (2011a) clarified the relationship between the CEV and the multivariate extreme value theory, whereas Das and Resnick (2011b) proposed a methodology to check for bivariate data whether the conditions for a CEV model are fulfilled and applied it to different data sets. The parameter estimation of the CEV model was discussed in Fougères and Soulier (2012). Recently, Das and Resnick (2014) proposed an extension of the CEV detection techniques from Das and Resnick (2011b) which is suitable for further models. Resnick and Zeber (2014) placed the approach of Heffernan and Tawn (2004) in a context based on transition kernels $K(x, A) := P(Y \in A|X = x)$ which specify the dependence between X and Y and compared the CEV model of Heffernan and Resnick (2007) with the original approach of Heffernan and Tawn (2004). Using the ideas of the CEV approach, Eastoe et al. (2014) proposed new nonparametric estimators for spectral measures and Pickands dependence functions which characterize multivariate extreme value distributions.

The above-mentioned studies on the CEV models start from assuming certain CEV model conditions and investigate their consequences and implications. However, there is also another direction of the CEV research which is focused on studying how to model the joint distributions of random vectors such that conditional limit statements can be deduced. This direction is further developed in this thesis, where we consider a class of distributions for random vectors (X, Y) such that for fixed ξ, ζ and $x \to \infty$ holds

$$P(X \leq \alpha(x) + \beta(x)\xi,\ Y \leq \gamma(x) + \delta(x)\zeta \mid X > x) \to G(\xi, \zeta) \tag{1.1}$$

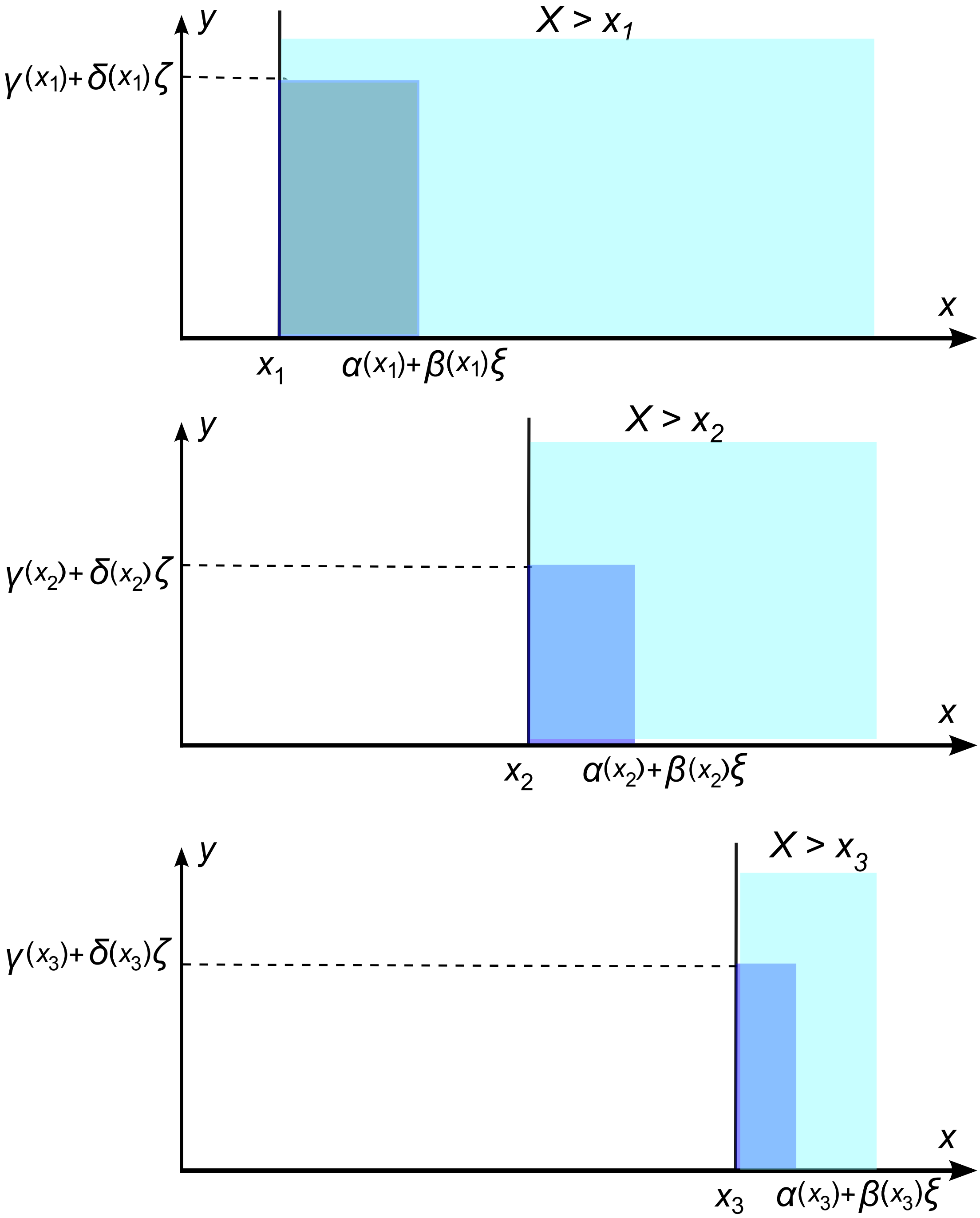


Figure 1.1: Illustration of conditional limit statement (1.1): half-planes $(x, \infty) \times (-\infty, \infty)$ (light-blue) and events $\{x_i < X \leq x_i + \beta(x_i)\xi, Y \leq \rho x_i + \delta(x_i)\zeta\}$ (dark-blue) for values $x_1 < x_2 < x_3$ with constant ξ and ζ.

for normalizing functions α, β, γ, δ and a non-degenerate distribution function G.

We illustrate the conditional limit statement (1.1) in our Figure 1.1. The half-planes $(x,\infty)\times(-\infty,\infty)$ are provided with coordinates ξ, ζ and are mapped onto each other by affine transformations with the normalizing functions, such that the probabilities of $(X,Y)\in(x,\alpha(x)+\beta(x)\xi]\times(-\infty,\gamma(x)+\delta(x)\zeta]$ under the condition $X>x$ become asymptotically equivalent for $x\to\infty$, i.e. that the ratio of probabilities corresponding to the dark-blue and the light-blue regions in all plots of Figure 1.1 remains asymptotically constant.

Some results of this kind were obtained in multivariate cases by Hashorva (2006), Balkema and Embrechts (2007), but primarily in a bivariate context by Berman (1983), Abdous et al. (2005), Fougères and Soulier (2010) and Hashorva (2009, 2012). All these works consider the class of elliptical and generalized distributions with a polar representation, on which we will have a closer look in the next section.

1.2 A generalization of elliptical distributions: Random vectors with polar representation

Now we consider several representations for bivariate random vectors (X,Y) which are useful for our analysis. Since we are interested in the asymptotic behavior for X becoming large, we consider $(X,Y)\in[0,\infty)\times(-\infty,\infty)$ on the right half-plane. Any bivariate random vector can be represented in *Euclidean polar coordinates*:

$$(X,Y)\stackrel{d}{=}A\cdot(\cos\Theta,\sin\Theta) \tag{1.2}$$

with Euclidean distance $A\geq 0$ and Euclidean angle $\Theta\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$. However, in many cases a representation as in (1.2) may not be appropriate for modeling because in general A and Θ are dependent, which occurs for instance even for the bivariate normal distribution with some non-vanishing correlation.

The popular class of *elliptical distributions* characterized by elliptical level lines of the joint density can be more conveniently described by

$$(X, Y) \stackrel{d}{=} R \cdot (\cos T, \rho \cos T + \sqrt{1-\rho^2} \sin T), \tag{1.3}$$

with the parameter $\rho \in (-1, 1)$ and stochastically independent $R \in [0, \infty)$ and T which is uniformly distributed on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Note that neither the radial component R nor the angular component T coincide with the Euclidean distance A resp. angle Θ (except for $\rho = 0$).
The class of elliptical distributions of (X, Y) covers among others the bivariate normal distribution with the radial component R possessing the density $h(r) = r \exp(-0.5r^2)$, the bivariate Student t-distribution with $h(r) = C \cdot r(1 + r^2/n)^{-(n+2)/2}$, the bivariate logistic distribution with $h(r) = C \cdot r \exp(-r^2)/(1 + \exp(-r^2))^2$ and the bivariate Kotz distribution with $h(r) = C \cdot r^{s-1} \exp(-r^s)$ for suitable normalizing constants C. Elliptical distributions are of particular interest for applications, as they can provide very different tail behavior and are commonly used in financial and risk models, cf. Hult and Lindskog (2002), Klüppelberg et al. (2007), Manner and Segers (2011).

Elliptically distributed random vectors have been intensively investigated with respect to their conditional limit behavior. For this purpose the radial component R describing the tail behavior of (X, Y) is assumed to be in the max-domain of attraction of some extreme value distribution, i.e. there exist sequences of normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that for the distribution function H of R it holds

$$\lim_{n\to\infty} H^n(a_n x + b_n) = K(x) \tag{1.4}$$

with a non-degenerate distribution function K. Then, due to the Fisher-Tippett Theorem (see e.g. Embrechts et al. 1997, Th. 3.2.3) K belongs to the type of one of the following distribution functions which are called *extreme value distributions*:

$$
\begin{array}{llll}
\text{Gumbel} & \Gamma(x) & = & \exp\left(-\mathrm{e}^{-x}\right), \quad x \in \mathbb{R}, \\[2ex]
\text{Fréchet} & \Phi_\alpha(x) & = & \begin{cases} 0, & x \leq 0 \\ \exp\left(-x^{-\alpha}\right), & x > 0 \end{cases} \quad \text{with } \alpha > 0, \\[2ex]
\text{Weibull} & \Psi_\alpha(x) & = & \begin{cases} \exp\left(-(-x)^{\alpha}\right), & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \text{with } \alpha > 0.
\end{array}
$$

These extreme value distributions exhibit quite distinct behavior. Among other properties, they differ in the right endpoint $x_E := \sup\{x : H(x) < 1\}$ of the distribution of R: For H being in the Weibull max-domain of attraction it holds $x_E < \infty$, in the Fréchet case it holds $x_E = \infty$, and in the Gumbel case we can either have $x_E < \infty$ or $x_E = \infty$. Hence, for the radial component R with an infinite right endpoint only the Fréchet or Gumbel max-domains of attraction with $x_E = \infty$ are appropriate.

The Fréchet case Φ_α comprises *heavy-tailed* distributions, like Pareto, Cauchy and Student t-distributions. R being in the Fréchet max-domain of attraction is equivalent to R having a regularly varying survival function $\overline{H} = 1 - H$ with the variation index $-\alpha$ (de Haan 1970, Th. 2.3.1), meaning that for all $\lambda > 0$ it holds:

$$
\frac{\overline{H}(\lambda x)}{\overline{H}(x)} = \frac{P\{R > \lambda x\}}{P\{R > x\}} \to \lambda^{-\alpha} \quad \text{for } x \to \infty\,. \tag{1.5}
$$

The Gumbel case Γ covers a large variety of distributions with very different tail behavior: *light-tailed* distributions like normal or exponential ones where the survival function $\overline{H}$ decreases at least exponentially fast as well as *mildly heavy-tailed* distributions like log-normal or heavy-tailed Weibull ones where $\overline{H}$ decreases faster than any power function but slower than an exponential function. Due to the richness of the Gumbel case, Embrechts et al. (1997) characterized it as "perhaps the most interesting among all maximum domains of attraction".

In contrast to the Fréchet case, the Gumbel case with $x_E = \infty$ implies that the survival function $\overline{H}$ is rapidly varying (Resnick 1987, p.53), meaning that:

$$\frac{\overline{H}(\lambda x)}{\overline{H}(x)} = \frac{P\{R > \lambda x\}}{P\{R > x\}} \to \lambda^{-\infty} := \begin{cases} \infty, & 0 < \lambda < 1 \\ 0, & \lambda > 1 \end{cases} \quad \text{for } x \to \infty\,. \tag{1.6}$$

More precisely, a random variable R with distribution function H and survival function $\overline{H} = 1 - H$ is in the Gumbel max-domain of attraction with $x_E = \infty$ if and only if $1/\overline{H}$ is Γ-varying, i.e. there exists a positive function ψ such that for all $z \in \mathbb{R}$ it holds:

$$\lim_{x\to\infty} \frac{\overline{H}(x + z\psi(x))}{\overline{H}(x)} = \lim_{x\to\infty} \frac{P\{R > x + z\psi(x)\}}{P\{R > x\}} = \mathrm{e}^{-z}. \tag{1.7}$$

Then we say that R resp. H is of type $\Gamma(\psi)$. The characterization (1.7) for distributions in the Gumbel max-domain of attraction was introduced by Gnedenko (1943, Th. 7). The important properties of ψ based on the Γ-variation are provided in the book of Geluk and de Haan (1987), e.g. ψ is unique up to asymptotic equivalence (i.e. a positive function ψ_2 is an auxiliary function for R of type $\Gamma(\psi_1)$ if and only if $\psi_1 \sim \psi_2$) and can be chosen differentiable with $\psi'(x) \to 0$ for $x \to \infty$. The auxiliary function ψ is Beurling slowly varying, i.e. $\psi(x + z\psi(x))/\psi(x) \to 1$ for $z \in \mathbb{R}$, $x \to \infty$. The important standard normal, exponential and log-normal distributions are of type $\Gamma(1/x)$, $\Gamma(1)$, $\Gamma(x/\ln x)$, the Weibull distribution with survival function $\overline{H}(x) = \exp(-cx^\beta)$, $c, \beta > 0$ is of type $\Gamma(x^{1-\beta}/(c\beta))$.

Furthermore, the auxiliary function ψ allows the *von Mises representation* for the radial component R of type $\Gamma(\psi)$:

$$\overline{H}(x) = P\{R > x\} = a(x) \cdot \exp\left(-\int_0^x \frac{1}{\psi(u)}\mathrm{d}u\right), \quad x \geq 0 \tag{1.8}$$

with $\lim_{x\to\infty} a(x) = a \in (0, \infty)$ (Balkema and de Haan 1972, p.1354, Resnick 1987, Prop. 1.4). The characterization for R in the Gumbel max-domain of attraction by the auxiliary function ψ as well as the von Mises representation of $\overline{H}$ based on ψ turn out to be important for our following analysis.

Now we consider the literature results for elliptical distributions in the CEV context. Berman (1983) proved for elliptical (X, Y) with R in the Gumbel max-domain of attraction that the conditional distribution of suitably normalized Y given $X > x$ converges weakly to the normal distribution. Moreover, Abdous et al. (2005) studied the behavior of the conditional probability $P(Y \leq y \,|\, X > x)$ for $x \to \infty$ while y is either a fixed boundary or becomes extreme as well. For this purpose, they considered elliptical random vectors (X, Y) and worked out the difference between the case where R is in the Gumbel resp. Fréchet max-domain of attraction. Based on their results it can be shown that in the first case X and Y are asymptotically independent, while in the second case X and Y are asymptotically dependent.

Furthermore, the popular class of elliptical distributions has been extended in the context of the CEV modeling. Balkema and Embrechts (2007) chose an n-dimensional geometric approach where the level lines of the joint density are compact and convex, but only locally elliptical. A more detailed discussion of their setting is provided in Chapter 4.

A recent approach for such generalizations is based on the idea to extend representation (1.3) for random vectors (X, Y) to even more broad *polar representations*

$$(X, Y) \stackrel{d}{=} R \cdot (u(T), v(T)) \tag{1.9}$$

with stochastically independent *polar components* $R \in [0, \infty)$ and T defined on some closed interval in $\mathbb{R}$, and quite arbitrary deterministic *coordinate functions* u and v, where u is bounded. The radial component R is usually assumed to be in the Gumbel max-domain of attraction. Fougères and Soulier (2010) investigated the conditional asymptotic behavior of random vectors (X, Y) with the polar representation (1.9), where R and T are independent, the level curve $(u(t), v(t))$ is locally convex and possesses a unique global maximal x-value in the (x, y)-plane. They pointed out that an arbitrarily small sector around this maximum determines the asymptotic behavior of (X, Y). Hashorva (2012) considered bivariate "scale mixture" random vectors $(X, Y) = R \cdot (U, V)$ for stochastically independent R and (U, V) without explicitly introducing an angular component. He analyzed the conditional limit behavior of (X, Y)

for the cases where U and V are deterministic functions of some other random variable ("model B") and where they are not ("model A").

Polar representations with their radial and angular components appear to be especially appropriate for our investigations: they cover the important elliptical and generalized distributions, moreover, they suit well to CEV analysis as the conditioning variable X corresponds to the radial component R which is shown in Theorem 2.1(i) in Chapter 2. Hence, we choose (1.9) together with the set of assumptions in Fougères and Soulier (2010) as a starting point for our analysis.

1.3 Aims of the thesis and major results

In this thesis we study the asymptotics of bivariate random vectors using a polar representation. In particular, we obtain novel statements for the conditional limit behavior of a bivariate random vector (X, Y) given that X becomes extreme. We contribute this research field mainly in the following directions: first, we introduce a geometric approach to these problems which is helpful for understanding the meaning of the technical assumptions and results; second, we develop an effective and intuitive representation for the required assumptions; third, we deduce several generalizations of previous limit theorems; fourth, we link our findings to those from the CEV theory. Next we briefly explain these contributions.

Geometric interpretations of the model assumptions permit a deeper insight and understanding. They can make underlying characteristics visible, such that further improvements and generalizations become possible. Geometric way of thinking is a recent promising direction in extreme value analysis, cf. e.g. Nolde (2010), Balkema et al. (2013). Inspired by the geometric approach in the book of Balkema and Embrechts (2007), in this thesis we consequently use a geometric language in order to illustrate the essential ideas and to facilitate the intuition behind the theoretical results and proofs.

In order to characterize the underlying assumptions, first note that a representation of the bivariate random vector (X, Y) by three variables R, $u(T)$ and $v(T)$ as in (1.9) makes the choice of these quantities ambiguous. To resolve this problem we reformulate the model via quantities which are independent from each other. For this purpose, we describe the curve $(u(t), v(t))$ locally by a single function l instead of the two coordinate functions u and v. Illustrated in Figure 1.2, the function l describes the geometry of the level curve $\{r = 1\}$ by displaying the horizontal distance to the vertical ray $\{x = 1\}$. Such a representation by a single function l is possible under the assumptions in Fougères and Soulier (2010) which require that u has a unique global maximum 1 at $t = 0$, increases strictly on $(-\epsilon, 0)$ and decreases strictly on $(0, \epsilon)$ and that v increases strictly on $(-\epsilon, \epsilon)$ for some $\epsilon > 0$.

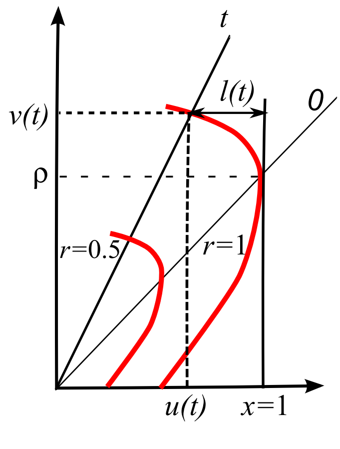


Figure 1.2: Polar representation: coordinate functions u, v form the level curve $\{r = 1\}$, i.e. the curve $(x, y) = (u(t), v(t))$, and the function l describes the horizontal distance between $\{r = 1\}$ and the vertical ray $\{x = 1\}$. The r-level curves possess their maximal x-value at the ray $\{y = \rho x\} = \{t = 0\}$ with $\rho := v(0)$.

Furthermore, we show that several technical requirements commonly met in the literature can be waived. In Section 2.5 we present both the original assumptions in Fougères and Soulier (2010) and how they can be reformulated. Altogether, we describe the polar model as effective as possible without loss of generality such that it becomes more convenient both in view of theoretical investigations and possible applications. More precisely, we characterize the model for random vectors $(X, Y) = R \cdot (u(T), v(T))$ by the following reformulated assumptions (cf. Section 2.3):

(A1) $R \geq 0$ and $T \in [-1, 1]$ are stochastically independent random variables;

(A2) R is of type $\Gamma(\psi)$ and T possesses a density g which is regularly varying at $t = 0$ with some variation index $\tau > -1$, i.e. $g(\lambda s)/g(s) \to \lambda^\tau$ for $s \to 0$;

(A3) the geometry of the curve $(u(t), v(t))$ is described in some neighborhood of $t = 0$ by a single function l which has a unique zero at $t = 0$ and its derivative l' is regularly varying at $t = 0$ with variation index $\kappa - 1$ for $\kappa > 1$.

These reformulated assumptions allow a better understanding of the problem of interest.

Next, in this thesis we provide several important generalizations of the results in the current literature. In Chapter 2, we resolve the important issue how to mitigate the assumptions on the coordinate functions u and v such that there remain only natural restrictions for the geometry of the "r-level curves", i.e. $t \mapsto (r \cdot u(t), r \cdot v(t))$. Up to now, only the case $\kappa > 1$ has been considered (in the special case of elliptical random vectors it holds $\kappa = 2$), i.e. only the case of (at least locally) convex r-level curves. We investigate the case $\kappa \leq 1$ which implies a qualitatively different asymptotic behavior, such that the original limit results according to the usual CEV statement (1.1) do not hold any longer. We succeed to solve this problem and obtain limit results by applying random norming (cf. Heffernan and Tawn 2004, Heffernan and Resnick 2007). The importance of this generalization is displayed in Figure 1.3 showing rather different scatter plots for $\kappa < 1$ in contrast to $\kappa > 1$. Furthermore, we remove the requirement on u to have a unique global maximum as well as allow for asymmetric behavior of the functions u and v. These generalizations make quite different forms of the r-level curves possible and, hence, permit a lot of freedom for describing the conditional asymptotic behavior of random vectors (X, Y) with a polar representation.

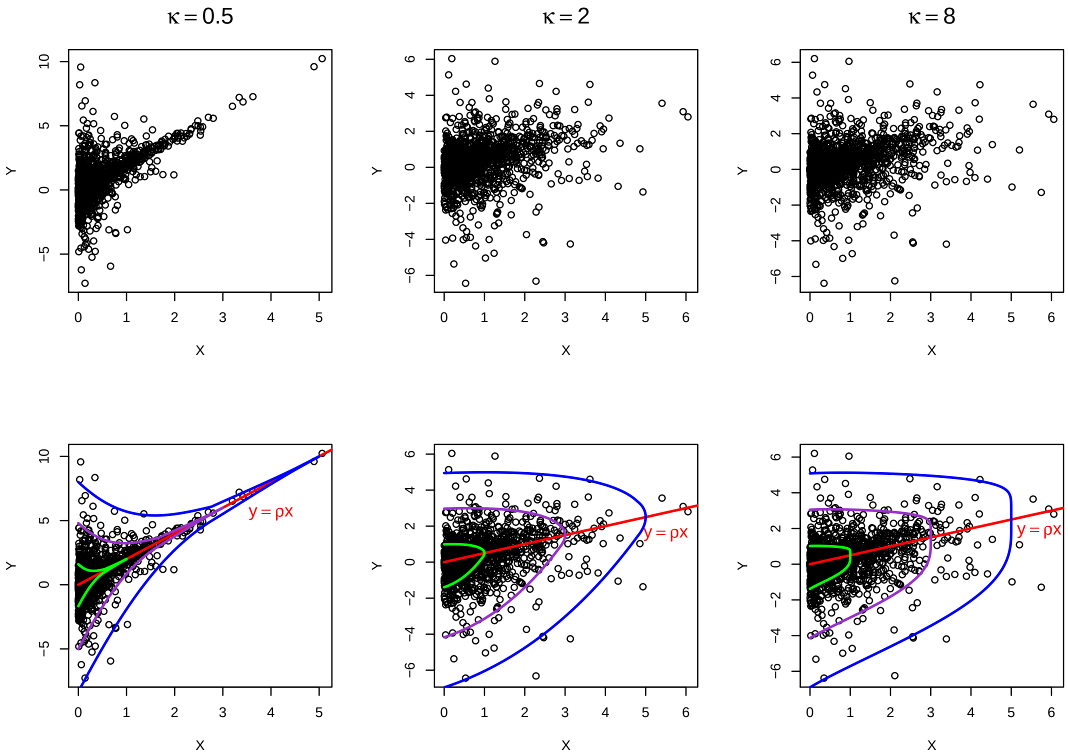


Figure 1.3: Comparison of random vectors with polar representation for different curve orders $\kappa = 0.5, 2, 8$ and the corresponding level curves $\{r = 1\}$, $\{r = 3\}$, $\{r = 5\}$: scatter plots are based on the same realizations of (R, T) with sample size $n = 2000$ where R is exponentially and T is uniformly distributed.

In Chapter 3 we suggest a further weakening of the assumptions by showing that no conditions on differentiability are necessary: We mitigate Assumption (A3) by permitting non-continuous functions l, u and v and by only assuming that l (and not l') is regularly varying with some variation index $\kappa > 0$ and deduce the corresponding conditional limit results for (X, Y). This extension requires a substantial change of proofs in contrast to those in Chapter 2, so that we apply appropriate methods from the probabilistic measure theory.
To sum it up, we work out the assumptions to be convenient and natural, permitting a lot of freedom for the geometry of the curves $(u(t), v(t))$. In particular, arbitrary curve orders $\kappa > 0$ as well as several global maxima of u are now permitted.

Another important point is that the usual stochastic independence assumption for the polar components R and T is rather rigid. This problem can be illustrated by the following argument: The stronger is the allowed dependence between R and T, the larger is the class of distributions possessing some representation $R \cdot (u(T), v(T))$; if R and T would be allowed to be arbitrarily dependent, then every distribution would possess a representation $R \cdot (u(T), v(T))$. In Chapter 4 we investigate how to weaken the independence assumption (A1) for R and T, such that the conditional limit theorems still can be deduced. For this purpose, we introduce a novel measure for the dependence structure and present convenient criteria for validity of limit theorems which possess geometrical meaning. Allowing some degree of dependence yields a flexible and stable model for bivariate random vectors to study their asymptotic behavior. Thus, our results also contribute to the established limit results for elliptical distributions by showing that they possess a certain stability with respect to deviations from the independence, which is particularly important in applications. Additionally, we compare our approach with a quite different one of Balkema and Embrechts (2007) which also allows a certain degree of dependence and analyze how far we extend their results.

Finally, we link our results to those from the CEV theory. This is done in Chapter 3, where we consider a univariate random variable X represented through a random pair

(R, T) and a deterministic function u as $X = R \cdot u(T)$. Under quite weak assumptions we deduce there a novel limit result in Theorem 3.1 for the polar components (R, T) given $X > x$, $x \to \infty$, which opens up new avenues for studying the conditional extreme behavior. In particular, we show that this theorem for the representation of the *single* random variable X permits us to obtain in an elegant and straightforward manner the conditional limit theorems for random *pairs* $(X, Y) = R \cdot (u(T), v(T))$ given that X becomes large.

We underscore the asymmetry between the variables X and Y in the CEV models by pointing out that the essential assumptions and results are primarily made on the conditioning variable X in terms of the pair (R, T). Following this approach, we both deduce new limit statements and recover (under considerably weaker assumptions) the results obtained previously in the literature. Thus, we provide another way of studying CEV models using polar representations.

As our limit theorem is (to the best of our knowledge) the only result focusing primarily on the asymptotic behavior of radial and angular components, it suggests some novel interesting opportunities which are also useful for extreme value problems beyond the CEV framework. For example, Demichel et al. (2015) investigate the asymptotic behavior of the diameter (i.e. the maximum interpoint distance) of a random cloud of n elliptical random vectors for $n \to \infty$. Their approach is based on the idea that the distance between two realized elliptical random vectors is large if they both possess a large norm and are located approximately on the opposite sides of the cloud. They exploit our Theorem 3.1 on the polar components for investigating the localization problem of random vectors with a large norm and, hence, for determining the limit distribution of the diameter.

1.4 Outline of the following chapters

In this section we give an outline of the thesis. For this purpose, we now introduce the structure of Chapters 2, 3 and 4 with the emphasis on the novel findings and results.

In Chapter 2 we study the asymptotic behavior of bivariate random vectors with polar representations given that one vector component becomes large. For this purpose we suggest an effective model which grounds on assumptions which are both very intuitive and natural due to the proposed geometric interpretations as well as more general than those in the previous model of Fougères and Soulier (2010). In particular, we obtain novel interesting results for situations where only random norming leads to a non-degenerate limit statement.

After providing an introduction and the necessary definitions in Sections 2.1 and 2.2, we first reformulate the assumptions in Fougères and Soulier (2010) in Section 2.3 in terms of functions and parameters such that they have a simpler structure, display more clearly the geometry of the curves $(u(t), v(t))$ and allow us to obtain some interesting generalizations. The equivalence of our assumptions with the original ones is shown in Section 2.5. In Theorem 2.1 we obtain the result of Fougères and Soulier (2010) under the reformulated assumptions. Further, in Propositions 2.1 and 2.2 we present two interesting modifications of this result.

In Section 2.4 we deduce the generalizations, whereby our form of the assumptions permits us to analyze cases which are not covered by the original set of assumptions. In Section 2.4.1 we discuss the case that both functions l (which describes the geometry of the r-level curves) and g (which is the density of T) are allowed to be regularly varying at $t = 0$ with different variation indices from below resp. from above of this point. We also weaken the assumption that u possesses a unique global maximum 1 such that it may now have $n > 1$ global maxima equal to 1 or may take the value 1 for all t in some interval $[t_1, t_2]$. These results are stated in Propositions 2.6, 2.7 and 2.8 which are placed in Section 2.6.

In Section 2.4.2 we investigate how the restriction $\kappa > 1$ can be dropped so that the r-level curves do not have to possess any longer vertical tangents at the ray $\{y = \rho x\} = \{t = 0\}$, but may form "cusps" with branches tangent to this ray. The fundamental difference between the cases $\kappa < 1$ and $\kappa > 1$ (for illustration see the plots in Figure 1.3) is analyzed there in detail. Moreover, we explain the essential distinctions between these cases and the consequences of these distinctions in a geometrical way using Figures 2.3, 2.4 and 2.6.
We obtain the following interesting results: In Proposition 2.3 we show that in the special case $\rho = 0$ the original conditional limit result is valid for all $\kappa > 0$. Considering the generic case $\rho \neq 0$ with the method of CEV as in (1.1) leads to a degenerate limit G for $\kappa < 1$ as it is shown in Theorem 2.2. Thus, the question arises: how to obtain non-degenerate limit results for this setting?
We present a novel approach in order to solve this problem: first, in Proposition 2.4 we deduce a limit statement using a linear transformation of Y. Then, in Theorem 2.3 we apply *random norming*, such that we obtain a non-degenerate conditional limit result where the bound on Y is evaluated at the actual value of X instead of the threshold value x for $x \to \infty$:

$$P(X \leq \alpha(x) + \beta(x)\xi,\ Y \leq \gamma(X) + \delta(X)\zeta \mid X > x) \to G(\xi, \zeta). \tag{1.10}$$

The method of random norming was implicitly introduced in the pioneering paper of Heffernan and Tawn (2004) in their extrapolation algorithm and further investigated by Heffernan and Resnick (2007, sect. 4). Our result in Theorem 2.3 appears to be useful also for applications as it is shown in Remark 2.8.
In Section 2.5.2 we explain the reasons why our reformulated assumptions are advantageous and why there exist no analogues to the generalizations in Proposition 2.4 and Theorem 2.3 if one bases on the assumptions in Fougères and Soulier (2010). The formal proofs of these results can be found in Section 2.7, while in Section 2.4 we present

the geometric intuition behind them. Particularly worthy to mention is the geometrical interpretation of the reason why the limit for (1.1) is degenerate whereas the limit for (1.10) is non-degenerate, cf. Figure 2.5 and Remark 2.9.

