Multivariate excess distributions

Guus Balkema
KdV Instituut voor Wiskunde
University of Amsterdam
The Netherlands

Paul Embrechts
Department of Mathematics
ETH Zürich
Switzerland

June 11, 2004

Abstract

This paper presents a continuous one parameter family of multivariate generalized Pareto distributions which describe the asymptotic behaviour of exceedances over linear thresholds. The one–dimensional theory has proved to be important in insurance, finance and risk management. The multivariate limit theory presented here is developed with similar applications in mind.

1 Introduction

Alternative titles for our paper could have been "Limit distributions for exceedances over linear thresholds" or "High risk market scenarios". We derive the limit for the conditional distribution of a random vector given that the vector takes values in a remote half space. The latter could for instance be interpreted as a portfolio value reflecting a high loss, hence "scenarios". In other applications the vector would correspond to an index, hence the usage of "market".

Though the basic theory presented in this paper is fairly mathematical, we often use the language of finance in order to highlight potential applications. This reflects our motivation for developing the theory. Other possible applications will be indicated below.

A major concern in financial risk management is that the market drifts off into some undesirable direction. One speaks of high risk to describe a large change in the state of the market within a short time span, a few days say. The desirability of the new state will obviously depend on one's own position. We take a global approach. In our terminology, high risk market behaviour describes the market conditional on it being in some region far out from the present position. For the sake of simplicity we take the regions to be closed half spaces. The market evolves according to some random process. As a first approximation this process is a Brownian motion in some high dimensional space with a given drift vector and non–singular covariance matrix. In this paper we take a non–dynamical approach and consider the state of the market at some fixed future date, say ten days ahead, as is typical for market risk management. The future state is described by a random vector Z in $\mathbf{R}^d$. The dimension d denotes the number of assets and may vary between ten (for a small investor) and several hundred or more for an index. Alternatively, d could correspond to a number of market sectors like energy, communications, biotechnology, food, banking and insurance, say.

Assume for the present that the state Z is described by a unimodal density. This may be a Gaussian density, but it could have fatter tails. For a given risk level, say $\alpha = 0.05$ or 0.01, there are many half spaces H such that $P\{Z \in H\} = \alpha$. Indeed under our assumption of a unimodal density, for any direction θ there exists a unique level q so that the half space $H = \{\theta \geq q\}$ of all vectors z with $\theta z \geq q$ satisfies $P\{Z \in H\} = \alpha$. For this paper, high risk market behaviour describes the market conditional on the event that the state lies in such a half space H. We shall henceforth use the term scenario to indicate a change in the underlying probability measure. In our case the new probability measure is obtained by conditioning Z to lie in the half space H, and we write Z^H for this conditional market scenario. We are interested in the asymptotic behaviour of the distribution of Z^H when the risk level $\alpha = P\{Z \in H\}$ tends to zero. The event $\{Z \in H\}$ may thus be termed rare or extreme. In this setting we make no assumptions on the direction of the half space. It will be shown that there exists a family of canonical limit laws which allows one to describe the asymptotic behaviour of these extreme market scenarios for a large class of densities.

A limit theory of the above type is of importance to handle questions in financial risk management and, for instance, non–life insurance. Examples of the former are discussed in Luethi and Studer [27] and Wüthrich [36]. For an example in insurance, see Juri and Wüthrich [24]. Other fields of application include biology, reliability and quality control.

For each dimension $d > 1$ we shall introduce a one parameter family of standardized multivariate generalized Pareto distributions (GPDs) on the upper half space

$$H_+ = \mathbf{R}^h \times [0, \infty) \qquad h = d - 1\,, \tag{1.1}$$

which extends the well known class of univariate GPDs on the half line $[0, \infty)$. Distributions which are linked by an affine change of coordinates are said to be of the *same type*. We are really interested in only one parameter, the shape parameter.

The results of this paper are exploratory and theoretical. For multivariate distributions the behaviour in infinity is much more complex than in the univariate case. One can hardly hope to fit this complexity in the straight jacket of a one–dimensional family of distribution types. The theory presented here should be regarded as a base line theory describing asymptotic behaviour in the most simple situation – exceedances over linear thresholds – under almost perfect conditions, a spherically symmetric density, or a density with convex level sets. Statistics is surprisingly powerful in situations where the data set is *bland* and does not exhibit any gross irregularities. In the multivariate setting such a data set may consist of a convex dark center surrounded by a halo of singletons with a decreasing intensity as one moves away from the center. If there are no striking features in the halo then we have a data set to which our theory may apply.

The paper is organized as follows. In Section 2 we recall the univariate GPDs as limit laws for high level exceedances. This notion is to be expanded in a truly multivariate setting. After developing the concept of a high risk scenario in Section 3, we investigate some basic results of the relevant (weak) limit theory. In Section 4 we show that, due to the set-up of our limit assumptions, the resulting laws will always exhibit a large degree of symmetry. The Sections 5–7 treat the important case of the Gauss-exponential high risk limit law, and gives examples of distributions in its domain of attraction. One may look at this case as some sort of central limit case capturing many interesting probability distributions. In Section 5 we introduce a class which extends the family of spherical Weibull densities $f_t(z) = e^{-\|z\|^t/t}/C_t$, $t > 0$. The function $e^{r^t/t}$ is replaced by a function which satisfies the von Mises tail condition for the univariate domain of the exponential limit law, and the spherical level sets are replaced by rotund (egg shaped) level sets. We show that such functions are integrable, and that the associated probability densities belong to the domain of the Gauss-exponential limit law. In Section 6 we multiply the density by a function L which is locally constant far out, and prove that the product is integrable and the associated probability distribution again belongs to the Gauss-exponential domain. In Section 7 we replace Lebesgue measure by a measure μ which locally behaves like Lebesgue measure far out and derive a similar result: the measures $fd\mu$ are finite and the associated probability distributions belong to the domain of the Gauss–exponential limit law. (In both cases we assume local convergence, and we obtain global results). In view of the rather thin tails of the underlying densities the results are not unexpected. However in the multivariate setting with a metric which is basically anisotropic the proofs are not quite trivial, and technical arguments are unavoidable. The reader is advised to skip proofs and lemmas in these three sections on a first reading. The last part of the paper, Sections 8–12, is of a more expository nature. Section 8 treats limit laws with heavy tails, and Section 9 limit laws with bounded support. In Section 10 the full family of multivariate GPD high risk limit laws is presented. Here all information about this family of multivariate probability distributions has been gathered for reference. Section 11 treats the limit behaviour when the half spaces diverge in a given direction. The paper ends with a section where some alternatives to the notion of high risk scenarios are discussed, and a short conclusion.

The aim of this paper is to open up a field of investigation which promises a rich theory with potenially important applications in diverse fields. The concepts underlying the theory are given in full generality. We then prove a number of results about the domains of attraction, concentrating on possible applications in finance. In addition we give a number of illustrative examples to indicate the scope of the theory. Towards the end some results are presented without going into the technicalities of their proofs. In future papers we shall give more details.

2 The univariate case

The intent of this paper is to sketch a multivariate theory for handling conditional extreme events. A full version of such a theory is well beyond the scope of this paper. In the univariate case the family of generalized Pareto distributions (GPDs) has turned out to be a very useful tool in a variety of different contexts. The basic idea there is that one should study the behaviour of a random variable Y conditional on it having exceeded a given threshold value q, where q is large so that the probability α that Y lies in the half line $H = [q, \infty)$ is small. If Y is interpreted as time, then one may think of $Y^H - q$ as the residual life time of Y beyond time q. When Y denotes the potential loss of some financial position at some fixed, future date, and q the Value-at-Risk, then $Y^H - q$ corresponds to the conditional excess loss above q. The latter terminology is also standard in reinsurance: there, the Value–at–Risk threshold corresponds to a so–called attachment point for an excess of loss treaty.

Let q tend towards the upper endpoint of the distribution of Y or let α go to zero. The conditional distributions, suitably scaled, may have a limit. The limit laws are the generalized Pareto distributions (GPDs). A GPD is a distribution on $[0, \infty)$ which (up to scaling) has a density of the form e^{-v}, or $t/(1+v)^{t+1}$ with $t > 0$, or $t(1-v)^{t-1}$ on $[0, 1)$ with $t > 0$. These limit laws are known as the *residual life* or *excess loss limit laws.* Replace t by the parameter $\tau = 1/t$. Then a change of scale yields a continuous one parameter family of distribution functions G_τ, $\tau \in \mathbf{R}$, on $[0, \infty)$, the *standardized* GPDs, defined by

$$1 - G_\tau(v) = (1 - \tau v)_+^{1/\tau} \qquad \tau \in \mathbf{R}\,. \tag{2.1}$$

By continuity $1 - G_0(v) = e^{-v}$. In the univariate case the GPDs form a family of limit laws which may be used to describe the asymptotic behaviour of exceedances. The theory has been applied in diverse fields to describe extreme situations. The limit theory leading to (2.1) was first derived in Balkema and de Haan [2] and Pickands [32]. The domains of attraction of the univariate limit laws have been characterized completely. They contain many of the classical probability densities, and hence most of the standard models in risk management fall within the scope of this theory. For a textbook treatment, see for instance Embrechts et al. [18].

Applications to finance and insurance are broadly discussed in Embrechts [17]. See the latter two references also for further reading. A freeware package EVIS for statistical analysis of extremes based on the GPDs is available from Alexander McNeil (www.math.ethz.ch/∼mcneil).

In finance, the Value–at–Risk (VaR) at level $1 - \alpha$ for α small ($\alpha = 0.05$ or 0.01 in the case of market risk) denotes, for a given portfolio, the level $q_\alpha > 0$ such that the event $\{Y \geq q_\alpha\}$, that the loss of the portfolio lies above the level q_α at the end of a given period (the holding time; 1– or 10–days, say) occurs with a probability α. See Crouhy et al. [14] and Jorion [23] for further details. There exist other measures of risk such as the expected shortfall, the expectation of $Y - q_\alpha$ given that $Y \geq q_\alpha$, or the conditional variance, $\mathrm{var}(Y \mid Y \geq q_\alpha)$, which give a better impression of the actual risk residing in the portfolio. See Artzner et al. [1] for an axiomatic theory in this context. All these measures are functionals of the distribution of the high risk scenario Y^H, where H is the half line $[q_\alpha, \infty)$. Ideally one would like to know the full conditional probability distribution of the value of the loss. If the distribution of Y lies in the domain of a GPD then one may replace the distribution of the overshoot $Y^H - q_\alpha$ by a suitable GPD and use the latter distribution to estimate the various measures of risk listed above. Such a procedure has to be handled with care, but it has the advantage that it is more robust than estimating the expectation or variance of the overshoot, and that one can use the results of a well developed statistical theory.

At this point we should say something about the following question: Why should the distribution of the loss Y, where $-Y$ describes the value of a given portfolio ten days hence, say, be assumed to lie in the domain of attraction of one of these generalized Pareto distributions? This is a philosophical rather than statistical concern. Note that the domain of attraction depends only on the behaviour of the upper tail of the distribution of Y, so in practice one can never tell whether a distribution which is only known from a finite sample will have an upper tail of the prescribed form. On the other hand there exist many examples where the theory has been successful in predicting future behaviour. We give two arguments which make this success plausible.

1) Underlying the apparently erratic behaviour of the market is an extremely

complex dynamical economic system, whose intricate behaviour we may never fully understand, but that system may generate variables whose probability distributions have densities with smooth tails. The classic example is the change in position of a pollen grain submitted to the bombardment of water molecules as described by Einstein and Smoluchowski and nicely discussed in Mazo [29]. This reasoning has also been imported in microeconomic theory. The position of the pollen grain has become the price of a stock, and the water molecule a trader, see Föllmer and Schweizer [21].

2) We are not interested whether the limit behaviour remains valid for $\alpha < 10^{-99}$. The assumption of the existence of a limit distribution for $\alpha \downarrow 0$ is purely a matter of mathematical convenience. Our basic assumption in this paper is that the distribution of the high risk scenarios Z^H and Z^K, suitably standardized, are close if $P\{Z \in H\}$ and $P\{Z \in K\}$ are small and the overlap of the two half spaces H and K is relatively large. In the univariate case the half spaces are half lines and the assumption implies that the distributions of the high risk scenarios translated to the half line $[0, \infty)$, and suitably scaled, are close to a common GPD.

The following example from the one–dimensional theory will be useful for future reference.

Example 2.1 The Gaussian variable Y with density $e^{-y^2/2}/\sqrt{2\pi}$ lies in the domain of the exponential GPD. Indeed let $\alpha \in (0,1)$. Choose $q = q_\alpha$ so that $P\{Y \geq q\} = \alpha$. Write $Y = q + V_q/q$. Then V_q has density

$$f_q(v) = e^{-(q+v/q)^2/2}/\sqrt{2\pi}q \propto e^{-v}e^{-v^2/2q^2} .$$

We may write $\{V_q \geq 0\} = \{Y \in H\}$ for $H = [q, \infty)$. Conditioning on this event we find that the normalized overshoot $V_q = q(Y^H - q)$ has density

$$g_q(v) = e^{-v}e^{-v^2/2q^2}/C \qquad v \geq 0\,, \quad \text{where } C = \int_0^\infty e^{-v}e^{-v^2/2q^2}\,dv\,.$$

It is clear that $C \uparrow 1$ for $q \uparrow \infty$. So $g_q(v) \to e^{-v}$ pointwise, and Scheffé's theorem ensures that the distribution of V_q converges weakly to the standard exponential distribution. This shows that the normal distribution lies in the domain of attraction of the exponential distribution. In other words, for a random variable having a normal distribution, excesses over a high threshold q have approximately an exponential distribution. For a more detailed discussion see Section 3.4 in Embrechts et al. [18]. ¶

3 Basic concepts

In this section we introduce the limit relation (3.1) below which is the basis of our investigations, and prove some general results. We shall first clarify more precisely some of our concepts in the context of standard multivariate models in finance.

In the multivariate case one uses a multivariate Brownian motion with a deterministic drift function and a non–singular covariance as a first order model to describe the dynamics of the log–price process. The drift vector and the covariance matrix may change in the course of time but over a period of a few days they may be assumed constant. In the literature, numerous generalizations of the Brownian model have been proposed. See for example Carr et al. [13], Levin and Tchernitser [26], and Eberlein [15]. For a global macroeconometric model, see Pesaran et al. [31].

In this paper the *market* is a random vector Z in the space $\mathbf{R}^d$. The coordinate, Z_i, $i = 1, \ldots, d$, may be interpreted as the logarithm of the price of the ith asset at a fixed future date. It is helpful to think of Z as having a multivariate density, Gaussian–like in the center, but with fatter tails. Typical examples include elliptical densities as discussed in Breymann et al. [9], Embrechts et al. [18], Bingham and Kiesel [4], and Hult and Lindskog [22]. The multivariate Student t and generalized hyperbolic distributions fall within the scope of our theory. See for instance Rosenberg and Schuermann [35] for an application of these distributions to risk management. We therefore think of a density whose level sets $\{f > c\}$ are ellipsoids, or more generally open convex sets. Heavy tails means that the univariate marginal tails $P\{\theta Z > t\}$ decrease like a negative power of t for $t \to \infty$ for every direction vector θ. Light tails means that the moment generating function $\xi \mapsto Ee^{\xi Z}$ exists on a neighbourhood of the origin. We shall pay particular attention to the intermediate region between light and heavy tails, referred to as *intermediate tails.*

The existence of a density is not essential to the theory which we shall develop;

see Section 7. We do however exclude the situation where the value of certain assets is a function of the value of other assets, such as is the case when besides certain financial underlyings, also derivatives (like options) written on those underlyings are included. Functional relations between the various components of the vector Z will force the distribution of Z to lie on a lower dimensional manifold in the space $\mathbf{R}^d$. Such distributions fail to meet the requirement that there should be no striking irregularities in the halo of the sample cloud underlying our analysis.

The density of the high risk scenario Z^H can of course be written down explicitly given the density of the vector Z. Unfortunately the formula gives little insight in the asymptotic behaviour of the distribution of Z^H, even in the univariate case, as may be seen in Example 2.1. For reference purposes, we state this trivial result as a proposition.

Proposition 3.1 *Suppose the random vector Z has density f. Let H be a closed half space so that $P\{Z \in H\} = \alpha > 0$. Then Z^H has density*

$$f^H(z) = 1_H(z) f(z)/\alpha \qquad z \in \mathbf{R}^d .$$

¶

We are interested in the asymptotic behaviour of the distribution of Z^H as the half space H drifts off and $P\{Z \in H\} > 0$ tends to zero. It is important to note that we do *not* impose conditions on the direction in which the half spaces diverge. It turns out that there exists an elegant multivariate version of the theory for exceedances over thresholds sketched in Section 2. It is our aim in this paper to describe the outlines of this basic theory where the sole condition is $0 < P\{Z \in H\} \to 0$. Since we assume that there exists a limit law, the asymptotic behaviour of Z^H is the same in all directions. This presupposes a certain degree of directional homogeneity for the underlying distribution. Even when the whole market, say all stocks in an index, does not satisfy the condition of directional homogeneity there may be segments, say energy or banking, which in first approximation do. The homogeneity is obviously present when the vector Z has a spherically symmetric distribution. One of the aims of this paper is to investigate this homogeneity condition. It will turn out that the condition of directional homogeneity is more restrictive in the case of heavy tails than for light tails. In the latter case it may be regarded as a smoothness condition on the underlying distribution.