In Chapter 3 we present a novel approach to extreme value analysis of bivariate random vectors with polar representations contributing to the research field of the CEV models. We show that the essential assumptions are primarily made on the conditioning variable X in terms of the pair (R, T) and determine the limit distribution for this pair. Then, using the obtained result, we deduce directly the conditional limit statements for (X, Y). Our analysis is presented in three steps:

Step 1: In Section 3.2 we begin with the univariate random variable X and state our Assumptions A and B on its representation $R \cdot u(T)$, which are quite weak as we omit conditions on differentiability and monotonicity for the coordinate function u usually made in the literature. The main result of this chapter is provided in Theorem 3.1 where we obtain a conditional limit result for the polar components (R, T) given $T > 0$, $X > x$ as $x \to \infty$. In Remark 3.2 we explain the intuition behind this theorem, whereas its proof is given in Section 3.5.

Step 2: In Section 3.3 we present how to remove the restriction $T > 0$ in Theorem 3.1 which corresponds to the originally one-sided concept of regular variation (cf. e.g. Geluk and de Haan 1987, Def. 1.1). For this purpose we state in Assumption C a two-sided version of Assumption B allowing asymmetric behavior on both sides of $t = 0$. We determine the conditional limit distribution of the variable sign(T) given $X > x$, $x \to \infty$ in Proposition 3.1. Next, we present two-sided extensions of Theorem 3.1 using the random norming in Theorem 3.2 and the classical deterministic norming in Theorem 3.3.

Step 3: Finally in Section 3.4 we introduce the second variable $Y = R \cdot v(T)$ and show that conditional limit theorems for the *bivariate* vector $(X, Y) = R \cdot (u(T), v(T))$ can be directly obtained by applying continuous mapping arguments to one of our The-

orems 3.1, 3.2 or 3.3 on the representation of the *single* variable X. For this purpose we investigate systematically several settings of our assumptions concerning the behavior of the functions u and v. In Corollary 3.1 we recover the result of Fougères and Soulier (2010) under weaker assumptions on the coordinate functions u and v. Then in Corollaries 3.2 and 3.3 we obtain new conditional limit statements for further settings of u and v which require different normalizations of Y and lead to different limit distributions.
Moreover, by imposing additional assumptions we manage to unite the three settings considered in Corollaries 3.1, 3.2 and 3.3 such that we obtain a single limit result: First, in Corollary 3.4 we obtain a result similar to Theorem 2.3 from Chapter 2 as a direct consequence of Theorem 3.1. Second, in Corollary 3.5 we investigate a more flexible form for the relation between v and u by allowing not only a linear relation but also a polynomial one.

To summarize, our three step approach provides a better understanding and systematization of conditional limit statements and gives new insights for studying the extreme behavior of random vectors using polar representations with respect to the CEV models. The essential result is obtained for the conditioning variable X in terms of the pair (R, T). Then conditional limit theorems for random vectors (X, Y) given $X > x, x \to \infty$ can be directly gained from this result on X.

In Chapter 4 we further generalize the model for polar representations $(X, Y) = R \cdot (u(T), v(T))$ by weakening the stochastic independence of the polar components R and T, which is assumed in the literature for elliptical and generalized distributions. This assumptions is rather restrictive particularly for applications. Our aim is to study which degree of dependence between R and T could be allowed such that conditional limit statements still remain valid. For this purpose we propose a novel geometric dependence measure which helps us to develop convenient criteria for limit results.

After some preliminaries in Sections 4.1 and 4.2, we discuss in Section 4.3 the classical extreme value model with independent polar components and summarize in Theorem 4.1 the results of Fougères and Soulier (2010, Th. 3.1) and of our Theorem 2.3. Next, we show in Theorem 4.2 that these limit results also hold for *dependent* polar components in case that the conditional distributions of R given $T = t$ have a similar tail behavior with asymptotically equivalent auxiliary functions ψ_t. The proof of this theorem is presented in Section 4.6.

In Section 4.4 we introduce our novel dependence measure such that we manage to obtain model-independent geometric criteria. For this purpose we quantify the dependence between the polar components R and T by comparing $R\cdot(u(T), v(T))$ with some *reference model* $\widetilde{R} \cdot (u(T), v(T))$ where $\widetilde{R}$ and T are independent. More precisely, the difference between the conditional distribution functions $H_t(r) = P(R \leq r \mid T = t)$ and the reference distribution function $\widetilde{H}(r) = P(\widetilde{R} \leq r)$ is measured by *shifts* $\delta_t(r)$. The geometrical meaning of our dependence measure is illustrated in Figure 4.1.
The use of this measure permits us to formulate in Theorem 4.3 the intuitive criteria for limit results for dependent polar components. In the important case of regularly varying auxiliary function for $\widetilde{H}$ we only have to require that the *relative shifts* $\delta_t'(r)$ vanish asymptotically for growing r uniformly in t; then Theorem 4.2 implies that the original limit statements still hold. Note that this criterion covers arbitrarily large absolute shifts $\delta_t(r)$. Alternatively, for the case that $\widetilde{H}$ cannot be chosen to possess a regularly varying auxiliary function, we amend our criterion by a small additional condition.

In a further step, we investigate how to extend our results permitting a stronger deviation from the independence of R and T with non-vanishing relative shifts, so that the auxiliary functions ψ_t are no more asymptotically equivalent. For this purpose we use the freedom in choosing a polar representation for the random vector (X, Y), which is discussed in Remark 4.3. Then we apply this freedom to change the polar representation of (X, Y) for “counterbalancing” the dependence of the initial radial and angular component. This gives us the opportunity to deduce limit results in Theorem 4.4 for relative shifts which do not vanish asymptotically but tend to a t-dependent limit.

The extensions of the polar extreme value model with the weakened independence assumption are illustrated by several examples, which also point out a geometric interpretation of the shifts $\delta_t(r)$ with respect to the level lines of the joint density of (X, Y), cf. Corollary 4.2.

Up to now, to the best of our knowledge, there is only one result provided in the book of Balkema and Embrechts (2007) which implicitly allows for some dependence of polar components. In particular, they introduced a class $\mathcal{L}$ of functions such that the joint density can be multiplied by them without changing the asymptotics. In Section 4.5 we compare our approach for weakening the independence assumption with the one of Balkema and Embrechts (2007). Theorem 4.5 analyzes cases where our Theorem 4.4 extends their result.

This thesis bases on the following three papers:

(1) On conditional extreme values of random vectors with polar representation.
Published in *Extremes* 17(2), 193-219 (2014)

(2) A conditional limit theorem for a bivariate representation of a univariate random variable and conditional extreme values.
With Philippe Barbe (CNRS, Paris, France), the early working paper is available at arXiv:1311.0540 (2013)

(3) Weakening the independence assumption on polar components: Limit theorems for generalized elliptical distributions.
To appear in *Journal of Applied Probability* (2015)

Chapter 2

Generalizations and geometrical analysis for the polar extreme value model [1]

2.1 Introduction

2.1.1 Conditional extreme value models

In many fields of applications it is necessary to model extreme events for which no direct, reliable experience exists, because they occur too rarely. Often one is interested in the tail of a joint distribution of two or more quantities, e.g. hurricanes (force and affected area) or financial market (shares from various sectors).

In the bivariate case this raises the question as to which events are declared as extreme, whether both components become extreme or only one of them. Another issue that appears is how to describe the dependence structure between the quantities.

An effective approach to deal with these problems is the conditional extreme-value (CEV) model based on the concept of Heffernan and Tawn (2004). It was de-

[1] The contents of this chapter are published in SEIFERT, M.I. (2014): On conditional extreme values of random vectors with polar representation. *Extremes* **17**(2), 193–219.

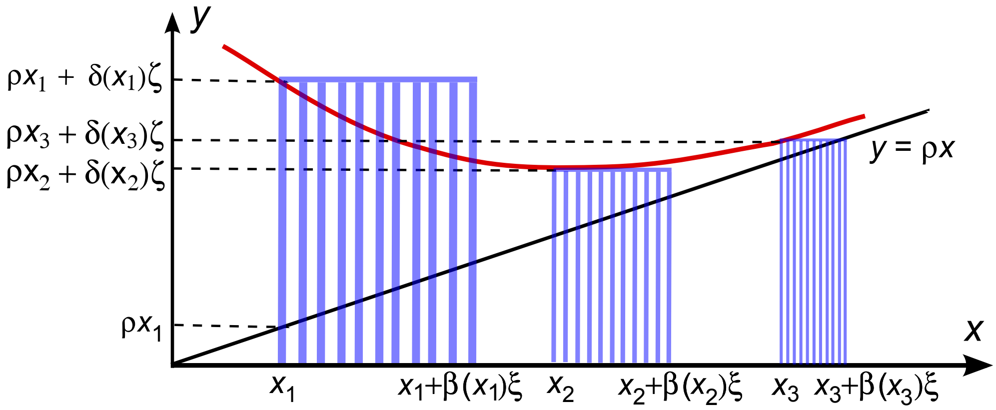


Figure 2.1: Events $\{x_i < X \leq x_i + \beta(x_i)\xi, Y \leq \rho x_i + \delta(x_i)\zeta\}$ for three different values x_1, x_2, x_3 with constant ξ and ζ and the (red) curve joining the edges of all these events, i.e. points $(x, y) = (x, \rho x + \delta(x)\zeta)$; see limit statement (2.2).

veloped and provided with a mathematical framework and terminology by Heffernan and Resnick (2007) and Das and Resnick (2011a). Moreover, distributions of a random vector (X, Y) could be studied such that

$$P(X \leq \alpha(x) + \beta(x)\xi,\ Y \leq \gamma(x) + \delta(x)\zeta \mid X > x) \to G(\xi, \zeta) \text{ for } x \to \infty \tag{2.1}$$

holds with normalizing functions α, β, γ, δ and a non-degenerate distribution function G; in this chapter we treat the special form

$$P(X \leq x + \beta(x)\xi,\ Y \leq \rho x + \delta(x)\zeta \mid X > x) \to G(\xi, \zeta) \text{ for } x \to \infty \tag{2.2}$$

with some $\rho \in \mathbb{R}$.

An advantage of the CEV model is that only one component has to belong to the domain of attraction of an extreme value distribution. Such asymmetry fits to several extreme situations. For example, in a flood the peak is of primary interest but the severity also depends on the values of volume or duration, cf. Yue (2001).

The limit statement (2.2) is illustrated in Figure 2.1: the half-planes $(x, \infty) \times (-\infty, \infty)$ are provided with coordinates ξ, ζ and are mapped onto each other by affine transformations with normalizing functions β and δ, such that the probabilities of $(X, Y) \in (x, x + \beta(x) \cdot \xi] \times (-\infty, \rho x + \delta(x) \cdot \zeta]$ under the condition $X > x$

become asymptotically equivalent for growing x with ξ and ζ remaining constant.

Beginning with the popular bivariate normal distribution, conditional limit statements such as (2.1) were investigated for distributions with the polar representation

$$(X, Y) \stackrel{d}{=} R \cdot (U, V), \tag{2.3}$$

where (U, V) and R are stochastically independent and U is bounded ($P\{|U| < c\} = 1$ for some $c > 0$).

Berman (1983) proved limit theorems, where X and Y have the form of trigonometric polynomials including elliptical distributions, i.e. the isolines of the joint density of (X, Y) are elliptical (see Section 2.3 below).

Balkema and Embrechts (2007) chose an n-dimensional geometric approach to generalize the class of elliptical distributions with no preferred direction. The level lines of the joint density are compact and convex, but are only locally elliptical. They also consider modifications of the joint density which – in two dimensions and the language of polar representations – imply that R and (U, V) are no longer independent. They have influenced the geometric language used in this chapter.

Barbe (2003) investigated the asymptotics for conditional distributions using differential geometric methods, whereby he also placed special emphasis on geometric analysis.

Balkema and Nolde (2010) and Nolde (2010) described multivariate distributions with homothetic densities. They deduced sufficient conditions under which the components of the underlying random vector are asymptotically independent.

Abdous et al. (2005) considered limit statements for elliptical random vectors and pointed out the difference between the case where R is in the Gumbel max-domain of attraction (e.g. for (X, Y) with normal or logistic distribution) and the case where R is in the Fréchet max-domain of attraction (e.g. for (X, Y) with Student t-distribution). From their results it follows that in the first case X and Y are asymptotically independent and in the second case they are asymptotically dependent. In the first case the limit $G(\xi, \zeta)$ is a product $G_1(\xi) \times G_2(\zeta)$ of its marginal distributions, in the second case it is not.

Fougères and Soulier (2010) presented quite general assumptions to obtain a conditional limit theorem for random vectors (X, Y) with the polar representation $R \cdot (u(T), v(T))$, where R and T are independent and R is in the Gumbel max-domain of attraction. They assumed that u has a unique global maximum 1 at t_0 and denoted the value of v at t_0 by ρ. They pointed out that the asymptotic behavior is determined by an arbitrarily small sector containing $y = \rho x$, i.e. $t \in (t_0 - \epsilon, t_0 + \epsilon)$ for any $\epsilon > 0$.

In a series of papers Hashorva analyzed and extended elliptical distributions where limit theorems can be derived (2006, 2007, 2009, 2010, 2012).
In the last of these papers he investigated two models:

A: (X, Y) with representation (2.3) where U and V are "unconstrained"; they are only connected by the assumption that for a certain family of events shrinking down to the "absorbing point" $\{U = 1, V = \rho\}$ the conditional probabilities possess a (3-parameter) limit function,

B: a generalization of the elliptical distribution, where the random variable V has the form $\rho U \pm f(U)$ with an arbitrary positive measurable function f.

2.1.2 Outline of the chapter

The purpose of this chapter is: first, to describe the polar model as effective as possible without loss of generality; second, to investigate generalizations with the aim to have only natural restrictions for the geometry of the "r-level curves", i.e. $t \mapsto (r \cdot u(t), r \cdot v(t))$; and last, to consequently use a geometric language and to illustrate the essential notions and ideas with several figures.

Models like those of Fougères and Soulier (2010) or Hashorva (2012, Model A) can raise the following problems:

- If one expresses the random vector (X, Y) by three variables R, U and V, some fixing of the freedom is needed to make invariant properties of (X, Y) manifest and to display what the extent of these models is.

- In very general cases the assumptions tend to become intricate and implicit.

In Section 2.2 we give definitions and notations for later use.

In view of the problems mentioned above, we reformulate the assumptions of Fougères and Soulier (2010) and their theorem in Section 2.3 in terms of functions and parameters, which imply no mutual restrictions on each other. The equivalence with the original assumptions is shown in Section 2.5.

In Section 2.4 we deduce generalizations, whereby our form of the assumptions permits us to analyze cases which are not covered by the original assumptions.

In Section 2.4.1 we discuss the case where the r-level curves have more than one global maximal x-value and related situations. Detailed results are collected in Section 2.6.

In Section 2.4.2 we investigate the case where the r-level curves are at their maximal x-value x_m no longer tangent to the vertical line $\{x = x_m\}$, but rather form "cusps" with branches tangent to the ray $y = \rho x$. The essential results are:
In the generic case $\rho \neq 0$, the usual method of CEV in (2.2) leads to a degenerate limit G; but we can obtain a non-degenerate limit theorem using random norming (cf. Heffernan and Resnick 2007). This result is not only theoretically instructive but also useful for applications, see Remark 2.8.

Section 2.7 contains the proofs of the theorems and propositions of Sections 2.3 and 2.4.

2.2 Preliminaries

We recall the notions of Γ-variation and of regular variation (Resnick 1987, Geluk and de Haan 1987) which will be used throughout the chapter.

Two functions are said to be asymptotically equivalent if $f(x)/g(x) \to 1$ for $x \to \infty$ (written $f \sim g$).

Definition 2.1. A positive random variable R with distribution function H and survival function $\overline{H} = 1 - H$ is in the Gumbel max-domain of attraction with infinite right endpoint (i.e. $\sup\{r : H(r) < 1\} = \infty$) if and only if $1/\overline{H}$ is Γ-varying, i.e. there exists a positive function ψ (called auxiliary function) such that for all $z \in \mathbb{R}$,

$$\lim_{x\to\infty} \frac{P\{R > x + z\psi(x)\}}{P\{R > x\}} = \lim_{x\to\infty} \frac{\overline{H}(x + z\psi(x))}{\overline{H}(x)} = \mathrm{e}^{-z}. \tag{2.4}$$

Then we say R resp. H is of type $\Gamma(\psi)$.

The auxiliary function ψ is unique up to asymptotic equivalence (i.e. if R is of type $\Gamma(\psi_1)$, then R is also of type $\Gamma(\psi_2)$ if and only if $\psi_1 \sim \psi_2$), satisfy $\lim_{x\to\infty} \psi(x)/x = 0$ and can be chosen differentiable (Geluk and de Haan 1987, Th. 1.28(ii), Cor. 1.29)

Example 2.1. The exponential distribution is of type $\Gamma(1)$, the normal distribution is of type $\Gamma(1/x)$, and the log-normal distribution is of type $\Gamma(x/\ln x)$. $\diamondsuit$

Definition 2.2. A measurable function $h : \mathbb{R} \to \mathbb{R}$ is said to be regularly varying at 0 with index $\alpha \in \mathbb{R}$ (written $h \in \mathrm{RV}_\alpha(0)$), if for all $\lambda > 0$:

$$\lim_{q\to 0} \frac{h(\lambda q)}{h(q)} = \lambda^\alpha. \tag{2.5}$$

A regularly varying function with index 0 is said to be slowly varying.

A function f is said to be $\mathrm{RV}_\alpha(t_0)$ for some $t_0 \in \mathbb{R}$, if $h(t) = f(t - t_0) \in \mathrm{RV}_\alpha(0)$.

We will need a one-sided version of regular variation (in particular in Section 2.4.1), which we define as follows:

A function g is said to be regularly varying at 0 from above (written $g \in \mathrm{RV}_\alpha^+(0)$) resp. from below ($g \in \mathrm{RV}_\alpha^-(0)$) with index $\alpha \in \mathbb{R}$, if for all $\lambda > 0$:

$$\lim_{q \downarrow 0} \frac{g(\lambda q)}{g(q)} = \lambda^\alpha \quad \text{resp.} \quad \lim_{q \uparrow 0} \frac{g(\lambda q)}{g(q)} = \lambda^\alpha. \tag{2.6}$$

Additionally we introduce following notion:
If $f \in \mathrm{RV}_\alpha(0)$ and fulfills

$$\lim_{q \downarrow 0} \frac{|f(q)|}{|f(-q)|} = 1,$$

we call it "infinitesimally symmetric" and write $f \in \mathrm{RV}_\alpha^s(0)$.

Remark 2.1.

(i) A function $f \in \mathrm{RV}_\alpha(0)$ can be represented as $f(t) = F(t) \cdot |t|^\alpha$ with $F \in \mathrm{RV}_0(0)$.

(ii) For $f \in \mathrm{RV}_\alpha(0)$ it follows $f(t) \neq 0$ for $t \in (-\epsilon, \epsilon) \setminus \{0\}$ for some $\epsilon > 0$. $\diamondsuit$

Example 2.2. The following functions are examples of slowly varying functions at 0:

(i) continuous functions F with $F(0) \neq 0$;

(ii) $F(t) = |\log(|t|)|^\alpha$ with $\alpha \in \mathbb{R}$;

(iii) step functions $F(t) = a$ for $t < 0$ and $F(t) = b$ for $t \geq 0$ with $a, b \neq 0$, $a \neq b$. Note that this F is not in $\mathrm{RV}_0^s(0)$. $\diamondsuit$

To describe the geometry of random vectors (X, Y) possessing a polar representation $R \cdot (u(T), v(T))$ with $R \in [0, \infty)$, $T \in [-1, 1]$ we introduce following notations:

r**-level curve** $\{r = r_0\} : t \mapsto (r_0 \cdot u(t), r_0 \cdot v(t))$,

t**-ray** $\{t = t_0\} : r \mapsto (r \cdot u(t_0), r \cdot v(t_0))$.

2.3 The polar extreme value model

Berman (1983) investigated the asymptotic behavior of elliptically distributed random vectors

$$(X,Y) \stackrel{d}{=} R \cdot (\cos\Theta, \rho\cos\Theta + \sqrt{1-\rho^2}\sin\Theta) \tag{2.7}$$

with $\rho \in (-1,1)$ and independent random variables R of type $\Gamma(\psi)$ with values in $[0,\infty)$ and Θ uniformly distributed on $(-\pi,\pi]$.
He (cf. also Abdous et al. 2005) presented for elliptical vectors (X,Y) with representation (2.7) a conditional limit statement (2.2) with the normalizing functions

$$\beta(x) = \psi(x), \quad \delta(x) = x\sqrt{1-\rho^2}\sqrt{\psi(x)/x} \tag{2.8}$$

and the limit distribution $G(\xi,\zeta) = \left(1-\mathrm{e}^{-\xi}\right)\Phi(\zeta)$, where Φ denotes the standard normal distribution function.

Fougères and Soulier (2010) extended this theorem to random vectors with a polar representation on the right half-plane ($X \geq 0$)

$$(X,Y) \stackrel{d}{=} R \cdot (u(T), v(T)), \tag{2.9}$$

where u is bounded and the random variables R and T fulfill:

Assumption 1.

(i) The radial component $R \in [0,\infty)$ is of type $\Gamma(\psi)$,

(ii) R and T are independently distributed.

Remark 2.2. The corresponding map

$$\begin{aligned} [0,\infty)\times[-1,1] &\to [0,\infty)\times\mathbb{R} \\ (r,t) &\mapsto (x,y) = (ru(t), rv(t)) \end{aligned}$$

has to be neither injective nor surjective. A polar representation is not unique: From (2.9) follows $(X,Y) \stackrel{d}{=} \hat{R}\cdot(\hat{u}(\hat{T}), \hat{v}(\hat{T}))$ with $\hat{R} = c\cdot R$, $\hat{u} = u/c$, $\hat{v} = v/c$, $c \in (0,\infty)$

and $\hat{T} = \varphi(T)$, where φ is a bijection of $[-1,1]$. In the following, c is chosen such that $\sup_{t\in[-1,1]} u(t) = 1$. $\diamondsuit$

Instead of the original Assumptions 2_{FS} and 3_{FS} of Fougères and Soulier (2010) we state following assumptions in three modifications A, B and C. In Section 2.5, Assumption 2_{FS} and 2 are compared.

Assumption 2. It holds

$$\begin{aligned} u(t) &= 1 - l(t) \text{ for } t \in [-1,1], \\ v(t) &= (t+\rho)\cdot(1-l(t)) \text{ with } \rho \in \mathbb{R} \text{ for } t \in (-\epsilon,\epsilon),\ \epsilon > 0, \end{aligned} \tag{2.10}$$

where $l : [-1,1] \to [0,1]$ has a unique zero at 0 and for some $\kappa > 0$ its derivative

2_{A}: l' is $\mathrm{RV}_{\kappa-1}(0)$,

2_{B}: l' is $\mathrm{RV}^s_{\kappa-1}(0)$,

2_{C}: $l'(t) = F(t)|t|^{\kappa-1}$, with continuous F and $F(0) = \kappa L_0$ for some $L_0 > 0$.

Remark 2.3. According to Remark 2.1(ii), $u'(t) \neq 0$ in some ϵ-neighborhood of zero and, hence, u increases strictly on $(-\epsilon, 0)$ and decreases strictly on $(0,\epsilon)$. As a consequence l possesses inverses $l_\pm^{-1} \in \mathrm{RV}^+_{1/\kappa}(0)$ and their derivatives $(l_\pm^{-1})' = 1/(l' \circ l_\pm^{-1})$ on $(-\epsilon, 0)$ and $(0,\epsilon)$ are $\mathrm{RV}^+_{1/\kappa-1}(0)$. As $l(0) = 0$, l is $\mathrm{RV}_\kappa(0)$. $\diamondsuit$

Assumption 3. T possesses a density $g : [-1,1] \to [0,\infty)$, which is

3_{A}: $\mathrm{RV}_\tau(0)$,

3_{B}: $\mathrm{RV}^s_\tau(0)$,

3_{C}: $g(t) = \tilde{F}(t)\cdot|t|^\tau$, where $\tilde{F}$ is continuous with $\tilde{F}(0) = G_0 > 0$,

for some $\tau > -1$.

As $l(0) = 0$ and g is integrable, $\kappa < 0$ and $\tau < -1$ are not allowed.

Fougères and Soulier treat only the case $\kappa > 1$ and calculate the probabilities of sets $\{X > x, Y > y\}$ for $y > 0$ by integrating the survival function $\overline{H}$ of R over the boundaries (parameterized by t). With $t_1 = y/x - \rho$ they obtain in the case $\rho > 0$ and $0 \leq t_1 < \epsilon$:

$$P\{X > x, Y > y\} = \int_{\epsilon_-}^{\epsilon} \overline{H}\left(\max\left(\frac{x}{u(t)}, \frac{y}{v(t)}\right)\right) g(t)\,\mathrm{d}t \; + \; rem$$

$$=: \underbrace{\int_{\epsilon_-}^{t_1} \overline{H}\left(\frac{y}{v(t)}\right) g(t)\,\mathrm{d}t}_{=:I} \; + \; \underbrace{\int_{t_1}^{\epsilon} \overline{H}\left(\frac{x}{u(t)}\right) g(t)\,\mathrm{d}t}_{=:J} + rem \tag{2.11}$$

with $\epsilon_- := -\min(\rho, \epsilon)$ (to ensure that the whole sector (ϵ_-, ϵ) is in the first quadrant) and some remainder rem. This is illustrated in Figure 2.2.

In the proof it is shown that the asymptotic behavior of I and J is determined by regularly varying functions k_I and k_J, where k_J has a smaller index, hence $I = o(J)$ for $x \to \infty$. This corresponds to the fact that in the points $(x, \rho x)$ the

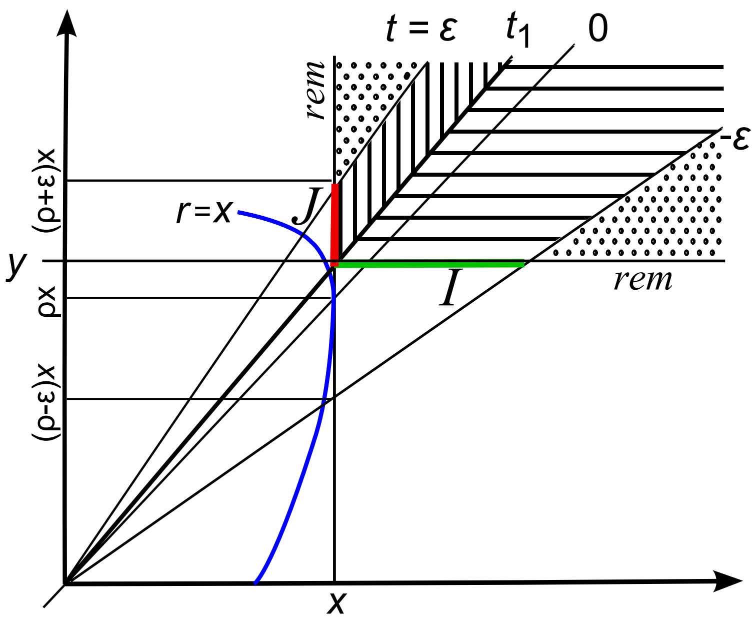


Figure 2.2: The event $\{X > x, Y > y\}$ and the integral domains in its boundary, cf. (2.11).

decrease of $\overline{H}$ is maximal in x-direction and minimal in y-direction, cf. also Figure 2.3(a) in Section 2.4.2.

The remainder rem decreases faster than any rational function as a consequence of Lemma 5.1 in Fougères and Soulier (2010). Hence, only an arbitrarily small sector $\{|t| < \epsilon\}$ determines the asymptotics for $x \to \infty$ and $y = \rho x + \delta(x)\zeta$.

With

$$\eta \;:=\; \frac{1+\tau}{\kappa}, \tag{2.12}$$

$$k \;:=\; k_+ + k_- \text{ with } k_\pm(q) := \pm\Gamma(\eta)\cdot q \cdot \left((l_\pm^{-1})' \cdot g \circ l_\pm^{-1}\right)(q) \in \mathrm{RV}_\eta^+(0), \tag{2.13}$$

$$h \;:=\; \begin{cases} \kappa^{-1/\kappa} \cdot l_+^{-1}, & \text{if } \zeta \geq 0, \\ -\kappa^{-1/\kappa} \cdot l_-^{-1}, & \text{if } \zeta < 0, \end{cases} \in \mathrm{RV}_{1/\kappa}^+(0) \tag{2.14}$$

we state:

Theorem 2.1 (Fougères and Soulier)**.**
Let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumptions 1, 2_B, 3_B with $\kappa > 1$. Then the following statements hold for all $\xi > 0$ and $\zeta \in \mathbb{R}$:

(i) *X is of type $\Gamma(\psi)$ and*

$$P\{X > x\} \sim k\left(\psi(x)/x\right) \cdot P\{R > x\},$$

(ii) $\lim\limits_{x\to\infty} P(Y \leq \rho x + x \cdot h\left(\psi(x)/x\right) \cdot \zeta \mid X > x) = G_{\kappa,\tau}(\zeta),$

(iii) $\lim\limits_{x\to\infty} P(X \leq x + \psi(x)\xi,\ Y \leq \rho x + x \cdot h\left(\psi(x)/x\right) \cdot \zeta \mid X > x)$
$= \left(1 - \mathrm{e}^{-\xi}\right) G_{\kappa,\tau}(\zeta),$

with

$$G_{\kappa,\tau}(\zeta) = \frac{1}{2 \cdot \kappa^{\eta-1}\Gamma(\eta)} \cdot \int_{-\infty}^{\zeta} \exp\left(-|s|^\kappa/\kappa\right)|s|^\tau \,\mathrm{d}s.$$

Remark 2.4. The reason for the choice of h in Theorem 2.1 shows up in the proof, see Remark 2.10 in Section 2.7.1. ♢

For the next proposition we assume that

$$p_{\pm} := \lim_{q\downarrow 0} \frac{k_{\pm}(q)}{k(q)}. \tag{2.15}$$

exist with $k_{\pm}$ from (2.13).

Proposition 2.1. *Let* $(X,Y) = R \cdot (u(T), v(T))$ *satisfy Assumptions 1, 2*$_A$*, 3*$_A$ *with* $\kappa > 1$. *Then the statements of Theorem 2.1 hold with*

$$G_{\kappa,\tau}(\zeta) = p_{-} + p_{\mathrm{sign}(\zeta)} \cdot \mathrm{sign}(\zeta) \cdot M_{\kappa,\tau}(|\zeta|)$$

where

$$M_{\kappa,\tau}(r) = \frac{1}{\kappa^{\eta-1}\Gamma(\eta)} \cdot \int_0^r \exp\left(-s^{\kappa}/\kappa\right) s^{\tau} \, \mathrm{d}s$$

is the regularized incomplete Gamma function $P(r^{\kappa}/\kappa, \eta)$.

Remark 2.5. The quantities p_{-}, $p_{+} \in [0,1]$, with $p_{-} + p_{+} = 1$, denote the weights with which the parts below resp. above the maximum of u contribute to the limit distribution. The example $k_{-}(q) = |\log(q)|q^{\eta}$ and $k_{+}(q) = q^{\eta}$ shows that the values 0 and 1 can occur.

Theorem 2.1 is a special case of Proposition 2.1 with $p_{-} = p_{+} = 0.5$. $\diamondsuit$

Proposition 2.2. *Under Assumptions 2*$_C$*, 3*$_C$ *instead of 2*$_B$*, 3*$_B$*, the functions* k *and* h *from Theorem 2.1 can be chosen as power functions:*

$$k(q) = \frac{2}{\kappa} \cdot \Gamma(\eta) \cdot G_0 \cdot \left(\frac{q}{L_0}\right)^{\eta}, \quad h(q) = \left(\frac{q}{\kappa L_0}\right)^{1/\kappa}.$$

Remark 2.6.

(i) If T possesses a positive, continuous density g (this implies that $\tau = 0$), then the limit statements of Theorem 2.1 are independent of the choice of g.

(ii) Elliptical random vectors (2.7) are special polar ones satisfying Assumptions 2$_C$, 3$_C$ (after having transformed the angular component Θ to T) with $\kappa = 2$, $\tau = 0$, $L_0 = \left(2(1-\rho^2)\right)^{-1}$ and $G_0 = \left(2\pi\sqrt{1-\rho^2}\right)^{-1}$.

Theorem 2.1(ii) and (iii) are fulfilled with the standard normal distribution Φ as limit distribution $G_{\kappa,\tau}$ and $h(q) = \sqrt{1-\rho^2}\sqrt{q}$. Statement (i) of Theorem 2.1 yields $P\{X > x\} \sim \sqrt{\psi(x)/(2\pi x)} \cdot P\{R > x\}$. $\diamondsuit$

2.4 Generalizations

In this section we will investigate how far one can weaken the conditions

(i) function u possesses a *unique* maximum 1,

(ii) function l is regularly varying with index $\kappa > 1$.

2.4.1 Polar representations with several maxima and related situations

In Assumptions 2 and 3 the function l has a unique zero, also l and g are regularly varying with the same index κ resp. τ on both sides. We want to extend this into three directions:

(I) l possesses n zeros at $t_1, \dots t_n$, $n \in \mathbb{N}$,
l and g are regularly varying at t_i with indices κ_i resp. τ_i, $i = 1, \dots n$;

(II) l possesses a unique zero at $t = 0$,
l and g are regularly varying at 0 from below with index κ_1 resp. τ_1, from above with index κ_2 resp. τ_2;

(III) l takes the value zero on the interval $[t_1, t_2]$,
l is regularly varying at t_1 from below with index κ_1 and at t_2 from above with index κ_2,
$\int_{t_1}^{t_2} g(t)\,\mathrm{d}t > 0$.

Analogously to Section 2.3, we define ρ_i by $\{y = \rho_i x\} = \{t = t_i\}$ and $\eta_i := (1+\tau_i)/\kappa_i$, $i = 1, 2$.

To case I: If one cancels the word “unique” in Assumption 2, several zeros of l and, hence, several global maxima of u are possible (local maxima smaller than 1 are already allowed).
An infinite number of global maxima possesses an accumulation point on the compact domain $[-1, 1]$ of T, where the local requirements of Assumption 2, in particular the monotonicity properties, cannot be fulfilled. Hence, only a finite number n of global maximum points is possible in case I.
Let denote $\eta^{\min} := \min(\eta_i | i = 1, \ldots, n)$. Generically, we have a unique $\eta_j = \eta^{\min}$, then this j-th maximum dominates and leads to a non-degenerate limit distribution, which has the form as in Theorem 2.1 resp. Proposition 2.1.
In the special case if $\eta_j = \eta^{\min}$ for several $j \in J \subset \{1, \ldots, n\}$, one obtains – depending on the choice of the bounds for Y – either asymptotically a step function (steps on ρ_j, $j \in J$) or a limit distribution with “point mass” in $\{-\infty\}$ and/or in $\{+\infty\}$, whereas the remainder is non-degenerate on $\mathbb{R}$.
In this special case for two maxima of same order η, we obtain a result similar to Example 2.2.2 in Fougères and Soulier (2010), where the mixture of two bivariate Gaussian vectors is considered (hence $\kappa_1 = \kappa_2 = 2$, $\tau_1 = \tau_2 = 0$, $\eta_1 = \eta_2 = 0.5$).

In case II we permit different variation indices in Assumption 2 and 3 on both sides, which can be interpreted as a limit case $t_1 = t_2$ of case I. For $\eta_1 \neq \eta_2$, the side with the smaller η_i dominates and leads to a non-degenerate limit distribution, which has the support either $(-\infty, 0]$ or $[0, \infty)$. Only for the special case $\eta_1 = \eta_2$ the limit distribution could have the support $(-\infty, \infty)$.

In case III the function l is no longer regularly varying from above at t_1 and from below at t_2, but it can be regarded as a limit case $1/\kappa = 0 = \eta$:
Here the quotient of the survival functions of X and R is asymptotically constant. The limit distribution is determined by the probability mass in the sector $[t_1, t_2]$.

The detailed assumptions and results are presented in Section 2.6.

2.4.2 Polar representations with cusps

If one drops the condition $\kappa > 1$ of Theorem 2.1, the r-level curves might not possess any longer vertical tangents at $y = \rho x$. For $\kappa = 1$ they form “edges”, for $0 < \kappa < 1$ they form “cusps” located on the ray $y = \rho x$.

For the following statements, we need not to specify the modification A, B or C for the Assumptions 2 and 3; the occurring limit distributions $G_{\kappa,\tau}$ are for B or C the same as in Theorem 2.1, for A the same as in Proposition 2.1.

Figures 2.3(a), (b) display the fundamental difference between the cases $\kappa > 1$ and $0 < \kappa < 1$: Near the ray $y = \rho x$ the values of the survival function $\overline{H}(r)$ have their maximal decrease in the case $\kappa > 1$ in the x-direction and in the case $0 < \kappa < 1$ in the directions orthogonal to the ray. Hence for $\rho \neq 0$ the coordinate directions lose the meaning they had in the case $\kappa > 1$.

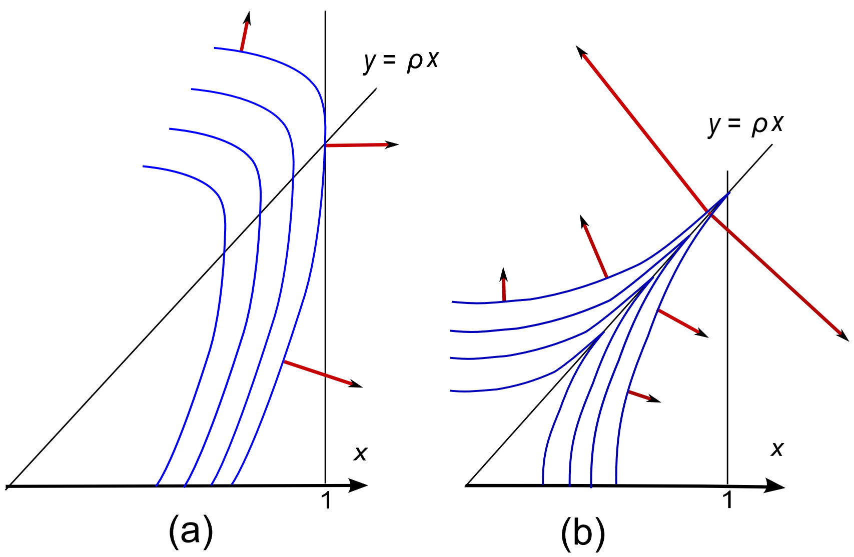


Figure 2.3: r-level curves: (a) for $\kappa > 1$, (b) for $\kappa < 1$. The arrows denote the maximal decrease of $\overline{H}(r)$.

In Assumption 2 the leading term in

$$\rho - v(t) = \rho l(t) - t + t \cdot l(t), t \in (-\epsilon, \epsilon) \tag{2.16}$$

is $-t$ not only for $\kappa > 1$, but also for $\rho = 0$, $\kappa > 0$. This is the reason why our proof of Theorem 2.1 in Section 2.7.1 implies for all $\kappa > 0$:

Proposition 2.3. *Let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumptions 1, 2 and 3 with $\rho = 0$. Then the statements of Theorem 2.1 resp. Proposition 2.1 hold.*

But for $0 < \kappa < 1$, $\rho \neq 0$ in (2.16) the leading term is $\rho \cdot l$, which has the variation index κ. Hence we have

$$\rho - v \sim \rho l = \rho \cdot (1 - u), \quad v' \sim -\rho l' = \rho u' \tag{2.17}$$

(see Figure 2.4), which suggests that I and J (cf. (2.11)) are asymptotically equivalent; the proof is given in Section 2.7.2.

In Figure 2.5 we illustrate the result for the case $\rho > 0$, $0 < \kappa < 1$, first for $\zeta = 0$ (i.e. $t_1 = 0$):

$$I_0 := \int_{\epsilon_-}^{0} \overline{H}\left(\frac{y}{v(t)}\right) g(t)\,\mathrm{d}t, \quad J_0 := \int_{0}^{\epsilon} \overline{H}\left(\frac{x}{u(t)}\right) g(t)\,\mathrm{d}t$$

$$J_0^- := \int_{\epsilon_-}^{0} \overline{H}\left(\frac{x}{u(t)}\right) g(t)\,\mathrm{d}t, \tag{2.18}$$

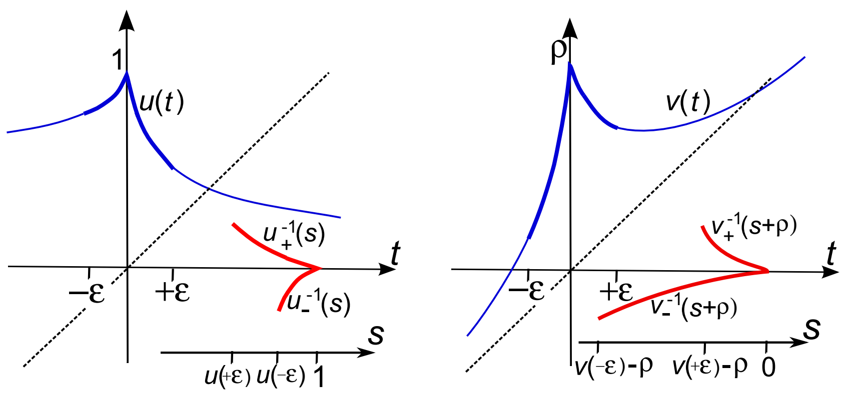


Figure 2.4: Functions u, v and their inverses for $\kappa < 1$.

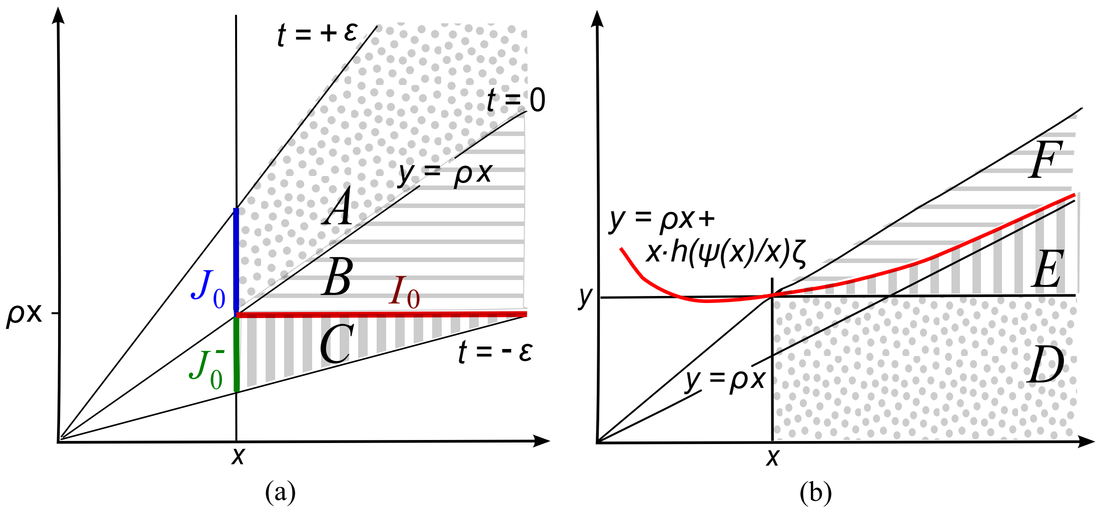


Figure 2.5: Representation of the events A, B, C with the domains of the integrals I_0, J_0, J_0^- for (2.18), (2.19) and of the events D, E, F (see Remark 2.9(i)).

where ϵ_- is defined as in (2.11).

For the following events

$$\begin{aligned} A &:= \{X > x,\ Y > \rho X\}, \\ B &:= \{X > x,\ \rho x < Y \leq \rho X\}, \\ C &:= \{X > x,\ Y \leq \rho x\}, \end{aligned} \tag{2.19}$$

it holds $P(A) \sim J_0, P(B) \sim I_0, P(B) + P(C) \sim J_0^-$.

- For $\kappa > 1$ we had $I = o(J)$ and, hence, $P(B) = o(P(A))$.
- For $0 < \kappa < 1$, in the case of "cusps", now we have $I_0 \sim J_0^-$ and, hence, $P(C) = o(P(B))$, i.e.

$$P\left(Y \leq \rho x \mid X > x\right) = \frac{P(C)}{P(A) + P(B) + P(C)} \to 0.$$

We obtain for $\zeta = 0$, and, consequently, for $\zeta < 0$:

$$P\left(Y \leq \rho x + x \cdot h(\psi(x)/x) \cdot \zeta \mid X > x\right) \to 0 \text{ for } x \to \infty. \tag{2.20}$$

That (2.20) also holds for $\zeta > 0$ is shown in Section 2.7.2, hence:

Theorem 2.2. *Let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumptions 1, 2 and 3 for $\rho \neq 0$ and $0 < \kappa < 1$. Then statement (i) of Theorem 2.1 holds and instead of statement (ii) one obtains a degenerate limit,*

$$\lim_{x \to \infty} P\left(Y \leq \rho x + x \cdot h(\psi(x)/x) \cdot \zeta \mid X > x\right) = 0$$

for $\zeta \in \mathbb{R}$.

The function h was defined in (2.14).

Remark 2.7. For $\kappa > 1$ Fougères and Soulier (2010, Comment (iii) in sect. 3) stated that the case $\rho \neq 0$ can be deduced from the case $\rho = 0$. We complement this by showing that this does not hold for $\kappa < 1$: For $\rho = 0$ but not for $\rho \neq 0$, there exists a non-degenerate limit distribution.
The reason why we cannot obtain for $0 < \kappa < 1$ the result for $\rho = 0$ (Proposition 2.3) as a limit of the result (Theorem 2.2) for $\rho \neq 0, \rho \to 0$, is that for $0 < \kappa < 1, \rho \neq 0$ the leading term $\rho \cdot l = \rho \cdot (1 - u)$ in (2.16) dominates on some neighborhood $(-\epsilon_\rho, \epsilon_\rho)$. For $\rho \to 0$, ϵ_ρ has to go to zero. $\diamondsuit$

To gain a non-degenerate limit result also for $\rho \neq 0$ one can transform $(X, Y) = R \cdot (u(T), v(T))$ to $(X, \tilde{Y}) = R \cdot (u(T), \tilde{v}(T))$ with $\tilde{Y} = Y - \rho X$, $\tilde{v}(t) = v(t) - \rho u(t)$ and $\tilde{\rho} = \tilde{v}(0) = 0$. From Proposition 2.3 it follows for all $\kappa > 0, \rho \in \mathbb{R}$:

Proposition 2.4. *Let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumptions 1, 2 and 3. Then,*

$$\begin{aligned}
&\lim_{x \to \infty} P(X \leq x + \psi(x)\xi,\ Y \leq \rho X + x \cdot h\left(\psi(x)/x\right) \cdot \zeta \mid X > x) \\
&= \left(1 - \mathrm{e}^{-\xi}\right) G_{\kappa,\tau}(\zeta),
\end{aligned}$$

where $G_{\kappa,\tau}$ is defined as in Theorem 2.1 resp. Proposition 2.1.

Further, we can deduce (cf. proof in the end of Section 2.7.1) another non-degenerate conditional limit theorem for *random norming* (cf. Heffernan and Resnick 2007, sect. 4), i.e. in formula (2.2) the normalizing functions for the bound on Y are not evaluated at the threshold value x but at X. This results in the following theorem valid for all $\kappa > 0, \rho \in \mathbb{R}$:

Theorem 2.3 (Random norming)**.**
Let $(X, Y) = R \cdot (u(T), v(T))$ *satisfy Assumptions 1, 2 and 3. Then,*

$$\begin{aligned}&\lim_{x\to\infty} P(X \le x + \psi(x)\xi,\ Y \le \rho X + X \cdot h\left(\psi(X)/X\right) \cdot \zeta \mid X > x)\\ &= \left(1 - \mathrm{e}^{-\xi}\right) G_{\kappa,\tau}(\zeta),\end{aligned}$$

where $G_{\kappa,\tau}$ *is defined as in Theorem 2.1 resp. Proposition 2.1.*

In Section 2.5.2 we will show that these non-degenerate limit results in Proposition 2.4 and Theorem 2.3 cannot be obtained under the original Assumption 2_{FS} in Fougères and Soulier (2010).

Remark 2.8. In their pioneering paper, Heffernan and Tawn (2004) proposed a multivariate regression model together with an application to an analysis of air pollution monitoring data. After having estimated the characteristic quantities, for a limit statement of type (2.1) they extrapolated the joint tail of (X, Y) to extreme values by

$$Y = \gamma(X) + \delta(X) \cdot Z \tag{2.21}$$

where Z is independent of X and has a non-degenerate distribution.
Let us assume that we already have estimated values $\hat{\rho}$, $\hat{\kappa}$, $\hat{\tau}$, $\hat{\psi}$ (for a discussion of such estimators see Fougères and Soulier (2012)).
In the case $\hat{\kappa} < 1$, $\hat{\rho} \neq 0$ a conditional limit statement of form (2.2) can only be gained with a degenerate limit (see Theorem 2.2).
On the other hand, Theorem 2.3 states for all $\kappa > 0$ that X is approximately exponentially distributed with location parameter x and scale function $\hat{\psi}$ and that Y can be

approximated for large X by the regression

$$Y = \hat{\rho}X + X \cdot h\left(\hat{\psi}(X)/X\right) \cdot Z$$

where Z is independent of X and possesses the non-degenerate distribution $G_{\hat{\kappa},\hat{\tau}}$.
Hence, Theorem 2.3 justifies the simulation algorithm:
One simulates X from an exponential distribution with location parameter x and scale function $\hat{\psi}$, samples Z from $G_{\hat{\kappa},\hat{\tau}}$ independently of X and extrapolates Y by $\hat{\rho}X + X \cdot h\left(\hat{\psi}(X)/X\right) \cdot Z$. ♢

Remark 2.9.

(i) In Figure 2.5(a) we confined ourselves to the special case $\zeta = 0$, where the statements of Proposition 2.4 and Theorem 2.3 coincide. In Figure 2.5(b) we consider the case $\zeta > 0$ with the events:

$$\begin{aligned}
D &:= \{X > x,\ Y \leq \rho x + xh\left(\psi(x)/x\right)\zeta\},\\
E &:= \{X > x,\ \rho x + xh\left(\psi(x)/x\right)\zeta < Y \leq \rho X + Xh\left(\psi(X)/X\right)\zeta\},\\
F &:= \{X > x,\ \rho X + Xh\left(\psi(X)/X\right)\zeta < Y \leq \rho X + xh\left(\psi(x)/x\right)\zeta\}.
\end{aligned}$$

These events correspond to the different bounds on Y: for Theorem 2.2 by D, for Proposition 2.4 by $D \cup E \cup F$ and for Theorem 2.3 by $D \cup E$.
The different limit results can be explained as follows:
For large x-values the y-values concentrate sharply along the ray $y = \rho x$, i.e. $\{t = 0\}$. Every sector-tail $\{|t| < \epsilon, r > r_0\}$ for any $\epsilon, r_0 > 0$ intersects for any $\zeta \in \mathbb{R}$ every E, but not necessarily D.

(ii) The densities $d_{\kappa,0}$ of $G_{\kappa,0}$ (i.e. for $\tau = 0$) are distinguished by their kurtosis, as it is illustrated in Figure 2.6. ♢

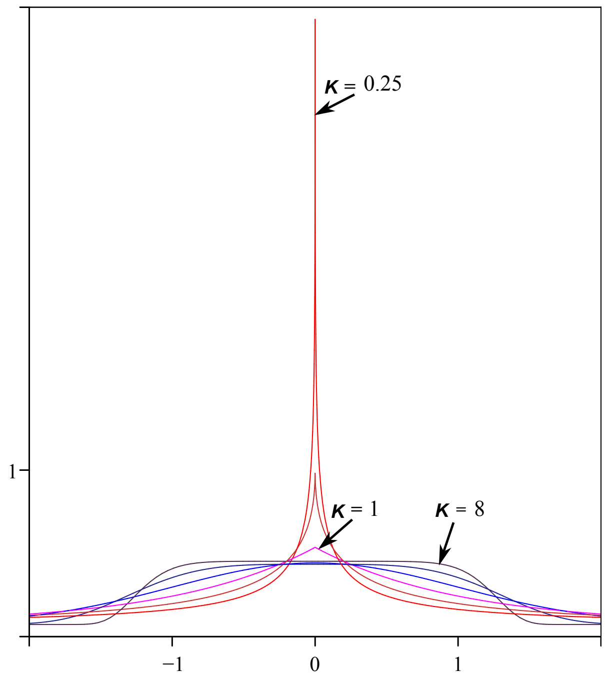


Figure 2.6: Density $d_{\kappa,0}$ of $G_{\kappa,0}$ for $\kappa = 0.25, 0.5, 1, 2, 4, 8$.

2.5 Comparison with the original model of Fougères and Soulier (2010)

2.5.1 The assumptions

In the paper of Fougères and Soulier (2010) the domain of the angular component T was $[0, 1]$, we use $[-1, 1]$ instead, because we want $t_0 = 0$ for simplicity without loss of generality.

Then the original assumptions are essentially:

Assumption 2_{FS}.

(I) The function $u : [-1, 1] \to [0, 1]$ is continuous, has a unique maximum 1 at a point $t_0 \in (-1, 1)$, and has an expansion $u(t_0 + t) = 1 - l(t)$, where l is decreasing from $[-\epsilon_0, 0]$ to $[0, \eta_-]$ and increasing from $[0, \epsilon_0]$ to $[0, \eta_+]$ for some

$\epsilon_0, \eta_-, \eta_+ > 0$, and regularly varying at 0 with index $\kappa > 0$. The generalized inverses $l_-^{\leftarrow} : [0, \eta_-] \to [-\epsilon_0, 0]$ and $l_+^{\leftarrow} : [0, \eta_+] \to [0, \epsilon_0]$ are absolutely continuous and their derivatives $(l_-^{\leftarrow})', (l_+^{\leftarrow})'$ are regularly varying at 0 with index $1/\kappa - 1$.

(II) The function v defined on $[-1, 1]$ is strictly increasing on $[t_0 - \epsilon_0, t_0 + \epsilon_0]$ with $v(t_0) = \rho$ and $w : t \mapsto v(t_0 + t) - \rho$ is regularly varying at 0 with index $\delta > 0$. Its generalized inverse $w^{\leftarrow}$ is absolutely continuous and its derivative $(w^{\leftarrow})'$ is regularly varying with index $1/\delta - 1$.

Analogously to Remark 2.3 one can show under Assumption 2_{FS} that
$l' = 1/((l_+^{\leftarrow})' \circ l) \in \mathrm{RV}_{\kappa-1}(0)$ and $w' \in \mathrm{RV}_{\delta-1}(0)$
exist on $(-\epsilon_0, \epsilon_0)$. Hence the function

$$b(t) := \frac{v(t)}{u(t)} - \rho = \frac{w(t - t_0) + l(t - t_0)\rho}{1 - l(t - t_0)}, \quad t \in (t_0 - \epsilon_0, t_0 + \epsilon_0) \tag{2.22}$$

can be differentiated for $t \in (t_0 - \epsilon_0, t_0 + \epsilon_0)$:

$$\begin{aligned} b'(t) &= \frac{w'(t - t_0) + \rho \cdot l'(t - t_0)}{1 - l(t - t_0)} + \frac{l'(t - t_0) \cdot (w(t - t_0) + \rho \cdot l(t - t_0))}{(1 - l(t - t_0))^2} \\ &= w'(t - t_0) + \rho \cdot l'(t - t_0) + f(t - t_0) \end{aligned}$$

for some $f \in \mathrm{RV}_{\kappa+\min(\kappa,\delta)-1}(0)$. For $\kappa > \delta$ or $\rho = 0$ the leading term of the derivative b' is w' which is positive, hence, there exists an $\epsilon \in (0, \epsilon_0]$ such that

- $b'(t) > 0$ for $t \in (t_0 - \epsilon, t_0 + \epsilon) \setminus \{t_0\}$,
- $|b(t_0 \pm \epsilon)| < 1$ (to ensure that $b(t) \in (-1, 1)$ for $t \in [t_0 - \epsilon, t_0 + \epsilon]$).

The function b can be extended from $(t_0 - \epsilon, t_0 + \epsilon)$ to $[-1, 1]$ as a C^1-diffeomorphism of $[-1, 1]$ and used for a reparametrization $\tilde{T} = b(T)$ of the angular component:

$$\begin{aligned} \tilde{u} = u \circ b^{-1} \quad &\text{and} \quad \tilde{l} = 1 - \tilde{u} \in \mathrm{RV}_{\tilde{\kappa}}(0) \text{ with } \tilde{\kappa} = \kappa/\delta, \\ \tilde{v} = v \circ b^{-1} \quad &\text{and} \quad \tilde{w} = \tilde{v} - \rho \in \mathrm{RV}_1(0), \\ \tilde{g} = (b^{-1})' \cdot g \circ b^{-1} \in \mathrm{RV}_{\tilde{\tau}}(0) \quad &\text{with} \quad \tilde{\tau} = (\tau + 1)/\delta - 1, \\ \kappa > \delta \Leftrightarrow \tilde{\kappa} > 1 \quad &\text{and} \quad \tau > -1 \Leftrightarrow \tilde{\tau} > -1. \end{aligned}$$

It holds $\tilde{v}\left(\tilde{t}\right) = \left(\tilde{t} + \rho\right)\tilde{u}\left(\tilde{t}\right)$ for $\tilde{t} \in (b(t_0 - \epsilon), b(t_0 + \epsilon))$.

Note that the extended $b(t)$ might not fulfill (2.22) for $\epsilon < |t - t_0| < \epsilon_0$.

Hence, we obtain:

Proposition 2.5. *Assumptions 2_{FS} and 2 are equivalent in the following sense:*
Any random vector possessing a polar representation fulfilling Assumption 2_{FS} with $\kappa > \delta > 0$ admits a polar representation satisfying Assumption 2 with $\kappa > 1$, which can be obtained by the transformation $t \mapsto b(t)$ described above.
On the other hand, a polar representation fulfilling Assumption 2 with $\kappa > 1$ is a special case of a polar representation satisfying Assumption 2_{FS} with $t_0 = 0$, $\delta = 1$ and $\epsilon_0 = \epsilon$.
In the case $\rho = 0$, Assumption 2_{FS} and 2 are equivalent for all $\kappa > 0$, $\delta > 0$.

The transformation $\tilde{t} = b(t)$ establishes the following functional dependence between u and v in some neighborhood of zero, and so permits the elimination of the parameter δ:

$$\tilde{t} + \rho = v\left(t\right)/u\left(t\right) = y/x = \tan\varphi,$$

where φ is the angle between the t-ray (i.e. the ray from $(0,0)$ through (x,y)) and the x-axis.

Fougères and Soulier (2010) require in the proof of their theorem that the quotient v/u is invertible, and therefore they have to choose an ϵ possibly smaller than ϵ_0 given in their Assumption 2_{FS}.

Example 2.3. (see Figure 2.7) The functions

$$u(t) = \frac{1}{1 + \tan^2\left(\frac{\pi}{2}t\right)}, \quad v(t) = 2 + \tan\left(\frac{\pi}{2}t\right)$$

fulfill Assumption 2_{FS} with $\kappa = 2$, $\delta = 1$, $\rho = 2$, $t_0 = 0$, $\epsilon_0 = 1$.
But $(v(t)/u(t))' < 0$ for $t \in (-1/2, -2/\pi \arctan(1/3)) \approx (-0.5, -0.2)$. So ϵ has to be chosen at most 0.2.

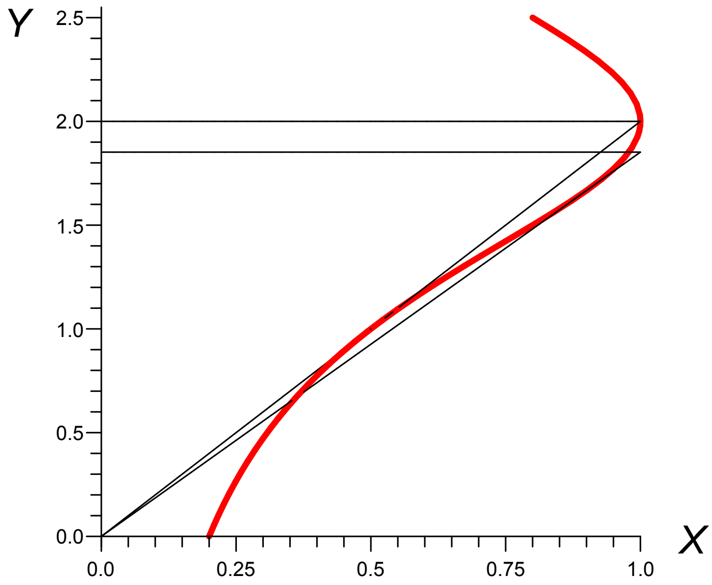


Figure 2.7: In Example 2.3 the function v increases along the curve $\{r = 1\}$, i.e. $(x, y) = (u(t), v(t))$, but the slope v/u decreases between $v = 1$ and $v = 5/3$ from 2 down to 1.85.

This demonstrates that even if v is monotonic on $[-1, 1]$ and u is monotonic on $[-1, t_0]$ and on $[t_0, 1]$, rays $\{(ru(t), rv(t)), r \geq 0\}$ might coincide for different values of t, e.g. for $t = 0$ and $t = -0.5$. The point $(x, y) = (1, 2)$ possesses polar representations with $(r, t) = (1, 0)$ and $(r, t) = (2, -0.5)$ as well; the map $(r, t) \mapsto (x, y)$ is not bijective. $\diamondsuit$

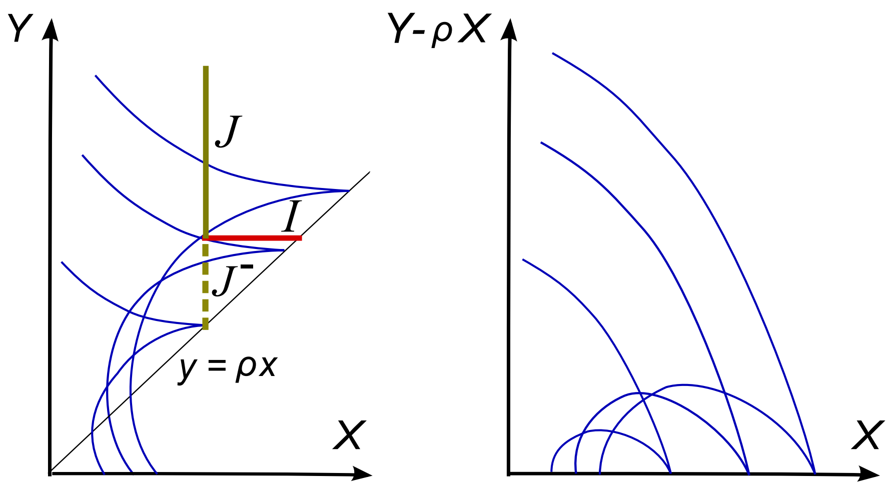
2.5.2 Polar representations with cusps under the original assumptions

In Section 2.4.2 we investigated polar random vectors under our Assumption 2 with $\kappa < 1$. Now we want to consider the analogue under Assumption 2_{FS} with $\kappa < \delta$.
For $\rho \neq 0$ the Assumptions 2 (with $\kappa < 1$) and 2_{FS} (with $\kappa < \delta$) are no longer equivalent. A function v according to Assumption 2_{FS} is monotonic and both branches of the r-level curves have tangents becoming horizontal at the cusps (see Figure 2.8(a)) in contrast to Assumption 2 with $\kappa < 1$ (see Figure 2.3(b)).
Though the map $(r, t) \mapsto (x, y)$ is not invertible in any neighborhood of a point $(x, \rho x)$

and it may not be wise to call R and T components of a representation of (X, Y), we still consider this case here:

Under Assumption 2_{FS} in the case $0 < \kappa < \delta$, we obtain for J the same result as for $\kappa > \delta$ (cf. proof of Theorem 2.1), but for I the domain of integration shrinks down to $\{0\}$ (in contrast to the case $\kappa > \delta$). On the other hand for $\kappa < \delta$ the function k_I now has a smaller index than k_J (cf. Section 2.3). So no comparison of I and J follows from this consideration. Calculations with MAPLE for $\kappa = 1/2$ and $\kappa = 1/3$ indicate that $J^- \sim I$ with $J^- = \int_{\epsilon_-}^{t_1} \overline{H}\,(x/u(t))\, g(t)\mathrm{d}t$ and one obtains a degenerate limit as in Theorem 2.2(ii).