The vector Z^H lives on the half space H. In order to compare distributions of different vectors Z^H, we map the half spaces H onto a common closed half space H_+, the upper half space, $\mathbf{R}^h \times [0, \infty)$ in (1.1). This may be achieved by rotating H so that the normal vector to ∂H, pointing into H, falls along the positive vertical axis. After rotation the half space has the form $\{z_d \geq q\}$ for some constant q. Now translate the half space along the vertical axis until it coincides with H_+. It is convenient to write $z = (x, y)$ and to think of $x = (z_1, \ldots, z_{d-1})$ as the horizontal component of the vector z, and $y = z_d$ as the vertical coordinate.

An *affine transformation* β is a map of the form $z \mapsto w = b + Bz$ where b is a vector in $\mathbf{R}^d$, and B an invertible linear transformation of $\mathbf{R}^d$. The inverse $w \mapsto z = B^{-1}(w - b) = a + Aw$ with $A = B^{-1}$ and $a = -Ab$, again is an affine transformation, and so is the composition of two affine transformations. We shall denote the group of all affine transformations on $\mathbf{R}^d$ by $\mathcal{A} = \mathcal{A}(d)$. The set of affine transformations which map the upper half space H_+ onto itself is denoted by $\mathcal{A}_+$. It is a subgroup of $\mathcal{A}$.

For any fixed half space H there are many affine transformations which map H_+ onto H. Indeed if β is such a transformations then so is $\beta \circ \alpha$ for any $\alpha \in \mathcal{A}_+$. If $P\{Z \in H\}$ is positive then Z^H is well defined and $W_H = \beta^{-1}(Z^H)$ is a vector on H_+. The inverse of the affine transformation β mapping H_+ into H transforms the high risk scenario Z^H into a vector W_H on H_+. We are interested in the case where it is possible to choose for each half space H, with $P\{Z \in H\} > 0$, an affine transformation β_H mapping H_+ onto H, so that the normalized high risk vectors converge in distribution to a non-degenerate random vector W on H_+:

$$W_H = \beta_H^{-1}(Z^H) \Rightarrow W \qquad 0 < P\{Z \in H\} \to 0. \tag{3.1}$$

Recall that a random vector is *degenerate* if it lives on a hyperplane.

The condition $P\{Z \in H\} \to 0$ is vacuous if Z for instance lives on a cube in $\mathbf{R}^3$, and the vertices carry positive mass. We want $P\{\theta Z \geq q\}$ to tend to zero for any direction θ when q increases to the upper endpoint of the distribution of the random variable θZ. Therefore the next condition on the distribution π of Z is natural:

$$\pi H > 0 \Rightarrow \pi \partial H < \pi H \tag{3.2}$$

for any closed half space H. Here ∂H denotes the boundary hyperplane of H. In $d = 1$, the condition states that the distribution function of Z is continuous in its endpoints; see Corollary 3.1.2 in Embrechts et al. [18] for the relevance of this condition for one–dimensional extreme value theory. Let us say a few words about the multivariate condition since it will be imposed throughout the paper. A geometric formulation of (3.2) may be given in terms of the convex support of the vector Z. The *convex support* is the intersection of all closed half spaces H such that $P\{Z \in H\} = 1$.

Lemma 3.2 *Condition (3.2) holds if and only if $P\{Z \in H\} = 0$ for any closed half space which is disjoint from the interior of the convex support of π.*

Proof Suppose $P\{Z \in H\}$ is positive. If H is disjoint from the interior of the convex support then $P\{Z \in H\} = P\{Z \in \partial H\}$. Conversely $P\{Z \in H\} = P\{Z \in \partial H\}$ for $H = \{\theta \geq c\}$ implies that Z lives on $\{\theta \leq c\}$ and hence the interior of the convex support lies in $\{\theta < c\}$ and is disjoint from H. ¶

The parametrization $H = \{\theta \geq c\}$, with θ in the unit sphere, S, and $c \in \mathbf{R}$, shows that the set $\mathcal{H}$ of closed half spaces is homeomorphic to the cylinder $S \times \mathbf{R}$. Condition (3.2) ensures that the set of half spaces which carry positive mass is open in $\mathcal{H}$. If we make the stronger assumption that π does not charge hyperplanes then the map $H \mapsto \pi^H$ is continuous on this open set. These two statements follow from the next proposition.

Proposition 3.3 *Let Z be a random vector with distribution π, and let H_n, $n \geq 0$, be closed half spaces converging to $H_0 = \{\theta \geq q\}$. If $P\{\theta Z > q\}$ is positive, then πH_n is positive eventually. If πH_0 is positive and $P\{\theta Z = q\} = 0$ then $\pi H_n \to \pi H_0$ and $Z^{H_n} \Rightarrow Z^{H_0}$.*

Proof Any point in the interior of H_0 will eventually lie in H_n. This also holds for any compact convex set K in the interior of H_0. Choose K so that πK is positive to obtain the first statement. To prove the second statement write $h_n = 1_{H_n}$ and observe that $h_n \to h_0$ holds π-a.e. Hence $p_n = P\{Z \in H_n\} = \int h_n d\pi$ converges to p_0. It remains to prove that for any bounded continuous function φ on $\mathbf{R}^d$ the expectations $E(h_n\varphi)(Z)$ converge. This also follows from Lebesgue's theorem on dominated convergence. ¶

We now return to relation (3.1), the basic limit relation of this paper.

Definition The vector W in (3.1) is called a *high risk limit scenario*, its non-degenerate distribution ρ a *high risk limit law* or *limit distribution for exceedances over linear thresholds*, and the original vector Z (or its distribution π) is said to *lie in the domain of attraction of* W (or ρ) provided the boundary condition (3.2) is satisfied. Note in particular that W is assumed to be non-degenerate and that the normalizing transformations β_H have to satisfy the condition $\beta_H(H_+) = H$. We write $Z \in \mathcal{D}(W)$, or $\pi \in \mathcal{D}(\rho)$. ¶

The random vectors Z^H, W_H and W are introduced as a notational convenience. The limit may be formulated in terms of weak convergence of probability measures as $\beta_H^{-1}(\pi^H) \to \rho$.

We are faced with two mathematical problems:

1) Determine the class of all high risk limit laws, and

2) For each limit law, determine its domain of attraction.

In their full generality, these two problems are still open. In this paper we shall give a partial answer to question 1 by introducing a one–parameter family of high risk limit laws. Partial answers to question 2 will also be given. Though our results are only partial in nature, we believe them to be sufficiently general to handle many of the standard situations faced with in practical problems in general, and in financial and insurance risk management in particular. In this context the next example is fundamental.

Example 3.4 If the market Z has a multivariate Gaussian distribution with a non–singular covariance matrix then a high risk limit law exists.

The argument is geometric. We begin by choosing the origin and coordinates so that Z has a standard Gaussian density $e^{-z^T z/2}/(2\pi)^{d/2}$. Let H be a closed half space $H = \{\theta \geq q\}$ for some unit vector θ and some constant $q \geq 0$. Because of symmetry, we may assume that H has the form $\{y \geq q\}$ where y is the vertical coordinate of $z = (x, y) \in \mathbf{R}^h \times \mathbf{R}$. Write $Z = (X, Y)$, and write $Y = q + V_q/q$ as in Example 2.1. Then $\{V_q \geq 0\} = \{Z \in H\}$. Condition on this event and set

$$\beta_H(u, v) = (u, q + v/q) \qquad w = (u, v) \in \mathbf{R}^h \times \mathbf{R}. \tag{3.3}$$

Then β_H maps H_+ onto $H = \{v \geq q\}$, and the components of $W_H = (U, V_q) = \beta_H^{-1}(Z^H)$ are independent. The vertical component V_q has density g_q of Example 2.1, and the horizontal component $U = X$ is standard Gaussian on $\mathbf{R}^h$. Hence $W_H \Rightarrow W = (U, V)$ where V is standard exponential and independent of the Gaussian vector U. ¶

The limit vector $W = (U, V)$ in Example 3.4 has a *standard Gauss-exponential* distribution. This is a high risk limit law. It has density

$$g_0(w) = e^{-v} e^{-u^T u/2}/(2\pi)^{h/2} \qquad w = (u, v) \in H_+ = \mathbf{R}^h \times [0, \infty). \tag{3.4}$$

The theory developed in this paper is geometric in the sense that it does not depend on the choice of coordinates. This is expressed in our next result.

Theorem 3.5 (Invariance) *Suppose Z lies in the domain of attraction of the high risk limit vector W. Let $\alpha \in \mathcal{A}$, $\beta \in \mathcal{A}_+$. Then $\beta(W)$ is a high risk limit vector and $\alpha(Z)$ lies in its domain of attraction.*

Proof Suppose $\beta_H^{-1}(Z^H) \Rightarrow W$. Set $\tilde{H} = \alpha(H)$ and $\tilde{Z} = \alpha(Z)$. Then clearly $\{\tilde{Z} \in \tilde{H}\} = \{Z \in H\}$. So these events have the same probability. Therefore $W_H = \beta_H^{-1} \circ \alpha^{-1}(\tilde{Z}^{\tilde{H}})$ and $W_H \Rightarrow W$ implies $\beta(W_H) \Rightarrow \beta(W)$ for $P\{\tilde{Z} \in \tilde{H}\} \to 0$. The condition that W is non-degenerate, and the boundary condition (3.2),are geometrical, and hence they also hold for $\tilde{W}$ and $\tilde{Z}$. ¶

High risk scenarios only look at the outer edges of a distribution. The asymptotic behaviour of the high risk scenarios Z^H for $0 < P\{Z \in H\} \to 0$ does not depend on the precise form of the probability distribution of Z on any compact subset of the interior of the convex support of the distribution. Similarly if two densities are asymptotically equal and one lies in the domain of attraction of a high risk limit l aw, then so does the other, with the same normalization. Since the support of our measures may be bounded we shall now first define more precisely what we mean by divergence and asymptotic equality in a multivariate setting.

Definition Let O be an open set in $\mathbf{R}^d$. We say that a sequence of points x_n in O *diverges in* O and write $x_n \to \partial$ (or $x_n \to \partial_O$) if any compact subset of O contains only finitely many terms of the sequence. Two non-negative functions f and g defined on O are *asymptotic on* O and we write

$$f(x) \sim g(x) \qquad x \to \partial$$

if for each $\epsilon > 0$ there exists a compact set $K \subset O$ so that

$$e^{-\epsilon} f(x) \leq g(x) \leq e^{\epsilon} f(x) \qquad x \in O \setminus K. \tag{3.5}$$

Theorem 3.6 (Asymptotics) *Let the distribution π lie in the domain of the high risk limit law ρ. Let O be the interior of the convex hull of the support of π. Suppose $\pi(O) = 1$. Let the probability measure μ on O have the form $d\mu = d\mu_0 + g d\pi$ where g is a Borel function on O with a finite limit $c > 0$ for $z \to \partial_O$, and $\mu_0(H)/\pi(H) \to 0$ for $0 < \pi(H) \to 0$. Then μ lies in the domain of ρ. Moreover the normalizing transformations for μ and π are the same.*

Proof Let $0 \leq \varphi \leq 1$ be a continuous function on H_+, and let $\epsilon \in (0,1)$. Choose $K \subset O$ compact so that $1 - \epsilon < g/c < 1 + \epsilon$ outside K, and choose $\delta > 0$ so small that $\mu_0(H)/\pi(H) < \epsilon c$ for $0 < \pi(H) < \delta$. Let Z have distribution π and let S have distribution μ. Assume $\beta_H^{-1}(Z^H) \Rightarrow W$ where W has distribution ρ, and β_H maps H_+ onto H. Let H be a closed half space disjoint from K with $0 < \pi(H) < \delta$. Set $\psi_H = \varphi \circ \beta_H^{-1}$. Then

$$(1-\epsilon) \int_H \psi d\pi \leq (1/c) \int_H \psi d\mu \leq (1+\epsilon) \int_H \psi d\pi + \epsilon \pi(H)$$

holds for both $\psi = \psi_H$ and $\psi \equiv 1$. On division the constant c drops out, and one has a simple bound on the difference between $E\psi_H(S^H)$ and $E\psi_H(Z^H)$. ¶

Write $H_n \to \partial_O$ if all half spaces H_n intersect the open set O, but any compact set $K \subset O$ only intersects H_n for finitely many indices n. If (3.2) holds one may formulate the limit condition $0 < P\{Z \in H_n\} \to 0$ in terms of the interior O of the convex support of the distribution of Z as

$$H_n \to \partial_O. \tag{3.6}$$

In applications the vector Z often has a density of the form f/C where f is a positive integrable function on an open set O in $\mathbf{R}^d$ and $C = \int_O f(z) dz$. The boundary condition (3.2) is satisfied. For convergence in (3.1) it suffices that the densities g_H of the normalized high risk scenarios $\beta_H^{-1}(Z^H)$ converge

$$g_H(w) = |\det \beta_H| f(\beta_H(w))/CP\{Z \in H\} \to g(w) \qquad 0 < P\{Z \in H\} \to 0$$

pointwise on $\{g > 0\}$ where g is a probability density on H_+ (by Scheffé's theorem and Fatou). In order to avoid the multidimensional integration needed to evaluate C and $P\{Z \in H\}$ one might consider convergence of the quotients

$$h_H(w) = \frac{(f \circ \beta_H)(w)}{(f \circ \beta_H)(0)} \to h(w) \qquad w \in \{h > 0\},\ H \to \partial_O \tag{3.7}$$

where h is an integrable non-negative function on H_+. It is clear that one needs extra conditions to conclude that Z lies in the domain of attraction of a high risk limit law with density proportional to h. Here we list a few such conditions.

Proposition 3.7 (Densities) *Let Z have continuous density f/C, and let h be a continuous non-negative integrable function on H_+ which does not vanish identically. Suppose $\beta_H \in \mathcal{A}$ maps H_+ onto H so that (3.7) holds. Set $p_H = \beta_H(0)$, and let g_H be the density of the normalized high risk vector $\beta_H^{-1}(Z^H)$, and ρ_H the distribution. Let W have density $g = h/C_0$.The following are equivalent for $P\{Z \in H\} \to 0$:*

1) Z lies in the domain of attraction of W;

2) the family of probability distributions ρ_H, $P\{Z \in H\} > 0$, is tight;

3) $g_H \to g$ pointwise on $\{g > 0\}$;

4) $h_H \to h$ in $\mathbf{L}^1$;

5) $\|h_H\|_1 \to \|h\|_1$, and

6) $|\det \beta_H| f(p_H)/P\{Z \in H\} \to 1/C_0$.

Proof First observe that the density g_H of the normalized high risk vector satisfies

$$\frac{g_H(w)}{g_H(0)} = \frac{|\det \beta_H|(f \circ \beta_H)(w)}{P\{Z \in H\} g_H(0)} \qquad \frac{g_H(w)}{g_H(0)} = \frac{(f \circ \beta_H)(w)}{(f \circ \beta_H)(0)}.$$

The implications 4) $\Rightarrow$ 5) $\Rightarrow$ 3) $\Rightarrow$ 1) $\Rightarrow$ 2) are obvious. The right hand equality above gives 3) $\Rightarrow$ 4) and 2) $\Rightarrow$ 3). The two equalities together with 3) and 4) imply 6) with $1/C = g(0)$. Together with 6) the two equalities give 5) with $C = \|h\|_1$. ¶

Condition 6) relates the exceedance probability $P\{Z \in H\}$ to the density f/C in the point p_H. If one of the conditions 1) – 6) holds, the density f can not have sharp peaks far out in O. Let us therefore now assume that f is a continuous function with convex level sets $\{f > c\}$ for $c > 0$. Such functions are also called *unimodal* or *quasiconcave*, see Boyd and Vandenberghe [8]. The domain of f, $O = \{f > 0\}$, is convex and open. The quotients h_H in (3.7) have convex level sets, and so has the limit h. This implies that h^q and f^q are integrable for all $q \geq 1$, and that $h_H^q \leq h_H$ outside some compact subset of H_+ for all half spaces H disjoint from some compact set $K \subset O$. Hence 5) holds for h_H^q for any $q \geq 1$ if it holds for $q = 1$. This proves

Theorem 3.8 (Power families) *Let the vector Z have density f/C where f is continuous on $\mathbf{R}^d$ with convex level sets $\{f > c\}$ for $c > 0$. Suppose (3.7) holds pointwise and in $\mathbf{L}^1$. Then for each $q > 1$ there exists a high risk limit law with density $g_q = h^q/C_q$. The function f^q is integrable and the corresponding density lies in the domain of attraction of g_q with the norming transformations β_H of (3.7).*

We conclude that the class of high risk limit laws may contain power families:

$$g_q(w) = g^q(w)/C_q \qquad q \in J$$

where g is a continuous function on H_+ with convex level sets $\{g > c\}$, and J is an interval in $\mathbf{R}$ with upper endpoint ∞. The class of Gauss-exponential densities on H_+ is invariant under positive power transformations: if (X, Y) has density g then $(X/\sqrt{q}, Y/q)$ has density proportional to g^q for any $q > 0$. In Section 8 we shall introduce the power family generated by $g(u, v) = 1/((1 + v)^2 + u^T u)$, the heavy tailed spherical Pareto limit laws, with $J = (d/2, \infty)$; and in Section 9 the power family generated by $g(u, v) = (1 - v - u^T u)_+$, the parabolic power limit laws, with bounded support, with $J = (-1, \infty)$. Both these power families are asymptotically Gauss–exponential for $q \to \infty$. The second family yields a singular limit law for $q \downarrow -1$. These high risk limit laws will be called the multivariate GPDs, or the *Pareto-parabolic distributions*. As far as we know there are no other high risk limit laws.