Under Assumption 2_{FS} there exists no analogue to Proposition 2.4 or Theorem 2.3: After the affine transformation $(X, \tilde{Y}) = (X, Y - \rho X)$ the corresponding function $\tilde{v} = v - \rho u$ is no longer monotonic and, hence, part (II) of Assumption 2_{FS} is violated (see Figure 2.8(b)).



(a) (b)

Figure 2.8: r-level curves for $0 < \kappa < \delta$: (a) under the original Assumption 2_{FS}, the integral domains are shown; (b) after the transformation $(X, Y) \mapsto (X, Y - \rho X)$.

2.6 Details to Section 2.4.1

Here we present the assumptions and results in detail for the generalizations in Section 2.4.1. For simplicity, we consider the case I for $n = 2$ and infinitesimally symmetric l and g. Remember, we defined $\eta_i := (1 + \tau_i)/\kappa_i$ for $i = 1, 2$.

Assumption 2$_\mathrm{m}$ (cases I, II and III). Let $l : [-1, 1] \to [0, 1]$ be a function with

(I) exactly two zeros at $t_1, t_2 \in (-1, 1)$, $t_1 < t_2$,

(II) exactly one zero at $t = 0$,

(III) $l(t) = 0$ exactly on $[t_1, t_2]$, $t_1 < t_2$,

and a derivative l', which for some $\kappa_i > 0$, $i = 1, 2$ is

(I) $\mathrm{RV}^s_{\kappa_1 - 1}(t_1)$ and $\mathrm{RV}^s_{\kappa_2 - 1}(t_2)$,

(II) $\mathrm{RV}^-_{\kappa_1 - 1}(0)$ and $\mathrm{RV}^+_{\kappa_2 - 1}(0)$,

(III) $\mathrm{RV}^-_{\kappa_1 - 1}(t_1)$ and $\mathrm{RV}^+_{\kappa_2 - 1}(t_2)$.

It holds $u(t) = 1 - l(t)$ for $t \in [-1, 1]$.
For some $\rho_i \in \mathbb{R}$, $\rho_1 < \rho_2$ resp. $\rho \in \mathbb{R}$ and $\epsilon > 0$ we have

(I) $v(t) = (\rho_i + t - t_i)(1 - l(t))$ for $t \in (t_i - \epsilon, t_i + \epsilon)$,

(II) $v(t) = (\rho + t)(1 - l(t))$ for $t \in (-\epsilon, \epsilon)$,

(III) $v(t) = (\rho_i + t - t_i)(1 - l(t))$ for $t \in (t_1 - \epsilon, t_1)$ and $t \in (t_2, t_2 + \epsilon)$,
$v(t) = \rho_1 + (t - t_1)/(t_2 - t_1) \cdot (\rho_2 - \rho_1)$ for $t \in [t_1, t_2]$.

Assumption 3$_\mathrm{m}$. The density $g : [-1, 1] \to [0, \infty)$ for $\tau_i > -1$, $i = 1, 2$ is

(I) $\mathrm{RV}^s_{\tau_1}(t_1)$ and $\mathrm{RV}^s_{\tau_2}(t_2)$,

(II) $\mathrm{RV}^-_{\tau_1}(0)$ and $\mathrm{RV}^+_{\tau_2}(0)$,

(III) $\int_{t_1}^{t_2} g(t)\, \mathrm{d}t > 0$.

Proposition 2.6 (case I). *Let $(X,Y) = R \cdot (u(T), v(T))$ satisfy Assumptions 1, 2_m(I) and 3_m(I) with $\kappa_i > 1$.*

(i) *Then the statements of Theorem 2.1 hold for*

- $\eta_1 < \eta_2$ *with* ρ_1, κ_1 *and* τ_1,
- $\eta_1 > \eta_2$ *with* ρ_2, κ_2 *and* τ_2.

(ii) *Let $\eta_1 = \eta_2 =: \eta$, then it holds:*

(i) *X is of type $\Gamma(\psi)$ and*

$$P\{X > x\} \sim (k_1 + k_2)\,(\psi(x)/x) \cdot P\{R > x\},$$

(ii)

$$\lim_{x\to\infty} P(Y \leq \lambda x \mid X > x) = \begin{cases} 0, & \text{if } \lambda < \rho_1, \\ p_-, & \text{if } \rho_1 < \lambda < \rho_2, \\ 1, & \text{if } \rho_2 < \lambda, \end{cases}$$

(iii) *for $\zeta \in \mathbb{R}$:*

$$\lim_{x\to\infty} P\left(Y \leq \rho_1 x + x \cdot h_1\left(\psi(x)/x\right) \cdot \zeta \mid X > x\right) = p_- \cdot G_{\kappa_1,\tau_1}(\zeta)$$

and

$$\lim_{x\to\infty} P\left(Y \leq \rho_2 x + x \cdot h_2\left(\psi(x)/x\right) \cdot \zeta \mid X > x\right) = p_- + p_+ \cdot G_{\kappa_2,\tau_2}(\zeta),$$

where k_i, h_i, G_{κ_i,τ_i}, $i = 1,2$ are defined analogously to Theorem 2.1 and $p_\pm$ as in (2.15).

Proposition 2.7. *Under assumptions analogously to Assumptions 2_C, 3_C we get:*

$$k(q) = 2 \cdot \Gamma(\eta) \cdot q^\eta \cdot (K_1 + K_2), \quad h_i(q) = \left(\frac{q}{\kappa_i L_{0,i}}\right)^{1/\kappa_i}$$

with $K_i := G_{0,i}/\kappa_i \cdot (L_{0,i})^{-\eta}$.
The weights are $p_i := K_i/(K_1 + K_2)$, $i = 1,2$.

Proposition 2.8 (case II). *Let $(X, Y) = R\cdot(u(T), v(T))$ satisfy Assumptions 1, 2_m(II), 3_m(II) with $\kappa_i > 1$. Then the following statements hold:*

X is of type $\Gamma(\psi)$ and

$$P\{X > x\} \sim k\left(\psi(x)/x\right) \cdot P\{R > x\}$$

holds with

$$k(q) = \begin{cases} k_-, & \text{if } \eta_1 < \eta_2, \\ k_+, & \text{if } \eta_1 > \eta_2, \\ k_- + k_+, & \text{if } \eta_1 = \eta_2, \end{cases}$$

where $k_\pm$ are defined in (2.13).

For all $\zeta \in \mathbb{R}$ it holds:

$$\lim_{x\to\infty} P(Y \leq \rho x + x \cdot h\left(\psi(x)/x\right) \cdot \zeta \mid X > x) = H(\zeta)$$

- *for $\eta_1 < \eta_2$ with $h = h_1$ and*

$$H(\zeta) = \begin{cases} 2 \cdot G_{\kappa_1,\tau_1}, & \text{for } \zeta < 0, \\ 1, & \text{for } \zeta \geq 0; \end{cases}$$

- *for $\eta_1 > \eta_2$ with $h = h_2$ and*

$$H(\zeta) = \begin{cases} 0, & \text{for } \zeta < 0, \\ 2 \cdot G_{\kappa_2,\tau_2} - 1, & \text{for } \zeta \geq 0; \end{cases}$$

- *for $\eta_1 = \eta_2 =: \eta$ with*

$$h = \begin{cases} h_1, & \text{for } \zeta < 0, \\ h_2, & \text{for } \zeta \geq 0, \end{cases}$$

$$H(\zeta) = \begin{cases} \frac{p_-}{\kappa_1^{\eta-1}\cdot\Gamma(\eta)} \int_{-\infty}^{\zeta} \exp\left(-\left|s\right|^{\kappa_1}/\kappa_1\right) \left|s\right|^{\tau_1}\, \mathrm{d}s, & \text{for } \zeta < 0, \\ p_- + \frac{p_+}{\kappa_2^{\eta-1}\cdot\Gamma(\eta)} \int_0^{\zeta} \exp\left(-s^{\kappa_2}/\kappa_2\right) s^{\tau_2}\, \mathrm{d}s, & \text{for } \zeta \geq 0, \end{cases}$$

where G_{κ_i,τ_i}, h_i, $i = 1, 2$ and $p_\pm$ are defined as in Theorem 2.1 resp. formula (2.15).

Proposition 2.9 (case III). *Let* $(X, Y) = R \cdot (u(T), v(T))$ *satisfy Assumptions 1, 2*$_m$*(III), 3*$_m$*(III) with* $\kappa_i > 1$. *Then the following statements hold:*

X *is of type* $\Gamma(\psi)$ *and*

$$P\{X > x\} \sim k \cdot P\{R > x\}$$

holds with $k = \int\limits_{t_1}^{t_2} g(t)\,\mathrm{d}t$.

For all $\xi > 0$ *and* $\zeta \in \mathbb{R}$ *it holds:*

$$\lim_{x\to\infty} P(Y \le \lambda x \mid X > x) = \begin{cases} 0, & \text{if } \lambda < \rho_1, \\ \frac{1}{k}\int_{t_1}^{t(\lambda)} g(t)\,\mathrm{d}t, & \text{if } \rho_1 < \lambda < \rho_2, \\ 1, & \text{if } \rho_2 < \lambda, \end{cases}$$

with $t(\lambda) = t_1 + (\lambda - \rho_1)/(\rho_2 - \rho_1) \cdot (t_2 - t_1)$.

2.7 Proofs

2.7.1 Proofs of Theorems 2.1, 2.3, and of Propositions 2.3 – 2.9

We will use the following lemma from Fougères and Soulier (2010, lemma 5.2).

Lemma 2.1. *Let* H *be a distribution function on* $[0, \infty)$ *of type* $\Gamma(\psi)$. *Let* $k \in \mathrm{RV}_\alpha(0)$ *with* $\alpha > -1$ *be bounded on compact subsets of* $(0, \infty]$. *Then*

$$\lim_{x\to\infty} \int\limits_a^b \frac{\overline{H}(x + z\psi(x))}{\overline{H}(x)} \frac{k(z\psi(x)/x)}{k(\psi(x)/x)} \mathrm{d}z = \int\limits_a^b \mathrm{e}^{-z} z^\alpha \mathrm{d}z$$

holds locally uniformly for $a \in [0, \infty),\ b \in (0, \infty]$.

Theorem 2.1, Proposition 2.3 and 2.4 can be proved as following:

For a simpler notation we will drop in $\psi(x)$ the argument x and write ψ in the following proof.

To determine the asymptotic behavior of I and J (see (2.11)), we perform following substitutions (for the justification of steps (II) and (III) see Remark 2.10):

(I) to apply Lemma 2.1, the argument of $\overline{H}$ has to be $x + z\psi$; so we substitute:

- in J: $t \mapsto z(t) := (l(t)/(1 - l(t))) \cdot (x/\psi)$
- in I: $t \mapsto z(t) := ((y/x)/v(t) - 1) \cdot (x/\psi)$

(II) as the limit distribution is obtained by keeping ζ fixed, we express $y = (\rho + t_1)x$ as $y = \rho x + \delta(x)\zeta = \rho x + x \cdot h(\psi/x) \cdot \zeta$,

(III) to obtain the form of a survival function, we finally substitute $z \mapsto s := \pm(\kappa z)^{1/\kappa}$ with $\operatorname{sign}(s) = \operatorname{sign}(\zeta)$.

Ad integral J: First we consider the case

1.a $\rho > 0, 0 \leq t_1 < \epsilon$

where $t_1 = y/x - \rho = h(\psi/x) \cdot \zeta$ with $h(q) = \kappa^{-1/\kappa} l_+^{-1}(q)$ (cf. (2.14)). Substitution $t \mapsto z(t)$ yields:

$$J = \frac{\psi}{x} \int\limits_{z(t_1)}^{z(\epsilon)} \overline{H}\,(x+z\psi)\, f_J \left(\frac{z\psi/x}{1 + z\psi/x} \right) \frac{\mathrm{d}z}{(1+z\psi/x)^2}, \tag{2.23}$$

where

$$f_J := \left(g \circ l_+^{-1}\right) \cdot \left(l_+^{-1}\right)' \in \mathrm{RV}_{(1+\tau)/\kappa - 1}(0). \tag{2.24}$$

It holds $z(\epsilon) \to \infty$ for $x \to \infty$.

As $l \in \mathrm{RV}_\kappa(0)$, and $l_+^{-1}(\psi/x) \to 0$, we obtain with (2.5):

$$\begin{aligned} z(t_1) &\sim l(t_1) \cdot x/\psi \\ &= \frac{\psi}{x} \frac{l\left(\kappa^{-1/\kappa} \zeta l_+^{-1}(\psi/x)\right)}{l\left(l_+^{-1}(\psi/x)\right)} \cdot \frac{x}{\psi} \sim |\zeta|^\kappa/\kappa. \end{aligned} \tag{2.25}$$

Lemma 2.1 yields:

$$\begin{aligned} J &\sim \overline{H}(x) \cdot \kappa^{(1+\tau)/\kappa-1} \cdot k_J(\psi/x) \int\limits_{\zeta^\kappa/\kappa}^{\infty} \frac{\overline{H}(x+z\psi)}{\overline{H}(x)} \cdot \frac{\tilde{k}(z\psi/x)}{\tilde{k}(\psi/x)} \mathrm{d}z \\ &\sim \overline{H}(x) \cdot \kappa^{(1+\tau)/\kappa-1} \cdot k_J(\psi/x) \int\limits_{\zeta^\kappa/\kappa}^{\infty} \mathrm{e}^{-z} z^{(1+\tau)/\kappa\,-1} \mathrm{d}z, \end{aligned} \tag{2.26}$$

where

$$\begin{aligned} \tilde{k}(q) &= f_J(q/(1+q)) \cdot (1+q)^{-2} \in \mathrm{RV}_{(1+\tau)/\kappa-1}(0), \\ k_J(q) &= \kappa^{1-(1+\tau)/\kappa} \cdot q \cdot \tilde{k}(q) \in \mathrm{RV}_{(1+\tau)/\kappa}(0). \end{aligned} \tag{2.27}$$

The final substitution $z \mapsto s$ results in:

$$J \sim \overline{H}(x) \cdot k_J(\psi/x) \cdot \int\limits_{\zeta}^{\infty} \exp\left(-|s|^\kappa/\kappa\right) |s|^\tau \,\mathrm{d}s \quad \text{for } x \to \infty. \tag{2.28}$$

The other cases can be traced back to 1.a:

1.b $\boldsymbol{\epsilon_- < t_1 < 0}$: As $\frac{\mathrm{d}z}{\mathrm{d}t}$ changes sign, we partition $J = \int_{t_1}^{\epsilon} \ldots \mathrm{d}t = \int_0^{\epsilon} \ldots \mathrm{d}t - \int_0^{t_1} \ldots \mathrm{d}t$ previous to substitution $t \mapsto z$, where in the second integral l_-^{-1}, k_- have to be chosen.

1.c $\boldsymbol{\epsilon \le t_1}$ or $\boldsymbol{t_1 \le \epsilon_-}$: Choose a new pair (x, y) with the same value of ζ, where x is large enough such that $\epsilon_- < h(\psi/x) < \epsilon$, then case 1.a resp. 1.b can be applied.

2. $\boldsymbol{\rho < 0}$: with $(X, -Y)$ this is case 1.

3. $\boldsymbol{\rho = 0}$: see case 1 for $t_1 \ge 0$ resp. case 2 for $t_1 < 0$.

Note that in Example 2.3 the value $t_1 = 0$ belongs to case 1.a and $t_1 = -0.5$ to case 1.c, in spite of the fact that both values describe the same ray in the (x, y)-plane.

Remark 2.10. If one chooses h as a function which goes to zero faster resp. slower than any function in $\mathrm{RV}^+_{1/\kappa}(0)$, then $z(t_1) \to 0$ resp. $z(t_1) \to \infty$ for every $\zeta \in \mathbb{R}$. In both cases, the resulting limit distribution $G(\xi, \zeta)$ would be degenerate. The special form of $h = \pm\kappa^{-1/\kappa} \cdot l_{\pm}^{-1}$ is chosen to obtain after steps (2.25) and (2.28) the limit $G_{\kappa,\tau}$ in a form such that $G_{2,0} = \Phi$. $\diamondsuit$

Ad integral *I*: It will be shown that

$$I = o(J) \text{ for } \kappa > 1 \text{ or } \rho = 0. \tag{2.29}$$

The first substitution $t \mapsto z(t)$ yields

$$I = \frac{\psi}{x}\frac{y}{x} \int\limits_{z(t_1)}^{z(\epsilon_-)} \overline{H}\,(x + z\psi)\, f_I \left(\frac{y/x}{1 + z\psi/x}\right) \frac{\mathrm{d}z}{(1 + z\psi/x)^2}, \tag{2.30}$$

where

$$f_I := \left(g \circ v^{-1}\right) \cdot \left(v^{-1}\right)' \in \mathrm{RV}_\tau(\rho). \tag{2.31}$$

It holds $z(\epsilon_-) \to \infty$ for $x \to \infty$.
Fougères and Soulier (2010) claim the following asymptotic form (for $\delta = 1$)

$$I \sim \overline{H}(x) \cdot \rho \cdot \zeta^\tau \cdot k_2\,(\psi/x) \int\limits_{|\zeta|^\kappa/\kappa}^{\infty} \mathrm{e}^{-z}\mathrm{d}z, \tag{2.32}$$

with $k_2 \in \mathrm{RV}_{1+\tau/\kappa}(0)$. Then $k_2/k_J \in \mathrm{RV}_{1-1/\kappa}(0)$ and, hence, $I = o(J)$ for $\kappa > 1$. They argue that $1 + z\psi/x$ in (2.30) can be replaced by 1 as $1 + z\psi/x \sim 1$. But this convergence is only locally uniform in z and does not hold on the upper bound $z(\epsilon_-)$ of the integral I:

$$1 + z(\epsilon_-)\frac{\psi}{x} = \frac{y/x}{v(\epsilon_-)} = \frac{\rho + h\,(\psi/x) \cdot \zeta}{v(\epsilon_-)} \sim \frac{\rho}{v(\epsilon_-)} > \frac{\rho}{v(0)} = 1$$

for $x \to \infty$ and constant $\zeta \in \mathbb{R}$. Actually we will show that the regularly varying function k_I determining the asymptotic behavior of I possesses different indices depending on the different cases of the parameters ρ, ζ and τ.
With

$$E := 1 + z\psi/x, \quad F := h\,(\psi/x) \cdot \zeta - \rho \cdot z \cdot \psi/x \tag{2.33}$$

we write

$$I = \overline{H}(x) \cdot \frac{\psi}{x} \cdot (\rho + h(\psi/x) \cdot \zeta) \int\limits_{z(t_1)}^{z(\epsilon_-)} \frac{\overline{H}(Ex)}{\overline{H}(x)} \cdot f_I(F/E + \rho) \cdot \frac{\mathrm{d}z}{E^2}.$$

In the generic case, Lemma 2.1 cannot be applied because of the form of $|F|^\tau$ (occurring in $f_I(F/E+\rho)$ with variation index τ); only in the special cases $\rho = 0$ or $\zeta = 0$ or $\tau = 0$ the asymptotic form of integral I can be calculated directly with this lemma.

For $\rho = 0, \zeta = 0$ in (2.11) we have $\epsilon_- = t_1$ and $I = 0$.
In the following two special cases we get asymptotic forms of I, which differ indeed in the variation indices from the result (2.32) of Fougères and Soulier:

(A) for $\rho = 0, \zeta \neq 0$ we obtain

$$I \sim \overline{H}(x) \cdot k_I(\psi/x) \cdot f_\zeta(x) \int\limits_{|\zeta|^\kappa/\kappa}^{\infty} \mathrm{e}^{-z}\mathrm{d}z,$$

where $k_I(q) = (1+q)^{-2-\tau} q^{1+(1+\tau)/\kappa} \in \mathrm{RV}_{1+(1+\tau)/\kappa}(0)$,

(B) for $\zeta = 0, \rho \neq 0$ or $\tau = 0, \rho \neq 0$ we have

$$I \sim \overline{H}(x) \cdot k_I(\psi/x) \cdot f_\zeta(x) \int\limits_{|\zeta|^\kappa/\kappa}^{\infty} \mathrm{e}^{-z} z^\tau \mathrm{d}z,$$

where $k_I(q) = (1+q)^{-2-\tau} q^{1+\tau} \in \mathrm{RV}_{1+\tau}(0)$.

Thereby $f_\zeta(x)$ denotes a mean value of $f_I(F/E+\rho)$.
In the generic case $\rho \neq 0, \tau \neq 0, \zeta \neq 0$ one can find bounds $\hat{I}$ with $I \leq \hat{I}$ such that the asymptotic behavior of $\hat{I}$ is determined by a function $k_{\hat{I}} \in \mathrm{RV}_{1+\tau/\kappa}(0)$.
In all cases, the function k_J has a smaller index than k_I resp. $k_{\hat{I}}$ for $\kappa > 1$ or $\rho = 0$.
This completes the proof of (2.29) and altogether this settles statement (ii) of Theorem 2.1.
A proof of Theorem 2.1(iii) is not presented in Fougères and Soulier (2010).

To deduce statement (iii) one has to subtract the following expression from $G_{\kappa,\tau}(\zeta)$:

$$\begin{aligned}&P(X > x + \psi(x)\xi,\ Y \le \rho x + x \cdot h\left(\psi(x)/x\right)\zeta \mid X > x)\\ = \ &\frac{P\{X > \tilde{x}\}}{P\{X > x\}} \cdot P(X > \tilde{x}, Y \le \rho\tilde{x} + \tilde{x} \cdot h\left(\psi(\tilde{x})/\tilde{x}\right)\tilde{\zeta} \mid X > \tilde{x}). \qquad (2.34)\end{aligned}$$

Thereby $\tilde{x} := x + \psi(x)\xi$ and $\tilde{\zeta}$ is implicitly defined by

$$\rho x + x \cdot h\left(\psi(x)/x\right) \cdot \zeta = \rho\tilde{x} + \tilde{x} \cdot h\left(\psi(\tilde{x})/\tilde{x}\right) \cdot \tilde{\zeta}, \qquad (2.35)$$

where $\tilde{\zeta}$ depends on x for fixed ξ, ζ. Since X is of type $\Gamma(\psi)$ (see Theorem 2.1(i)), the first factor of (2.34) converges to $\exp(-\xi)$. The second factor is asymptotically equivalent to $G_{\kappa,\tau}(\zeta)$ for $x \to \infty$, if $\tilde{\zeta} \to \zeta$ for $x \to \infty$. Rewriting (2.35) we obtain with $\lambda_x := \psi(x + \psi(x)\xi)/\psi(x) \cdot (1 + \psi(x)/x \cdot \xi)^{-1}$:

$$\zeta - \underbrace{\left(1 + \frac{\psi(x)}{x}\xi\right) \cdot \frac{h(\lambda_x \cdot \psi(x)/x)}{h(\psi(x)/x)}}_{\to 1} \cdot \tilde{\zeta} = \underbrace{\rho\xi\frac{\psi(x)/x}{h\left(\psi(x)/x\right)}}_{\to 0}. \qquad (2.36)$$

The convergence on the left hand of (2.36) holds, since the convergence of (2.5) is uniformly in λ (cf. Resnick 1987, Proposition 0.5), ψ is Beurling slowly varying (cf. Resnick 1987, Lemma 1.3), i.e. $\psi(x + \psi(x)\xi)/\psi(x) \to 1$ for $x \to \infty$, and $\lambda_x \to 1$ for $x \to \infty$. The right hand of (2.36) converges to zero for $\kappa > 1$ (i.e. h has a variation index $1/\kappa < 1$) or $\rho = 0$. Consequently one gets $\tilde{\zeta} \to \zeta$ for $x \to \infty$.
Therefore Theorem 2.1, Proposition 2.3 and, hence, 2.4 are proved.

The proof of Propositions 2.6, 2.8 and 2.9 can be dealt analogously to Theorem 2.1.

To deduce Theorem 2.3 from Proposition 2.4 we need a further step:
If $x \le X \le x + \psi(x)\xi$, then there exists a $\vartheta \in [0,1]$ with $X = x + \vartheta\psi(x)\xi$. With the normalizing function $\delta(x) = x \cdot h\left(\psi(x)/x\right)$ and $\delta(X) = \delta(x + \vartheta\psi(x)\xi)$ we obtain

$$\frac{\delta(X)}{\delta(x)} = \left(1 + \vartheta\frac{\psi(x)}{x}\xi\right) \cdot \frac{h(\lambda_x \cdot \psi(x)/x)}{h(\psi(x)/x)} \to 1$$

uniformly for fixed $\xi > 0$, because $\psi(x)/x \to 0$, the convergence of (2.5) is uniformly in λ, ψ is Beurling slowly varying and $\lambda_x := \psi(x + \vartheta\psi(x)\xi)/\psi(x) \cdot (1 + \vartheta\psi(x)/x\xi)^{-1} \to 1$ for $x \to \infty$.

$\square$

2.7.2 Proof of Theorem 2.2

Under Assumptions 2 in the case $0 < \kappa < 1$ the asymptotic form of J is still given by (2.28).

Since we will distinguish I by the value of ζ, we write I_ζ. As in the proof of Theorem 2.1 we write ψ instead of $\psi(x)$ in the following for a simpler notation.

$\zeta \leq 0$: The integral I_0 can directly be evaluated asymptotically. It holds:

$$I_0 = \overline{H}(x)\frac{\psi}{x}\rho \int\limits_0^{z(\epsilon_-)} \frac{\overline{H}\,(x + z\psi)}{\overline{H}(x)} \tilde{f}_I^- \left(\frac{\rho}{1 + z\psi/x}\right) \frac{\mathrm{d}z}{(1 + z\psi/x)^2}$$

with

$$\tilde{f}_I^\pm := \left(g \circ v_\pm^{-1}\right) \cdot \left(v_\pm^{-1}\right)' \in \mathrm{RV}_{(1+\tau)/\kappa-1}(\rho) \tag{2.37}$$

as an analogue to f_I in (2.31), but now with the same variation index as f_J in (2.24). Analogously to J it follows with Lemma 2.1 that

$$I_0 \sim \overline{H}(x) \cdot \kappa^{(1+\tau)/\kappa-1} \tilde{k}_I\,(\psi/x) \cdot \int\limits_0^\infty \mathrm{e}^{-z} z^{(\tau+1)/\kappa-1} \mathrm{d}z \tag{2.38}$$

for $\tilde{k}_I^-(q) = \rho\kappa^{1-(1+\tau)/\kappa} \tilde{f}_I^-(\rho/(1+q)) \cdot q/(1+q)^2$.

Since from (2.17) follows that $l_-^{-1}(s) \sim v_-^{-1}(\rho - \rho s)$, we have $\tilde{k}_I^- \sim k_J^-$ where

$$k_J^-(q) := -\kappa^{1-(1+\tau)/\kappa} \left(g \circ l_-^{-1}(q) \cdot \left(l_-^{-1}\right)'\right) (q/(1+q)) \cdot q/(1+q)^2 \tag{2.39}$$

is analogously defined to k_J in (2.27). Hence, we have $I_0 \sim J_0^-$.

This proves (2.20) for $\zeta \leq 0$.

$\zeta > 0$: In the expression (cf. (2.33))

$$F = h(\psi/x)\cdot\zeta - \rho\cdot z\cdot\psi/x = \rho\cdot z\cdot\psi/x\cdot(z_0/z - 1), \tag{2.40}$$

with its zero $z_0 := h(\psi/x)\cdot\zeta/\rho\cdot(\psi/x)^{-1}$, now the term which is independent of ζ dominates. It holds $\lim_{x\to\infty} z_0 = 0$, since h has a variation index $1/\kappa < 1$. In the substitution $t \mapsto z(t) := ((y/x)/v(t) - 1)\cdot(x/\psi)$ the derivative $\frac{\mathrm{d}z}{\mathrm{d}t}$ changes its sign at $t = 0$, we partition

$$\begin{aligned} I_\zeta &= I^+ + I^- \\ &= \int_0^{t_1} \overline{H}\left(\frac{y}{v(t)}\right) g(t)\mathrm{d}t + \int_{\epsilon_-}^{0} \overline{H}\left(\frac{y}{v(t)}\right) g(t)\mathrm{d}t \\ &= \overline{H}(x)\frac{\psi}{x}\left(\rho + h(\psi/x)\zeta\right)\cdot\left(\int_0^{z(t_1)} M^+\mathrm{d}z + \int_0^{z(\epsilon_-)} M^-\mathrm{d}z\right) \end{aligned}$$

with $M^\pm := \overline{H}(Ex)/\overline{H}(x)\cdot\tilde{f}_I^\pm(F/E + \rho)\cdot E^{-2}\mathrm{d}z$, where $\tilde{f}_I^\pm$, E and F are defined in (2.37) and (2.33).

By choosing x large enough we can assume that $z_0 < \sqrt{z_0} < z(t_1)$. We partition $I^+ = I_1^+ + I_2^+ = \ldots\int_0^{\sqrt{z_0}} M^+\mathrm{d}z + \ldots\int_{\sqrt{z_0}}^{z(t_1)} M^+\mathrm{d}z$ to isolate the zero z_0 of F. For $z \geq \sqrt{z_0}$ it holds

$$1 > 1 - z_0/z \geq 1 - z_0/\sqrt{z_0} = 1 - \sqrt{z_0} \sim 1$$

for $x\to\infty$, which implies that $1 - z_0/z \to 1$ uniformly in z in the domain of I_2^+. Hence the integrand of I_2^+ converges uniformly to the integrand of I_0 and so analogously to (2.38)

$$I_2^+ \sim \overline{H}(x)\kappa^{(1+\tau)/\kappa - 1}\cdot k_J(\psi/x)\cdot\int_0^{z(t_1)} \mathrm{e}^{-z} z^{(1+\tau)/\kappa - 1}\mathrm{d}z,$$

where $z(t_1) \sim |\zeta|^\kappa/\kappa$. As the domain $(0, \sqrt{z_0})$ of I_1^+ shrinks down to $\{0\}$, it can be shown that $I_1^+ = o(I_2^+)$ even if F occurs with a negative power (i.e. if $(1+\tau)/\kappa < 1$).

The integral I^- can be calculated analogously to I_2^+ with the boundary $z(\epsilon_-) \to \infty$. Hence,

$$\begin{aligned} I_\zeta &\sim \overline{H}(x)\kappa^{(1+\tau)/\kappa-1} \cdot \Bigg[k_J^-\left(\psi/x\right) \cdot \int\limits_0^\infty \mathrm{e}^{-z} z^{(1+\tau)/\kappa-1} \mathrm{d}z \\ &\qquad\qquad + k_J\left(\psi/x\right) \cdot \int\limits_0^{\zeta^\kappa/\kappa} \mathrm{e}^{-z} z^{(1+\tau)/\kappa-1} \mathrm{d}z \Bigg] \\ &\sim J_\zeta^- \end{aligned}$$

with k_J, k_J^- from (2.27) resp. (2.39) and $J_\zeta^- = \int_{\epsilon_-}^{t_1} \overline{H}\left(x/u(t)\right) g(t)\,\mathrm{d}t$.
For the choice of $h(q) = \kappa^{-1/\kappa} l_+^{-1}(q)$ it holds:

$$\begin{aligned} P\{X > x, Y > \rho x + x \cdot h(\psi/x) \cdot \zeta\} &\sim J_\zeta + I_\zeta \sim J_\zeta + J_\zeta^- \\ &\sim P\{X > x\} \end{aligned}$$

and we get (2.20) also for $\zeta > 0$.