4 Symmetries of the limit measure

The symmetry of the Gauss-exponential density in (3.4) may seem to be due to the spherical symmetry in the Gaussian distribution in Example 3.4. We shall argue below that a large symmetry group is characteristic for high risk limit laws.

Let us first point out a peculiarity of the Gauss-exponential high risk limit law. The Gaussian vector Z lies in the domain of attraction of the Gauss-exponential vector W, but W does not belong to its own domain of attraction. Indeed take half spaces $H_n = \{y \leq 1/n\}$. Then the high risk vector W^{H_n}, blown up by a factor n in the vertical direction, converges in distribution to the vector (U, T) where U is standard Gaussian on $\mathbf{R}^{d-1}$ and T is uniformly distributed on the interval $[0, 1]$, and independent of the horizontal component U. By the convergence of types theorem, the sequence of high risk vectors W^{H_n} can not be normalized to converge to the Gauss–exponential vector W.

In order to understand better why the Gauss-exponential density should be a high risk limit density we introduce the infinite measure ρ^* which has the density (3.4) on the whole space $\mathbf{R}^d$. Let H be the closed half space of all points $z = (x, y)$ above the graph of an affine function $y = \varphi(x)$

$$H = \{y \geq \varphi(x)\} \qquad \varphi(x) = c_0 + c_1 x_1 + \cdots + c_{d-1} x_{d-1},\ c_0, \ldots, c_{d-1} \in \mathbf{R}. \quad (4.1)$$

Then $\rho^*(H)$ is finite and $\rho_H = 1_H d\rho^* / \rho^*(H)$ is a probability measure which lives on H. Below we shall see that there exists an affine transformation γ_H which maps H_+ onto H so that ρ_H is the image of the standard Gauss-exponential distribution ρ on H_+. The measure ρ^* on $\mathbf{R}^d$ has the property that for any closed half space H with finite measure the corresponding probability measure is of the Gauss-exponential type.

Let $\mathcal{S}$ denote the symmetry group of ρ^*. This is the set of all affine transformations β so that $\beta(\rho^*) = c\rho^*$ for some positive constant $c = c_\beta$. Obviously $\mathcal{S}$ contains the group of all rotations around the vertical axis because of the symmetry of the horizontal Gaussian component and the independence of the vertical component. It also contains the translations along the vertical axis since the density of ρ^* is a product $g(u)e^{-v}$ and the density e^{-v} on $\mathbf{R}$ is semi-invariant under translations. However there is an additional symmetry. The level curves of the density are paraboloids, and these are most symmetric figures. By an affine change of coordinates one can transfer the top of the paraboloid to any point on the paraboloid. Let b denote a vector in $\mathbf{R}^{d-1}$. Consider the affine transformation β given by

$$\beta(u,v) = (u - b, v + b^T u - b^T b/2) \qquad (u,v) \in \mathbf{R}^{d-1} \times \mathbf{R}.$$

It maps the origin into the point below $u = -b$ on the paraboloid $v = -u^T u/2$. Under the inverse transformation, the density g_0 of ρ^* transforms into

$$g(x,y) = g_0(x - b, y + b^T x - b^T b/2) = e^{-y} e^{-b^T x} e^{b^T b/2} e^{-(x-b)^T (x-b)/2} / (2\pi)^{(d-1)/2}.$$

The right hand side is exactly $g_0(x,y)$. So ρ^* is invariant under the affine transformation β^{-1}, and hence also under β.

The collection of all half spaces H of the form (4.1) is a d-dimensional space. The symmetry group $\mathcal{S}$ of ρ^* is transitive: it can map any element of this space into any other element. We formulate our results as a proposition.

Proposition 4.1 *For any half space H for which $\rho^* H$ is finite, there exists an affine transformation γ_H mapping H_+ onto H which sends the probability measure ρ on H_+ into the probability measure $1_H d\rho^*/\rho^* H$. The transformation γ_H is a symmetry of ρ^*: $\gamma_H(\rho^*) = c_H \cdot \rho^*$ with $c_H = \rho^* H$. Conversely any symmetry σ of ρ^* has the property that $\sigma(\rho)$ is the probability measure $1_H d\rho^*/\rho^* H$ on $H = \sigma(H_+)$.*

Let us now try to explain why the measure ρ^* should have such a large symmetry group. We turn to the basic limit relation (3.1). Let Z with distribution π lie in the domain of the Gauss-exponential law. For half spaces H far out the distribution of Z^H will look like the Gauss-exponential law if we choose appropriate coordinates. Let H_0 be such a half space far out. Choose coordinates so that $H_0 = H_+$. Then $\pi^{H_0} \approx \rho$. There is a d-dimensional set of half spaces H close to H_0, in the sense that H and H_0 have approximately the same mass and have a large relative overlap. Take such a half space. For simplicity assume $\pi H = \pi H_0$. Then $\pi^H \approx \beta(\rho)$ for $\beta = \beta_H$ mapping H_0 onto H by (3.1). So $\rho \approx \beta(\rho)$ holds on $H \cap H_0$ since π^H and π^{H_0} agree on the overlap. In the limit this yields identities $\rho^* = \beta(\rho^*)$ for a large collection of affine maps β. The argument derives from the fact that for high risk limit laws convergence has to hold for all half spaces H, whatever their direction, as long as $0 < P\{Z \in H\} \to 0$. We may therefore expect other high risk limit laws to have similar large symmetry groups.

In the univariate case, the Radon measure ρ^* has a statistical interpretation. The sample point process from a standard Gaussian distribution, properly normalized, converges to a Poisson point process with intensity e^{-v} on $\mathbf{R}$ as the number of points in the sample increases without bound. In the multivariate case a similar result holds for samples from a standard Gaussian distribution on $\mathbf{R}^d$. The limiting Poisson point process has intensity $e^{-u^Tu/2}e^{-v}$ on $\mathbf{R}^h \times \mathbf{R}$, see Brown and Resnick [10] or Eddy [16]. This Poisson point process describes asymptotically the local behaviour around the outer edge of a large sample from a Gaussian distribution. The measure ρ^* links the asymptotic theory of high risk scenarios to the asymptotic description of the local behaviour of the halo of large sample clouds. High risk scenarios thus enable us to handle multivariate extremes in a more geometric way than the classical theory of coordinatewise maxima.

5 The Gauss-exponential domain, rotund sets

In this section we shall introduce a structured family of distributions in the domain of attraction of the Gauss-exponential limit law. This family will be enlarged in the next two sections.

In Example 3.4 it was shown that the domain of the Gauss–exponential high risk limit law contains the multivariate Gaussian densities. We shall extend this result. In first instance we consider spherical Weibull densities:

$$f_t(z) = e^{-\|z\|^t/t}/C, \qquad t > 0.$$

For each $t > 0$ the level sets $\{f_t > c\}$ of the density f_t are concentric open balls. A change of coordinates transforms the balls into ellipsoids. The level sets of the limit density (3.4) are parallel parabolas $v < c - u^Tu/2$ in (u, v)-space. Locally a circle looks like a parabola. Indeed this holds for any convex set provided the boundary is smooth and the curvature is positive in each point. A typical example of such a set is an egg. We shall therefore also consider probability densities whose level sets are concentric eggs.

Let us now introduce more formally the convex sets with which we shall be concerned. Assume that D is a bounded open convex set in $\mathbf{R}^d$ containing the origin. Write $\mathbf{R}^d = \mathbf{R}^h \times \mathbf{R}$ and let D_0 be the (bounded open convex) image of D under the projection $(x, y) \to x$. The upper boundary of D may be described by a concave function $\gamma_+ : D_0 \to \mathbf{R}$. We say that the boundary of D is C^2 in a point $p = (x, \gamma_+(x))$, and has *positive curvature*, if the function γ_+ is C^2 in x with negative definite second derivative $\gamma_+''(x)$. The map $x \mapsto (x, \gamma_+(x))$ describes the structure of the compact manifold ∂D locally. The parametrization γ_+ will come in handy below, but there is an alternative global description of ∂D which is more convenient.

Since D contains the origin there exists a unique function $n = n_D : \mathbf{R}^d \to [0, \infty)$, the *gauge function of* D, which satisfies the two conditions:

$$\{n < 1\} = D; \qquad n(rz) = rn(z) \qquad z \in \mathbf{R}^d,\ r \geq 0.$$

If D is the unit ball then n is the Euclidean norm. We do not assume that D is symmetric. Hence n need not be a norm. The conditions on D do however ensure that n is continuous and convex. Assume that n is C^2. Since n is linear on rays the second derivative can not be positive definite. That explains why we need the function n^2 in the definition below.

Definition A *rotund* set in $\mathbf{R}^d$ is a bounded open convex set which contains the origin and whose gauge function n is C^2 on $\mathbf{R}^d \setminus \{0\}$. In addition the second derivative of n^2 is positive definite in each point $z \neq 0$. ¶

The implicit function theorem may be used to show that rotundity of D is equivalent to ∂D being a compact C^2 manifold with positive curvature in each point. In order to investigate the boundary around a point $p \in \partial D$ it is convenient to choose coordinates so that $p = (0, 1)^T$. A linear map $A : \mathbf{R}^h \times \mathbf{R} \to \mathbf{R}^d$ is called an *initial transformation for* p if $A(0, 1)^T = p$, $\{y = 1\}$ is tangent to $A^{-1}D$ in $(0, 1)^T$, and if the gauge function $n \circ A$ of $A^{-1}D$ satisfies $(n \circ A)(x, 1) - 1 \sim x^T x/2$ for $x \to 0$. The linear map A is unique up to a rotation around the vertical axis. Locally it may be chosen to depend continuously on p. Hence the set J of all initial transformations over $p \in \partial D$ is compact.

We can now state the main result of this paper.

Theorem 5.1 *Let n be the gauge function of a rotund set D in $\mathbf{R}^d$ and let ψ be a*

C^2 function on an interval $[0, t_\infty)$, with $t_\infty \in (0, \infty]$, which satisfies the following conditions

1) $\psi(t) \to \infty$ for $t \to t_\infty$;

2) $\psi'(t) > 0$ for $t \in (0, t_\infty)$, and

3) $(1/\psi')'(t) \to 0$ for $t \to t_\infty$.

Then the function $e^{-\varphi}$, with $\varphi = \psi \circ n$, is integrable, and the corresponding probability density $f = e^{-\varphi}/C$ lies in the domain of attraction of the Gauss-exponential limit law. (If t_∞ is finite we set $f = 0$ outside the open convex set $t_\infty D$.)

There exist affine transformations β_H so that $\beta_H(H_+) = H$ and

$$(f \circ \beta_H)(w)/(f \circ \beta_H)(0) \to e^{-v} e^{-u^T u/2} \qquad w = (u, v) \in H_+,\ 0 < P\{Z \in H\} \to 0$$

uniformly on compact subsets of H_+, and in $\mathbf{L}^1$.

This theorem is a consequence of a more technical proposition which we formulate below. Note that there are two ingredients in the density. The convex set D determines the shape of the level sets; the function ψ determines how fast the tails of the density decrease. The smoothness and curvature conditions on the boundary of the set D ensure that the level sets of the densities of the high risk scenarios Z^H asymptotically are parabolic; the conditions on ψ ensure that the parabolas corresponding to levels e^{-n} are asymptotically equidistant. The conditions on ψ are related to the well-known von Mises conditions for the domain of the exponential limit law in the univariate case: A positive random variable whose tail function $T = 1 - F$ is asymptotic to $e^{-\psi}$ in its upper endpoint, t_∞, lies in the domain of attraction of the exponential law,

$$T(t + v/\psi'(t))/T(t) \to e^{-v} \qquad v \geq 0,\ t \to t_\infty$$

see (5.2) below.

The asymptotic behaviour of ψ is well known from univariate extreme value theory. The four limit relations below contain useful information.

Set $a(t) = 1/\psi'(t)$. The derivative a' vanishes in t_∞ by condition 3). If t_∞ is finite then a vanishes in t_∞ by condition 1) and $a(t) = o(t_\infty - t)$ in t_∞. If $t_\infty = \infty$ then $a(t) = o(t)$ and hence $t\psi'(t) \to \infty$, which implies that $t^m e^{-\psi(t)} \to 0$ for any $m \geq 1$. This ensures that $e^{-\varphi}$ is integrable.

Let $t_n \to t_\infty$ from below and let (s_n) be bounded. By the arguments above $a(t_n + s_n a(t_n))$ is well defined eventually, and

$$a(t_n + s_n a(t_n))/a(t_n) = 1 + (a(t_n + s_n a(t_n)) - a(t_n))/a(t_n) \to 1 \qquad (5.1)$$

since a' vanishes in t_∞. A positive function which satisfies $a(t + sa(t))/a(t) \to 1$ is said to be self–neglecting or Beurling slowly varying, see Bingham et al. [6]. Now observe that, by the mean value theorem,

$$\psi(t + sa_t) - \psi(t) = sa_t/a(t + \sigma a_t) \to s \qquad t \to t_\infty \qquad (5.2)$$

uniformly on bounded s-intervals.

The inverse function $\bar{\psi}$ is defined on $[\psi(0), \infty)$, increases from 0 to t_∞, is C^2 on $(\psi(0), \infty)$, and $\bar{\psi}'(s) = a(t)$ for $s = \psi(t)$. The relation $a' \to 0$ in t_∞ translates into $(\log \bar{\psi}')' \to 0$ in ∞. Hence

$$\bar{\psi}'(s + u)/\bar{\psi}'(s) \to 1 \qquad s \to \infty \qquad (5.3)$$

uniformly on bounded u-intervals. The mean value theorem then gives

$$\frac{\bar{\psi}(s + u) - \bar{\psi}(s)}{\bar{\psi}'(s)} \to u \qquad \frac{\bar{\psi}(s + u) - \bar{\psi}(s)}{\bar{\psi}(s + 1) - \bar{\psi}(s)} \to u \qquad s \to \infty \qquad (5.4)$$

uniformly on bounded u-intervals.

Lemma 5.2 *Let ψ satisfy the conditions of Theorem 5.1. Given $\epsilon > 0$, there exists $t_\epsilon < t_\infty$ so that for any $t \geq t_\epsilon$ one has the inequalities $q_n \leq e^{n\epsilon}/\epsilon$, $n \geq 0$, for the sequence (q_n) defined by $\psi(t + q_n a_t) = \psi(t) + n$.*

Proof The limit relation $(\bar{\psi}(s+2)-\bar{\psi}(s+1))/(\bar{\psi}(s+1)-\bar{\psi}(s)) \to 1$ for $s \to \infty$ implies $q_{n+1}-q_n \leq e^\epsilon(q_n-q_{n-1})$ for $t \geq t_0$ and hence $q_n-q_0 \leq (e^{n\epsilon}-1)/(e^\epsilon-1)(q_1-q_0)$. Now use $q_1 = (\bar{\psi}(s+1)-\bar{\psi}(s))/a(t) \to 1$ for $s = \psi(t) \to \infty$, which implies $q_1 \leq (e^\epsilon-1)/\epsilon$ for $t \geq t_1$. ¶

In our proof below we need a simple global bound on a class of convex sets D.

Lemma 5.3 *Let D be an open convex set in $\mathbf{R}^h \times \mathbf{R}$ which is disjoint from the half space $\{y \geq 1\}$ and whose boundary contains the vertical unit vector $(0,1)^T$. Let $\delta \in (0,2)$, and set $\epsilon = \delta^2/16$. If*

$$(x, y) \in D, \; \|x\| \leq \delta \;\Rightarrow\; y < 1 - x^T x/4 \qquad (5.5)$$

then

$$(x,y) \in D \cap H_+ \;\Rightarrow\; y < 1 - \epsilon x^T x. \tag{5.6}$$

Proof By radial symmetry we may assume that D is a subset of the plane $\mathbf{R}^2$. The line L through the points $(0,1)$ and $(\delta, 1-\delta^2/4)$ intersects the x-axis in $4/\delta$ as does the parabola $y = 1 - \epsilon x^2$. The convex set $D \cap H_+$ lies below L for $x > \delta$, and hence certainly below the parabola. It lies below the parabola on $[0,\delta]$ since $\epsilon < 1/4$. ¶

Proposition 5.4 *Let the conditions of Theorem 5.1 hold except that the condition that $D \subset \mathbf{R}^h \times \mathbf{R}$ is rotund is replaced by*

$$n_D(x,1) - 1 \sim x^T x/2 \qquad x \to 0. \tag{5.7}$$

Let $H^t = \{y \geq t\}$ and let $(X_t, Y_t) = Z^{H^t}$. Set

$$a_t = 1/\psi'(t) \qquad b_t = \sqrt{t a_t} \qquad t \in (0, t_\infty). \tag{5.8}$$

Then

$$W_t = \beta_t^{-1}(Z^{H^t}) = (X_t/b_t, (Y_t - t)/a_t) \Rightarrow W \qquad t \to t_\infty$$

where W is standard Gauss-exponential.

If in addition the function $x \mapsto n_D(x,1)$ is C^2 on a neighbourhood of 0 in $\mathbf{R}^h$, then for any sequence of half spaces $H_n = \{\theta_n \geq t_n\}$, with $\theta_n \to (0,1)$ and with $P\{Z \in H_n\} \to 0$ the sequence of high risk vectors Z^{H_n} is asymptotically Gauss–exponential, and the normalized densities converge in $\mathbf{L}^1$.

Proof The proof consists of three parts. The vector W_t has density $\propto e^{-\varphi_t}$ where

$$\varphi_t(u,v) = \varphi(b_t u, t + a_t v) - \varphi(0,t).$$

1) We first prove that $\varphi_t(u,v) \to u^T u/2 + v$ uniformly on bounded subsets of $\mathbf{R}^d$ for $t \to t_\infty$.

Write φ_t as a sum of two terms $A_t = \varphi(b_t u, t^*) - \varphi(0, t^*)$ with $t^* = t + a_t v$ and $B_t = \psi(t + a_t v) - \psi(t)$. Then $B_t \to v$ by (5.2), and convergence is uniform on bounded v-sets. By homogeneity $n(b_t u, t^*) = t^* n(b_t u/t^*, 1)$ and we may use (5.7) to write A_t as

$$\psi(t^* + (b_t^2 u^2/t^*)((1+o(1))/2)) - \psi(t^*) = \psi(t^* + a_{t^*} u^2 (1+o(1))/2) - \psi(t^*) \to u^2/2$$

for $t \to t_\infty$, since $b_t^2 = ta_t \sim t^* a_{t^*}$, see (5.1). The limit relation also holds if we replace the variable t by a sequence $t_n \to t_\infty$ and u by $u_n \to u_0$. Hence convergence is uniform on bounded u-sets.

2) We next prove that $e^{-\varphi_t}$ converges in $\mathbf{L}^1(H_+)$ to the function $e^{-u^T u/2} e^{-v}$. For $\mathbf{L}^1$-convergence we have to study the level sets of the normalized densities on H_+. These have the form

$$F_{t,q} = H_+ \cap \{\varphi_t < \psi(t + qa_t) - \psi(t)\} \qquad t, q > 0.$$

For fixed $q > 0$ the sets $F_{t,q}$ tend to the parabolic cap $\{0 \le v + u^T u/2 < q\}$ since $\varphi_t(w) \to v + u^T u/2$ and $\psi(t + qa_t) - \psi(t) \to q$ by (5.2). Note that $F_{t,q}$ is contained in the slice $\{0 \le v \le q\}$ since $\varphi_t(0, q) = \psi(t + qa_t) - \psi(t)$. We shall prove that there exists $t_{00} < t_\infty$ and $q_0 > 0$ so that $F_{t,q} \subset C_q$ for $t \in [t_{00}, t_\infty)$ and $q \ge q_0$, where C_q is the cylinder segment

$$C_q = \{(u, v) \mid \|u\| \le q, 0 \le v \le q\} \subset H_+.$$

Let $\epsilon > 0$, and let $t \in [t_\epsilon, t_\infty)$ with t_ϵ as in Lemma 5.2. Choose q_n so that $\psi(t + q_n a_t) = \psi(t) + n$. Then $q_n \le e^{\epsilon n}/\epsilon$ for $n = 1, 2, \dots$ by Lemma 5.2 and $e^{-\varphi_t} \le e^{-n}$ outside F_{t,q_n}, and hence outside the compact set C_{q_n} for $t \ge t_{00}$ and $q_n \ge q_0$. We conclude that the functions $e^{-\varphi_t}$ are uniformly bounded outside the compact set C_{q_0} by the step function τ with value e^{-m} on $C_{q_{m+1}} \setminus C_{q_m}$, $m \ge 0$. Since the set C_q has volume $|C_q| = B_h q^d$ where B_h is the volume of the unit ball in $\mathbf{R}^h$, we see that $|C_{q_n}| \le M_0 e^{\epsilon d n}$ and hence the step function τ is integrable if we choose $\epsilon < 1/d$.

It remains to prove the inclusion $F_{t,q} \subset C_q$ for $t \ge t_{00}$ and $q \ge q_0$.

There exists $\delta_0 \in (0, 2)$ so that (5.5) holds for $\delta = \delta_0$. This follows since the concave function γ_+ which parametrizes the upper boundary of D satisfies the relation $1 - \gamma_+(x) \sim x^T x/2$ for $x \to 0$ by (5.7). Set $\epsilon_0 = \delta_0^2/16$. By Lemma 5.3 the set $D \cap H_+$ lies below the parabola $y = 1 - \epsilon_0 x^T x$, and hence $rD \cap H_+$ lies below the parabola $y = r - \epsilon_0 x^T x/r$ for $r > 0$. Let $r = t + a_t q > t$. Then $F_{t,q}$ is the image of $rD \cap \{y \ge t\}$ under the map α_t^{-1} where $\alpha_t(u, v) = (b_t u, t + a_t v)$ with $b_t^2 = ta_t$. Hence $F_{t,q}$ lies below the parabola $v = q - \epsilon_0 u^T ut/r$. This parabola intersects the coordinate plane $\{v = 0\}$ in a circle with radius R where R satisfies

$$q + (a_t/t)q^2 = \epsilon_0 R^2.$$

If $t \geq t_{00}$ then $a_t/t \leq \epsilon_0/2$. If moreover $q \geq 2/\epsilon_0$ then $R \leq q$, which gives $F_{t,q} \subset C_q$.
3) Finally we prove the second part of the proposition.

Let C be the closed cone $\|x\| \leq \delta y$ and assume $\delta > 0$ is so small that $x \mapsto n(x,1)$ is C^2 on $\|x\| < 2\delta$ and the second derivative is positive definite for $\|x\| \leq \delta$. (We know that $n_{xx}(0) = I$ by condition 3) and that n_{xx} is continuous on a neighbourhood of the origin.)

Let K be the compact set of initial linear maps $A \in J$ mapping $(0,1)^T$ onto p for $p \in C \cap \partial D$. Recall from calculus that $(n \circ A)''(z)(u,v) = n''(Az)(Au, Av)$ for $z \in C \setminus \{0\}$ and $u, v \in \mathbf{R}^d$. Set $n_A(x) = (n \circ A)(x)$. Then $(A, x) \mapsto n''_A(x)$ is continuous in $(A, x) \in K \times \{\|x\| \leq \delta\}$. Since $n''_A(0) - I$ vanishes identically on K there exists for any $\epsilon > 0$ a constant $\delta_\epsilon > 0$ so that

$$\|n''_A(x) - I\| < \epsilon \qquad \|x\| < \delta_\epsilon,\ A \in K.$$

Choose a continuous increasing function $\eta : [0, \delta] \to [0, \infty)$ so that $\eta(0) = 0$ and $\eta(\delta_\epsilon) \geq \epsilon$ for $\epsilon > 0$. Then

$$f_0(\|x\|) \leq n_A(x) \leq f_1(\|x\|) \qquad \|x\| \leq \delta,\ A \in K$$

if we choose $f''_0 = 1 - \eta$ and $f''_1 = 1 + \eta$ with the initial conditions $f_i(0) = f'_i(0) = 0$ for $i = 0, 1$. Let $F(x, y) = y f_0(x)$ for $y > 0$ and define γ implicitly on a disk $\|x\| < \delta_2$ by $F(x, \gamma(x)) = 1$. Then $\gamma_+(x) \leq \gamma(x) \leq 1$ holds for $\|x\| < \delta_2$, where γ_+ is the concave function which parametrizes the upper boundary of D, and $1 - \gamma(x) \sim x^T x/2$. Together with Lemma 5.3 this implies that $H_+ \cap A^{-1}D$ lies below a parabola $y = 1 - \epsilon x^T x$ uniformly in $A \in K$. Hence (with the obvious notation) $W_A^t \Rightarrow W$ for $t \to t_\infty$ uniformly for $A \in K$. ¶

Remark 5.5 The normalizing transformations β_H in Theorem 5.1 are only determined up to asymptotic equality and modulo rotations around the vertical axis. If $\sigma_H(u, v) = (Su, v)$ for some $S \in \mathcal{O}(h)$, the orthogonal group on $\mathbf{R}^h$, and if β'_H maps H_+ onto H and $\beta_H^{-1}\beta'_H \to \mathrm{id}$, then $\gamma_H^{-1}(Z^H) \Rightarrow W$, where $\gamma_H = \beta'_H \circ \sigma_H$, and γ_H maps H_+ onto H. We find it convenient to choose

$$\beta_H = \beta_t \circ A_H \qquad \beta_t(u, v) = (b_t u, t + a_t v) \qquad a_t = 1/\psi'(t),\ b_t = \sqrt{t a_t}. \tag{5.9}$$

The constants a_t and b_t were introduced in (5.8). We choose $t > 0$ so that the half space H is tangent to the rotund set tD. (We assume $0 \notin H$.) By strict convexity of rotund sets there exists a unique point $p = p_H$ so that $p \in H$, $t = n(p)$, and H is disjoint from $\{n < t\}$. The map $A = A_H : \mathbf{R}^h \times \mathbf{R} \to \mathbf{R}^d$ is an initial linear map. It maps $(0,t)^T$ into p, the half space $H^t = \{y \geq t\}$ onto H, and the gauge function $n \circ A$ of $A^{-1}(D)$ satisfies $(n \circ A)(1,x) - 1 \sim x^T x/2$ for $x \to 0$. The initial map A is unique modulo rotations. Locally it may be choosen to depend continuously on the direction of the half space H. The compact set J of all initial maps was introduced above. It is a non-trivial fibre bundle over the group $\mathcal{O}(h)$.

The proof of Theorem 5.1 now is straightforward. Let H_n be a sequence of half spaces so that $P\{Z \in H_n\} \to 0$. Since the unit sphere is compact we may assume that the direction of H_n converges to some unit vector θ_0. The initial map associated with the half space with direction θ_0 introduces coordinates (x,y) so that Proposition 5.4 applies: $n_D(x,1) - 1 \sim x^T x/2$ for $x \to 0$ and n_D is C^2 on a neighbourhood of $(x,y) = (0,1)$. In the new coordinates the direction of H_n converges to the vertical direction $(0,1)$, and the second part of Proposition 5.4 ensures that $\beta_n^{-1}(Z^{H_n}) \Rightarrow W$ where W has a standard Gauss–exponential distribution and the affine normalization β_n maps H_+ onto H_n.

6 Flat functions

In this section we extend the class of densities of Theorem 5.1 of the form

$$f_0 = e^{-\psi \circ n}/C_0. \tag{6.1}$$

The density f_0 above satisfies the limit relation

$$(f_0 \circ \beta_H)(w)/(f_0 \circ \beta_H)(0) \to e^{-v-u^T u/2} \tag{6.2}$$

uniformly on compact w-sets in $\mathbf{R}^d$, and in $\mathbf{L}^1(H_+)$, with β_H as defined in (5.9). Let f be a continuous positive density on $O = \{f_0 > 0\}$ which satisfies the same limit relation with the same normalizations β_H. Then one may write $f = Lf_0$ where L is a continuous positive function on O which satisfies

$$(L \circ \beta_H)(w)/(L \circ \beta_H)(0) \to 1 \qquad H \to \partial_O \tag{6.3}$$

uniformly on compact w-sets.

Relation (6.3) states that far out in O the function L locally behaves like a positive constant. Such a function will be called *flat*. The condition (6.3) is weaker than asymptotic equality. If L satisfies (6.3) uniformly on compact sets then $f = Lf_0$ satisfies (6.2) uniformly on compact sets. If Lf_0 is integrable, then one may define the probability density $f_L = Lf_0/C_L$, and ask whether the associated probability distribution lies in the domain of the Gauss-exponential law. In this section we shall show that that is the case. In addition we shall give a simple sufficient condition for (6.3) in terms of partial derivatives. As a concrete application we shall show that the multivariate hyperbolic distributions belong to the domain of attraction of the Gauss-exponential high risk limit law.

We start by looking at densities of the form $L(z)f_0(z)/C$ where f_0 is a Gaussian density and L a continuous function which behaves locally as a positive constant for $\|z\| \to \infty$. The question is, how much can one alter a Gaussian density while retaining the high risk limit behaviour (with the original normalizations β_H).

Example 6.1 Let f_0 be an arbitrary Gaussian density on $\mathbf{R}^d$. For $z \in \mathbf{R}^d$ write $z = r\theta$ where $r = \|z\|$ and θ is a unit vector. The function rf_0 is integrable and the corresponding density lies in the domain of the Gauss-exponential distribution. This is also true for $r^c f_0$ for any $c > 0$, and for $Q^c f_0$ where Q is a positive quadratic function on $\mathbf{R}^d$. The function $e^{\sqrt{r}} f_0$ normalized to a probability density also lies in the domain of attraction, and so does $\chi(\theta) f_0$ for any continuous positive function χ on the unit sphere. Proofs are given below. ¶

In the first part of this section we consider a class $\mathcal{L}$ of continuous positive functions L on $\mathbf{R}^d$ which satisfy a growth condition which may be regarded as the multivariate additive version of slow variation:

$$L(z+w)/L(z) \to 1 \qquad \|z\| \to \infty, \quad w \in \mathbf{R}^d. \tag{6.4}$$

If $\lambda : \mathbf{R}^d \to \mathbf{R}$ is a C^∞ function whose partial derivatives vanish in infinity then $L = e^\lambda$ satisfies (6.4). The converse is also true. Continuity of L implies that if (6.4) holds pointwise it holds uniformly on bounded w-sets, see Bingham et al. [6]. Let $\lambda = \log L$, and define $\lambda_0 = \pi_0 * \lambda$ as the convolution of λ with a C^∞ probability density π_0 with compact support. Then

$$\lambda_0(z) - \lambda(z) = \int \pi_0(w)(\lambda(z-w) - \lambda(z))dw \to 0$$

and the partial derivative of λ_0 of any order in a point z_0 is the convolution of the difference $\lambda(z) - \lambda(z_0)$ with the corresponding partial derivative of π_0. Hence it vanishes in infinity. Thus we have constructed a C^∞ function λ_0 whose partial derivatives of all orders vanish in infinity such that e^{λ_0} is asymptotic to L.

In the univariate case the functions

$$r, \quad r^7, \quad e^{\sin\sqrt{r}}, \quad e^{\sqrt{r}}, \quad e^{-\sqrt{r}}, \quad r^5 e^{\sqrt{r}\sin r^{1/3}}$$

all satisfy the functional relation

$$f(r+s)/f(r) \to 1 \qquad r \to \infty \tag{6.5}$$

for each $s \in \mathbf{R}$. Functions in $\mathcal{L}$ satisfy the relation $L(\theta r + \theta s)/L(\theta r) \to 1$, $r \to \infty$, in every direction $\theta \in S$.

In the multivariate setting one would like to know how the limit relations in different directions are coordinated. The answer is surprising. Given any countable family F of continuous positive functions f on $(0, \infty)$ which satisfy (6.5), for instance the family

$$\alpha r^\gamma e^{\beta r^\theta \sin(r^\eta)} \qquad \alpha, \beta, \gamma, \theta, \eta \in \mathbf{Q},\ \alpha, \theta > 0, \eta \geq 0, \theta + \eta < 1,$$

there exists a function $L \in \mathcal{L}$ with the property: for each $f \in F$ there is a dense set S_f of directions θ in the unit sphere S such that

$$L(\theta r)/f(r) \to 1 \qquad r \to \infty,\ \theta \in S_f,\ f \in F.$$

The construction will be given elsewhere. Here we only want to warn the reader that the functions L which satisfy (6.4) are not as tame as they may seem.

Theorem 6.2 *Let the vector $Z \in \mathbf{R}^d$ have density $f_0(z) = e^{-n(z)^2/2}/C$ where n is the gauge function of a rotund set D, and let β_H satisfy (5.9). Suppose $L \in \mathcal{L}$. Then the product $f = Lf_0$ is integrable and satisfies (6.2) uniformly on bounded w-sets, and in $\mathbf{L}^1(H_+)$. In particular the random vector with density Lf_0/C_L lies in the domain of attraction of the Gauss-exponential law with the same normalization as f_0, and the normalized densities converge in $\mathbf{L}^1$.*

Proof We give only a sketch since we shall prove a more general result in Theorem 6.4 below. The function $\varphi = n^2/2$ has second derivative $nn'' + n'(n')^T$ which is positive definite and homogeneous of degree zero. So $(\varphi - \lambda)''$ is positive definite outside a compact set if we choose λ, with $e^\lambda \sim L$, to be C^∞ with partial derivatives of all orders vanishing in infinity. This means that the densities g_H of the normalized high risk scenarios are strongly unimodal. Hence Lf is integrable, and pointwise convergence in (6.2) implies convergence in $\mathbf{L}^1$.

By (5.9) we may write $\beta_H = \beta_t \circ A^{-1}$ for an initial linear transformation $A \in J$. Since J is compact, for any ball B^r the diameters of the ellipsoids $\beta_H(B^r)$ with center $p_H = \beta_H(0) \to \infty$ are uniformly bounded, and hence (6.3) holds uniformly on compact w-sets. This yields (6.2) for f_L. ¶

We shall now prove a more general result for the densities $f_0 = e^{-\psi \circ n}/C_0$ introduced in Theorem 5.1. Instead of strong unimodality our proof makes use of the fact that $f_0^{1/2}$ is integrable and the corresponding density belongs to the domain of attraction of a high risk limit law. These conditions are satisfied since $c\psi$ satisfies the same three conditions as ψ for all $c > 0$.

First we have to say more precisely what we mean by a flat function, see (6.3). We shall give a formal definition which can also be used in the case of heavy tailed limit laws, or limit laws with bounded support, see Sections 8 and 9. We shall also have to ponder on the relation between the convergence relation $(f \circ \beta_H)(w)/(f \circ \beta_H)(0) \to h(w)$ uniformly on compact sets, see (6.1) and weak convergence of the associated probability distributions on H_+.