By that, Theorem 2.2 is shown.

□

Chapter 3

A new approach for conditional limit theorems in the CEV model [1]

3.1 Introduction

In this chapter we present a novel approach for analyzing the conditional limit behavior of a bivariate random vector (X, Y) given that $X > x$ becomes extreme. Having deduced the results for a univariate random variable $X = R \cdot u(T)$, we give a new insight into established conditional limit theorems for random vectors with a polar representation $(X, Y) = R \cdot (u(T), v(T))$ and we generalize them (weakening of the assumptions as well as extending the classes of admissible Y). This study contributes to the current research in the field of the *conditional extreme value (CEV) models*, which were introduced in the pioneering paper of Heffernan and Tawn (2004) and further developed by Heffernan and Resnick (2007) and Das and Resnick (2011a, 2011b).

Polar representations $(X, Y) = R \cdot (u(T), v(T))$ form natural extensions of elliptically distributed bivariate random vectors where the stochastically independent components R and T can be interpreted as generalizations of radius and angle, respectively.

[1]The contents of this chapter are based on Barbe, P. and Seifert, M.I. (2013): A conditional limit theorem for a bivariate representation of a univariate random variable and conditional extreme values. arXiv:1311.0540.

Assuming the distribution of R to be in the Gumbel max-domain of attraction, conditional limit results for such random vectors (X, Y) given $X > x$, $x \to \infty$ are obtained among others by Berman (1983), Fougères and Soulier (2010), Hashorva (2012).

The analysis in this chapter is presented in three steps:

1. We start with a *univariate* random variable $X = R \cdot u(T)$ for independent components R and T, whereby the function u takes its unique global maximum 1 at $t = t_0$ and $1 - u$ is regularly varying at $t = t_0$, and deduce a conditional limit theorem for (R, T) given $X > x$, $T > t_0$ and $x \to \infty$.

2. Then this theorem, corresponding to the originally one-sided concept of regular variation, proved for $T > t_0$ will be extended to all T.

3. Finally, these results for the representation of the single variable X are exploited in order to obtain in a simple and elegant way several conditional limit statements for a *bivariate* random vector $(X, Y) = R \cdot (u(T), v(T))$ given that X becomes extreme.

The novelty of this chapter is to show that a conditional limit theorem for (X, Y) given $X > x$, $x \to \infty$ can be gained by considering solely the representation of the single variable X in terms of the pair (R, T). Applying methods from the probability measure theory permits us to set the underlying assumptions quite weak and convenient, such that they are natural for the problem and not forced by a certain method of proof. Previous results of Fougères and Soulier (2010) as well as from Chapter 2 are recovered under weaker assumptions omitting conditions on differentiability and monotonicity for u and v. Furthermore, we deduce generalizations for different cases concerning the relation between u and v covering a considerable larger set of random vectors with representation $R \cdot (u(T), v(T))$.

Our three step approach also underscores the asymmetry between the variable Y and the conditioning variable X in the CEV models: we point out that the essential assumptions and results are primarily made on X from which conditional limit statements for (X, Y) can be directly deduced. Hence, our results provide a new insight for studying the extremal behavior of random vectors with polar representation with respect to the CEV models.

This chapter is organized as follows. In Section 3.2 we provide the main result of this chapter in Theorem 3.1: we get a conditional limit result for the components (R, T) of $X = R \cdot u(T)$ given $T > t_0$, $X > x$ as $x \to \infty$ under quite weak assumptions on T and u with R in the Gumbel max-domain of attraction (Assumptions A and B). In Section 3.3 we show how the restriction $T > t_0$ can be removed and we deduce two-sided extensions (Theorems 3.2 and 3.3) allowing asymmetric behavior on both sides of t_0 (Assumptions A and C). In Section 3.4 we derive conditional limit theorems for $(X, Y) = R \cdot (u(T), v(T))$ given that X is large by applying continuous mapping arguments to the limit results on (R, T) from Theorem 3.1. Our approach makes it possible to investigate systematically different cases for the behavior of the functions u and v as it is illustrated further in Section 3.4. The proof of Theorem 3.1 with many technical details is provided in Section 3.5, whereas all other proofs are placed immediately after the corresponding statements.

3.2 Limit results for the representation of a univariate random variable

Let the random variable X be represented as $X = R \cdot u(T)$. We study the conditional limit behavior of properly normalized (R, T) given $X > x$ as x tends to infinity.

The generalized radius R possesses a domain of form $[a, \infty)$ resp. (a, ∞), a typical choice is $a = 0$. Let H denote the distribution function of R and $\overline{H} = 1 - H$ its survival function. The domain of the generalized angle T is some closed interval on $\mathbb{R}$. Let g denote the density of T, and $\mathbb{1}$ the indicator function.

We make the following four assumptions on the representation $X = R \cdot u(T)$.

Assumption A.1. The random variables R and T are stochastically independent.

Assumption A.2. The survival function $\overline{H}$ of R is of type $\Gamma(\psi)$.

This means that there exists a positive function ψ (called *auxiliary function*) such that for any $z \in \mathbb{R}$ it holds

$$\lim_{x \to \infty} \frac{\overline{H}\left(x + z\psi(x)\right)}{\overline{H}(x)} = e^{-z}\,.$$

This property is equivalent to H belonging to the Gumbel max-domain of attraction (de Haan 1970, Th. 2.5.1) with infinite right endpoint: $\sup\{x : H(x) < 1\} = \infty$. The auxiliary function ψ is unique up to asymptotic equivalence, and it holds $\psi(x) = o(x)$ for $x \to \infty$.

Assumption B.1. There exists some t_0 such that $u(t_0) = 1$ and for any $\epsilon > 0$ it holds $\sup_{t-t_0>\epsilon} u(t) < 1$. Moreover, the function

$$\tilde{u}(s) := u(t_0) - u(t_0 + s)$$

is regularly varying at $0+$ with variation index $\kappa > 0$.

This means that for any $\lambda > 0$ it holds

$$\lim_{s \to 0+} \frac{\tilde{u}(\lambda s)}{\tilde{u}(s)} = \lambda^{\kappa}\,.$$

For the notation convenience we put in this chapter a tilde for functions which are regularly varying at $0+$ (resp. at $0-$ in the following section), hence $\tilde{u}$ corresponds to l used in the previous Chapter 2.

Remark 3.1. The first part of Assumption B.1 asserts that on the right of t_0, the function u has a unique global maximum 1 at t_0 and that for $u(t)$ to be close to 1, we must have t close to t_0. For continuous u, we would simply have to assume that $u(t_0) = 1$ and $u(t) < 1$ for $t > t_0$. $\diamondsuit$

Since $\tilde{u}$ is regularly varying with positive index and $\psi(x) = o(x)$ at infinity, there exists a positive function ϕ such that

$$\tilde{u} \circ \phi(x) \sim \frac{\psi(x)}{x} \tag{3.1}$$

as x tends to infinity, and $\lim_{x\to\infty} \phi(x) = 0$. Hence ϕ is an asymptotic inverse function of the right branch of $\tilde{u}$ applied to $\psi(x)/x$. It can be chosen as $\phi(x) = \tilde{u}_+^{\leftarrow}(\psi(x)/x)$, where the generalized right inverse $\tilde{u}_+^{\leftarrow}$ is regularly varying at $0+$ with index $1/\kappa$ (Bingham et al. 1987, Th. 1.5.12). This function ϕ will play an important role in our following results.

We also use the notation $\tilde{g}(s) := g(t_0 + s)$ with the density g of T, and assume:

Assumption B.2. $T - t_0$ possesses a density $\tilde{g}$ which is regularly varying at $0+$ with index $\tau > -1$.

Since $\tilde{g}$ is locally integrable, its variation index must be at least -1. If $\tilde{g}$ is positive and continuous at 0, then τ vanishes.

The main result of this chapter is the following conditional limit theorem for (R, T) given $X > x$ and $T > t_0$, as x tends to infinity.

Theorem 3.1. *Let $X = R \cdot u(T)$ fulfill Assumptions A.1, A.2, B.1 and B.2. Then the conditional distribution of*

$$(R_x, T_x) := \left(\frac{R - x}{\psi(x)}, \frac{T - t_0}{\phi(x)} \right) \tag{3.2}$$

given $X > x$ and $T > t_0$ converges weakly, as x tends to infinity, to the distribution characterized by the density with respect to the Lebesgue measure:

$$\frac{\kappa}{\Gamma\left(\frac{1+\tau}{\kappa}\right)} t^{\tau} e^{-r}\, \mathbb{1}\{\, 0 < t < r^{1/\kappa}\,\}. \tag{3.3}$$

Furthermore, for $x \to \infty$ it holds that:

$$P\{X > x\,,\, T > t_0\} \sim \frac{1}{\kappa} \Gamma\left(\frac{1+\tau}{\kappa}\right) \phi(x) \cdot \tilde{g} \circ \phi(x)\, \overline{H}(x)\,. \tag{3.4}$$

Remark 3.2. A formal proof of Theorem 3.1 is presented in Section 3.5. Here we want to give an intuition behind the proof and explain why (3.2) is a proper normalization for (R, T).

We have $X = R \cdot u(T)$. A consequence of the survival function $\overline{H}$ being of type $\Gamma(\psi)$ is that for every $\alpha > 1$ it holds $P\{R > \alpha x\} = o\left(P\{R > x\}\right)$ for $x \to \infty$. Next, since u is at most 1, it follows: If $X > x$ for some large x, we should expect R to be about x and $u(T)$ about 1, and hence T about t_0.

Let us normalize R to R_x, such that $R = x + \psi(x) \cdot R_x$ with ψ introduced in Assumption A.2, and normalize T to T_x in an analogue way, such that $T = t_0 + \phi(x) \cdot T_x$ for some function ϕ which tends to 0 at infinity.

This gives us the representation

$$X = R \cdot u(T) = (x + \psi(x)R_x) \cdot u\left(t_0 + \phi(x)T_x\right) . \tag{3.5}$$

The function u can be analyzed in some neighborhood around t_0 by applying $\tilde{u}(s) = u(t_0) - u(t_0 + s)$. Since this function is regularly varying at $0+$ with index κ and $u(t_0) = 1$, and T_x remains bounded in probability, we have

$$u\left(t_0 + \phi(x)T_x\right) = u(t_0) - \tilde{u}\left(\phi(x)T_x\right) = 1 - (T_x)^{\kappa}\, \tilde{u} \circ \phi(x)(1 + o(1)) .$$

Because $\psi(x) = o(x)$ and R_x remains bounded in probability, it follows from representation (3.5):

$$X = R \cdot u(T) = x + \psi(x)R_x - (T_x)^{\kappa}\, x \cdot \tilde{u} \circ \phi(x)(1 + o(1)) .$$

This result points out that in order to find the limit behavior of R and T conditioned on $X > x$, i.e. for R_x and T_x to contribute to X, we should have $\psi(x)$ and $x \cdot \tilde{u} \circ \phi(x)$ of the same order of magnitude (otherwise, one of the terms would dominate the other one, and either R_x or T_x would be lost in the asymptotic). In fact, this is the reason to choose the normalizing function ϕ of T as in (3.1). Consequently, we obtain

$$X = x + \psi(x) \cdot \left(R_x - (T_x)^{\kappa}\left(1 + o(1)\right)\right) \tag{3.6}$$

and the condition that $X > x$ translates into $R_x > (T_x)^\kappa$, which is reflected in the indicator function appearing in the limit density (3.3). $\diamondsuit$

Due to the property $\phi(x) \sim \tilde{u}^{\leftarrow}(\psi(x)/x)$, it holds

$$(\phi \cdot \tilde{g} \circ \phi)(x) \sim (\tilde{u}^{\leftarrow} \cdot \tilde{g} \circ \tilde{u}^{\leftarrow})\,(\psi(x)/x) \tag{3.7}$$

as x tends to infinity. The function $\tilde{u}^{\leftarrow} \cdot \tilde{g} \circ \tilde{u}^{\leftarrow}$ is regularly varying with positive index $1/\kappa + \tau \cdot 1/\kappa = (1+\tau)/\kappa$, so is the whole expression in terms of the argument $\psi(x)/x$. This property will play an important role in the following considerations, e.g. in the consequences of Assumption C.3 as well as in the proof of Theorem 3.1.

3.3 Two-sided extensions

For many applications, limit statements conditioning only on the event $X > x$ are of interest, so that the condition $T > t_0$ should be removed. Under two-sided assumptions on the behavior of $\tilde{u}$ and $\tilde{g}$ near t_0, such extensions present no conceptual difficulty. In this section, to illustrate this assertion, we present two such extensions of Theorem 3.1, relying on the following two-sided versions of Assumptions B.1 and B.2.

Assumption C.1. There exists some t_0 such that $u(t_0) = 1$ and for any $\epsilon > 0$ it holds $\sup_{|t-t_0|>\epsilon} u(t) < 1$. Moreover, the function $\tilde{u}$ is regularly varying at $0-$ and at $0+$ with positive indices κ_- and κ_+, respectively.

The second part of Assumption C.1 signifies that for any given sign σ in $\{-,+\}$ and any $\lambda > 0$ it holds

$$\lim_{s \to 0+} \frac{\tilde{u}(\sigma \lambda s)}{\tilde{u}(\sigma s)} = \lambda^{\kappa_\sigma}\,.$$

Assumption C.2. The density $\tilde{g}$ of $T - t_0$ is regularly varying at $0-$ and at $0+$ with variation indices $\tau_- > -1$ and $\tau_+ > -1$, respectively.

Equipped with these two-sided assumptions, we define for each sign σ a positive function ϕ_σ, analogously to (3.1), such that

$$\tilde{u}\left(\sigma\phi_\sigma(x)\right) \sim \frac{\psi(x)}{x}$$

as x tends to infinity, and $\lim_{x\to\infty}\phi_\sigma(x) = 0$. Hence, $-\phi_-$ and ϕ_+ are asymptotic inverse functions of the two branches of $\tilde{u}$ applied to $\psi(x)/x$.

In order to describe the contributions of both sides of t_0 to the asymptotic behavior of (R, T), we further suppose:

Assumption C.3. For any sign σ, the limit

$$p_\sigma := \lim_{x\to\infty} \frac{(\phi_\sigma \cdot \tilde{g}(\sigma\phi_\sigma))(x)}{(\phi_- \cdot \tilde{g}(-\phi_-))(x) + (\phi_+ \cdot \tilde{g}(\phi_+))(x)}$$

exists.

Remark 3.3.

(i) Both p_- and p_+ are nonnegative and their sum is 1. They represent the contribution of the events $T < t_0$ and $T > t_0$ to the limiting conditional distribution of $T - t_0$ given $X > x$.

(ii) In the general case that $(1+\tau_-)/\kappa_- \neq (1+\tau_+)/\kappa_+$, Assumption C.3 is always fulfilled, where the weights p_σ take the values 0 resp. 1. If, however, both p_- and p_+ do not vanish, then it holds $(1+\tau_+)/\kappa_+ = (1+\tau_-)/\kappa_-$.
This follows by consideration of $\phi_\sigma \cdot \tilde{g}(\sigma\phi_\sigma)$ as a regularly varying function of $\psi(x)/x$ with index $(1+\tau_\sigma)/\kappa_\sigma$, cf. (3.7).
For $\kappa_- = \kappa_+ =: \kappa$ and $\tau_- = \tau_+ =: \tau$, Assumption C.3 does not have to be fulfilled, but it is fulfilled for (approximately) polynomial functions $\phi_\sigma(x) = f_1(\psi(x)/x) \cdot (\psi(x)/x)^{1/\kappa}$ and $\tilde{g}(s) = f_2(s) \cdot |s|^\tau$ where f_1 and f_2 are continuous with $f_1(0), f_2(0) \in (0, \infty)$.

$\diamondsuit$

We define a random sign $\mathcal{S}$ whose distribution is

$$P\{\mathcal{S}=\sigma\} := \frac{\frac{p_\sigma}{\kappa_\sigma}\Gamma\left(\frac{1+\tau_\sigma}{\kappa_\sigma}\right)}{\frac{p_-}{\kappa_-}\Gamma\left(\frac{1+\tau_-}{\kappa_-}\right)+\frac{p_+}{\kappa_+}\Gamma\left(\frac{1+\tau_+}{\kappa_+}\right)}, \qquad \sigma\in\{-,+\}. \tag{3.8}$$

The following consequence of Theorem 3.1 is central to our two-sided extension. It is also of importance to understand how the results in the next section, stated under one-sided assumptions and an extra condition $T > t_0$, can be extended with two-sided assumptions and no restriction $T > t_0$.

Proposition 3.1. *Under Assumptions A.1, A.2, C.1–C.3, the conditional distribution of*

$$S := \begin{cases} +, & T > t_0 \\ -, & T \le t_0 \end{cases} \tag{3.9}$$

given $X > x$ converges weakly as x tends to infinity to the distribution of $\mathcal{S}$ defined in (3.8).

Proof. The second assertion of Theorem 3.1 implies that for any sign σ it holds

$$P\{X > x\,,\, S=\sigma\} \sim \frac{1}{\kappa_\sigma}\Gamma\left(\frac{1+\tau_\sigma}{\kappa_\sigma}\right)\phi_\sigma(x)\cdot\tilde{g}\left(\sigma\phi_\sigma(x)\right)\overline{H}(x)$$

as x tends to infinity. Then the proposition follows from the formula

$$P\left(S=\sigma \mid X > x\right) = \frac{P\{X>x\,,\,S=\sigma\}}{P\{X>x\,,\,S=-\}+P\{X>x\,,\,S=+\}}. \qquad \square$$

Next, we define a random pair $(\mathcal{R},\mathcal{T}_{\mathcal{S}})$ whose conditional distribution given $\mathcal{S}=\sigma$ has the following density with respect to the Lebesgue measure

$$\frac{\kappa_\sigma}{\Gamma\left(\frac{1+\tau_\sigma}{\kappa_\sigma}\right)} t^{\tau_\sigma}e^{-r}\,\mathbb{1}\{\,0<t<r^{1/\kappa_\sigma}\,\}. \tag{3.10}$$

Theorem 3.2. *Under Assumptions A.1, A.2, C.1–C.3, the conditional distribution of*

$$\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi_S(x)}\right)$$

given $X > x$ converges weakly as x tends to infinity to the distribution of $(\mathcal{R},\mathcal{S}\mathcal{T}_{\mathcal{S}})$.

The density of the limit distribution can be written explicitly as

$$\sum_{\sigma\in\{-,+\}} |t|^{\tau_\sigma} e^{-r} \frac{p_\sigma \, \mathbb{1}\{\, |t|^{\kappa_\sigma} < r \,:\, \sigma t > 0\,\}}{\frac{p_-}{\kappa_-}\Gamma\left(\frac{1+\tau_-}{\kappa_-}\right) + \frac{p_+}{\kappa_+}\Gamma\left(\frac{1+\tau_+}{\kappa_+}\right)}\,.$$

Proof. For any Borel subset A of $\mathbb{R}^2$, we have

$$\begin{aligned} & P\left(\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi_S(x)}\right) \in A \,\middle|\, X > x\right) \\ = & \sum_{\sigma\in\{-,+\}} P\left(\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi_S(x)}\right) \in A \,\middle|\, X > x, S = \sigma\right) \cdot P\left(S = \sigma \,|\, X > x\right). \end{aligned} \tag{3.11}$$

Theorem 3.1 implies that the conditional distribution of

$$\left(\frac{R-x}{\psi(x)}, \sigma\frac{T-t_0}{\phi_\sigma(x)}\right)$$

given $X > x$ and $S = \sigma$ converges weakly to that of a random variable $(\mathcal{R}, \mathcal{T}_\sigma)$ whose density with respect to the Lebesgue measure is

$$\frac{\kappa_\sigma}{\Gamma\left(\frac{1+\tau_\sigma}{\kappa_\sigma}\right)} t^{\tau_\sigma} e^{-r}\, \mathbb{1}\{\, 0 < t < r^{1/\kappa_\sigma}\,\}\,.$$

Combining Proposition 3.1 and equation (3.11), we obtain that the conditional distribution of

$$\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi_S(x)}\right)$$

given $X > x$ converges weakly to that of $(\mathcal{R}, S\mathcal{T}_S)$. $\square$

In Theorem 3.2 the second component $T - t_0$ is normalized by $\phi_S(x)$, where S is a random variable (*random norming*). However, it is also possible to normalize $T - t_0$ by the *deterministic* function

$$\phi_* := \phi_+ + \phi_-$$

for which we assume:

Assumption C.4. For any sign σ, the limit

$$q_\sigma := \lim_{x\to\infty} \frac{\phi_\sigma}{\phi_*}(x)$$

exists.

Remark 3.4. The weights q_- and q_+ fulfill the properties analogously to those of p_- and p_+ in Remark 3.3, except for the variation indices κ_σ which take here the role of $(1+\tau_\sigma)/\kappa_\sigma$ in Remark 3.3(ii).
If the generalized angle T is assumed to be uniformly distributed (as a typical choice), then the weights p_σ and q_σ coincide. $\diamondsuit$

We can now state the following result with standard *deterministic norming*.

Theorem 3.3. *Under Assumptions A.1, A.2, C.1–C.4, the conditional distribution of*

$$\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi_*(x)}\right)$$

given $X > x$ converges weakly as x tends to infinity to the distribution of $(\mathcal{R}, q_{\mathcal{S}}\mathcal{S}\mathcal{T}_{\mathcal{S}})$.

The limit density can be obtained similarly to the case of Theorem 3.2.

Proof. Given Proposition 3.1 and the definition of p_σ, the conditional distribution of the random variable $\phi_{\mathcal{S}}(x)/\phi_*(x)$ given $X > x$ converges weakly to that of $q_{\mathcal{S}}$, and this convergence holds jointly with the conditional convergence of

$$\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi_{\mathcal{S}}(x)}\right) .$$

The result follows. $\square$

3.4 Conditional limit results for a bivariate random vector

The purpose of this section is to present a new approach for analyzing the conditional limit behavior of bivariate vectors (X, Y) given $X > x$, $x \to \infty$. We point out that limit results for $(X, Y) = R \cdot (u(T), v(T))$ could be deduced directly from Theorem 3.1, i.e. by considering only the representation of the single variable X.

For $Y = R \cdot v(T)$ we define

$$\tilde{v}(s) = v(t_0) - v(t_0 + s)$$

similarly to $\tilde{u}$, and assume:

Assumption B.3. The function $\tilde{v}$ is regularly varying at $0+$ with index $\delta > 0$ and $\rho := v(t_0)$.

Analogously to Remark 3.2 below Theorem 3.1, where we pointed out the influence of the distributions of

$$R_x = \frac{R - x}{\psi(x)} \qquad \text{and} \qquad T_x = \frac{T - t_0}{\phi(x)}$$

on those of X, now we investigate Y in a similar line:

$$\begin{aligned} Y &= R \cdot v(T) \\ &= (x + \psi(x)R_x)\, v\,(t_0 + \phi(x)T_x) \\ &= (x + \psi(x)R_x)\,(\rho - \tilde{v}\,(\phi(x)T_x)) \\ &= \rho x + \rho\psi(x)R_x - (x + \psi(x)R_x)\,\tilde{v}\,(\phi(x)T_x) \end{aligned}$$

Taking into account that $\tilde{v}$ is regularly varying, $\psi(x) = o(x)$, and both R_x and T_x remain bounded in probability, we obtain for nonnegative T_x:

$$Y = \rho x + \rho\psi(x)R_x - T_x^\delta x\tilde{v} \circ \phi(x)\,(1 + o(1))\ . \tag{3.12}$$

Under the conditions $X > x$ and $T > t_0$, Theorem 3.1 asserts that (R_x, T_x) converges in distribution to some $(\mathcal{R}, \mathcal{T})$ whose density with respect to the Lebesgue measure is given by (3.3). Using the Skorohod-Dudley-Wichura theorem (see Dudley 2002, Th. 11.7.2), we can choose versions of R_x and T_x which converge almost surely to $(\mathcal{R}, \mathcal{T})$ on the events $\{\, X > x\,, T > t_0\,\}$. Then we obtain, under the conditional distribution given $X > x$ and $T > t_0$, that

$$Y \stackrel{\mathrm{d}}{=} \rho x + \rho\psi(x)\mathcal{R}\,(1 + o(1)) - \mathcal{T}^\delta x\tilde{v} \circ \phi(x)\,(1 + o(1))$$

as x tends to infinity. With equation (3.1) it holds under the conditions $X > x$ and $T > t_0$:

$$\frac{Y-\rho x}{\psi(x)} \stackrel{\mathrm{d}}{=} \rho\mathcal{R}\left(1+o(1)\right) - \frac{\tilde{v}\circ\phi(x)}{\tilde{u}\circ\phi(x)}\mathcal{T}^{\delta}\left(1+o(1)\right) , \tag{3.13}$$

$$\frac{Y-\rho x}{x\cdot\tilde{v}\circ\phi(x)} \stackrel{\mathrm{d}}{=} \rho\frac{\tilde{u}\circ\phi(x)}{\tilde{v}\circ\phi(x)}\mathcal{R}\left(1+o(1)\right) - \mathcal{T}^{\delta}\left(1+o(1)\right) . \tag{3.14}$$

In the following, we investigate for which cases concerning the behavior of $\tilde{u}$ and $\tilde{v}$ which of the normalizations (3.13) or (3.14) is appropriate.
Recall that $\tilde{u}$ is regularly varying with index κ and $\tilde{v}$ is regularly varying with index δ, and that we have $\mathcal{R} > \mathcal{T}^{\kappa}$ almost surely.

Remark 3.5. If we want to waive the condition $T > t_0$, we need to introduce the random sign S from (3.9) and follow the approach provided in Section 3.3. Identity (3.12) becomes, with rather obvious notation,

$$Y = \rho x + \rho\psi(x)R_x\left(1+o(1)\right) - |T_x|^{\delta_S}\, x\cdot\tilde{v}\left(S\phi_S(x)\right)\left(1+o(1)\right) .$$

Two-sided extensions of the following bivariate conditional limit theorems cause no principal problem, but one needs to discuss the behavior of $\tilde{v}$ on both sides of 0, both in terms of regular variation and sign. As such a discussion does not seem to bring further insights, we remain focused on the one-sided results. $\diamondsuit$

We now show that our Theorem 3.1 for the univariate variable $X = R\cdot u(T)$ allows us to deduce conditional limit theorems for the bivariate vector (X, Y) in a very concise manner. For this purpose we investigate systematically several settings of our assumptions for the limit behavior of the quotient $\tilde{u}/\tilde{v}$ which are covered in the following Cases 1 to 3. They differ in normalization and limit distribution of Y. After that, in Cases 4 and 5, we present some opportunities (by imposing additional assumptions) to deduce limit results covering all the Cases 1 to 3 together in a single limit statement.

Case 1. We assume that

$$\lim_{s\to 0+} \rho\cdot\tilde{u}(s)/\tilde{v}(s) = 0 . \tag{3.15}$$

This case covers the result of Fougères and Soulier (2010) as they assumed $\delta < \kappa$. Note that it is also satisfied for $\rho = 0$. Up to the conditioning on $T > t_0$ which can be removed by using the arguments as in the previous section, the following result has been proved under stronger assumptions in Fougères and Soulier (2010, Th. 3.1). In particular, they imposed monotonicity conditions on v and u (such that v has to be increasing near t_0, and u is increasing on the left and decreasing on the right of t_0), and regular variation conditions on the derivatives of the generalized inverses $\tilde{u}_-^{\leftarrow}$, $\tilde{u}_+^{\leftarrow}$ and $\tilde{v}^{\leftarrow}$. The following corollary shows that these conditions could be omitted.

Corollary 3.1. *Let $(X, Y) = R \cdot (u(T), v(T))$ fulfill Assumptions A.1, A.2, B.1–B.3, and (3.15). Then the conditional distribution of*

$$\left(\frac{X - x}{\psi(x)}, \frac{Y - \rho x}{x \cdot \tilde{v} \circ \phi(x)}\right)$$

given $X > x$ and $T > t_0$ converges weakly to that of $(\mathcal{R} - \mathcal{T}^{\kappa}, -\mathcal{T}^{\delta})$ as x tends to infinity, where $(\mathcal{R}, \mathcal{T})$ possesses the density given in equation (3.3) from Theorem 3.1 with respect to the Lebesgue measure.

Proof. It follows directly from (3.6) for the first and (3.14) for the second component of the random vector, respectively. □

Now we deduce bivariate conditional limit statements for further cases, which are not covered by Fougères and Soulier (2010). These generalizations can be obtained as direct consequences of Theorem 3.1.

Case 2. Assume that

$$\lim_{s \to 0+} |\tilde{u}(s)/\tilde{v}(s)| = +\infty\,. \tag{3.16}$$

For instance, this is the case if $\delta > \kappa$.

Corollary 3.2. *Let $(X, Y) = R \cdot (u(T), v(T))$ fulfill Assumptions A.1, A.2, B.1–B.3, and (3.16). Then the conditional distribution of*

$$\left(\frac{X - x}{\psi(x)}, \frac{Y - \rho x}{\psi(x)}\right)$$

given $X > x$ *and* $T > t_0$ *converges weakly to that of* $(\mathcal{R} - \mathcal{T}^\kappa, \rho\mathcal{R})$ *as* x *tends to infinity, where* $(\mathcal{R}, \mathcal{T})$ *possesses the density given in equation (3.3) from Theorem 3.1 with respect to the Lebesgue measure.*

Proof. It follows directly from (3.6) and (3.13). $\square$

Note that when ρ vanishes, Corollary 3.2 yields a limit distribution with degenerate second marginal. This means that in the conditional distribution given $X > x$ and $T > t_0$, it holds $Y = o_P(\psi(x))$ as x tends to infinity.

Case 3. Assume that

$$\lim_{s\to 0+} \tilde{u}(s)/\tilde{v}(s) = C \in \mathbb{R}\,. \tag{3.17}$$

Corollary 3.3. *Let* $(X, Y) = R \cdot (u(T), v(T))$ *fulfill Assumptions A.1, A.2, B.1–B.3, and (3.17). Then the conditional distribution of*

$$\left(\frac{X - x}{\psi(x)}, \frac{Y - \rho x}{x \cdot \tilde{v} \circ \phi(x)}\right)$$

given $X > x$ *and* $T > t_0$ *converges weakly to that of* $(\mathcal{R} - \mathcal{T}^\kappa, C\rho\mathcal{R} - \mathcal{T}^\delta)$*, where* $(\mathcal{R}, \mathcal{T})$ *possesses the density given in equation (3.3) from Theorem 3.1 with respect to the Lebesgue measure.*

Proof. It follows directly from (3.6) and (3.14). $\square$

For $C = 0$ we are in Case 1.

For $C \neq 0$, then Corollary 3.3 also asserts that the conditional distribution of

$$\left(\frac{X - x}{\psi(x)}, \frac{Y - \rho x}{\psi(x)}\right)$$

given $X > x$ and $T > t_0$ converges weakly to that of $(\mathcal{R} - \mathcal{T}^\kappa, \rho\mathcal{R} - \mathcal{T}^\delta/C)$.

The limit case $C = \infty$ restates the result in Case 2.

The Cases 1, 2, and 3, considered above, are distinguished by the limit behavior of the quotient $\tilde{v}/\tilde{u}$ and, hence, by the relation between the variation indices κ and δ. For

$\kappa > \delta$ we are in Case 1, for $\kappa < \delta$ in Case 2, and $\kappa = \delta$ may yield either any of the Cases 1, 2 or 3 or none of these cases (because for $\kappa = \delta$ the quotient $\tilde{v}/\tilde{u}$ is regularly varying at $0+$ with index 0). Depending on the specific cases, the proper normalization for Y could be either of the form in (3.13) or in (3.14).

The next question to investigate is whether it is possible to have a unified normalization for Y so that its conditional distribution converges. The answer is given by our Theorems 2.2 and 2.3 in Chapter 2: There does not exist a non-degenerate limit result with a unified normalization for Y using the standard threshold normalization, i.e. for the normalization functions of Y evaluated at the threshold value x. However, the problem can be solved with *random norming*, which has been studied in Heffernan and Tawn (2004) and Heffernan and Resnick (2007), involving the random value X. For this purpose, we fix the relation between u and v. In Case 4 we consider a linear relation, in Case 5 a polynomial relation.

Case 4. Assume now that

$$v(t) = (t - t_0 + \rho) \cdot u(t) \ \text{ in some neighborhood of } t_0. \tag{3.18}$$

The following result is similar to those obtained in Theorem 2.3, but is proven now under weakened assumptions (up to the conditioning on $T > t_0$).

Corollary 3.4. *Let $(X, Y) = R \cdot (u(T), v(T))$ fulfill Assumptions A.1, A.2, B.1, B.2, and (3.18). Then the conditional distribution of*

$$\left(\frac{X - x}{\psi(x)}, \frac{(Y/X) - \rho}{\phi(x)} \right)$$

given $X > x$ and $T > t_0$ converges weakly to that of $(\mathcal{R} - \mathcal{T}^\kappa, \mathcal{T})$ as x tends to infinity, where $(\mathcal{R}, \mathcal{T})$ possesses the density given in equation (3.3) from Theorem 3.1 with respect to the Lebesgue measure.

Proof. We have

$$Y = Rv(T) = X\frac{v}{u}(T) = X(T - t_0 + \rho)\,.$$

Thus, it holds

$$\frac{(Y/X) - \rho}{\phi(x)} = T_x \, .$$

The result follows because T_x converges in distribution to $\mathcal{T}$ given $X > x$. $\square$

In many situations, as e.g. in Fougères and Soulier (2010), the assumption about v and u as in (3.18) can be met by a suitable reparametrization of T without loss of generality (which is shown in Section 2.5).
The linear form (3.18) is chosen to be suitable in view of theoretical investigations. However, it may be of interest to consider more flexible forms for the relation between v and u removing the need of a reparametrization, which we investigate now in Case 5.

Case 5. The form (3.18) can be generalized in the following way. Define the function θ by the relation

$$v(t) = \theta(t)u(t) \tag{3.19}$$

and assume that θ is n times differentiable with some nonnegative integer n and that for its derivatives it holds

$$\theta^{(j)}(t_0) = 0 \text{ if } j = 1, 2, \ldots, n-1, \quad \text{and} \quad \theta^{(n)}(t_0) \neq 0 \, . \tag{3.20}$$

In the Taylor expansion of $(v/u)(t) - \rho$ at $t = t_0$, the first non-vanishing Taylor coefficient is of order n.

Corollary 3.5. *Let $(X, Y) = R \cdot (u(T), v(T))$ fulfill Assumptions A.1, A.2, B.1, B.2, (3.19), and (3.20). Then the conditional distribution of*

$$\left(\frac{X - x}{\psi(x)}, \frac{(Y/X) - \theta(t_0)}{\phi(x)^n} \right)$$

given $X > x$ and $T > t_0$ converges weakly to that of $(\mathcal{R} - \mathcal{T}^\kappa, \mathcal{T}^n \theta^{(n)}(t_0)/n!)$ as x tends to infinity, where $(\mathcal{R}, \mathcal{T})$ possesses the density given in equation (3.3) from Theorem 3.1 with respect to the Lebesgue measure.

Proof. Using (3.19), we have

$$Y/X = \theta(T) = \theta(t_0 + \phi(x)T_x)$$

as x tends to infinity. Since $\phi(x)$ tends to 0 as x tends to infinity, Taylor formulas and the convergence in distribution of T_x to $\mathcal{T}$ yield

$$Y/X \stackrel{\mathrm{d}}{=} \theta(t_0) + \phi(x)^n \mathcal{T}^n \theta^{(n)}(t_0)/n!\,(1 + o(1))$$

as x tends to infinity, which is the result. □

Of course, one could extend this form (3.19) further by assuming that $\theta(t_0 + s) - \theta(t_0)$ is regularly varying at $0+$ as well as for many other cases of interest.
The presented cases illustrate the freedom to choose the type of function v for Y.

3.5 Proof of Theorem 3.1

This proof consists of two steps, convergence and tightness, which are disguised as asymptotic analysis of some integrals.

We will use repeatedly that, since $u(t_0) = 1$, it holds

$$u(t) = 1 - \tilde{u}(t - t_0)\,.$$

Step 1. Convergence. Let f be a nonnegative continuous function on $\mathbb{R}^2$, whose support is a compact subset of $(\mathbb{R} \setminus \{\, 0 \,\})^2$. Consider the integral

$$I(x) := \int f\left(\frac{r - x}{\psi(x)}, \frac{t - t_0}{\phi(x)}\right) \mathbb{1}\{\, ru(t) > x\,,\, t > t_0 \,\} g(t)\, \mathrm{d}H(r)\, \mathrm{d}t\,.$$

This integral is

$$\mathrm{E}\biggl(f\left(\frac{R - x}{\psi(x)}, \frac{T - t_0}{\phi(x)}\right) \mathbb{1}\{\, X > x\,,\, T > t_0 \,\}\biggr)\,,$$

that is, the conditional expectation given $X > x$ and $T > t_0$ multiplied by $P\{ X > x\,, T > t_0 \}$. The change of variables consisting in substituting r for $(r-x)/\psi(x)$ and t for $(t-t_0)/\phi(x)$ yields

$$\begin{aligned} I(x) \;&=\; \int f(r,t)\, \mathbb{1}\{\, (x + r\psi(x))\, u\,(t_0 + t\phi(x)) > x\,,\, t > 0 \,\} \\ &\qquad\qquad \phi(x) \cdot \tilde{g}\,(t\phi(x))\; \mathrm{d}H\,(x + r\psi(x))\; \mathrm{d}t\,. \end{aligned} \tag{3.21}$$

Since f has compact support which excludes the 0-coordinates, this integral is in fact an integral over a compact subset of $\mathbb{R}^2$ which excludes $r = 0$ and $t = 0$. Since r and t are now in a compact set which excludes 0, the regular variation properties of the various functions yield

$$u\,(t_0 + t\phi(x)) = 1 - \tilde{u}\,(t\phi(x)) = 1 - t^{\kappa} \cdot \tilde{u} \circ \phi(x)\,(1 + o(1))$$

and

$$\tilde{g}\,(t\phi(x)) = t^{\tau} \cdot \tilde{g} \circ \phi(x)\,(1 + o(1))$$

as x tends to infinity, and both $o(1)$ are uniform in t such that (r,t) is in the support of f, again, because we excluded the axis of $\mathbb{R}^2$. Thus, since $\psi(x) = o(x)$, we have

$$\begin{aligned} (x + r\psi(x))\, u\,(t_0 + t\phi(x)) \;&=\; (x + r\psi(x))\,(1 - \tilde{u}\,(t\phi(x))) \\ &=\; x + r\psi(x) - x \cdot (1 + o(1))\,\tilde{u}\,(t\phi(x)) \\ &=\; x + r\psi(x) - x t^{\kappa} \cdot \tilde{u} \circ \phi(x)\,(1 + o(1))\;. \end{aligned}$$

Hence applying the definition (3.1) of ϕ and that the auxiliary function ψ is positive, for the indicator function in (3.21) it holds:

$$\begin{aligned} &\mathbb{1}\{\, (x + r\psi(x))\, u\,(t_0 + t\phi(x)) > x\,,\, t > 0 \,\} \qquad (3.22)\\ =\;& \mathbb{1}\{\, r\psi(x) - t^{\kappa} x \cdot \tilde{u} \circ \phi(x)\,(1 + o(1)) > 0\,,\, t > 0 \,\} \\ =\;& \mathbb{1}\{\, \psi(x) \cdot (r - t^{\kappa}\,(1 + o(1))) > 0\,,\, t > 0 \,\} \\ =\;& \mathbb{1}\{\, r > t^{\kappa}\,(1 + o(1))\;,\, t > 0 \,\}\;. \end{aligned}$$

If x is large, the previous display shows that the indicator function in (3.22) can be sandwiched between functions

$$\mathbb{1}\{\, r > (1 \pm \epsilon) t^{\kappa} \,,\, t > 0 \,\}\,,$$

for some $\epsilon > 0$. That allows us to sandwich $I(x)$ between integrals of the form

$$I_{\pm\epsilon}(x) := \int f(r,t)\ \mathbb{1}\{\, r > (1 \pm \epsilon) t^{\kappa} \,,\, t > 0 \,\}\, \phi(x) \cdot \tilde{g} \circ \phi(x) t^{\tau} \,\mathrm{d}t \,\mathrm{d}H\left(x + r\psi(x)\right), \tag{3.23}$$

provided x is large enough, thus for positive ϵ and x large enough,

$$(1 - \epsilon) I_{-\epsilon}(x) \le I(x) \le (1 + \epsilon) I_{\epsilon}(x)\,. \tag{3.24}$$

The measure $\mathrm{d}H(x + \psi(x)r)/\overline{H}(x)$ converges vaguely to a measure with density e^{-r} with respect to the Lebesgue measure. Note that we are using vague convergence of measure, so that r has to remain in a compact set, which is why we took f having a compact support with respect to both variables r and t. Consequently, we obtain

$$\begin{aligned} &\lim_{x \to \infty} \frac{I_{\pm\epsilon}(x)}{\phi(x) \cdot \tilde{g} \circ \phi(x) \overline{H}(x)} \\ &= \int f(r,t)\ \mathbb{1}\{\, r > (1 \pm \epsilon) t^{\kappa} \,,\, t > 0 \,\}\, t^{\tau} \,\mathrm{d}t \, e^{-r} \,\mathrm{d}r \end{aligned} \tag{3.25}$$

as x tends to infinity. Since ϵ is arbitrary, combining (3.24) and (3.25) yield

$$\lim_{x \to \infty} \frac{I(x)}{\phi(x) \cdot \tilde{g} \circ \phi(x) \overline{H}(x)} = \int f(r,t)\ \mathbb{1}\{\, r > t^{\kappa} \,,\, t > 0 \,\}\, t^{\tau} \,\mathrm{d}t \, e^{-r} \,\mathrm{d}r\,.$$

Step 1+1/2. Refinement. In Step 1, the function f is supported in $(\mathbb{R} \setminus \{\, 0 \,\})^2$. To prove vague convergence of the distribution as distribution on $\mathbb{R}^2$, we need to allow for compact support in the entire $\mathbb{R}^2$, not excluding the axes. To make this extension, it suffices to show that there is no mass accumulation along the axes $\{\, 0 \,\} \times \mathbb{R}$ and $\mathbb{R} \times \{\, 0 \,\}$. Thus, setting

$$J_{1,\epsilon}(x) := P\left\{ \frac{|R - x|}{\psi(x)} \le \epsilon\,,\ R \cdot u(T) > x\,,\, T > t_0 \right\}$$

and

$$J_{2,\epsilon}(x) := P\left\{\frac{|T-t_0|}{\phi(x)} \le \epsilon\,,\ R\cdot u(T) > x\,,\ T > t_0\right\}\,,$$

we need to prove that for $j = 1, 2$,

$$\lim_{\epsilon\to 0}\limsup_{x\to\infty}\frac{J_{j,\epsilon}}{\phi(x)\cdot\tilde{g}\circ\phi(x)\overline{H}(x)} = 0\,. \tag{3.26}$$

To do this, we have, for x large enough, that $J_{1,\epsilon}(x)$ is at most

$$\begin{aligned}
&P\left\{|R-x| \le \epsilon\psi(x)\,,\ u(T) > \frac{x}{x+\epsilon\psi(x)}\,,\ T > t_0\right\}\\
&\quad\le \left(\overline{H}\left(x-\epsilon\psi(x)\right) - \overline{H}\left(x+\epsilon\psi(x)\right)\right) P\left\{u(T) > 1 - 2\epsilon\frac{\psi(x)}{x}\,,\ T > t_0\right\}\\
&\qquad\qquad \le \overline{H}(x)(e^{\epsilon} - e^{-\epsilon})\,(1+o(1))\,P\left\{\tilde{u}(T-t_0) < 2\epsilon\frac{\psi(x)}{x}\,,\ T > t_0\right\}\,,
\end{aligned}$$

where the last inequality comes from $\overline{H} \in \Gamma(\psi)$, the definition of $\tilde{u}$ and that $u(t_0) = 1$. But since $\tilde{u}$ is regularly varying with index κ,

$$\tilde{u}\left((2\epsilon)^{1/\kappa}\phi(x)\right)\ \sim\ 2\epsilon\cdot\tilde{u}\circ\phi(x)\ \sim\ 2\epsilon\cdot\psi(x)/x \tag{3.27}$$

as x tends to infinity. Consequently, for x large enough,

$$\begin{aligned}
& P\left\{\tilde{u}(T-t_0) < 2\epsilon\frac{\psi(x)}{x}\,,\ T > t_0\right\}\\
\le\quad & P\left\{\tilde{u}(T-t_0) < \tilde{u}\left((4\epsilon)^{1/\kappa}\phi(x)\right)\,,\ T > t_0\right\}\\
\le\quad & P\left\{|T-t_0| < (8\epsilon)^{1/\kappa}\phi(x)\,,\ T > t_0\right\}\,,
\end{aligned}$$

where the last inequality comes from the fact that a regularly varying function with positive index is asymptotically equivalent to an increasing function, see Resnick (1987, Prop. 0.8(vii)).

Note that for any θ positive,

$$P\left\{0 < T - t_0 \le \theta\phi(x)\right\} \sim \phi(x)\cdot\tilde{g}\circ\phi(x)\int_0^{\theta} y^{-\tau}\,\mathrm{d}y\,, \tag{3.28}$$

as x tends to infinity, because this probability is

$$\int_0^{\theta\phi(x)} \tilde{g}(s)\,\mathrm{d}s = \phi(x)\cdot\tilde{g}\circ\phi(x)\int_0^{\theta}\frac{\tilde{g}\left(s\phi(x)\right)}{\tilde{g}\circ\phi(x)}\,\mathrm{d}s$$

and $\tilde{g}$ is regularly varying with index $\tau > -1$. Thus, combining the various bounds, we have, for x large enough,

$$J_{1,\epsilon}(x) \leq 2\overline{H}(x)(e^{\epsilon} - e^{-\epsilon})\phi(x) \cdot \tilde{g} \circ \phi(x) \int_0^{(16\epsilon)^{1/\kappa}} y^{-\tau}\, \mathrm{d}y$$

and this proves (3.26) for $j = 1$.

To prove (3.26) for $j = 2$, we see that for x large enough,

$$\begin{aligned} J_{2,\epsilon}(x) &\leq P\{0 < T - t_0 \leq \epsilon\phi(x)\,,\, R > x\} \\ &= P\{0 < T - t_0 \leq \epsilon\phi(x)\}\, \overline{H}(x)\,. \end{aligned}$$

Then we use (3.28) to bound $J_{2,\epsilon}(x)$, establishing (3.26) for $j = 2$.

Combined with Step 1, this shows that for any nonnegative continuous compactly supported function f on $\mathbb{R}^2$

$$\begin{aligned} &\mathrm{E}\left(\mathrm{f}\left(\frac{\mathrm{R}-\mathrm{x}}{\psi(\mathrm{x})}, \frac{\mathrm{T}-\mathrm{t}_0}{\phi(\mathrm{x})}\right) \mathbb{1}\{\,\mathrm{X} > \mathrm{x}\,,\, \mathrm{T} > \mathrm{t}_0\,\}\right) \\ &\qquad \sim \phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x) \int f(r,t)\; \mathbb{1}\{\, r > t^{\kappa}\,,\, t > 0\,\} t^{\tau}\, \mathrm{d}t\, e^{-r}\, \mathrm{d}r \end{aligned}$$

as x tends to infinity. By writing any continuous function as the sum of its positive and negative part, this still holds for any continuous and compactly supported function on $\mathbb{R}^2$.

Step 2. Tightness. We now show that $(R-x)/\psi(x)$ and $(T-t_0)/\phi(x)$ are tight random variables under the conditional probability given $X > x$ and $T > t_0$. For this purpose, given Step 1 and anticipating the conclusion of the proof, we need to show that

$$\lim_{r\to\infty} \limsup_{x\to\infty} \frac{P\left\{\frac{|R-x|}{\psi(x)} > r\,,\, R \cdot u(T) > x\,,\, T > t_0\right\}}{\phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x)} = 0\,, \tag{3.29}$$

and

$$\lim_{t\to\infty} \limsup_{x\to\infty} \frac{P\left\{\frac{|T-t_0|}{\phi(x)} > t\,,\, R \cdot u(T) > x\,,\, T > t_0\right\}}{\phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x)} = 0\,, \tag{3.30}$$

We will examine the different cases obtained when 'removing' the absolute values in (3.29) and (3.30).

At first, to prove statement (3.29) we consider for positive r the probabilities:

(1) $P_{1,r}(x) := P\left\{R > x + r\psi(x)\,,\; R\cdot u(T) > x\,,\; T > t_0\right\}\,,$

(2) $P_{2,r}(x) := P\left\{R < x - r\psi(x)\,,\; R\cdot u(T) > x\,,\; T > t_0\right\}\,.$

Ad (1): For x large enough, $P_{1,r}(x)$ is at most

$$\overline{H}\left(x + r\psi(x)\right) P\left\{u(T) > \frac{x}{x + r\psi(x)}\,,\; T > t_0\right\}$$
$$\sim \overline{H}(x)e^{-r}P\left\{\tilde{u}(T - t_0) < r\frac{\psi(x)}{x}\left(1 + o(1)\right)\,,\; T > t_0\right\}\,.$$

As in Step 1+1/2, using (3.28), this is of order at most

$$\overline{H}(x)\phi(x)\cdot g\circ\phi(x)e^{-r}\int\limits_0^{(4r)^{1/\kappa}} y^{-\tau}\,\mathrm{d}y\,.$$

Thus,

$$\lim_{r\to\infty}\limsup_{x\to\infty}\frac{P_{1,r}(x)}{\phi(x)\cdot\tilde{g}\circ\phi(x)\overline{H}(x)} = 0\,.$$

Ad (2): When x is large enough, $\psi(x)$ is well defined and positive. In that range of x, since $|u| \le 1$, we cannot have $R\cdot u(T) > x$ while having $R < x - r\psi(x)$. Thus $P_{2,r}(x) = 0$ whenever x is large enough.

Now to prove the second statement (3.30) we consider the probability

(3) $P_{3,t}(x) := P\left\{T > t_0 + t\phi(x)\,,\; R\cdot u(T) > x\,,\; T > t_0\right\}\,,$

It holds

$$P_{3,t}(x) = \int\limits_t^{\infty}\overline{H}\left(\frac{x}{u\left(t_0 + \phi(x)s\right)}\right)\phi(x)\cdot\tilde{g}\left(s\phi(x)\right)\,\mathrm{d}s\,. \tag{3.31}$$

We may assume that t is greater than 1. Let η be a (small) positive real number and let ϵ be small enough so that Potter's bounds (cf. Bingham et al. 1987, Th. 1.5.6)

$$\tilde{u}\left(s\phi(x)\right) \geq \frac{1}{2}\tilde{u} \circ \phi(x) s^{\kappa-\eta}$$

and

$$\tilde{g}\left(s\phi(x)\right) \leq 2\tilde{g} \circ \phi(x) s^{\tau+\eta} \tag{3.32}$$

apply on the range $1 \leq t \leq s \leq \epsilon/\phi(x)$. Then we have, on that range of s (provided ϵ was chosen small enough),

$$\begin{aligned} \frac{1}{u\left(t_0 + s\phi(x)\right)} = \frac{1}{1 - \tilde{u}\left(s\phi(x)\right)} &\geq 1 + \frac{1}{4}\tilde{u}\left(s\phi(x)\right) \\ &\geq 1 + \frac{1}{8}\tilde{u} \circ \phi(x) s^{\kappa-\eta}\,. \end{aligned} \tag{3.33}$$

Referring to part of the integral (3.31), using the definition of ϕ, (3.32) and (3.33), we have then

$$\int_t^{\epsilon/\phi(x)} \overline{H}\left(\frac{x}{u\left(t_0 + s\phi(x)\right)}\right)\phi(x) \cdot \tilde{g}\left(s\phi(x)\right) \,\mathrm{d}s$$
$$\leq 2\phi(x) \cdot \tilde{g} \circ \phi(x) \int_t^{\epsilon/\phi(x)} \overline{H}\left(x + \frac{1}{16}\psi(x) s^{\kappa-\eta}\right) s^{\tau+\eta}\,\mathrm{d}s$$

Using the first statement of Lemma 5.1 in Fougères and Soulier (2010) (note we can take $C = 2$ in that Lemma, which we do here), this upper bound is at most

$$4\phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x) \int_t^{\infty} \frac{s^{\tau+\eta}}{\left(1 + (s^{\kappa-\eta}/16)\right)^p}\,\mathrm{d}s\,, \tag{3.34}$$

where p is taken large enough so that the integral converges.

We now work on the easy part of the integral (3.31), namely, that for s between $\epsilon/\phi(x)$ and ∞. Given how this integral was obtained, this part corresponds to $T > t_0 + \epsilon$, and it is at most (again, provided we choose ϵ small enough)

$$P\left\{Ru(t_0 + \epsilon/2) > x\right\} = \overline{H}\left(\frac{x}{u(t_0 + \epsilon/2)}\right)\,. \tag{3.35}$$

We now claim that if $c > 1$ (think of c as $1/u(t_0 + \epsilon/2)$), then

$$\overline{H}(cx) = o\left(\overline{H}(x)\phi(x) \cdot \tilde{g} \circ \phi(x)\right) . \tag{3.36}$$

Indeed, using the second statement of Lemma 5.1 in Fougères and Soulier (2010), for any positive p we have

$$\overline{H}(cx) \leq \left(\frac{\psi(x)}{x}\right)^p \overline{H}(x)$$

provided x is large enough (note that we can take $C = 1$ in their inequality: it suffices to divide their p by 2 and see that their C times $(\psi(x)/x)^{p/2}$ tends to 0 and is less than 1 for x large enough). Thus, to prove (3.36), we have to show that for any p large enough it holds:

$$\left(\frac{\psi(x)}{x}\right)^p = o\left(\phi(x) \cdot \tilde{g} \circ \phi(x)\right) . \tag{3.37}$$

Since we may view $\phi(x) \cdot \tilde{g} \circ \phi(x)$ as a regularly varying function of $\psi(x)/x$ with index $(1+\tau)/\kappa$, see (3.7), equation 3.37 is obviously fulfilled for all $p > (1+\tau)/\kappa$.

Now, combining (3.35) and (3.36), we obtain that, referring to part of (3.31)

$$\int_{\epsilon/\phi(x)}^{\infty} \overline{H}\left(\frac{x}{u\left(t_0 + \phi(x)s\right)}\right) \tilde{g}\left(\phi(x)s\right) \phi(x) \,\mathrm{d}s = o\left(\phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x)\right)$$

as x tends to infinity. Combined with (3.34), and referring to (3.31) this shows that

$$\limsup_{x\to\infty} \frac{P_{3,t}(x)}{\phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x)} \leq 4 \int_t^{\infty} \frac{s^{\tau+\eta}}{(1 + (s^{\kappa-\eta}/16))^p} \,\mathrm{d}s \,,$$

Hence, it holds

$$\lim_{t\to\infty} \limsup_{x\to\infty} \frac{P_{3,t}(x)}{\phi(x) \cdot \tilde{g} \circ \phi(x)\overline{H}(x)} = 0 \,.$$

and the proof of (3.30) is completed.

To conclude the proof of Theorem 3.1, we combine Steps 1, 1+1/2 and 2, and obtain that

$$
\begin{aligned}
& P\{X > x\,,\, T > t_0\} \\
\sim\quad & \phi(x)\cdot\tilde{g}\circ\phi(x)\overline{H}(x)\int \mathbb{1}\{\, r > t^{\kappa}\,,\, t > 0\,\} t^{\tau} e^{-r}\,\mathrm{d}t\,\mathrm{d}r \\
\sim\quad & \phi(x)\cdot\tilde{g}\circ\phi(x)\overline{H}(x)\,\frac{1}{\kappa}\,\Gamma\left(\frac{1+\tau}{\kappa}\right)
\end{aligned}
$$

as x tends to infinity. Then Step 2 implies that the conditional distribution of

$$\left(\frac{R-x}{\psi(x)}, \frac{T-t_0}{\phi(x)}\right)$$

given $X > x$ and $T > t_0$ is tight, and Step 1 proves that it converges to the limit given in Theorem 3.1. $\square$

Chapter 4

Weakening the independence assumption on polar components: Limit theorems for generalized elliptical distributions [1]

4.1 Introduction

Analyzing and predicting extreme events for random vectors is of particular interest in numerous applications. In contrast to the univariate case, there are different ways to define the term *extreme event* by a requirement that one or several or even all of the vector components have to be large simultaneously. An effective approach for the analysis of multivariate extreme values was introduced by Heffernan and Tawn (2004): They examined the distribution tail of a random vector in terms of the conditional distribution given that one of the components of the random vector becomes large. This approach was further developed and extended to the *conditional extreme value (CEV) model* by Heffernan and Resnick (2007) and Das and Resnick (2011a).

[1] The contents of this chapter will appear in SEIFERT, M.I. (2015): Weakening the independence assumption on polar components: Limit theorems for generalized elliptical distributions. *Journal of Applied Probability*.

Conditional limit statements for elliptical and more general random vectors possessing a polar representation $(X,Y) = R\cdot(u(T),v(T))$ with radial component R and angular component T were intensively investigated among others by Berman (1983), Abdous et al. (2005), Fougères and Soulier (2010), Hashorva (2012) and in Chapter 2 in this thesis. The latter three works show that rather weak and local assumptions on the coordinate functions u and v are sufficient to derive limit statements. But one assumption made is very rigid, namely that R and T are stochastically independent. This requirement is global and unstable, such that its validity is assumed to hold even in regions which are not important for the limit behavior.

A possibility to weaken this independence assumption would underscore a certain stability of the polar extreme value model with respect to the above mentioned limit results, which is of particular interest for statistical inferences in applications. Hence, a natural question remains open: How much can we deviate from the stochastic independence of the polar components R and T, such that we still obtain a conditional limit result for $(X,Y) = R\cdot(u(T),v(T))$? Up to now, to the best of our knowledge, there is only one result in this direction presented in the book of Balkema and Embrechts (2007) which does not explicitly use polar representations.

In this chapter we introduce a novel approach for weakening the independence assumption. After definitions in Section 4.2, we present in Section 4.3 the extreme value model with independent polar components where the radial component R belongs to the Gumbel max-domain of attraction with some auxiliary function ψ. We unite two important results from Fougères and Soulier (2010) and our main result from Chapter 2 in Theorem 4.1 and deduce a new Theorem 4.2 now for *dependent* polar components which states: the limit results still hold if the conditional distributions of R given $T = t$ have a similar tail behavior with asymptotically equivalent auxiliary functions ψ_t.

In order to verify such a condition in empirical applications, we develop convenient and model-independent geometric criteria. We describe the dependence between R and T by comparing $R\cdot(u(T),v(T))$ with some *reference model* $\widetilde{R}\cdot(u(T),v(T))$

where $\widetilde{R}$ and T are independent. The difference between the conditional distribution functions $H_t(r) = P(R \le r \mid T = t)$ and the reference distribution function $\widetilde{H}(r) = P(\widetilde{R} \le r)$ is measured by *shifts* $\delta_t(r)$. In Section 4.4 we show in Theorem 4.3 that the limit results still hold for *relative shifts* $\delta_t(r)/r$ vanishing asymptotically for $r \to \infty$. Furthermore, we deduce in Theorem 4.4 limit results for relative shifts which tend to a t-dependent limit, so that the auxiliary functions ψ_t are no more asymptotically equivalent. In Section 4.5 we compare our approach for weakening the independence assumption with Balkema and Embrechts' one. Theorem 4.5 analyzes cases where Theorem 4.4 extends their results.

4.2 Preliminaries

First, we give the definitions and important properties of the regular and of Γ-variation (Resnick 1987, Geluk and de Haan 1987, de Haan and Ferreira 2006). All considered functions are assumed to be Lebesgue measurable. Two functions f and g are said to be asymptotically equivalent if $f(x)/g(x) \to 1$ for $x \to \infty$ (written $f \sim g$).

Definition 4.1. An eventually positive function $f : (0, \infty) \to \mathbb{R}$ is said to be regularly varying at infinity with index $\alpha \in \mathbb{R}$ (written $f \in \mathrm{RV}_\alpha(\infty)$), if for all $\lambda > 0$ holds

$$\lim_{x \to \infty} \frac{f(\lambda x)}{f(x)} = \lambda^\alpha. \tag{4.1}$$

A function $g : \mathbb{R} \to \mathbb{R}$ is said to be regularly varying at t_0 with index $\alpha \in \mathbb{R}$ (written $g \in \mathrm{RV}_\alpha(t_0)$) if for all $\lambda > 0$ holds

$$\lim_{s \to 0} \frac{g(t_0 + \lambda s)}{g(t_0 + s)} = \lambda^\alpha. \tag{4.2}$$

If $g \in \mathrm{RV}_\alpha(t_0)$ fulfills

$$\lim_{s \downarrow 0} \frac{|g(t_0 + s)|}{|g(t_0 - s)|} = 1,$$

we call it *infinitesimally symmetric* and write $g \in \mathrm{RV}^s_\alpha(t_0)$.

Remark 4.1.

(i) The convergence in (4.1) and (4.2) is locally uniform in λ.

(ii) For $g \in \mathrm{RV}_\alpha(t_0)$ it follows $g(t) \neq 0$ for $t \in (t_0 - \epsilon, t_0 + \epsilon) \setminus \{t_0\}$ for some $\epsilon > 0$.

(iii) $f \in \mathrm{RV}_\alpha(\infty)$ implies $\lim_{x\to\infty} f(x) = 0$ for $\alpha < 0$, $\lim_{x\to\infty} f(x) = \infty$ for $\alpha > 0$.

(iv) If f is eventually positive and

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)}$$

exists, is finite and positive for all λ in a set of positive Lebesgue measure, then $f \in \mathrm{RV}_\alpha(\infty)$ for some $\alpha \in \mathbb{R}$ (Th. 1.2 in Geluk and de Haan 1987, Th. B.1.3 in de Haan and Ferreira 2006). $\diamondsuit$

Definition 4.2. A non-decreasing function f is said to be $\Gamma(\psi)$-varying (with a positive *auxiliary function* ψ) if for all $z \in \mathbb{R}$ holds

$$\lim_{x\to\infty} \frac{f(x + z\psi(x))}{f(x)} = \mathrm{e}^z. \tag{4.3}$$

We say that a random variable R on $[0, \infty)$ and its distribution function H resp. survival function $\overline{H} = 1 - H$ are of type $\Gamma(\psi)$, if $1/\overline{H}$ is $\Gamma(\psi)$-varying, i.e. if for all $z \in \mathbb{R}$ holds

$$\lim_{r\to\infty} \frac{P\{R > r + z\psi(r)\}}{P\{R > r\}} = \lim_{r\to\infty} \frac{\overline{H}(r + z\psi(r))}{\overline{H}(r)} = \mathrm{e}^{-z}. \tag{4.4}$$

Remark 4.2.

(i) The auxiliary function ψ is unique up to asymptotic equivalence, i.e. a positive function ψ_2 is an auxiliary function for R of type $\Gamma(\psi_1)$ if and only if $\psi_1 \sim \psi_2$. It can be chosen differentiable satisfying $\lim_{r\to\infty} \psi'(r) = 0$ (Geluk and de Haan 1987, Th. 1.28(ii), Cor. 1.29).