Definition Let Z lie in the domain of attraction of W, i.e. $\beta_H^{-1}(Z^H) \Rightarrow W$ for $0 < P\{Z \in H\} \to 0$. Assume that the high risk limit vector W has density g on H_+, that $\{g > 0\}$ is a convex set, open in H_+, and that g is continuous on $\{g > 0\}$. A function L is *flat for* Z if L is defined, positive and continuous on the interior O of the convex support of Z, or on $O \setminus K$ for some compact subset K of O, if $P\{Z \in O\} = 1$, and if (6.3) holds uniformly on compact subsets of $\{g > 0\}$. ¶

Remark 6.3 If L and L_0 are flat for Z then so are LL_0 and L^t for $t \in \mathbf{R}$.

We can now formulate the main result of this section.

Theorem 6.4 *Let Z have density $f_0 = e^{-\psi\circ n}/C_0$ as in Theorem 5.1. If L is flat for Z then the product Lf_0 is integrable and the random vector with density Lf_0/C lies in the domain of attraction of the Gauss-exponential distribution.*

The density $f_0 = e^{-\psi\circ n}/C_0$ of Theorem 5.1 has the following properties: f_0 is continuous on $\mathbf{R}^d$, $f_0(z) < f_0(0)$ for $z \neq 0$, the sets $\{f_0 > c\}$, $0 < c < f(0)$, are strictly convex, and for each point p with $f_0(p) = c$ there is a unique closed half space H_p containing p and disjoint from $\{f_0 > c\}$. The collection of all densities f with these properties will be denoted by $\mathcal{F}_0$.

For $f \in \mathcal{F}_0$ there is a one to one correspondence between closed half spaces H which intersect the domain $O = \{f > 0\}$ and do not contain the origin, and points $p \in O \setminus \{0\}$, given by $H = H_p$. If f is the density of a random vector Z then the condition $0 < P\{Z \in H\} \to 0$ is equivalent to $H = H_p$ and $p \to \partial_O$. This one to one correspondence is used in the proof of our next result.

Proposition 6.5 (Basic inequality) *Let Z have density $f = e^{-\varphi} \in \mathcal{F}_0$. Let $O = \{f > 0\}$, and for each $p \in O \setminus \{0\}$ let $\beta_p \in \mathcal{A}$ map H_+ onto H_p and 0 into p. Assume that*

$$h_p(w) = (f \circ \beta_p)(w)/f(p) \to h(w) \qquad p \to \partial_O \tag{6.6}$$

uniformly on compact subsets of H_+, and in $\mathbf{L}^1$, for some continuous integrable positive function h on H_+. Let W have density $g = h/C$ on H_+.

Assume that $L = e^\lambda$ is positive and continuous on O, and

$$L_p(w) = (L \circ \beta_p)(w)/L(p) \to 1 \qquad p \to \partial_O,\ w \in H_+ \tag{6.7}$$

uniformly on $\{h \geq c\}$ for some $c < 1$.

Then Z lies in the domain of attraction of the high risk limit vector W, L is flat for Z and for all $\epsilon > 0$ there exists $\delta > 0$ so that

$$|\lambda(z) - \lambda(p)| < \epsilon + \epsilon(\varphi(p) - \varphi(z)) \qquad z \in H_p,\ 0 < f(p) < \delta. \tag{6.8}$$

Proof Let $\epsilon > 0$. Write $K = \{h \geq c\} \subset H_+$. First observe that the function h has convex level sets $\{h > t\}$ on H_+ since this holds for the functions h_p. Uniform convergence $h_p \to h$ on K implies that there exists $\delta_0 > 0$ so that $h_p(w) < \sqrt{c}$ for $h(w) = c$ and $f(p) < \delta_0$. Hence

$$f(p) < \delta_0, z \in H_p \cap O, f(z)/f(p) \geq \sqrt{c} \ \Rightarrow\ \beta_p^{-1}(z) \in K.$$

By assumption there exists $\delta > 0$ so that

$$f(p) < \delta, z \in H_p \cap O, \beta_p^{-1}(z) \in K \;\Rightarrow\; |\lambda(z) - \lambda(p)| < \epsilon.$$

For $p \in O \setminus \{0\}$, $z \in H_p \cap O$, choose points $p_0, \ldots, p_n$ on the line segment $[p, z]$ so that $p_0 = p$, $f(p_k) = \sqrt{c} f(p_{k-1})$ for $k = 1, \ldots, n$, $f(z) \geq \sqrt{c} f(p_n)$, and set $p_{n+1} = z$. Now suppose $f(p) < \delta_1 = \delta_0 \wedge \delta$. Then $h(\beta_{p_k}^{-1}(p_{k+1}) \geq c$ for $k = 0, \ldots, n$. Hence

$$|\lambda(z) - \lambda(p)| < \epsilon + \epsilon_0(\varphi(z) - \varphi(p)) \qquad \epsilon_0 = \epsilon / \log(1/\sqrt{c}), z \in H_p \cap O.$$

¶

Proof (of Theorem 6.4) Flatness implies (6.1) for Lf uniformly on compact subsets of H_+. The basic inequality gives $L(z)/L(p) \leq 2(f(p)/f(z))^\epsilon$ for $z \in H_p$, $f(p) < \delta$. Hence Lf is integrable if f^c is integrable for some $c < 1$. If ψ satisfies the three condtions of Theorem 5.1, then so does $c\psi$ for any $c > 0$. Take $c = 1/2$. So $e^{-\psi \circ n}/C_c$ and f_0 both lie in the domain of the Gauss-exponential law. We may use the same normalizations by Theorem 3.8 on power families, and we may choose the normalizations β_p of f_0 by the convergence of types theorem. Hence we have $\mathbf{L}^1$ convergence of the quotients $h_p^{1/2}$. By dominated convergence $((Lf) \circ \beta_p)(w)/(Lf)(p) \to h$ in L^1. ¶

In the remainder of this section we develop the theory of flat functions.

Let us first give a geometric formulation of condition (6.3). Let B be the open unit ball, and B^r the ball of radius r. A given positive continuous function $L = e^\lambda$ on O is flat if it is asymptotically constant on the ellipsoids $E_p = \beta_p(B)$ where $\beta_p = \beta_{H_p}$ is the normalization defined in (5.9).

The ellipsoids $E_p^r = \beta_p(B^r)$, $p \in O \setminus \{0\}$, determine a Riemannian structure on $O \setminus \{0\}$, and hence a metric d. Roughly the distance between two points p and q in O is $2nr$ for $r > 0$ small, where $n = n_r$ is the number of ellipsoids E_z^r needed to form a chain from p to q. The function e^λ is flat if $\lambda(p) - \lambda(q)$ vanishes for $p \to \partial_O$ and $d(p, q)$ is bounded.

If Z has density $f = e^{-\psi \circ n}/C$ as in Theorem 5.1 then by (5.9) the ellipsoids E_p have the form $A \circ \beta_t(B)$ where A is an initial map and $\beta_t(u, v) = (b_t u, t + a_t v)$ with $a_t = 1/\psi'(t)$ and $b_t = \sqrt{t a_t}$. Since $a_t = o(t)$ for $t \to t_\infty$ we see that

$$a_t << b_t << t \qquad t \uparrow t_\infty.$$

Hence the diameter of the ellipsoids $\beta_t(B)$ vanishes for $t \to t_\infty$ if t_∞ is finite, and else it is $o(t)$. Since the set J of initial maps is compact, this also holds for the ellipsoids E_p. Thus the ellipsoids E_p are like buttons attached to the surface of the rotund set tD where $t = n(p)$.

Example 6.6 (Functions which are flat for all densities $f_0 = e^{-\psi \circ n}/C_0$)

1) If t_∞ is finite then $O = \{f_0 > 0\}$ is a bounded set. Let L be a continuous positive function, defined on an open neighbourhood V of ∂O. The restriction of L to $O \cap V$ is flat.

2) Let χ_0 be a positive continuous function on the unit sphere S in $\mathbf{R}^d$. Then $z \mapsto \chi(z) = \chi_0(z/\|z\|)$ on $\mathbf{R}^d \setminus \{0\}$ is homogeneous of degree zero. and flat for all f_0 as in (6.1). So is the function $z \mapsto r = \|z\|$. ¶

Recall that the standardized generalized multivariate hyperbolic distribution has density

$$f(z) \propto (1 + z^T z)^{c_1} K_{\lambda - d/2}(c_2\sqrt{1 + z^T z})e^{\gamma z} \qquad z \in \mathbf{R}^d. \tag{6.9}$$

Here K_ν denotes a modified Bessel function of the third kind with index ν, and the constants $c_1 \in \mathbf{R}$, $\lambda \in \mathbf{R}$, $c_2 > 0$, and $\gamma \in \mathbf{R}^d$ with $\|\gamma\| < c_2$ are parameters, with $c_1 = (\lambda - d/2)/2$.

Proposition 6.7 *The domain of the Gauss-exponential limit law contains the hyperbolic densities (6.9).*

Proof The asymptotic behaviour of the Bessel function does not depend on the index. For any $\nu \in \mathbf{R}$

$$K_\nu(t) \sim \sqrt{\pi/2t}e^{-t} \qquad t \to \infty.$$

Since $e^{-c\sqrt{1+z^T z}} \sim e^{-cr}$ for $\|z\| = r \to \infty$ we see that

$$f(z) \sim c_0 r^{2c_1 - 1/2} e^{-c_2 r} e^{\gamma z} = c_0 r^{2c_1 - 1/2} e^{-n(z)} = L(z) e^{-n(z)} \qquad r = \|z\| \to \infty$$

where n is the gauge function of the rotund set $D = \{z \in \mathbf{R}^d \mid c_2^2 z^T z < 1 + \gamma z\}$, an ellipsoid, which is excentric for $\gamma \neq 0$, and L is flat for e^{-n} by Remark 6.3 and the example above. ¶

The ellipsoids associated with f_0 have a simple geometric description in terms of the level sets of f_0. Translate the ellipsoid E_p along the ray through p towards the origin until the half space H_p is tangent to $E_p - (p - q)$. The center q of the shifted ellipsoid satisfies $f_0(q)/f_0(p) \to e$ for $p \to \partial_O$; the boundary of the shifted ellipsoid and the surface $\{f_0 = f_0(p)\}$ agree up to terms of the second degree around the point p. These two conditions determine the ellipsoids E_p asymptotically. They are an obvious translation of the fact that the unit sphere and $\partial(A^{-1}D)$ osculate in $(0,1)^T$ for all initial maps $A \in J$.

For the density $e^{-v-u^Tu/2}$ of the Radon measure ρ_0^* one may define a family of ellipsoids $E_w^* = \sigma_w(B)$ where σ_w is the symmetry of ρ_0^* mapping the origin into w, see Section 4. This family has the same geometric description as above. The limit relation $h_p \to e^{-v-u^Tu/2}$ also holds for the first and second derivatives of $-\log h_p$ for $p \to \partial_O$ and this ensures that $\beta_p^{-1}(E_q) \to E_w^*$ for $q = \beta_p(w)$, $p \to \partial$.

The ellipsoids E_w^* along a given vertical line are translations of each other (since the symmetry group $\mathcal{S}$ of ρ_0^* contains the translations). By symmetry we also have the equivalence $p \in E^{*r}_q \iff q \in E^{*r}_p$ for all $r > 0$ if p and q lie on the same parabolic surface $v + u^Tu/2 = c$. These relations hold asymptotically for the family of ellipsoids E_p^r. Hence we have:

Proposition 6.8 *For $\epsilon > 0$ there exists $\delta_0 > 0$ so that for $f_0(p) \in (0, \delta_0]$*

$$\begin{aligned} E_p^c &\subset \{e^{-c-\epsilon} f_0(p) < f_0 < e^{c+\epsilon} f_0(p)\} \qquad c \in (0,1] \\ p \in E_q^{3/2} &\;\; \& \;\; f_0(q) = f_0(p) \;\Rightarrow\; q \in E_p^2. \end{aligned} \tag{6.10}$$

Proof Set $t = n(p)$. Then by (5.3)

$$\psi(t) - c - \epsilon < \psi(t - ca_t) < \psi(t) = f_0(p) < \psi(t + ca_t) < \psi(t) + c + \epsilon.$$

In the limit the unit ball fits between the parabolic surfaces $v = 1 - u^Tu/2$ and $v = -1 - u^Tu/2$. In the point $(0,1)$ the sphere and paraboloid osculate. For $c < 1$ eventually $E_p^c \subset \{n < t + ca_t\}$ since the curvature of the ellipsoid in the common boundary point is larger for the ellipsoid E_p^c than for the surface $\{n = t + ca_t\}$. ¶

Functions $L = e^\lambda \in \mathcal{L}$ might be called flat for the collection of balls $p+B$, $p \in \mathbf{R}^d$, for $\|p\| \to \infty$, or for the Euclidean metric, since $\lambda(p) - \lambda(q)$ vanishes for $\|p\| \to \infty$ if $\|q - p\| < 1$. For a function which is flat for f_0 the role of the balls $p + B$ is taken over by the ellipsoids $E_p = \beta_p(B)$, for $p \to \partial_O$, and the role of the Euclidean metric by the Riemannian metric on $O \setminus \{0\}$ induced by these ellipsoids. For the class L there was a simple characterization in terms of partial derivatives: if λ is C^1 and the first order partial derivatives of λ vanish in infinity then $L \in \mathcal{L}$. We shall now formulate a similar sufficient condition for flat functions. Since the ellipsoid $\beta_t(B)$ is the image of the unit ball under the simple transformation $(u, v) \mapsto (b_t u, t + a_t v)$ it is centered in $(0, t)^T$, has semi-axis a_t in the vertical direction, and b_t in the horizontal direction. So our conditions are

$$\lambda_r(p) = o(1/a_t) = o(\psi'(t)) \qquad \lambda_t(p) = o(1/b_t) = o(\sqrt{\psi'(t)/t}) \qquad t \to t_\infty. \quad (6.11)$$

Here λ_r is the radial partial derivative and λ_t any tangential partial derivative along the boundary of the rotund set tD. It is convenient to define $\lambda_T(p)$ as the maximum of the directional derivative $|\partial_u \lambda(p)|$ over all unit vectors u parallel to ∂H_p. Note that (6.11) depends on the behaviour of ψ, which is obvious, but also on the shape of D, since that determines the direction of the tangent vectors.

Example 1) Suppose Z has a standard Gaussian distribution. Then $n(z) = \|z\|$ and $\psi(t) = t^2/2$. Hence $a_t = 1/t$ and $b_t = 1$. In particular all $L \in \mathcal{L}$ are flat, and so are the functions e^r, $\exp(r^c)$ with $c \in [1, 2)$, and for instance $r^7 e^{2r^{4/3} \sin \sqrt{r}}$. Flat functions may also depend on the direction $\theta = z/r$, but if L increases like e^r, or faster, along a ray $z = r\theta_0$ then the condition in the tangential direction imposes a strict discipline on the behaviour of the function L along different rays: The asymptotic equality $L(r\theta) \sim L(r\theta_0)$ for $r \to \infty$ has to hold uniformly in all unit vectors θ.

2) Suppose Z has density $e^{-n(z)}/C$ where n is the Gauge function of the rotund set D. Now $a_t \equiv 1$ and $b_t = \sqrt{t}$. The function $e^{\sqrt{r}}$ is flat for Z only if D is a ball.¶

Theorem 6.9 *Let Z have density $f = e^{-\psi \circ n}/C$ as in Theorem 5.1. Let K be a compact subset of $O = \{f > 0\}$ containing the origin, and let $\lambda : O \setminus K \to \mathbf{R}$ be a C^1 function whose partial derivatives satisfy (6.11). Then $L = e^\lambda$ is flat for Z.*

Proof First observe that the functions $z \mapsto n(z)$ and $z \mapsto \psi'(n(z))$ are asymptotically constant on the ellipsoid E_p for $p \to \partial_O$ by the first relation in (6.10) since t and $\psi'(t)$ are asymptotically constant on intervals over which ψ increases by a fixed amount. This follows from the three conditions on ψ in theorem 5.1. (Use $a_t = o(t)$ and (5.2), and (5.3).)

One can go from p to any point q in E_p by first going to the point p' on the ray through p with $n(p') = n(q)$ and then moving to q along the surface $\{n = n(q)\}$. Let $n(p) = t$ and suppose $E_p = \beta_t(B)$. Then $\|p' - p\| \leq a_t$ and the length of the curve from p' to q is bounded by $b_t + 2a_t$ since the curve may be represented by a decreasing function. Hence the oscillation of λ over E_p vanishes for $p \to \partial_O$. This remains true in the general case $E_p = (A \circ \beta_t)(B)$ since the set J of initial maps A is compact. ¶

There is another reason for being interested in the family of ellipsoids $E_p = \beta_p(B)$. From the univariate limit theory for exceedances and extremes it is known that the normalization coefficients give important information about the limit distribution and its domain. One might try to develop a theory of regular variation based on quotients $\beta_p^{-1}\beta_q$ for q so close to p that E_p and E_q intersect, and $p \to \partial_O$. The role of the power function $x \mapsto x^r$ is now taken over by the group $\mathcal{S}$ of symmetries of the measure ρ^* associated with the limit distribution. Unfortunately the structure of the set J of initial maps makes it impossible to choose the transformations β_p to depend continuously on p for $p \in O \setminus \{0\}$ in dimension $d > 2$. The ellipsoids $E_p = \beta_p(B)$ with β_p defined in (5.9) depend continuously on p and may be regarded as geometric objects reflecting the affine transformations β_p.