(ii) R is of type $\Gamma(\psi)$ if and only if R is in the Gumbel max-domain of attraction with infinite right endpoint: $\sup\{r : H(r) < 1\} = \infty$.

(iii) The von Mises representation: H is of type $\Gamma(\psi)$ if and only if it fulfills

$$\overline{H}(r) = 1 - H(r) = a(r)\,\exp\left(-\int\limits_0^r 1/\psi(u)\,\mathrm{d}u\right) \tag{4.5}$$

with $\lim_{r\to\infty} a(r) = a \in (0,\infty)$ (Resnick 1987, Prop. 1.4). $\diamondsuit$

4.3 From independent to dependent polar components

We consider bivariate random vectors $(X,Y) \in [0,\infty) \times (-\infty,\infty)$ on the right half-plane, since we are interested in the asymptotic behavior for X becoming large. (X,Y) can be represented in *Euclidean polar coordinates* $(X,Y) \stackrel{d}{=} A \cdot (\cos\Theta, \sin\Theta)$ with Euclidean distance $A \geq 0$ and Euclidean angle $\Theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
The popular class of *elliptical distributions* is described conveniently by:

$$(X,Y) \stackrel{d}{=} R \cdot (\cos T, \rho\cos T + \sqrt{1-\rho^2}\sin T), \qquad \rho \in (-1,1) \tag{4.6}$$

with stochastically independent R and T, where T is uniformly distributed.
More generally, we investigate random vectors which possess a *polar representation*

$$(X,Y) \stackrel{d}{=} R \cdot (u(T), v(T)) \tag{4.7}$$

with *polar components* R and T and quite arbitrary *coordinate functions* u and v.

Such elliptical and generalized distributions were intensively investigated with respect to their conditional limit behavior. A detailed overview about this research field is given in Chapter 1. In this chapter, we start at random vectors with polar representation (4.7) fulfilling the following three assumptions.

Assumption 1. (i) R and T take values in $[0,\infty)$ resp. in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$;
(ii) (R,T) possesses a positive, continuous joint density f_{RT};
(iii) there exists a diffeomorphism τ of $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ with derivative $\tau' > 0$ such that $\Theta = \tau(T)$ for the Euclidean angle Θ.

Assumption 2. (i) R is of type $\Gamma(\psi)$; (ii) R and T are stochastically independent.

Assumption 3.

$$\begin{aligned} u(t) &= u_{\max} - l(t) \text{ for } t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \text{ with some } u_{\max} \in (0, \infty), \\ v(t) &= \tan(t) \cdot u(t) \text{ for } t \in (t_0 - \epsilon_0, t_0 + \epsilon_0) \text{ with some } \epsilon_0 > 0, \end{aligned} \tag{4.8}$$

where $l : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [0, u_{\max}]$ has a unique zero at $t_0 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and its derivative l' is $\mathrm{RV}^s_{\kappa-1}(t_0)$ for some $\kappa > 0$. We denote $\rho := (v/u)(t_0) = \tan(t_0)$.

According to Remark 4.1(ii), in some ϵ-neighborhood $(t_0 - \epsilon, t_0 + \epsilon) \setminus \{t_0\}$ it holds $l' \neq 0$ and, hence, $u' \neq 0$. Thus, u increases strictly on $(t_0 - \epsilon, t_0)$ and decreases strictly on $(t_0, t_0 + \epsilon)$. As a consequence, l possesses two branches of inverses $l_\pm^{-1} \in \mathrm{RV}_{1/\kappa}(0)$ on $(-\epsilon, 0)$ resp. $(0, \epsilon)$. Assumption 3 implies that l is $\mathrm{RV}^s_\kappa(t_0)$ and that u has the unique global maximum $u_{\max}$ at $t = t_0$.

Remark 4.3. There is much freedom to select a polar representation (4.7).

The angular component T simply labels the rays $y = \gamma x$, $\gamma \in \mathbb{R}$ in the (x, y)-plane. We specify T (in Assumption 3) to coincide locally with the Euclidean angle Θ and assume (in Assumption 1) a diffeomorphism τ between T and Θ, which enables us to get the density f_{XY} of (X, Y) by using $f_{XY}(a\cos(\vartheta), a\sin(\vartheta)) = f_{A\Theta}(a, \vartheta)/a$:

$$f_{XY}(ru(t), rv(t)) = \frac{f_{RT}(r, t)}{r \cdot (u^2(t) + v^2(t)) \cdot \tau'(t)}, \tag{4.9}$$

as $a^2 = r^2 \cdot (u^2(t) + v^2(t))$ and $\vartheta = \tau(t)$.

For the radial component R only a linear scaling is allowed:
$(X, Y) = R \cdot (u(T), v(T))$ possesses also the polar representation

$$(X, Y) = R^* \cdot (u^*(T), v^*(T))$$

with $R^* = c \cdot R$ and $u^* = u/c$, $v^* = v/c$.
If $c \in (0, \infty)$ is a *constant*, then this is a global rescaling of R, often done to get $\max u^* = 1$ as in the representations (4.6).

If c is a *non-constant* function of T, then this changes the dependence structure between the polar components; in Theorem 4.4 we take advantage of this possibility. $\diamondsuit$

With $k = \kappa^{-1/\kappa} \cdot l_+^{-1}$ for $\zeta \geq 0$, and $k = -\kappa^{-1/\kappa} \cdot l_-^{-1}$ for $\zeta < 0$, as well as

$$G(\zeta) = \frac{1}{2\kappa^{1/\kappa-1}\Gamma(1/\kappa)} \int_{-\infty}^{\zeta} \exp\left(-\left|s\right|^{\kappa}/\kappa\right) \mathrm{d}s, \tag{4.10}$$

we state results of Fougères and Soulier (2010, Th. 3.1) for variation indices $\kappa > 1$ and of our Theorem 2.3 from Chapter 2 for $\kappa > 0$ using *random norming* (cf. Heffernan and Resnick 2007) with the bound on Y evaluated not at the threshold x but at the actual value X:

Theorem 4.1 (Independent polar components)**.**
Let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumption 1, 2, and 3. Then for all $\xi > 0$, $\zeta \in \mathbb{R}$, and $\kappa > 1$ it holds

$$\lim_{x\to\infty} P(X \leq x + \psi(x)\xi,\, Y \leq \rho x + xk\left(\psi(x)/x\right)\zeta \mid X > x) = \left(1 - \mathrm{e}^{-\xi}\right) \cdot G(\zeta), \tag{4.11}$$

and for arbitrary $\kappa > 0$ it holds

$$\lim_{x\to\infty} P(X \leq x + \psi(x)\xi,\, Y \leq \rho X + Xk\left(\psi(X)/X\right)\zeta \mid X > x) = \left(1 - \mathrm{e}^{-\xi}\right) \cdot G(\zeta). \tag{4.12}$$

Note that the radial component R influences the limit statements of Theorem 4.1 *only* by its tail behavior characterized by the auxiliary function ψ.

Now we deduce a generalization of Theorem 4.1, where R and T do not have to be stochastically independent anymore. We assume the conditional distribution functions

$$H_t(r) = P\left(R \leq r \mid T = t\right), \quad r \in (0, \infty),\ t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{4.13}$$

to be of type $\Gamma(\psi_t)$ with asymptotically equivalent auxiliary functions ψ_t, i.e. there exists some ψ with $\psi_t \sim \psi$ for all t. Then the distinction among the H_t is captured by the a_t in the von Mises representation, cf. Remark 4.2(i), (iii).

Theorem 4.2 (Dependent polar components with similar conditional tails)**.**
Let $(X,Y) = R\cdot(u(T),v(T))$ satisfy Assumptions 1 and 3. Instead of Assumption 2, let T have a positive, continuous marginal density and $\overline{H}_t(r) = a_t(r)\ \exp\left(-\int_0^r 1/\psi(u)\,\mathrm{d}u\right)$ with $a_t(r) \to a_t > 0$ uniformly in t for $r \to \infty$. Then the limit statements (4.11) and (4.12) hold.

The proof of Theorem 4.2 is provided in Section 4.6.

Remark 4.4. The distribution of (X,Y) according to Theorem 4.2 might differ substantially from those with independent polar components, even in the asymptotic region, see Example 4.2 with Figure 4.2. $\diamondsuit$

4.4 Geometric dependence measure and criteria for limit theorems

Here we present criteria formulated in terms of the distributions (not using the auxiliary functions) which allow us to apply Theorem 4.2. We describe the dependence between R and T by comparing $(X,Y) = R\cdot(u(T),v(T))$ with a *reference model* $(\widetilde{X},\widetilde{Y}) = \widetilde{R}\cdot(u(T),v(T))$. Hereby, (X,Y) fulfills only Assumptions 1 and 3, but $(\widetilde{X},\widetilde{Y})$ fulfills also Assumption 2, in particular $\widetilde{R}$ and T are independent. We denote quantities with respect to $\widetilde{R}$ by a tilde: distribution function $\widetilde{H}$, density $\widetilde{h}$ and auxiliary function $\widetilde{\psi}$, and joint densities $f_{\widetilde{R}T}$ and $f_{\widetilde{X}\widetilde{Y}}$.

The distance between the corresponding distributions $\widetilde{H}$ of $\widetilde{R}$ and H_t from (4.13) is measured by $\delta : \left[-\frac{\pi}{2},\frac{\pi}{2}\right] \times (0,\infty) \to \mathbb{R}$, $(t,r) \mapsto \delta_t(r)$ with:

$$\delta_t = \left(\widetilde{H}^{\leftarrow} - H_t^{\leftarrow}\right) \circ H_t \quad \Leftrightarrow \quad H_t(r) = \widetilde{H}(r+\delta_t(r)). \tag{4.14}$$

The asymptotics of $\widetilde{H}$ should correspond to that of the H_t. Besides that, the choice of $\widetilde{H}$ is free; we assume for later considerations on densities that $\widetilde{h}$ is monotonically

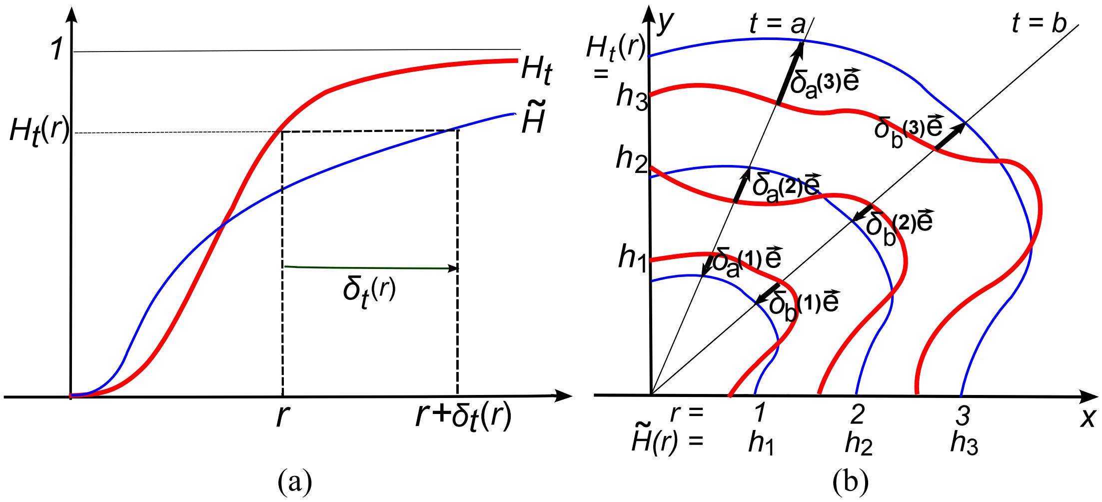


Figure 4.1: Meaning of δ_t: (a) distribution functions $\widetilde{H}$ (blue), H_t (red), (b) their level lines.

decreasing. Note that (R, T) as well as $(\widetilde{R}, T)$ fulfill Assumption 1(ii), hence, $\delta_t(r)$ is continuous in t and continuously differentiable in r, and the H_t possess densities h_t.

Figure 4.1(a) shows the meaning of $\delta_t(r)$. Figure 4.1(b) illustrates δ_t in the (x, y)-plane (considering r as a function of x, y) as a radially directed vector field $\delta_t(r)\,\vec{e}$ with $\vec{e} = \frac{1}{r} \cdot (x, y) = (u(t), v(t))$. The sets $\{r = r_1\} = \{\widetilde{H}(r) = f_1\}$ with $f_1 = \widetilde{H}(r_1)$ do not coincide with $\{H_t(r) = f_1\}$ but differ from them by the shifts $\delta_t(r_1)\,\vec{e}$.

Each of the following two assumptions guarantees that all H_t are of type Γ with asymptotically equivalent auxiliary functions.

Assumption $\widetilde{\mathbf{2}}_{\mathbf{A}}$. (i) $\widetilde{\psi} \in \mathrm{RV}_\alpha(\infty)$; (ii) $\delta_t'(r) \to 0$ uniformly in t for $r \to \infty$.

In the standard cases of elliptical vectors with $\widetilde{R}$ of type $\Gamma(\widetilde{\psi})$, e.g. bivariate normal, Kotz and logistic distributions, the auxiliary function $\widetilde{\psi}$ is regularly varying and, hence, fulfills part (i) of Assumption $\widetilde{2}_{\mathrm{A}}$. However, if $\widetilde{H}$ cannot be chosen such that it possesses a regularly varying $\widetilde{\psi}$, then we alternatively assume:

Assumption $\widetilde{2}_{\mathrm{B}}$. (i) $\delta_t(r) \cdot \widetilde{\psi}'(r)/\widetilde{\psi}(r) \to 0$ for $r \to \infty$; (ii) $\delta_t'(r) \to 0$ uniformly in t for $r \to \infty$.

Note that Assumptions $\widetilde{2}_{\mathrm{A}}$ and $\widetilde{2}_{\mathrm{B}}$ do not exclude shifts $\delta_t(r) \to \infty$, see Example 4.2.

Geluk and de Haan (1987) investigated the class of Γ-varying functions and provided their major properties. One result (in their proof of Prop. 1.31(3)) is:

Lemma 4.1. *Let $\widetilde{H}$ be a distribution function of type $\Gamma(\widetilde{\psi})$ and w a differentiable function with $w' \in \mathrm{RV}_\beta(\infty)$, $\beta > -1$. Then the composition $\widetilde{H} \circ w$ is of type $\Gamma((\widetilde{\psi} \circ w)/w')$.*

Under Assumption $\widetilde{2}_{\mathrm{A}}$(ii) resp. $\widetilde{2}_{\mathrm{B}}$(ii) for (4.14), this lemma with $w_t(r) = r + \delta_t(r)$, i.e. $w_t' \in \mathrm{RV}_0(\infty)$, shows that all H_t are of type $\Gamma(\psi_t)$ with

$$\psi_t(r) = \frac{\widetilde{\psi}(r + \delta_t(r))}{1 + \delta_t'(r)} \sim \widetilde{\psi}(r + \delta_t(r)). \tag{4.15}$$

With part (i) of Assumption $\widetilde{2}_{\mathrm{A}}$ resp. $\widetilde{2}_{\mathrm{B}}$ this gives:

Proposition 4.1. *Under Assumption $\widetilde{2}_A$ resp. $\widetilde{2}_B$, the conditional distribution functions H_t defined in (4.13) are of type $\Gamma(\widetilde{\psi})$ for all $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.*

Proof. Under Assumption $\widetilde{2}_{\mathrm{A}}$ it follows with Remark 4.1(i) and $\lambda_t(r) := 1 + \delta_t(r)/r \to 1$ for $r \to \infty$:

$$\frac{\widetilde{\psi}(r + \delta_t(r))}{\widetilde{\psi}(r)} = \frac{\widetilde{\psi}(\lambda_t(r) \cdot r)}{\widetilde{\psi}(r)} \to 1.$$

Under Assumption $\widetilde{2}_{\mathrm{B}}$ it follows for $r \to \infty$:

$$\begin{aligned} \frac{\widetilde{\psi}(r + \delta_t(r))}{\widetilde{\psi}(r)} &= \exp\left(\ln \widetilde{\psi}(r + \delta_t(r)) - \ln \widetilde{\psi}(r)\right) \\ &\sim \exp\left(\delta_t(r)(\ln \widetilde{\psi}(r))'\right) = \exp\left(\delta_t(r) \cdot \widetilde{\psi}'(r)/\widetilde{\psi}(r)\right) \to 1. \end{aligned}$$

Consequently, in both cases we have $\psi_t(r) \sim \widetilde{\psi}(r)$ for all $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and $r \to \infty$. Remark 4.2(i) yields that $\widetilde{\psi}$ is an auxiliary function for all H_t. $\square$

Example 4.1. We consider $\widetilde{\psi}(r) = f(r)\exp(-\gamma r^\tau)$ for $\gamma, \tau > 0$ with $f' \in \mathrm{RV}_{\alpha-1}(\infty)$ and $\alpha \in \mathbb{R}$ as an example for a not regularly varying auxiliary function.
Let $\delta_t'(r) \to 0$ for $r \to \infty$. With $\lambda_t(r) := 1 + \delta_t(r)/r$ and (4.15) one obtains:

$$\frac{\psi_t(r)}{\widetilde{\psi}(r)} \sim \frac{\widetilde{\psi}(\lambda_t(r)\cdot r)}{\widetilde{\psi}(r)} = \frac{f(\lambda_t(r)\cdot r)}{f(r)} \cdot \exp\left(-\gamma r^\tau \cdot [(\lambda_t(r))^\tau - 1]\right).$$

It holds

$$\psi_t(r) \sim \widetilde{\psi}(r) \quad \Leftrightarrow \quad -\gamma r^\tau \cdot ((\lambda_t(r))^\tau - 1) = -\gamma r^\tau \tau \frac{\delta_t(r)}{r} + o\left(\frac{\delta_t(r)}{r}\right) \to 0,$$

or equivalently, $\delta_t(r) = o(r^{1-\tau})$, i.e. the bounding condition (i) for δ_t in Assumption $\widetilde{2}_{\mathrm{B}}$ is *just* fulfilled:

$$\delta_t(r) \cdot \widetilde{\psi}'(r)/\widetilde{\psi}(r) = \delta_t(r) \cdot \left(f'(r)/f(r) - \gamma\tau r^{\tau-1}\right) \to 0.$$

Thus, weakening Assumption $\widetilde{2}_{\mathrm{B}}$ is not possible, as it would violate $\psi_t \sim \widetilde{\psi}$. $\diamondsuit$

Lemma 4.2. *Under Assumption $\widetilde{2}_A$(ii) resp. $\widetilde{2}_B$(ii) for (4.14), the functions $a_t(r)$ from the von Mises representation of $H_t(r)$ converge to $\widetilde{a}$ of $\widetilde{H}$ uniformly in t for $r \to \infty$.*

Proof. Assumption $\widetilde{2}_{\mathrm{A}}$(ii) resp. $\widetilde{2}_{\mathrm{B}}$(ii) implies the existence of $D := \max(|\delta_t'(r)|)$ as well as of a monotonic sequence $\xi_n \to \infty$, such that for all t and $r > \xi_n$ it holds: $|\delta_t'(r)| \leq D2^{-n}$, hence $|\delta_t(r)| \leq |\delta_t(\xi_n)| + D2^{-n} \cdot (r - \xi_n)$. Thus, all graphs of $|\delta_t|$ lie below a polygon of lines between ξ_n and ξ_{n+1} with slopes $D2^{-n} \to 0$ and, consequently, $\delta_t(r)/r \to 0$ uniformly in t for $r \to \infty$. Hence for a_t and $\widetilde{a}$ from the von Mises representations of H_t resp. $\widetilde{H}$ it holds: $a_t(r) = \widetilde{a}(r(1 + \delta_t(r)/r)) \to \widetilde{a}$ uniformly in t for $r \to \infty$. $\square$

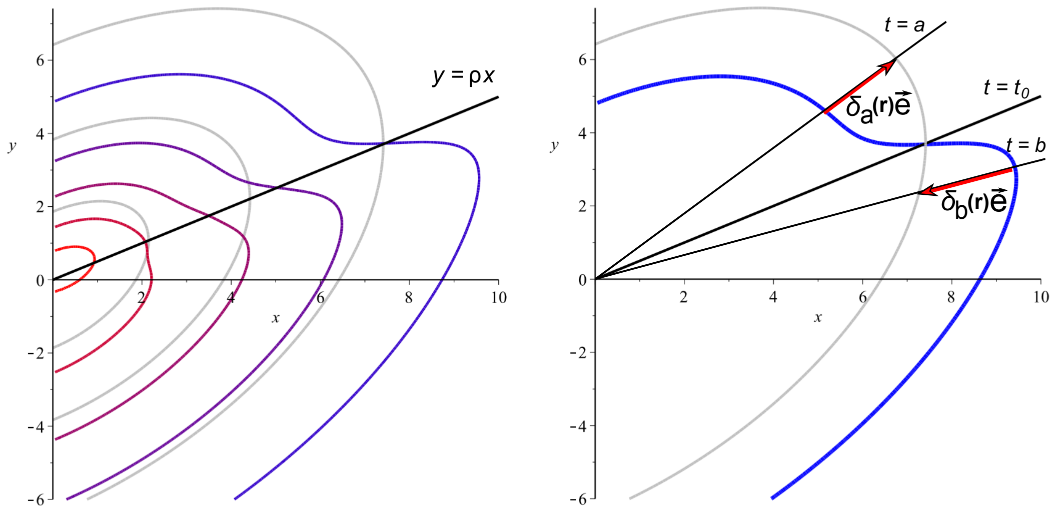
Proposition 4.1 and Lemma 4.2 show that for $\delta_t(r)$ fulfilling Assumption $\widetilde{2}_{\mathrm{A}}$ resp. $\widetilde{2}_{\mathrm{B}}$ we can apply Theorem 4.2 and obtain:

Theorem 4.3 (Dependent polar components, vanishing relative shifts)**.**
Let the reference model $(\widetilde{X}, \widetilde{Y})$ fulfill Assumptions 1 and 2 (with $f_{\widetilde{R}T}$ and $\widetilde{\psi}$), and let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumptions 1, $\widetilde{2}_A$ resp. $\widetilde{2}_B$, and 3. Then the limit statements (4.11) and (4.12) hold.

Example 4.2. We start with the elliptical normal distribution with correlation $\rho = 0.5$ as the reference model $(\widetilde{X}, \widetilde{Y})$, i.e. $\widetilde{H}(r) = 1 - \exp(-r^2/2)$. We choose the shifts:

$$\delta_t(r) = \sqrt{\frac{r^3 \sin^2(t)}{1 + r^2 \sin^2(t)}} = \sqrt{r} \cdot \sin(\arctan(r \sin(t)))$$

fulfilling the assumptions of Theorem 4.3. As $t_0 = 0$ and $\delta_0(r) = 0$, $\widetilde{H}$ coincides with H_0. Figure 4.2(a) compares the level lines of the joint density f_{XY} of (X, Y) with the dependent polar components to those of the reference density $f_{\widetilde{XY}}$ (gray ellipses) of $(\widetilde{X}, \widetilde{Y})$. Along some t-rays, the distance between the level lines of f_{XY} and $f_{\widetilde{XY}}$ becomes arbitrarily large as $\delta_t(r) \to \infty$, which is *not* excluded by $\delta_t'(r) \to 0$. We can also see that the level lines do not possess their maximal x-values on the ray $y = \rho x$ any longer, not even along any other fixed ray. However, Theorem 4.3 verifies that the limit results (4.11) and (4.12) hold with unchanged ρ, which means that the asymptotic behavior of (X, Y) is still determined by an arbitrarily small sector around the ray $\{y = \rho x\} = \{t = t_0\}$. $\diamondsuit$



(a) (b)

Figure 4.2: For Example 4.2: level lines of the density of (X, Y) illustrating (a) Theorem 4.3 and (b) Corollary 4.2(i).

Now we generalize Theorem 4.3 by considering relative shifts $\delta'_t(r)$ which do not vanish asymptotically but tend to a t-dependent limit, i.e. the auxiliary functions ψ_t are no longer asymptotically equivalent. For this purpose, we construct another polar representation for (X, Y) with a new radial component R^* counterbalancing the T-dependence and make the following assumptions.

Assumption 2*. (i) $\widetilde{\psi} \in \mathrm{RV}_\alpha(\infty)$, $\alpha < 1$; (ii) $\delta'_t(r) \to c(t) - 1$ uniformly in t for $r \to \infty$ with some function $c : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to (0, \infty)$.

Assumption 3*. Let $u : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [0, 1]$, $v : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to \mathbb{R}$ be such that $u^* := u/c$ and $v^* := v/c$ fulfill Assumption 3 (with corresponding quantities $u^*_{\max}$, l^*, t^*_0, κ^*, ρ^*).

Proposition 4.2. *Under Assumption 2* for (4.14), the conditional distribution functions H_t are of type $\Gamma(\psi_t)$ with*

$$\psi_t(r) \sim c^{\alpha-1}(t) \cdot \widetilde{\psi}(r) \textit{ for } r \to \infty.$$

Proof. With Lemma 4.1 and $\lambda_t(r) := 1 + \delta_t(r)/r \to c(t)$ for $r \to \infty$ it follows:

$$\psi_t(r) = \frac{\widetilde{\psi}(r + \delta_t(r))}{1 + \delta'_t(r)} \sim \frac{\widetilde{\psi}(\lambda_t(r) \cdot r)}{\widetilde{\psi}(r)} \cdot \frac{\widetilde{\psi}(r)}{c(t)} \sim c^\alpha(t) \cdot \frac{\widetilde{\psi}(r)}{c(t)}. \qquad \square$$

Remark 4.5. Assumption 2* is the counterpart to Assumption $\widetilde{2}_{\mathrm{A}}$ for $\widetilde{\psi} \in \mathrm{RV}_\alpha(\infty)$. There exists no analogue to Assumption $\widetilde{2}_{\mathrm{B}}$ for $\widetilde{\psi} \notin \mathrm{RV}_\alpha(\infty)$ because: If $\delta'_t \to c(t) - 1$ with a continuous non-constant $c(t)$ and it holds $\psi_t \sim a(t)\widetilde{\psi}$, $a(t) \in (0, \infty)$ as in Proposition 4.2, then $\widetilde{\psi}$ has to be regularly varying.

This can be shown by the following argument: We have $\delta_t(r) = (c(t) - 1) \cdot r + o(r)$, hence with Lemma 4.1 it follows: $\psi_t(r)/\widetilde{\psi}(r) \sim (1/c(t)) \cdot \widetilde{\psi}(c(t) \cdot r)/\widetilde{\psi}(r) \sim a(t)$. Remark 4.1(iv) implies that $\widetilde{\psi}$ is regularly varying.

What happens for not regularly varying $\widetilde{\psi}$ can be seen in this example:
for $\widetilde{H}(r) = 1 - \exp(1 - \mathrm{e}^r)$, $\widetilde{\psi}(r) = \mathrm{e}^{-r}$, the quotient

$$\widetilde{\psi}(c(t)r)/\widetilde{\psi}(r) = \exp(-(c(t) - 1)r)$$

tends to zero or to infinity for $c(t) \neq 1$. $\diamondsuit$

Now we construct for $R{\cdot}(u(T), v(T))$ a new polar representation $R^*{\cdot}(u^*(T), v^*(T))$ keeping the angular component T but changing the radial component as follows:

Proposition 4.3. *Let be $R^* = c(t) \cdot R$ for $T = t$. Then, under Assumptions 2* and 3* the distributions*

$$H_t^*(r) = P(R^* \leq r \mid T = t)$$

are of type $\Gamma(\widetilde{\psi})$ for all $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Proof. Lemma 4.1 with $w_t(r) = r/c(t)$ yields that $H_t^* \in \Gamma(\psi_t^*)$ with auxiliary functions $\psi_t^*(r) = c(t)\psi_t\left(r/c(t)\right)$. Under Assumptions 2*, 3* and with Proposition 4.2 we have: $\psi_t^*(r)/\widetilde{\psi}(r) = c(t)c^{\alpha-1}(t) \cdot \widetilde{\psi}(r/c(t))/\widetilde{\psi}(r) \sim c^{\alpha}(t)c^{-\alpha}(t) = 1$. Thus, all $H_t^* \in \Gamma(\widetilde{\psi})$. □

Lemma 4.3. *Under Assumption 2*(ii) for (4.14), the functions $a_t^*(r)$ from the von Mises representation of $H_t^*(r)$ converge to $\widetilde{a}$ of $\widetilde{H}$ uniformly in t for $r \to \infty$.*

Proof. $H_t^*(r) = H_t(r/c(t))$ implies that $\delta_t^*(r) = (1/c(t) - 1)r + \delta_t(r/c(t))$. We decompose $\delta_t(r) = {}_1\delta_t(r) + {}_2\delta_t(r)$ with ${}_1\delta_t(r) = (c(t) - 1)r$ and ${}_2\delta_t'(r) \to 0$. Then Lemma 4.2 implies that both $\delta_t(r)/r \to c(t) - 1$, $\delta_t^*(r)/r \to 0$ and, hence, $a_t^*(r) \to \widetilde{a}$ uniformly in t for $r \to \infty$. □

Proposition 4.3 and Lemma 4.3 show that for $(X, Y) = R^* \cdot (u^*(T), v^*(T))$ we can apply Theorem 4.2 and obtain:

Theorem 4.4 (Dependent polar components, finite relative shifts)**.**
Let the reference model $(\widetilde{X}, \widetilde{Y})$ fulfill Assumptions 1 and 2 (with $f_{\widetilde{R}T}$ and $\widetilde{\psi}$), and let $(X, Y) = R \cdot (u(T), v(T))$ satisfy Assumption 1, 2, and 3*. Then the limit statements (4.11) and (4.12) hold with ρ^*, κ^* and l^*.*

In the situation of Theorem 4.4, we demand Assumption 3 only for the *new* coordinate functions u^* and v^* so that the limit statements depend on their parameters.

Given a polar random vector as in Section 4.3 (fulfilling Assumption 1, 2, 3) for the original u and v, how much can we deviate from the independence of R and T such that the *original* limit statements remain valid? The next corollary gives an answer.

Corollary 4.1. *Let $(X, Y) = R\cdot(u(T), v(T))$ have coordinate functions u and v fulfilling Assumption 3 and dependent polar components fulfilling Assumptions 1 and 2* with $(c-1) \in \mathrm{RV}^s_\beta(t_0)$ for some $\beta > \kappa$ and $u_{\max} = u^*_{\max}$. Then, the limit statements (4.11) and (4.12) hold with the original parameters ρ and κ.*

Proof. The condition $(c-1) \in \mathrm{RV}^s_\beta(t_0)$ implies that $(1-1/c) \in \mathrm{RV}^s_\beta(t_0)$. Hence:

$$l^* = u^*_{\max} - u^* = u^*_{\max} - \frac{u}{c} = \left(u_{\max} - u + \left(1 - \frac{1}{c}\right) \cdot u\right) \in \mathrm{RV}^s_\kappa(t_0). \qquad \square$$

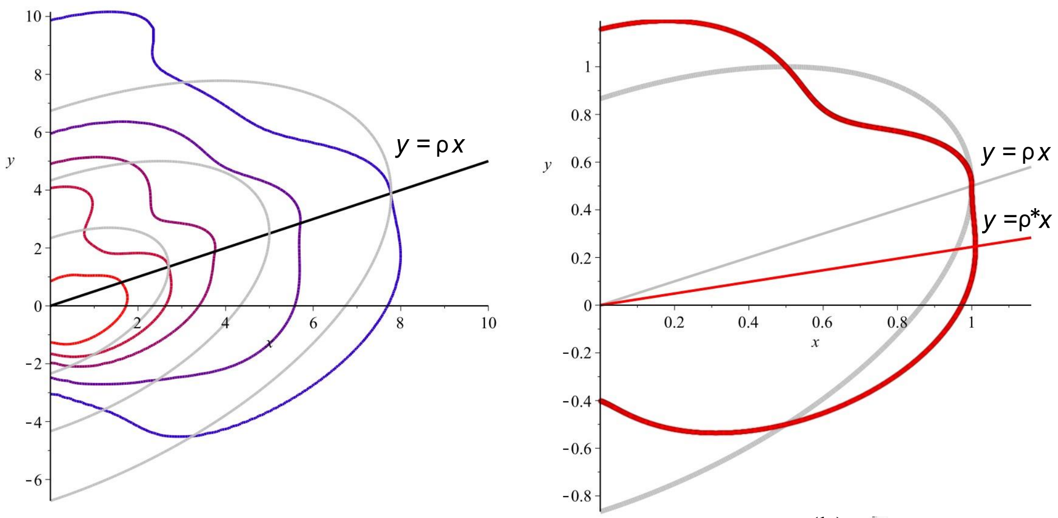
Example 4.3. Here we present an example for a polar random vector (X, Y) according to Theorem 4.4 and choose the same reference model as in Example 4.2, but now with

$$\delta_t(r) = r \cdot (c(t) - 1) + \frac{\sin(r) \cdot t}{\sqrt[4]{r+2}}, \qquad c(t) = \frac{t}{2}\sin(-|3t|) + 1.$$

Since $(c-1) \in \mathrm{RV}^s_1(0)$, the criterion of Corollary 4.1 is not fulfilled. The limit statements of Theorem 4.4 hold for $\kappa^* = 2$, $u^*_{\max} = 1.00995$, $t^*_0 = -0.28984$, $\rho^* = 0.24413$. Figure 4.3(a) displays – analogously to Figure 4.2(a) – level lines of the joint density f_{XY} and those of the reference density $f_{\widetilde{XY}}$ (gray ellipses). In Figure 4.3(b) the curve $(u^*(t), v^*(t))$ (red line) is contrasted to the reference curve $(u(t), v(t))$ (gray line). Note that the reference curve coincides with one of the level lines of $f_{\widetilde{XY}}$, while the curve $(u^*(t), v^*(t))$ in general does not coincide with any level line of f_{XY}. $\diamondsuit$

The functions δ_t introduced as the shifts between the distributions H_t of R and $\widetilde{H}$ of $\widetilde{R}$ also provides information about the density f_{XY} of (X, Y). We compare:

$$\begin{aligned} f_{XY}&: \quad \text{density of } (X, Y) = R \cdot (u(T), v(T)) \\ f_{\widetilde{XY}}&: \quad \text{density of } (\widetilde{X}, \widetilde{Y}) = \widetilde{R} \cdot (u(T), v(T)). \end{aligned} \tag{4.16}$$



(a) (b)

Figure 4.3: For Example 4.3: (a) level lines of the density of (X, Y) illustrating Theorem 4.4, (b) curve $(u^*(t), v^*(t))$ (red) in contrast to $(u(t), v(t))$ (gray).

An interpretation of statement (i) in the following corollary is given in Figure 4.2(b): In the context of Theorem 4.3, $\delta_t(r)$ displays the asymptotic distance between the level lines of f_{XY} and $f_{\widetilde{XY}}$ measured by the radial component R (not by the Euclidean distance).

Corollary 4.2.

(i) *Let (X, Y) and $(\widetilde{X}, \widetilde{Y})$ fulfill the assumptions of Theorem 4.3. Then for $r \to \infty$ it holds*

$$f_{XY}\big(r \cdot (u(t), v(t))\big) \sim f_{\widetilde{XY}}\big((r + \delta_t(r)) \cdot (u(t), v(t))\big).$$

(ii) *Analogously, under the assumptions of Theorem 4.4 it holds for $r \to \infty$*

$$\begin{aligned} f_{XY}\big(r \cdot (u(t), v(t))\big) \quad &\sim \quad c^2(t) \cdot f_{\widetilde{XY}}\big((r + \delta_t(r)) \cdot (u(t), v(t))\big) \\ &= f_{\widetilde{XY}}\big((r + \delta_t(r)) \cdot (u^*(t), v^*(t))\big). \end{aligned}$$

Examples 4.2 and 4.3 point out the meaning of Theorems 4.3 and 4.4 for given shifts $\delta_t(r)$. But in concrete situations, one usually starts from the (estimated) distributions and calculates the shifts, which is considered in the following example.

Example 4.4. We start with two distributions, the normal distribution with t-dependent variance σ_t^2 and a distribution with a quicker tail decay. The first one has a survival function ${}_1\overline{H}_t(r) = \exp\left(-r^2/(2\sigma_t^2)\right)$ where σ_t is continuous in t, the second one has a survival function ${}_2\overline{H}_t(r) = \exp\left(-f_t(r)\right)$ where $f : [-1,1] \times (0,\infty) \to [0,\infty)$ is continuous in t and differentiable in r fulfilling $f_t(0) = 0$, $f_t'(r) > 0$ and $f_t(r)/r^2 \to \infty$ for $r \to \infty$. We consider their mixture

$$H_t(r) = 1 - a(t)\exp\left(-\frac{r^2}{2\sigma_t^2}\right) - b(t)\exp(-f_t(r)),$$

where a and b are positive continuous functions with $a + b = 1$.

The natural choice of the reference model is $\widetilde{H}(r) = 1 - \exp(-r^2/2)$, and we derive the corresponding shifts from $\widetilde{H}(r + \delta_t(r)) = H_t(r)$:

$$-(r + \delta_t(r))^2/2 = \ln \overline{H}_t(r) \quad \Rightarrow \quad \delta_t(r) = -r + \sqrt{-2\ln \overline{H}_t(r)} \geq -r. \qquad (4.17)$$

For $r \to \infty$ it holds

$$\begin{aligned}
\overline{H}_t(r) &= \exp\left(-\frac{r^2}{2\sigma_t^2} + \ln a(t)\right)\left[1 + \exp\left(-f_t(r) + \frac{r^2}{2\sigma_t^2} + \ln\left(\frac{b(t)}{a(t)}\right)\right)\right] \\
&= \exp\left(-\frac{r^2}{2\sigma_t^2} + \ln a(t)\right) \cdot \left[1 + \frac{b(t)}{a(t)} \cdot \exp\left(-r^2\left(\frac{f_t(r)}{r^2} - \frac{1}{2\sigma_t^2}\right)\right)\right] \\
&\sim \exp\left(-\frac{r^2}{2\sigma_t^2} + \ln a(t)\right)
\end{aligned}$$

and, hence, with (4.17) we have

$$\delta_t(r) \;\sim\; -r + \sqrt{r^2/\sigma_t^2 - 2\ln a(t)} \;\sim\; r \cdot \left(-1 + \left(1 - \sigma_t^2 \ln a(t)/r^2\right)/\sigma_t\right).$$

For $\sigma_t^2 = 1$ the shifts $\delta_t(r) \sim -\ln a(t)/r$ fulfill the assumptions from Theorem 4.3. For $\sigma_t^2 \neq 1$ it holds $\delta_t(r) \sim r \cdot (1 - \sigma_t)/\sigma_t$ according to Theorem 4.4 with $c(t) = 1/\sigma_t$. $\diamondsuit$

Remark 4.6. The result in Example 4.4 is as expected: The component of distribution with the quicker tail decay is asymptotically negligible, the t-dependence of the variance can be removed with a change of the radial component. Theorems 4.2, 4.3 and 4.4 give a safe mathematical basis for such plausibility arguments. $\diamondsuit$

Remark 4.7. If $\delta_t' \in \mathrm{RV}_\beta(\infty)$ with $\beta > 0$ then no scale transformation of R permits to apply Theorem 4.2 as Lemma 4.1 shows. If $\delta_t' \notin \mathrm{RV}_\beta(\infty)$ there exist cases with $\psi_t \sim \psi$ for all t although $\delta_t'(r)$ does not converges to zero, as the following example shows: $\overline{H}_t(r) = \exp(-r - t\sin(r)/\pi)$ are of type $\Gamma(1)$ for all t, but the shifts with respect to $\widetilde{H} = H_0$ yield $\delta_t'(r) = t\cos(r)/\pi$ not converging to zero or to some other limit. $\diamondsuit$

4.5 Comparison with Balkema and Embrechts' approach

Now we contrast our geometric approach for weakening the independence assumption to those of Balkema and Embrechts (2007). They do not explicitly use polar representations, however, in the bivariate case their model can be reformulated in terms of a polar representation for (X, Y) with $\kappa = 2$ and $R \in \Gamma(\psi)$. Balkema and Embrechts (2007) transfer the limit result in Theorem 9.1, p. 137 for some random vector with density $f_{\widetilde{XY}}$ to another one in Theorem 11.2, p. 158 with density

$$f_{XY}(ru(t), rv(t)) = Q(r,t) \cdot f_{\widetilde{XY}}(ru(t), rv(t)), \tag{4.18}$$

where Q is in $\mathcal{L}$, meaning that

$$\mathcal{Q}(r,t) := \frac{Q(r + r_0, t)}{Q(r,t)}, \qquad \lim_{r\to\infty} \mathcal{Q}(r,t) = 1 \tag{4.19}$$

for all $r_0 > 0$ and all t. Under Assumption 1, Q is also the quotient $f_{RT}/f_{\widetilde{R}T}$, cf. (4.9).

Our results extends those of Balkema and Embrechts (2007) for the bivariate case in the sense that we consider more general polar distributions with arbitrary $\kappa > 0$. But even for $\kappa = 2$, we cover cases not included by Balkema and Embrechts, i.e. with $\lim_{r\to\infty} \mathcal{Q}(r,t) \neq 1$, as it is shown in the following theorem.

Theorem 4.5 (Ratio of densities)**.**

Let (X,Y) fulfill the assumptions of Theorem 4.4, then for Q from (4.18) we get the following results depending on the variation index α of the auxiliary function $\widetilde{\psi}$ and the values of the function c from Assumption 2*:

(a) for $\alpha < 0$ (i.e. $\widetilde{\psi}(r) \to 0$) it holds:

$$\lim_{r\to\infty} \mathcal{Q}(r,t) = \begin{cases} 0, & \text{for } t \text{ with } c(t) < 1 \\ 1, & \text{for } t \text{ with } c(t) = 1 \\ \infty, & \text{for } t \text{ with } c(t) > 1\,; \end{cases}$$

(b) for $\alpha > 0$ (i.e. $\widetilde{\psi}(r) \to \infty$) it holds:

$$\lim_{r\to\infty} \mathcal{Q}(r,t) = 1\,;$$

(c) for $\alpha = 0$:

- *if $\widetilde{\psi}(r) \to 0$, then (a) holds,*
- *if $\widetilde{\psi}(r) \to \infty$, then (b) holds,*
- *if $\widetilde{\psi}(r) \to k \in (0,\infty)$, then it holds:*
  $\lim_{r\to\infty} \mathcal{Q}(r,t) = \exp((1-1/c(t))\cdot r_0/k)$ *for t with $c(t) \neq 1$ and*
  $\lim_{r\to\infty} \mathcal{Q}(r,t) = 1$ *for t with $c(t) = 1$.*

Proof. With (4.9) we get:

$$Q(r,t) = \frac{f_{XY}(ru(t),rv(t))}{f_{\widetilde{XY}}(ru(t),rv(t))} = \frac{f_{RT}(r,t)}{f_{\widetilde{R}T}(r,t)} = \frac{h_t(r)}{\widetilde{h}(r)} = (1+\delta_t'(r))\cdot\frac{\widetilde{h}(r+\delta_t(r))}{\widetilde{h}(r)}.$$

Hence, we have:

$$\begin{aligned} \mathcal{Q}(r,t) &= \frac{1+\delta_t'(r+r_0)}{1+\delta_t'(r)}\frac{\widetilde{h}(r+r_0+\delta_t(r+r_0))}{\widetilde{h}(r+r_0)}\frac{\widetilde{h}(r)}{\widetilde{h}(r+\delta_t(r))} \\ &\sim \frac{\widetilde{h}(r)}{\widetilde{h}(r+\delta_t(r))}\frac{\widetilde{h}(r+r_0+\delta_t(r+r_0))}{\widetilde{h}(r+r_0)} \end{aligned}$$

$$\sim \exp\left(\int_r^{r+\delta_t(r)} 1/\widetilde{\psi}(u)\,\mathrm{d}u\right) \cdot \exp\left(-\int_{r+r_0}^{r+r_0+\delta_t(r+r_0)} 1/\widetilde{\psi}(u)\,\mathrm{d}u\right). \quad (4.20)$$

For the last step, we exploit that $1/(1-\widetilde{H})$ and $1/\widetilde{h}$ are $\Gamma(\widetilde{\psi})$-varying due to the assumed monotony of $\widetilde{h}$ (with l'Hospital in (4.4), $\widetilde{\psi}' \to 0$). Thus, one can apply the von Mises representation (4.5) for $\widetilde{h}$. As $\widetilde{\psi} \in \mathrm{RV}_\alpha(\infty)$ with $\alpha < 1$ (Assumption 2*), it holds

$$S(x) := \int_0^x 1/\widetilde{\psi}(u)\mathrm{d}u \in \mathrm{RV}_{1-\alpha}(\infty)$$

with a positive variation index, hence:

$$\begin{aligned} \exp\left(\int_r^{r+\delta_t(r)} 1/\widetilde{\psi}(u)\,\mathrm{d}u\right) &= \exp\left(S(r)\cdot\left(\frac{S((1+\delta_t(r)/r)\cdot r)}{S(r)} - 1\right)\right) \\ &\sim \exp\left(S(r)\cdot(c^{1-\alpha}(t)-1)\right). \end{aligned}$$

Inserting the last expression in (4.20) we finally get:

$$\begin{aligned} \mathcal{Q}(r,t) &\sim \exp\left((c^{1-\alpha}(t)-1)\cdot(S(r)-S(r+r_0))\right) \\ &= \left(\exp\left(c^{1-\alpha}(t)-1\right)\right)^{S(r)-S(r+r_0)}. \end{aligned} \quad (4.21)$$

The limit of (4.21) results from the behavior of S: for $\alpha < 0$ the variation index of S is $1-\alpha > 1$, for $\alpha > 0$ it is $1-\alpha < 1$, for $\alpha = 0$ it is $1-\alpha = 1$. □

Remark 4.8.

(i) For the situation considered in Theorem 4.3 with $\delta_t' \to 0$, i.e. $c \equiv 1$, the density quotient Q is in $\mathcal{L}$ as in Balkema and Embrechts (2007).

(ii) For the more general situation considered in Theorem 4.4, the case distinction in Theorem 4.5 corresponds to the tail behavior of $\widetilde{H}$:
If $1-\widetilde{H}$ decreases at least exponentially fast ("light tail"), we are in case (a) or (c) and generally $Q \notin \mathcal{L}$.
If $1-\widetilde{H}$ decreases slower than any exponential but faster than any power function ("mildly heavy tail"), we are in case (b) and $Q \in \mathcal{L}$.

(iii) In Example 4.4 (for $\sigma_t^2 \neq 1$), the corresponding Q is not in $\mathcal{L}$, the quotient $\mathcal{Q}(r,t)$ tends to zero if $\sigma_t^2 < 1$ resp. to infinity if $\sigma_t^2 > 1$. Thus, this example is not covered by the theorem of Balkema and Embrechts (2007).

$\diamondsuit$

Our approach to measure the dependence between the polar components R and T with the shifts $\delta_t(r)$ is intuitive and bases primarily on distributions and not on densities. Both criteria for $\delta_t(r)$ provided in Theorem 4.3 resp. 4.4 possess a geometrical interpretation, while Balkema and Embrechts (2007) “warn the reader that functions from the class $\mathcal{L}$ are not as tame as they may seem” (p. 158).

To sum it up, describing random vectors using a polar representation $R\cdot(u(T), v(T))$ permits a lot of freedom in modeling the asymptotic behavior as it requires only weak and local assumptions on u and v. Allowing a certain dependence between R and T, we show validity of the limit results, which is of importance for empirical applications.

4.6 Proof of Theorem 4.2

To prove Theorem 4.2 we follow the strategy of the proof for independent R and T, cf. Section 2.7.1. The additionally required steps are presented in the following.
Under the assumptions of Theorem 4.2 for any $p > 0$, $\alpha > 1$ there exist t-independent constants C_p, $C_{p,\alpha}$ such that for x large enough and $z \geq 0$:

$$\frac{\overline{H}_t(x + z\psi(x))}{\overline{H}_t(x)} \leq C_p \cdot (1+z)^{-p}, \qquad \frac{\overline{H}_t(\alpha x)}{\overline{H}_t(x)} \leq C_{p,\alpha} \cdot (\psi(x)/x)^p, \tag{4.22}$$

and for $\breve{k} \in \mathrm{RV}_\alpha(0)$ with $\alpha > -1$ bounded on compact subsets of $(0, \infty]$ it holds locally uniformly in $d \geq 0$:

$$\lim_{x\to\infty} \int_d^\infty \frac{\overline{H}_{t(z)}(x + z\psi(x))}{\overline{H}_{t(z)}(x)} \cdot \frac{\breve{k}(z\psi(x)/x)}{\breve{k}(\psi(x)/x)} \,\mathrm{d}z = \int_d^\infty \mathrm{e}^{-z} z^\alpha \,\mathrm{d}z. \tag{4.23}$$

For $\overline{H}$ instead of $\overline{H}_t$, this is shown in Fougères and Soulier (2010) as lemma 5.1 and 5.2. So it remains to prove that the constants can be chosen independent of t.

The condition $a_t(x) \to a_t$ uniformly in t implies that for all $A > 1$ and x large enough it holds $1/\sqrt{A} \leq a_t(x)/a_t \leq \sqrt{A}$ and $1/\sqrt{A} \leq a_t(x + z\psi(x))/a_t \leq \sqrt{A}$. Thus, we have:

$$\frac{1}{A} \leq \frac{\overline{H}_t(x+z\psi(x))}{\overline{H}_t(x)} \cdot \exp\left(\int_x^{x+z\psi(x)} \frac{1}{\psi(u)}\mathrm{d}u\right) = \frac{a_t(x+z\psi(x))}{a_t(x)} \leq A. \tag{4.24}$$

Now we sketch the proof of Theorem 4.2. The probability of a set $B_{xy} := \{X > x, Y > y\}$ with $y > 0$ is calculated by integrating the conditional survival function $\overline{H}_t$ over the boundary of B_{xy} parameterized by $t \in (t_-, \pi/2)$, where $t_- := \tau^{-1}(0)$ with τ from Assumption 1(iii) and g denoting the continuous and positive marginal density of T:

$$P\{X > x, Y > y\} = \int_{t_-}^{\pi/2} \overline{H}_t\left(\max\left(\frac{x}{u(t)}, \frac{y}{v(t)}\right)\right) g(t)\mathrm{d}t \tag{4.25}$$

$$= \int_{t_0-\epsilon}^{t_1(x)} \overline{H}_t\left(\frac{y}{v(t)}\right) g(t)\mathrm{d}t + \int_{t_1(x)}^{t_0+\epsilon} \overline{H}_t\left(\frac{x}{u(t)}\right) g(t)\mathrm{d}t + \int_{|t-t_0|>\epsilon} \overline{H}_t\left(\max\left(\frac{x}{u(t)}, \frac{y}{v(t)}\right)\right) g(t)\mathrm{d}t$$

$$=: I(x) + J(x) + rem(x)$$

with an arbitrary $\epsilon \in (0, t_-)$, $y = \rho x + xk(\psi(x)/x)\zeta$ with k from (4.11), and $t_1(x) := (v/u)^{-1}(y/x) = \arctan(\rho + k(\psi(x)/x)\zeta) \to \arctan(\rho) = t_0$ for $x \to \infty$.

We treat the case $\rho > 0, t_0 \leq t_1(x) < t_0 + \epsilon$, from which the other cases can be deduced as in the proof of Theorem 2.1 in Section 2.7.1. For dependent R and T the Mean Value Theorem is required:

$$J(x) = \overline{H}_{q(x)}(x) \int_{t_1(x)}^{t_0+\epsilon} \left(\overline{H}_t\left(x/u(t)\right)/\overline{H}_t(x)\right) g(t)\;\mathrm{d}t =: \overline{H}_{q(x)}(x)L(x) \tag{4.26}$$

for some mean value $q(x) \in (t_1(x), t_0 + \epsilon)$. For $x \to \infty$ it holds $J(x) \sim \overline{H}_{t_0}(x)L(x)$, as we show later, cf. (4.29).

Analogously to the proof presented in Section 2.7.1 (with $\tau = 0$), a substitution $t \mapsto z$ is made such that the argument of $\overline{H}$ becomes $x + z\psi(x)$ and (4.23) can be applied, such that we have for $x \to \infty$:

$$J(x) \sim \overline{H}_{t_0}(x)\,\kappa^{1/\kappa-1}\,k_J(\psi(x)/x)\int\limits_{|\zeta|^\kappa/\kappa}^{\infty} \mathrm{e}^{-z} z^{1/\kappa-1}\mathrm{d}z, \tag{4.27}$$

$$I(x) \sim \overline{H}_{t_0}(x)\,k_I(\psi(x)/x)\,f_\zeta(x)\int\limits_{|\zeta|^\kappa/\kappa}^{\infty} \mathrm{e}^{-z}\mathrm{d}z \tag{4.28}$$

with $k_J \in \mathrm{RV}_{1/\kappa}(0)$ and $k_I \in \mathrm{RV}_{1+1/\kappa}(0)$ for $\rho = 0$ resp. $k_I \in \mathrm{RV}_1(0)$ for $\rho \neq 0$, and some bounded function f_ζ. Since for $\kappa > 1$ as well as for $\rho = 0$ the variation index of k_J is smaller than that of k_I, and it follows $I(x) = o(J(x))$ for $x \to \infty$.

Since u has a unique global maximum 1 at t_0, for all $\epsilon > 0$ exists an $\eta_\epsilon \in (0, 1)$ with $u(t) < 1-\eta_\epsilon$ for all $|t-t_0| > \epsilon$. The Mean Value Theorem for some $|\breve{q}-t_0| > \epsilon$ yields:

$$rem(x) \le \int\limits_{|t-t_0|>\epsilon} \overline{H}_t\left(\frac{x}{u(t)}\right) g(t)\mathrm{d}t \le \overline{H}_{\breve{q}}\left(\frac{x}{1-\eta_\epsilon}\right)\int\limits_{|t-t_0|>\epsilon} g(t)\mathrm{d}t \le \overline{H}_{\breve{q}}\left(\frac{x}{1-\eta_\epsilon}\right).$$

The second statement of (4.22) implies that for $\alpha > 1, p > 0$ and for all q there exists a $C_{q,p,\alpha}$ with $\overline{H}_q(\alpha x)/\overline{H}_{t_0}(x) \le C_{q,p,\alpha}\cdot(\psi(x)/x)^p$. Choosing $\alpha = 1/(1-\eta_\epsilon)$, it follows $rem(x) = o(\overline{H}_{t_0}(x)\cdot(\psi(x)/x)^p)$ for all $p > 0$ and, hence, $rem(x) = o(J(x))$ for $x \to \infty$.

Consequently, for $\kappa > 1$ and $\rho = 0$, J determines the asymptotics of (4.25). The proof of Theorem 4.2 can be completed as in Section 2.7.1, where it is also shown how to deduce the statement (4.12) for random norming from the case $\rho = 0$.

Now we prove $J(x) \sim \overline{H}_{t_0}(x)L(x)$, cf. (4.26). For any $\epsilon_1 \in (0, \epsilon)$ we can decompose:

$$J(x) = \int\limits_{t_1(x)}^{t_0+\epsilon_1} \overline{H}_t\left(\frac{x}{u(t)}\right) g(t)\mathrm{d}t + \int\limits_{t_0+\epsilon_1}^{t_0+\epsilon} \overline{H}_t\left(\frac{x}{u(t)}\right) g(t)\mathrm{d}t =: J_1(x) + J_2(x). \tag{4.29}$$

Analogously to $rem(x) = o(J(x))$ as proved above, it follows that $J_2(x) = o(J_1(x))$ and, hence, $J(x) \sim J_1(x) = \overline{H}_{q_1(x)}(x)L(x)$ for some mean value $q_1(x) \in (t_0, t_0 + \epsilon_1)$. Since this holds for any arbitrarily small ϵ_1, it follows $J(x) \sim \overline{H}_{t_0}(x)L(x)$. □

Bibliography

Abdous, B., Fougères, A.-L., Ghoudi, K.: Extreme behavior for bivariate elliptical distributions. *Can. J. Stat.* 33(3), 317-334 (2005)

Balkema, G., Embrechts, P.: *High risk scenarios and extremes. A geometric approach.* Zurich Lectures in Advanced Mathematics. Zurich: European Mathematical Society (2007)

Balkema, G., Embrechts, P., Nolde, N.: The shape of asymptotic dependence. Springer Proceedings in Mathematics and Statistics, special volume. *Prokhorov and Contemporary Probability Theory* 33, 43-67 (2013)

Balkema, A.A., Haan, de, L.: On R. von Mises' condition for the domain of attraction of $\exp(-\mathrm{e}^{-x})$. *Ann. Math. Statist.* 43(4), 1352-1354 (1972)

Balkema, G., Nolde, N.: Asymptotic independence for unimodal densities. *Adv. Appl. Probab.* 42(2), 411-432 (2010)

Barbe, P.: Approximation of integrals over asymptotic sets with applications to probability and statistics. ArXiv:0312132 (2003)

Barbe, P., Seifert, M.I.: A conditional limit theorem for a bivariate representation of a univariate random variable and conditional extreme values. The early working paper is available at ArXiv:1311.0540 (2013)

Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J.: *Statistics of extremes.* Wiley Series in Probability and Statistics, Chichester (2004)

Berman, S.M.: Sojourns and extremes of Fourier sums and series with random coefficients. *Stoch. Process. Their Appl.* 15(3), 213-238 (1983)

Bingham, N.N., Goldie, C.M., Teugels, J.: *Regular Variation.* Cambridge University Press, Cambridge (1987)

Billingsley, P.: *Probability and Measure.* third ed. Wiley, New York (1995)

Billingsley, P.: *Convergence of Probability Measures.* second ed. Wiley, New York (1999)

Coles, S.G., Tawn, J.A.: Modelling extreme multivariate events. *J. R. Statist. Soc. B* 53, 377-392 (1991)

Das, B., Mitra, A., Resnick, S.I.: Living on the multi-dimensional edge: Seeking hidden risks using regular variation. *Adv. Appl. Probab.* 45(1), 139-163 (2013)

Das, B., Resnick, S.I.: Conditioning on an extreme component: Model consistency and regular variation on cones. *Bernoulli* 17(1), 226-252 (2011a)

Das, B., Resnick, S.I.: Detecting a conditional extreme value model. *Extremes* 14(1), 29-61 (2011b)

Das, B., Resnick, S.I.: Generation and detection of multivariate regular variation and hidden regular variation. ArXiv:1403.5774 (2014)

Demichel, Y., Fermin, A.-K., Soulier, P.: The diameter of an elliptical cloud. *Electron. J. Probab.* 20, no. 27, 1-32 (2015)

Dudley, R.M.: *Real Analysis and Probability.* Cambridge University Press, Cambridge (2002)

Eastoe, E.F., Heffernan, J.E., Tawn, J.A.: Nonparametric estimation of the spectral measure, and associated dependence measures, for multivariate extreme values using a limiting conditional representation. *Extremes* 17(1), 25-43 (2014)

Embrechts, P., Klüppelberg, C., Mikosch, T.: *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin (1997)

Falk, M., Hüsler, J., Reiss, R.-D.: *Laws of Small Numbers: Extremes and Rare Events.* second ed. Birkhäuser, Basel (2004)

Fougères, A.-L., Soulier, P.: Limit conditional distributions for bivariate vectors with polar representation. *Stoch. Models* 26(1), 54-77 (2010)

Fougères, A.-L., Soulier, P.: Estimation of conditional laws given an extreme component. *Extremes* 15(1), 1-34 (2012)

Geluk, J.L., Haan, de, L.: *Regular variation, extensions and Tauberian theorems.* CWI Tract 40. Stichting Mathematisch Centrum voor Wiskunde en Informatica, Amsterdam (1987)

Gnedenko, B.: Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math* 44(3), 423-253 (1943)

Gumbel, E.J.: *Statistics of extremes.* Columbia University Press, New York (1958)

Gumbel, E.J.: *Bivariate exponential distributions. J. Amer. Statist. Assoc.* 55, 698-707 (1960)

Haan, de, L.: *On Regular Variation and its Application to the Weak Convergence of the Sample Extremes.* Mathematisch Centrum, Amsterdam (1970)

Haan, de, L., Ferreira, A.: *Extreme Value Theory. An Introduction.* Springer, New York (2006)

Haan, de, L., Resnick, S.I.: Limit theory for multivariate sample extremes. *Z. Wahrsch. verw. Gebiete* 40, 317-337 (1977)

Hashorva, E.: Gaussian approximation of conditional elliptic random vectors. *Stoch. Models* 22(3), 441-457 (2006)

Hashorva, E.: Asymptotic properties of type I elliptical random vectors. *Extremes* 10(4), 175-206 (2007)

Hashorva, E.: Conditional limiting results for Type I polar distributions. *Extremes* 12(3), 239-263 (2009)

Hashorva, E.: On the residual dependence index of elliptical distributions. *Stat. Probab. Lett.* 80(13-14), 1070-1078 (2010)

Hashorva, E.: Exact tail asymptotics in bivariate scale mixture models. *Extremes* 15(1), 109-128 (2012)

Heffernan, J.E., Resnick, S.I.: Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.* 17(2), 537-571 (2007)

Heffernan, J.E., Tawn, J.A.: A conditional approach for multivariate extreme values. *J. R. Statist. Soc. B* 66(3), 497-546 (2004)

Hult, H., Lindskog, F.: Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv. Appl. Probab.* 34(3), 587-608 (2002).

Klüppelberg, C., Kuhn, K., Peng, L.: Estimating the tail dependence of an elliptical distribution. *Bernoulli* 13(1), 229-251 (2007)

Ledford, A., Tawn, J.: Statistics for near independence in multivariate extreme values. *Biometrika* 83(1), 169-187 (1996)

Ledford, A., Tawn, J.: Modelling dependence within joint tail regions. Statistics for near independence in multivariate extreme values. *J. R. Statist. Soc. B* 59(2), 475-499 (1997)

Ledford, A., Tawn, J.: Concomitant tail behaviour for extremes. *Adv. in Appl. Probab.* 30(1), 197-215 (1998)

Lindskog, F., Resnick, S.I., Roy, J.: Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probab. Surv.* 11, 270-314 (2014)

Manner, H., Segers, J.: Tails of correlation mixtures of elliptical copulas. *Insur. Math. Econ.* 48, 153-160 (2011)

McNeil, A., Frey, R., Embrechts, P.: *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, New Jersey (2005)

Maulik, K., Resnick, S.I.: Characterizations and examples of hidden regular variation. *Extremes* 7(1), 31-67 (2004)

Nolde, N.: *The Analysis of Extremes in Multivariate Models: A Geometric Approach.* PhD dissertation, ETH Zürich (2010)

Resnick, S.I.: *Extreme values, regular variation, and point processes.* Volume 4 of Applied Probability. A Series of the Applied Probability Trust. Springer, New York (1987, Reprint 2008)

Resnick, S.I.: Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* 5(4), 303-336 (2002)

Resnick, S.I.: *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York (2007)

Resnick, S.I., Zeber, D.: Transition kernels and the conditional extreme value model. *Extremes* 17(2), 263-287 (2014)

Seifert, M.I.: On conditional extreme values of random vectors with polar representation. *Extremes* 17(2), 193-219 (2014)

Seifert, M.I.: Weakening the independence assumption on polar components: Limit theorems for generalized elliptical distributions. To appear in *J. Appl. Prob.* (2015)

Tawn, J.A.: Modelling multivariate extreme value distributions. *Biometrika* 77(2), 245-253 (1990)

Yue, S.: A Bivariate Extreme Value Distribution Applied to Flood Frequency Analysis. *Nord. Hydrol.* 32(1), 49-64 (2001)