7 Flat measures

So far our life was simple because the vectors Z had a well-behaved density f, yielding the limit relation

$$(f \circ \beta_p)(w)/f(p) \to e^{-v-u^T u/2} \qquad p \to \partial_O,\ O = \{f > 0\}$$

uniformly on compact sets of $\mathbf{R}^d$ (and in $\mathbf{L}^1(H_+)$). Such simple densities are in accordance with our basic assumption that the underlying sample cloud is bland, consisting of a dark convex central region surrounded by a homogeneous halo. A distribution with a density of the form $f_0 = e^{-\psi \circ n}/C_0$ as in Theorem 5.1, or of the form $f = Lf_0/C$ with L flat for f_0 may fit such a data set well.

Our basic limit relation (3.1) is phrased in terms of weak convergence. Hence the theory developed in this paper also should allow discontinuous densities or even discrete probability distributions.

Example Suppose the random vector $Z \in \mathbf{Z}^d$ has distribution $P\{Z = k\} \sim f(k)$ for $\|k\| \to \infty$ where $f = Lf_0$ with $f_0 = e^{-\psi \circ n}/C_0$, and L flat for f_0. Assume $t_\infty = \infty$ and $\psi'(t)$ vanishes for $t \to \infty$. This implies that the ellipsoids E_p will contain arbitrary large balls as $\|p\| \to \infty$. So asymptotically the counting measure on the integer lattice will behave like Lebesgue measure on these large ellipsoids. Does this imply that Z lies in the domain of attraction of the Gauss-exponential limit law? ¶

In this section we prove a theorem which will handle the question above.

Let Z have a probability distribution of the form

$$d\pi = f_0 d\mu$$

where μ is a roughening of Lebesgue measure for the family of ellipsoids $E_p = \beta_p(B)$ associated with the density $f_0 = e^{-\psi \circ n}/C_0$. What this means will be explained below. Let us assume that the normalizations $\beta_p = \beta_{H_p}$ associated with f_0 in (5.9) may be used to normalize the distribution π^H of the high risk scenario Z^H

$$\beta_H^{-1}(d\pi^H)(w) = (f_0 \circ \beta_H)(w)\beta_H^{-1}(d\mu)/C_H \to e^{-v-u^T u/2} dw/(2\pi)^{h/2}$$

weakly for $H = H_p$, $p \to \partial$. We also have

$$f_0 \circ \beta_H(w)/f_0 \circ \beta_H(0) \to e^{-v-u^T u/2}$$

uniformly on compact sets by Theorem 5.1. We thus see that

$$\beta_H^{-1}(d\mu)/C'_H \to d\lambda \qquad C'_H = (2\pi)^{h/2} C_H/(f_0\beta_H)(0)$$

where λ is Lebesgue measure on H_+ and $\to$ denotes vague convergence. Similarly the last limit relation together with the second implies the first in the sense of vague convergence.

Let us call a Radon measure μ_0 on $\mathbf{R}^d$ a *roughening* of Lebesgue measure for the Euclidean norm if there exists a countable partition $\mathcal{F}$ of $\mathbf{R}^d$ into bounded Borel sets F such that $\mu_0 F/\lambda F \to 1$ and such that the diameter of F_p vanishes for $\|p\| \to \infty$

where F_p is the element of the partition $\mathcal{F}$ containing the point p. The measure μ_0 translated over z will converge vaguely to Lebesgue measure for $\|z\| \to \infty$ in any direction. Using this simple concept as a guide we now define the asymptotic relation we are really interested in:

Definition A Radon measure μ on $O = \{f_0 > 0\}$ is called a *roughening of Lebesgue measure for* f_0 if there exists a countable partition $\mathcal{F}$ of O into Borel sets F such that

1) $\lambda F > 0$ and $\mu F/\lambda F \to 1$, and

2) for all $\epsilon > 0$ there exists a compact set $K \subset O$ so that $F_p \in E_p^\epsilon$ for $p \in O \setminus K$.

The measure μ is called *flat for* f_0 if 1) above is replaced by the condition

1') $\lambda F > 0$ and $\mu F(p)/\lambda F(p) \sim L(p)$ for $p \to \partial_O$ for a function L, flat for f_0. ¶

We can now state our theorem.

Theorem 7.1 *Suppose* $f_0 = e^{-\psi \circ n}/C_0$ *satisfies the conditions of Theorem 5.1. Let* L *be flat for* f_0*, and let* μ *be a measure on* $O = \{f_0 > 0\}$ *which is flat for* f_0*. Set* $f = Lf_0$*. Then* $fd\mu$ *is a finite measure and the corresponding probability measure* $fd\mu/C$ *lies in the domain of attraction of the Gauss-exponential limit law with the normalizations* β_H *of* f_0 *given in (5.9).*

We need to do some preliminary work to prove this result.

First we define the canonical probability measures γ_t on the sets $\{n = t\}$. If D is the unit ball then γ_t is the uniform distribution on the sphere of radius t.

Let X be uniformly distributed on D. Set $N = n(X)$. Then the vector $\Xi = X/N$ lies in ∂D and is independent of N. Let γ denote its distribution. Then γ is a probability distribution on $\mathbf{R}^d$ which lives on ∂D. The definition of γ does not depend on the coordinates, as long as the origin is kept fixed. For any linear map the distribution $A(\gamma)$ of $A\Xi$ is the natural distribution on the boundary of the rotund set AD. We shall write γ_t for the distribution of $t\Xi$ on $\{n = t\}$. The probability measures γ_t, $t > 0$, describe the conditional distribution of Lebesgue measure given n. These may also be obtained by a limiting argument, taking shells whose width decreases to zero. By homogeneity, for any Borel set $F \subset \partial D$

$$\gamma(F) = \lambda\{tz \mid z \in F, s \leq t < 1\}/\lambda(D \setminus sD) \qquad s \in [0, 1).$$

As in the case of polar coordinates, for any non–negative Borel function g_0 on $[0, \infty)$ and any Borel set $F \subset \mathbf{R}^d$

$$\int_F g \circ n d\lambda = (|D|/d) \int_0^\infty t^{d-1} \gamma_t(F) g_0(t) dt.$$

Lemma 7.2 *Let $H_0 = \{\zeta_p \geq 1\}$ be the half space tangent to D in $p \in \partial D$. For $\sigma > 0$ define $C_p^\sigma = \partial D \cap \{\zeta_p \geq 1 - \sigma\}$. There is an $M > 1$ so that*

$$\bar{\gamma}_p(\sigma) = \gamma(C_p^{2\sigma})/\gamma(C_p^\sigma) \leq M \qquad p \in \partial D, \ \sigma > 0.$$

Proof There exists a constant $\sigma_0 \geq 1$ so that $C_p^{\sigma_0} = \partial D$ for all $p \in \partial D$. Hence $\bar{\gamma}_p(\sigma) \equiv 1$ for $\sigma \geq \sigma_0$. Rotundity implies $\gamma(C_p^\sigma) \sim A(p)\sigma^{h/2}$ for $\sigma \to 0$ for some constant $A(p) > 0$. The asymptotic equality is uniform in p since one may normalize the caps C_p^σ to converge to the parabolic cap $Q_+ = \{0 \leq v = 1 - u^T u\}$ uniformly in p, see Proposition 9.1. Hence $\bar{\gamma}_p(\sigma) \to 2^{h/2}$ for $\sigma \to 0$ uniformly in $p \in \partial D$ for $\sigma \to 0$. By continuity the function is bounded. ¶

Our next lemma relates the family of ellipsoids associated with $f_0 = e^{-\psi \circ n}/C_0$ to half spaces. Recall that H_p is the half space tangent to $\{n < n(p)\}$ in p.

Lemma 7.3 *There exists $t_0 < t_\infty$ so that for any points $p, q \in \{n = t\}$, $t \in [t_0, t_\infty)$*

$$q \notin E_p^2 \ \Rightarrow \ E_q^{1/3} \cap H_p = \emptyset.$$

Proof Given q, the coordinates $\beta = \beta_q : \mathbf{R}^h \times \mathbf{R} \to \mathbf{R}^d$ map the origin into q, the half space $\{y \geq 0\}$ onto H_q, and the open ball B^r of radius r around the origin onto the ellipsoid E_q^r. The image $Q_q = \beta_q^{-1}(tD)$ converges to the paraboloid $Q = \{y < -x^T x/2\}$ for $q \to \partial$.

Let $\tilde{Q}$ be the convex hull of the union of Q_q and $B^{1/3}$. Note that in dimension $d = 2$ the tangent line to Q in $(1, -1/2)$ passes through $(0, 1/2)$ and is tangent to $B^{1/\sqrt{8}}$. This implies that the convex hull $\tilde{Q}$ of $Q \cup B^{1/3}$ agrees with Q outside $B^{\sqrt{5}/2}$, also for $d > 2$. Pointwise convergence for convex C^1 functions on an open set U implies convergence of the derivatives uniformly on compact subsets of U. Hence $\tilde{Q}_q$ agrees with Q_q outside the ball $B^{3/2}$ for $n(q) \geq t_0$ for some $t_0 < t_\infty$. Since $\tilde{Q}_q$ is convex the half space H tangent to Q_q in a point p outside $B^{3/2}$ is disjoint from $\tilde{Q}_q$ and ipso facto from $B^{1/3}$. Now use (6.10) to obtain the desired result. ¶

The basic argument is contained in the sequence of inequalities in the proof of the next result.

Proposition 7.4 *Let μ be a roughening of Lebesgue measure for $f_0 = e^{-\psi \circ n}/C_0$ on $O = \{f_0 > 0\} = \{n < t_\infty\}$ with partition $\mathcal{F}$. Let $g = e^{-\chi \circ n}$ on O, where χ is a C^1 function on $[0, t_\infty)$, and let g vanish off O. If $\chi'(t) = O(\psi'(t))$ for $t \uparrow t_\infty$ then there exists a compact set $K \subset O$, and a constant M_0 (depending only on the shape of the rotund set D) so that*

$$I_\mu = \int_{H \cap \{f_0 \leq f_0(p_0)/3c\}} g d\mu \leq M_0 \int_{H \cap \{f_0 \leq f_0(p_0)/2c\}} g d\lambda = M_0 I_\lambda \tag{7.1}$$

for any half space $H = H_{p_0}$ disjoint from K and for any $c \geq 4$.

Proof Write $t_0 = n(p_0)$, $H_0 = \{\xi \geq t_0\}$, and let $\{f_0 = 3f_0(p_0)/2c\} = \{n = t_1\}$ define t_1. Introduce the intermediate terms

$$I_\mu \leq \sum \int_F g d\mu \leq \sum 4 \int_F g d\lambda \leq 4 \int_{\bigcup \tilde{C}_t} g d\lambda \leq 4M \int_{\bigcup C_t} = 4M I_\lambda.$$

Here the sum extends over all $F \in \mathcal{F}$ which intersect the set $H \cap \{f_0 \leq f_0(p)/3c\}$; the union is taken over all $t \geq t_1$; $C_t = \{n = t\} \cap H$; and $\tilde{C}_t$ is the enlarged cap, twice the height of C_t:

$$\tilde{C}_t = \{z \in \{n = t\} \mid \xi z \geq t_0 - s_0\} \qquad t_0 + s_0 = \max\{\xi z \mid z \in C_t\}.$$

There are five relations to prove. The final relation is an equality which holds by definition of t_1. The fourth inequality holds by Lemma 7.2 above. The first inequality is obvious since $g \geq 0$. The second inequality holds since $\mu F \leq 2\lambda F$ for almost all F, and also $\lambda F \max_F g(z) \leq 2 \int_F g d\lambda$. This last inequality holds since $\psi'(n) a_{n(p)} \to 1$ uniformly on E_p for $p \to \infty$ by (6.10) and (5.3). This implies that the oscillation of χ over B_p^ϵ is $O(\epsilon)$ for $p \to \partial$.

It remains to prove the third inequality. Let t_0 be so large that $F_p \subset E_p^{1/6}$ for $n(p) \geq t_0$. Let $c \geq 4$, and let $F \in \mathcal{F}$ intersect the set $H \cap \{f_0 \leq f_0(p_0)/3c\}$. We have to prove: 1) $F \subset \{n \geq t_1\}$; and 2) $F \cap \{n = t\} \subset \tilde{C}_t$ for $t \geq t_1$.

1) By assumption F contains a point z with $f_0(z) \leq f_0(p_0)/3c$, and $F \subset E_z^{1/6}$. By (6.10) this implies $F \subset \{f_0 < f_0(z)/2c\}$ if t_0 is sufficiently large.

2) Observe that $a_{t_1} \sim a_{t_0}$ since $(t_1 - t_0)/a_{t_0} \to \log 8$. Suppose $n(q) = t \geq t_1$ and $q \notin \tilde{C}_t$. The half space $\{\xi \geq t\}$ is tangent to $\{n = t\}$ in the point p_t. Assume

$c = 4$ and $t = t_1$ and write $p_1 = p_t$. Then $B^2_{p_1} \subset \{\xi \geq s_1\}$ with $s_1 = t_1 - 2a_{t_1}$. Since $\log 8 > 2$ we find $s_1 > t_0$ eventually and $B^2_{p_t} \subset H$. Because $t \mapsto t - 2a_t$ is increasing this also holds for $t > t_1$. So $q \notin B^2_{p_t}$. By Lemma 7.3 the ellipsoid $E_q^{1/3}$ does not intersect $\{\xi \geq t\}$. By construction of $\tilde{C}_t$ the ellipsoid $E_p^{1/6}$ does not intersect H. Contradiction. ¶

Proof (of Theorem 7.1) We may assume that μ is a roughening of Lebesgue measure, using Remark 6.3. Set $h_p(w) = (f \circ \beta_p)(w)/f(p)$ and define h_{0p} similarly in terms of f_0. Then $h_p \leq 2h_{0p}^{1/2}$ on H_+ for $f_0(p) \leq \delta_0$ by the basic inequality, Proposition 6.5. The proposition above with $g = f_0^{1/2}$ shows that $f d\mu$ is a finite measure. Vague convergence of $(f \circ \beta_p)d(\beta_p^{-1}(\mu))/C_p \to e^{-v-u^Tu/2}d\lambda$ follows from vague convergence $\beta_p^{-1}(d\mu)/|\det(\beta_p)| \to d\lambda$. The proposition above with $g = h_{0p}^{1/2}$ and c large then establishes weak convergence as in the proof of Theorem 6.4. ¶

Univariate exceedances are handled effectively by tail functions $T(y) = 1 - F(y) = P\{Y > y\}$. It is known that any distribution in the domain of the exponential limit law is tail asymptotic to a distribution with density $f_0 = e^{-\psi}$ where ψ satisfies the three conditions of Theorem 5.1.

Proposition 7.5 *Let T_0 be the tail function with density $f_0 = e^{-\psi}$ where ψ satisfies the three conditions of Theorem 5.1. Let π be a distribution on $(0,\infty)$ with tail function T. Define the measure μ on $(0,t_\infty)$ by $d\pi = f_0 d\mu$. Then μ is a roughening of Lebesgue measure with respect to the intervals $E_t = (t - a_t, t + a_t)$, with $a_t = 1/\psi'(t)$, if and only if T is asymptotic to T_0 in t_∞.*

Proof Observe that $T_0(t) \sim a_t f_0(t)$. Set $\beta_t(v) = t + a_t v$. Then

$$\frac{\beta_t(d\pi)}{T(t)} = \frac{f_0 \circ \beta_t}{f_0(t)} \frac{\beta_t(d\mu)}{T(t)/f_0(t)} \qquad t \uparrow t_\infty.$$

Since the first factor on the right tends to e^{-v} we see that the left side converges vaguely to $e^{-v}dv$ if and only if the right side converges vaguely to Lebesgue measure. Let M denote a distribution function of μ. Then $M(t+va_t) - M(t) \to v$ if and only if $T(t)/f_0(t) \sim a_t$; and $T(t) \sim T_0(t)$ for $t \uparrow t_\infty$ if and only if μ is a roughening of Lebesgue measure with respect to the intervals $E_t = (t - a_t, t + a_t)$, $t \in (0, t_\infty)$. ¶

The multivariate tail function $T(\theta, t) = P\{\theta Z > t\}$, $t > 0$, $\theta \in S$, determines the distribution since it determines the characteristic function. Asymptotics in infinity of the tail function are reflected in the asymptotics in the origin of the characteristic function. Basrak et al. [5] have investigated the behaviour of T for heavy tailed distributions on $\mathbf{R}^d$. For the lighter tailed case one may use 6) in Theorem 3.7 to link the asymptotic behaviour of the tail function and the density. In general in the multivariate setting densities are more convenient objects to work with than tail functions.

The Sections 5–7 present a fairly complete picture of the asymptotics of $\beta_H^{-1}(Z^H)$ in the case of a Gauss-exponential limit law when the normalizing transformations have a product form, $\beta_H = A \circ \beta_t$, where the first factor depends on the direction of H and the second on the distance to the origin (in terms of the gauge function n). This situation holds if the density f has concentric rotund level sets of the same shape. In principle it is possible that the shape of the level sets $\{f > c\}$ varies as $c \downarrow 0$ without converging to a limit. An analysis of this more general set-up would be of theoretical interest, but for practical applications the model developed in this paper should suffice to describe the behaviour of a distribution in the region of interest, $0.001 < P\{Z \in H\} < 0.05$, say.

8 Heavy tailed distributions

This section contains some examples of distributions for which the high risk limit law has heavy tails.

Theorem 8.1 *Let Z in $\mathbf{R}^d$ have a spherically symmetric density $f(z) = f_0(\|z\|)$. If f_0 varies regularly with exponent $-(t+d)$ for some $t > 0$ then Z lies in the domain of attraction of a high risk limit vector $W = (U, V)$ on H_+, with density*

$$g(w) = \frac{1/C}{((1+v)^2 + u^T u)^{(t+d)/2}} \qquad w = (u, v) \in \mathbf{R}^h \times [0, \infty), \tag{8.1}$$

where

$$C = (\pi^{h/2}/t)(\Gamma((t+1)/2)/\Gamma((t+d)/2)). \tag{8.2}$$

Proof Recall that regular variation means that, for $s > 0$, $f_0(rs)/f_0(s) \to 1/r^{t+d}$ as $s \to \infty$, for fixed $r > 0$. This implies

$$h_s(w) = \frac{f(sw)}{f(sp_0)} \to h(w) = \frac{1}{\|w\|^{d+t}} \qquad s \to \infty \tag{8.3}$$

where p_0 is any unit vector.

By a well known inequality, see Bingham et al. [6], there exists a constant $s_0 > 0$ so that

$$f_0(rs)/f_0(s) \leq 2/r^{t/2+d} \qquad s \geq s_0,\ r \geq 1.$$

This ensures convergence of the integrals

$$\int_{\|z\|\geq 1} \frac{f(sz)dz}{f_0(s)} = b(d)\int_1^\infty \frac{r^{d-1}f_0(rs)}{f_0(s)}dr \to b(d)\int_1^\infty \frac{dr}{r^{t+1}} = \frac{b(d)}{t}$$

where $b(d) = 2\pi^{d/2}/\Gamma(d/2)$ is the area of the surface of the unit ball in $\mathbf{R}^d$. Let W_s denote the vector Z/s conditioned to lie outside the unit ball. Then W_s has density

$$g_s(w) = c_s\frac{f(sw)}{f_0(s)} \to \frac{b(d)}{t\|w\|^{t+d}} \qquad s \to \infty,\ \|w\| \geq 1.$$

Let H be the closed half space $\{\theta \geq s\}$. By symmetry we may assume that θ is the vertical unit vector. Then Z^H/s is distributed like W_s conditioned to lie in the half space $\{v \geq 1\}$, and $Z^H/s - \theta \Rightarrow W$ where W lives on $\{v \geq 0\}$ with the density g above. The constant C is the value of the integral

$$\int_0^\infty\int_{\mathbf{R}^{d-1}} \frac{dvdu}{((1+v)^2+u^Tu)^{(t+d)/2}} = \int_0^\infty \frac{dv}{(1+v)^{t+1}}\int_0^\infty \frac{b(h)r^{d-2}}{(1+r^2)^{(t+d)/2}}.$$

The second integral yields a Beta function, see (10.1). ¶

Definition The distribution with density g on H_+ defined in (8.1) and (8.2) is called a *spherical Pareto distribution with exponent* t. ¶

The vertical component V and the horizontal component U are not independent. The vertical component has a Pareto distribution, $P\{V > v\} = 1/(1+v)^{t+1}$. The vector U has a spherically symmetric density on $\mathbf{R}^h$ which is proportional to $J_{t+h}(\|u\|)/\|u\|^{t+h}$, where

$$J_c(s) = \int_{1/s}^\infty \frac{dy}{(1+y^2)^{(c+1)/2}} = \int_{1/\sqrt{1+s^2}}^1 (1-r^2)^{c/2-1}dr \qquad c > 0, s \geq 0. \tag{8.4}$$

For spherical Pareto distributions there also exists an extension to an infinite Radon measure. It is convenient to use a normalization for this Radon measure ρ^* which is better adapted to the basic limit relation (8.3). Let ρ^* denote the measure on $\mathbf{R}^d \setminus \{0\}$ with density $1/\|w\|^{d+t}$. Then all closed half spaces H which do not contain the origin have finite mass, and the associated probability measure $d\rho_H = 1_H d\rho^*/\rho^*(H)$

is the image of the spherical Pareto limit distribution ρ under a map from H_+ onto H.

It is interesting to compare the behaviour of a sample from the standard Gaussian distribution on $\mathbf{R}^d$ and from a Student distribution. For the Gaussian distribution a sample of size n will form a black cloud of radius $r_n \sim \sqrt{2\log n}$ with a halo on a scale of $1/r_n$; for the Student distribution the sample has no central black region. If one normalizes the sample from the Student distribution by scaling one obtains a Poisson point process on $\mathbf{R}^d$ whose mean measure ρ^* has density $1/\|w\|^{t+d}$. This point process describes the whole sample. In the case of a Gaussian sample the Poisson point process with intensity $e^{-u^Tu/2}e^{-v}$ describes the local behaviour at the edge of the central black region.

Let us briefly look at flat functions for heavy tails. Recall that a positive continuous function L on $\mathbf{R}^d$ is flat for Z if $L \circ \beta_H(w)/L \circ \beta_H(0) \to 1$ uniformly on compact w-sets in H_+. In our case this means that L is asymptotically constant on rings $R_s = \{2s \leq \|z\| \leq 4s\}$ for $s \to \infty$ since L is asymptotically constant on the intersection of the ring R_s with any half space tangent to the ball $B(0,s)$. So L is asymptotic to a spherically symmetric function $L_0(\|z\|)$ where $L_0 : [0,\infty) \to (0,\infty)$ varies slowly in infinity. Flat functions do not help to extend the class of densities f introduced in Theorem 8.1.

For roughening Lebesgue measure the diameter of the sets F_p in the partition $\mathcal{F}$ should be $o(\|p\|)$ for $\|p\| \to \infty$. In particular, the counting measure on the lattice $\mathbf{Z}^d$ is a roughening of Lebesgue measure for the densities f in Theorem 8.1. We shall not enter into details here.

For random vectors Z there is a rich limit theory for Z^s, the vector Z conditioned to lie outside the ball $B(0,s)$ of radius s, with the obvious limit relation

$$Z^s/s \Rightarrow X \qquad s \to \infty. \tag{8.5}$$

In polar coordinates one may write $X = \Theta R$ with $R = \|X\| \geq 1$ and Θ a random element of the unit sphere. In Brozius and de Haan [11] it is shown that R has a Pareto distribution, $P\{R > r\} = 1/r^t$ for $r \geq 1$, for some parameter $t > 0$. The vector Θ may have any distribution on the unit sphere, and is independent of R.

If we condition X to lie in some half space H_θ tangent to the unit ball in θ, we obtain a limit distribution ρ_θ for high risk scenarios Z^{H_n} with half spaces

$H_n = \{\theta_n \geq s_n\}$ where $s_n \to \infty$ and $\theta_n \to \theta$. Here the high risk scenarios in different directions converge to different limit laws which depend continuously on the direction θ.

Now start with a density $f = f_0 \circ n$ where n is the gauge function of a bounded convex open set D containing the origin, and f_0 varies regularly. Then (8.3) becomes

$$h_s(w) = f(sw)/f(sp_0) \to h(w) = 1/n(w)^{d+t} \qquad s \to \infty,$$

where p_0 is a point on the boundary of D. Suppose the restrictions of h to the closed half spaces H_θ determine the same probability type. Then the level sets of $h1_{H_\theta}$ have the same shape (modulo affine transformations), and the sets $D \cap H$ have the same shape for every half space $H = \{\theta \geq 0\}$. This symmetry condition implies that D is an ellipsoid.

In the case of heavy tails we have two models for describing extremal behaviour: the description in terms of high risk scenarios by conditioning on half spaces, and the description (8.5) where the vector is conditioned to lie outside a ball whose radius s tends to infinity. The choice of the appropriate model depends on the sample set and the application. Recall that the high risk limit theory developed in this paper was presented as a base line model for determining the asymptotic behaviour of a multivariate distribution. If the polar coordinates have a natural interpretation and the irregularities in the halo are clear and persistent for $s \to \infty$, and not due to elliptic level sets, then the more versatile second model is appropriate. If the irregularities fade out in infinity the simpler first model is relevant. In the first model the limit distribution is determined by one real parameter $t > 0$, the tail exponent; in the second model there is an additional probability distribution on the unit sphere which has to be determined. The example below shows that the high risk limit theory is not just a special case of the more complex second model.

Example of a continuous density f on the plane which vanishes on the first and third quadrants, which is strictly positive on the interior of the second and fourth quadrants, and which lies in the domain of a spherical Pareto law.

By a rotation over $\pi/4$ we may assume that f vanishes for $|y| \geq |x|$ and is positive for $|y| < |x|$. The function $h(x, y) = 0 \vee (|x| - |y|) \wedge 1$ has this property.

Introduce the family of increasing ellipses

$$E_q : \left(\frac{x}{qe^q}\right)^2 + \left(\frac{qy}{e^q}\right)^2 < 1 \qquad q \in [1, \infty).$$

The function

$$f_0(x, y) = h(x, y)/e^{3q} \qquad (x, y) \in \partial E_q \quad q \geq 1$$

is well defined on the complement of E_1, and is integrable, since E_n lies inside the disk D_n of radius $e^{4n/3}$ eventually, and $\sum |D_n|/e^{3n}$ converges. The function f_0 may be adapted on an ellipse E_{q_0} so as to become a continuous probability density f which vanishes on $A_0 = \{|x| \leq |y|\}$ and is positive on $A_1 = \{|x| > |y|\}$.

Write $E_q = \beta_q(D)$ where D is the open unit disk and $\beta_q = e^q\text{diag}(q, 1/q)$. So β_q^{-1} maps E_q onto D. It also maps A_0 onto the thin wedge $\{|y| \geq q^2|x|\}$ and A_1 onto the complement of this wedge. Let the vector Z have density f. For $q \geq q_0$ the density g_q of $W_q = \beta_q^{-1}(Z)$ is constant on the unit circle except for a neighbourhood of diameter $O(1/q^2)$ around the points $(0, 1)$ and $(0, -1)$ where it is less. The image of E_{q+s} under α_q^{-1} is an ellipse with semi axes $e^s(1 + s/q)$ and $e^s/(1 + s/q)$, which is close to a disk with radius e^s for q large, and contained in a disk of radius $e^{4s/3}$ for all $s \geq 0$, $q \geq 1$. So

$$e^{3q}g_q(u, v) \to 1/r^3 \qquad r^2 = u^2 + v^2$$

pointwise for $u \neq 0$ and also in $\mathbf{L}^1$ on $r \geq \epsilon$ for any $\epsilon > 0$. This shows that Z lies in the domain of attraction of the spherical Pareto law with exponent $t = 1$. ¶

9 Bounded support

Densities with bounded support play only a minor role in risk theory. However the associated theory is simple and gives insight in the ideas underlying high risk limit theory, in particular in the role of rotund sets.

Let D be a rotund set in $\mathbf{R}^d$. If Z is uniformly distributed on D then for any half space H intersecting D the high risk vector Z^H is uniformly distributed on the cap $D \cap H$. Now suppose the volume $|D \cap H|$ tends to zero. Since the boundary ∂D has a continuously varying positive curvature, which is approximately constant on the cap when $|D \cap H|$ is small, the high risk vector Z^H, properly normalized, will converge in distribution to a random vector W which is uniformly distributed on the parabolic cap $Q_+ = Q \cap H_+$, where Q is the paraboloid

$$Q = \{(u,v) \mid v \leq 1 - u^T u\} \subset \mathbf{R}^h \times \mathbf{R}. \tag{9.1}$$

Proposition 9.1 *Let D be a rotund set. There exist affine transformations β_H mapping H_+ onto H such that $\beta_H^{-1}(D) \to Q$ for $0 < |D \cap H| \to 0$, where Q is the paraboloid in (9.1).*

Proof First assume $(0,1)^T$ is a boundary point of D, the half space $\{y \geq 1\}$ is disjoint from D, and the parametrization γ_+ of the upper boundary satisfies $1 - \gamma_+(x) \sim x^T x$ for $x \to 0$. Define

$$\alpha_\sigma(u,v) = (\sqrt{\sigma}u, 1 - \sigma + \sigma v) = (x,y) \qquad \sigma \in (0,1).$$

The upper boundary of the convex set $\alpha_\sigma^{-1}(D)$ has parametrization γ_σ which satisfies

$$1 - \gamma_\sigma(u) = (1 - \gamma_+(\sqrt{\sigma}u))/\sigma \to u^T u \qquad \sigma \to 0$$

uniformly on bounded u-sets. This proves the limit relation for half spaces H of the form $\{y \geq t\}$.

In order to establish the limit for half spaces H diverging in an arbitrary direction we use the argument in the final part of the proof of Theorem 5.1. ¶

As a corollary we see that the uniform distribution on the parabolic cap Q_+ is a high risk limit law, whose domain of attraction contains the uniform distribution on any rotund set. The Poisson point process on the open paraboloid Q is the limit in law of the normalized finite point processes $\beta_{H_n}^{-1}(N_n)$, where N_n is a sample of n points from the uniform distribution on D, and H_n are closed half spaces so that $|H_n \cap D| \sim |D|/n$. The Poisson point process on Q thus describes the local behaviour close to the boundary of the sample cloud for large samples.

Let us briefly discuss some issues related to these limit results.

1) A density f on D which lies in the domain of attraction of the uniform distribution on Q_+ need not be constant. Suppose $L = e^\lambda$ where $\lambda : D \to \mathbf{R}$ is C^1 and the radial and tangential derivatives, see Section 6, satisfy

$$\lambda_r = o(1/s) \qquad \lambda_t = o(1/\sqrt{s}) \qquad s = s(z) = (1 - n(z)) \vee 0 \to 0. \tag{9.2}$$

Here $n = n_D$ is the gauge function, and s the corresponding tent function on D. The function L is flat for the uniform distribution on D in the sense that

$$L \circ \beta_H(w)/L \circ \beta_H(0) \to 1 \qquad 0 < P\{Z \in H\} \to 0$$

holds uniformly on compact w-sets in Q_+. The condition on the radial derivative implies that on any ray the function L is slowly varying as one approaches the boundary of D. One can prove that L is integrable and that the density L/C lies in the domain of attraction of the uniform distribution on Q_+.

2) Let D be rotund and let γ denote the natural conditional distribution on ∂D, see Section 7. As for the case of the uniform distribution on D one can prove for the distribution γ on ∂D that it lies in the domain of a high risk limit law, (and so does $Ld\gamma/C$ for any continuous positive function L on ∂D), with the same normalizations β_H. The limit law is singular with respect to Lebesgue measure on $\mathbf{R}^d$. It is the uniform distribution on the upper boundary of Q_+: the limit vector $W = (U, V)$ has the form $V = 1 - U^T U$ where U is uniformly distributed on the disk $\|u\| < 1$ in $\mathbf{R}^h$.

3) The uniform distribution on Q_+ and on the upper boundary of Q_+ are only two possible limit laws. Let s be the tent function on the rotund set D, $s = (1-n)_+$, see (9.2). If the random vector Z on D has density $f(z) \propto s(z)^{c-1}$ for some $c > 0$ then Z lies in the domain of attraction of the vector W on Q_+ with density $q(w)^{c-1}/C_c$ where $q(u, v) = 1 - u^T u - v$ measures the vertical distance from $w = (u, v)$ to the upper boundary of Q_+. This remains true if Z has density Lf/C where $L = e^\lambda$, and λ satisfies (9.2).

4) If L/s is integrable then L/Cs lies in the domain of the uniform distribution on the upper boundary of Q_+.

Definition The high risk limit distributions on Q_+ and on the upper boundary of Q_+ are called the *parabolic power laws.* ¶

For rotund sets the behaviour in all boundary points is the same asymptotically. Now let us see what happens if D has a vertex in a point p_0.

Assume $D = (0, 1)^d$ is the unit cube in $\mathbf{R}^d$ and $p_0 = 0$. Consider half spaces $H = H_\xi = \{\xi \le 1\}$ with $\xi = (x_1, \ldots, x_d) \in (1, \infty)^d$. Then the closure of $H \cap D$ is the convex hull Σ_ξ of the points $0, x_1 e_1, \ldots, x_d e_d$, where $e_1, \ldots, e_d$ is the standard basis of $\mathbf{R}^d$. If Z is uniformly distributed over D then Z^H is uniformly distributed over the simplex Σ_ξ. The linear transformation $\beta_\xi = \mathrm{diag}(x_1, \ldots, x_d)$ maps the unit simplex Σ with vertices $0, e_1, \ldots, e_d$ onto Σ_ξ. Hence β_ξ^{-1} maps the high risk vector Z^{H_ξ} onto the vector W which is uniformly distributed on the unit simplex Σ. The corresponding Radon measure ρ^* is Lebesgue measure on the open positive quadrant $(0, \infty)^d$. Its symmetry group $\mathcal{S}$ consists of all positive diagonal linear maps. For any half space H with finite ρ^* mass there exists an element $\sigma_H \in \mathcal{S}$ mapping Σ onto $H \cap [0, \infty)^d = \Sigma_\xi$ so that the probability measure $1_H d\rho^*/\rho^*(H)$ is the image of ρ, the uniform distribution on Σ, under σ_H. There are other Radon measures on $(0, \infty)^d$ which are semi-invariant under the group $\mathcal{S}$. These are the measures with density $x_1^{c_1} \cdots x_d^{c_d}$, with $c_i > -1$ for $i = 1, \ldots, d$. These measures, restricted to Σ, and normalized, are possible limit laws for Z^H when Z is a vector on the open cube D with a non-uniform distribution, and the half spaces H have the form $\{\xi \leq 1\}$ with $\xi \in (0, \infty)^d$.

In this situation there is a finite dimensional family of limit laws which holds for half spaces whose normal points into the open negative orthant. We might speak of a local limit law. The associated Radon measure ρ^* on $(0, \infty)$ has a large group of symmetries. The high risk scenarios Z^H all describe the behaviour of Z in the neighbourhood of one particular boundary point. In that respect the limit behaviour is not very interesting. On the other hand the problem of describing the asymptotic behaviour of the convex hull of a sample of size n from the uniform distribution in a polygon in $\mathbf{R}^2$ has attracted considerable attention since Rényi and Sulanka [33], see for instance Cabo and Groeneboom [12]. The model is of some interest for finance since prices are by nature positive.

There is a variant where the vertex has a different structure: D is the cap of a cone $C = \{\|u\| < v\}$. In this case ρ^* is Lebesgue measure and the symmetry group $\mathcal{S}$ is the Lorentz group acting on the cone C of future events in relativity theory. Let $q = v - \|u\|$ denote the vertical distance to the boundary ∂C. The densities q^{c-1} with $c > 0$ are semi-invariant under $\mathcal{S}$. For an algebraic treatment of invariant measures on cones, see Faraut and Korányi [20].

There also exist local limit laws for heavy tailed distributions. Set $q(w) = v + u^T u$ for $w = (u, v) \in \mathbf{R}^d$, and let ρ^* be the Radon measure on the complement of the closed paraboloid $\{q \leq 0\}$ with density $1/q^{t+(d+1)/2}$ for some $t > 0$. The set $O = \{v > 1 - u^2/2\}$ has finite measure $C = \rho^*(O)$. Let Z have distribution ρ^*/C on this set. For each direction θ, with $\theta_d > 0$, the half space $H = \{\theta \geq t\}$ is contained in O for $P\{Z \in H\}$ sufficiently small, and the corresponding high risk vectors Z^H all have the same type. Hence weak convergence in (3.1) for $0 < P\{Z \in H\} \to 0$, $\theta_d > 0$, is trivial.

We conjecture that for large dimension d there are many finite dimensional families of local limit laws, and that the only global limit laws are the GPDs.

10 The Pareto–parabolic distributions

It is time now to introduce the complete class of *multivariate generalized Pareto distributions*, GPDs. These distributions may also be called the *Pareto–parabolic high risk limit distributions*. This section is meant for reference rather than for detailed reading. Below we shall give for every dimension $d > 1$

1) the standard multivariate GPD's π_τ as a continuous one-parameter family;
2) the power families of spherical Pareto, Gauss–exponential, and parabolic power distributions on H_+ in a simple form;
3) the associated spherical probability distributions μ_τ on $\mathbf{R}^h$, and
4) the underlying infinite Radon measures ρ^*_τ with their symmetry groups.

We shall use the notation:

$$h = d - 1 \geq 1 \qquad W = (U, V) \in H_+ = \mathbf{R}^h \times [0, \infty) \qquad \nu = 1/|\tau| \text{ for } \tau \neq 0.$$

For the Gauss-exponential limit law the representation in 1) and 2) is the same, and the corresponding spherical probability distribution in 3) is the standard multivariate Gaussian distribution. The associated Radon measure ρ^*_0 has density $e^{u^T u/2} e^{-v}$. A Gaussian density in dimension $d = 2, 3, 4, \ldots$ is determined by $5, 9, 14, \ldots$ parameters; a Gauss-exponential density on $H_+ \subset \mathbf{R}^d$ by $4, 8, 13, \ldots$ parameters. The group $\mathcal{A}_+(d)$ has dimension d^2, see Section 11, and the symmetry group $\mathcal{O}(h)$ on the horizontal hyperplane has dimension $h(h-1)/2$, which yields $(d^2 + 3d - 2)/2$ parameters for the Gauss-exponential densities on H_+.

In the univariate case the exponential distribution forms the central distribution in the family of GPDs linking the heavy tailed Pareto distributions and the power laws with finite upper endpoint. The multivariate situation is similar. In all cases the vertical component of the high risk limit vector $W = (U, V)$ has a univariate GPD. We shall use the shape parameter τ of this univariate law, see (2.1), to classify the multivariate distributions. In all cases the multivariate GPD has cylinder symmetry with respect to the vertical axis. This means that aside from the shape parameter there are essentially only two other parameters, the scale parameters for the horizontal and for the vertical component. For the standardized distributions we choose the vertical scale parameter so that the component V has a standard univariate GPD, see (2.1). The horizontal scale parameter then is determined asymptotically for $\tau \to 0$ by the continuity condition in $\tau = 0$.

In the univariate case the shape parameter τ varies over the whole real line; in the multivariate case the parameter τ varies over the set $[-2/h, \infty)$.

1) The Gauss-exponential distribution π_0 is the central term in a continuous one parameter family of high risk limit distributions π_τ, $\tau \geq -2/h$, on the upper half space H_+ in $\mathbf{R}^d$, the standard multivariate GPDs. For $\tau > -2/h$ the distribution π_τ has a density $g_\tau(u, v)$ of the form

$$g_\tau(u,v) = \begin{cases} \left((1+\tau v)^2 + \tau u^T u\right)^{-1/2\tau - d/2} / C_\tau & \tau > 0 \\ e^{-(v + u^T u/2)} / (2\pi)^{h/2} & \tau = 0 \\ \left(1 + \tau v + \tau u^T u/2\right)_+^{-1/\tau - 1 - h/2} / C_\tau & -2/h < \tau < 0\,, \end{cases}$$

where the constants C_τ have the value

$$C_\tau = \begin{cases} (\nu\pi)^{h/2}\Gamma((\nu+1)/2)/\Gamma((\nu+1+h)/2) & \tau > 0 \\ (2\nu\pi)^{h/2}\Gamma(\nu - h/2)/\Gamma(\nu) & -2/h < \tau < 0\,. \end{cases}$$

For $\tau = -2/h$ the probability measure π_τ is singular. It lives on the parabolic cap

$$\{2v = h - u^T u\} \cap \{v \geq 0\} \subset \mathbf{R}^d$$

and projects onto the uniform distribution on the disk $\{u^T u < h\}$ in the horizontal coordinate plane. So in the case $\tau = -2/h$ one may write the limit vector as $W = (U, V)$ where the vector U is uniformly distributed on the centered disk of radius $\sqrt{h}$ in $\mathbf{R}^h$, and the vertical coordinate $V = h - U^TU/2$ is a function of $U \in \mathbf{R}^h$.

For each $\tau \geq -2/h$ the vertical coordinate V has a standard univariate GPD with parameter τ, and the horizontal component U has a spherically symmetric density. For $\tau > 0$ this density does not have a simple form, see (8.4); for $\tau < 0$ the vector U lives on the ball with radius $\sqrt{2\nu}$ in $\mathbf{R}^h$ with density proportional to $(1 - u^Tu/2\nu)_+^{\nu+h/2}$.

The expectation exists for $\tau < 1$; the variance for $\tau < 1/2$:

$$E(U,V) = \frac{1}{1-\tau}(0,1) \qquad \mathrm{var}(U,V) = \frac{1}{1-\tau}\mathrm{diag}(1,\dots,1,\frac{1}{(1-\tau)(1-2\tau)}).$$

A spherical Student distribution on $\mathbf{R}^d$ with ν degrees of freedom has density proportional to $1/(1 + v^2/\nu + u^Tu/\nu)^{(d+\nu)/2}$. So the conditional distribution of U given $\nu + V^2 = c$ for a Student distribution with ν degrees of freedom is the same as the condtional distribution of U given $(1+V)^2 = c$ for the GPD with parameter $\tau = 1/\nu$.

The expressions for g_τ for positive and negative parameter values τ differ. For $d > 1$ the family of multivariate GPDs is continuous in τ, but no longer analytic.

2) The positive and negative parameter values determine two power families of limit densities.

For $\tau > 0$ the vector $(X,Y) = (U/\sqrt{\nu}, V/\nu)$ has a spherical Pareto density proportional to

$$1/((1+y)^2 + x^Tx)^{(\nu+d)/2}.$$

This density was treated in Section 8.

Let the vector S in $\mathbf{R}^h$ have density proportional to $1/(1 + s^Ts)^{(\nu+d)/2}$. Then $(1+y)S$ has density proportional to $1/((1+y)^2+s^Ts)^{(\nu+d)/2}$. This is the conditional density of X given $Y = y$. So we see that (X,Y) is distributed like $((1+Y)S, Y)$ where Y is independent of S.

For $\tau \in (-2/h, 0)$ the vector $(X,Y) = (U/\sqrt{2\nu}, V/\nu)$ has a parabolic power density proportional to

$$(1 - y - x^Tx)_+^{q-1} \qquad \nu = q + h/2.$$

Here $P\{Y \geq y\} = (1-y)_+^\nu$ and X has density proportional to $(1 - x^Tx)_+^q$.

Let $S \in \mathbf{R}^h$ have a spherical beta density, proportional to $(1 - s^T s)_+^{q-1}$, $q > 0$. Then $\sqrt{(1-y)}S$ has density proportional to $(1 - y - s^T s)_+^{q-1}$ for $0 \leq y < 1$. So (X, Y) is distributed like $(\sqrt{(1-Y)}S, Y)$ with S independent of Y. Note that S and X both have a spherical beta density, but the exponents differ. This also holds for the boundary case $\tau = -h/2$, where S is uniformly distributed over the boundary of the unit disk in $\mathbf{R}^h$, and X is uniformly distributed over the disk.

The value of the integrals may be computed from

$$\int_0^\infty \frac{r^{2a-1}dr}{(1+r^2)^{b+a}} = \int_0^1 r^{2a-1}(1-r^2)^{b-1}dr = \frac{1}{B(a,b)} = \frac{\Gamma(a)\Gamma(b)}{2\Gamma(a+b)}. \tag{10.1}$$

3) Introduce the family of spherical probability distributions μ_τ on $\mathbf{R}^h$ with densities proportional to

$$1/(1 + \tau s^T s)^{(h+1+\nu)/2}, \ \tau > 0; \qquad (1 + \tau s^T s/2)_+^{\nu - 1 - h/2}, \ \tau \in (-2/h, 0).$$

By continuity μ_0 is the standard Gauss distribution on $\mathbf{R}^h$, and $\mu_{-2/h}$ is the uniform distribution on the sphere of radius $\sqrt{h}$ in $\mathbf{R}^h$. Note that μ_τ for $\tau > 0$ is *not* the spherical Student distribution with ν degrees of freedom. (It lives on $\mathbf{R}^h$, not $\mathbf{R}^d$.)

Theorem 10.1 (Structure) *Let S have distribution μ_τ on $\mathbf{R}^h$, and let V have a standard univariate GPD on $[0, \infty)$ with df G_τ, see (2.1). Assume that S and V are independent. Set $W = ((1 + \tau V)S, V)$ if $\tau \geq 0$ and $W = (\sqrt{(1+\tau V)}S, V)$ if $-2/h \leq \tau \leq 0$. Then W has the standard multivariate GPD π_τ on H_+.*

Proof For $\tau > -2/h$ this follows from the density of the conditional distribution of U given $V = v$ where (U, V) has density g_τ above, see the arguments under 2). For $\tau = -2/h$ it follows from the fact that U is distributed on a sphere of radius $\sqrt{2\nu - v}$ in $\mathbf{R}^h$ if $V = v$, just as $\sqrt{(1 + \tau v)}S$. ¶

Remark 10.2 For $\tau = -2/h$ there is a second representation, $W = (U, U^T U/2)$, where U is uniformly distributed over the ball of radius $\sqrt{2\nu}$ in $\mathbf{R}^h$.

Theorem 10.3 (Projection) *Let $W = (U_1, \ldots, U_h, V)$ have distribution π_τ on H_+, and let $S = (S_1, \ldots, S_h)$ have distribution μ_τ on $\mathbf{R}^h$. Then $(U_1, \ldots, U_m, V)$ has distribution π_τ on $\mathbf{R}^m \times [0, \infty)$ for $m = 0, \ldots, h$, and $(S_1, \ldots, S_m)$ has distribution μ_τ on $\mathbf{R}^m$.*

Proof The densities proportional to $1/(1+s^Ts)^{(\nu+d)/2}$ or to $(1-s^Ts)_+^{\nu-1-h/2}$ are stable under orthogonal projections as is seen by integrating out the variable s_1. This then also holds for the distributions μ_τ for $\tau \neq 0$ and $\tau > -2/h$, and (by continuity) also for the two exceptional values of τ. Now apply the Theorem 10.1 to obtain the result for the GPDs. ¶

4) In Section 4 we extended the Gauss-exponential limit law on H_+ to a Radon measure on the whole space $\mathbf{R}^d$. This was also done for the other high risk limit laws. Adapt the normalization so as to achieve maximal simplicity. We obtain the measures

ρ^*_τ on $\mathbf{R}^d \setminus \{0\}$, with density $1/r^{t+d}$ for the spherical Pareto limit laws, $\tau > 0$;

ρ^*_0 on $\mathbf{R}^d$, with density $e^{-v}e^{-u^Tu/2}$ for the Gauss-exponential limit law;

ρ^*_τ on the paraboloid $Q = \{v + u^Tu < 0\}$, with density q^{t-1} for the parabolic power laws, with $q(u,v) = -(v+u^Tu)$ for $-2/h < \tau < 0$, and

$\rho^*_{-2/h}$ on ∂Q is Lebesgue measure on $\mathbf{R}^h$ lifted to the parabolic surface $v+u^Tu = 0$ for the parabolic power law with $\tau = -2/h$.

Recall that the symmetry group $\mathcal{S}$ of a measure μ is the set of all $\sigma \in \mathcal{A}$ for which there exists a constant $c_\sigma > 0$ so that $\sigma(\mu) = c_\sigma\mu$. In Section 4 it was shown that the symmetry group of the measure ρ^* on $\mathbf{R}^d$ with density $e^{-v}e^{-\|u\|^2/2}$ is generated by the vertical translations $(u,v) \mapsto (u, v+t)$, $t \in \mathbf{R}$, the orthogonal transformations which leave the points on the vertical axis in their place, and the parabolic translations

$$(u,v) \mapsto (u-b, v-b^Tu+b^Tb/2) \qquad b \in \mathbf{R}^{d-1}.$$

The symmetry group for the parabolic power measures with densities q^{t-1} on the paraboloid Q is the same except that we replace the vertical translations by the maps $(u,v) \mapsto (cu, c^2v)$, $c > 0$, which map Q onto itself. This symmetry group corresponds to the group of Euclidean transformations on the horizontal plane, augmented by the scalar transformations $u \mapsto cu$ with $c > 0$. It is also the symmetry group of the singular measure on ∂Q. The symmetry group of the spherical Pareto measures with densities $1/r^{t+d}$ on $\mathbf{R}^d \setminus \{0\}$ is different. It is generated by the group of orthogonal transformations on $\mathbf{R}^d$ together with the scalar maps $w \mapsto cw$, $c > 0$.

Proposition 10.4 *Let $\mathcal{S}$ be the symmetry group of the infinite Radon measure ρ^*_τ*

associated with a Pareto-parabolical high risk limit law π_τ on H_+. For $\tau = 0$ let $H_0 = H_+$, for $\tau > 0$ let $H_0 = \{v \geq 1\}$ and for $\tau < 0$ let $H_0 = \{v \geq -1\}$. In each case there exists a vertical translation mapping H_0 onto H_+ so that the measure ρ^*_τ restricted to H_0, and normalized, maps onto π_τ. For each closed half space H with $0 < \rho^*(H) < \infty$ there is a symmetry σ mapping H onto H_0.

Proof As for Proposition 4.1. ¶

11 Horizontal half spaces

In this section we examine the situation where the half spaces H diverge in a fixed direction. Insight in this simple situation helps to understand the case where convergence has to hold for half spaces diverging in arbitrary directions. For simplicity we consider horizontal half spaces $H^t = \{y \geq t\}$.

Write $Z^t = (X_t, Y_t)$ for the vector Z conditioned to lie in H^t. We assume that there exists a non-degenerate vector $W = (U, V) \in H_+$ and affine transformations β_t mapping H_+ onto H^t so that

$$\beta_t^{-1}(X_t, Y_t) \Rightarrow W \qquad t \uparrow y_\infty \tag{11.1}$$

where y_∞ is the upper endpoint of the distribution of Y, and $P\{Y = y_\infty\} = 0$.

Affine transformations α on $\mathbf{R}^d$ may be represented by a matrix in dimension $d + 1$ with top row $(1, 0, \dots, 0)$ by adding a zeroth component with the value 1 to each vector. If β maps H_+ onto a horizontal half space it has a block representation of size $1 + h + 1$ of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ p & A & q \\ b & 0 & a \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ p + Bx + qy \\ b + ay \end{pmatrix} \in \mathbf{R} \times \mathbf{R}^h \times \mathbf{R}. \tag{11.2}$$

Here p and q vectors in $\mathbf{R}^h$, A is an invertible matrix of size h, and $a > 0$ and b are real numbers. These maps form a group $\mathcal{A}^*(d)$. From the representation we see that the dimension of the group $\mathcal{A}_+(d)$, where $b = 0$, is d^2.

For β as in (11.2) let $\tilde{\beta}$ map the vertical component $y \in \mathbf{R}$ into $b + ay$. Then $\beta \mapsto \tilde{\beta}$ is a homomorphism of the group $\mathcal{A}^*(d)$ onto the two-dimensional group of positive affine transformations on $\mathbf{R}$. This implies

Theorem 11.1 *If (11.1) holds then*

$$\tilde{\beta}_t^{-1}(Y_t) = (Y_t - t)/a_t \Rightarrow V \qquad t \uparrow y_\infty$$

where Y_t is the vertical component Y, conditioned to the half line $[t, \infty)$.

We mention two simple corollaries.

Corollary 11.2 *If Z lies in the domain of a high risk limit vector $W = (U, V)$ then $\theta Z = (\theta_1 Z_1 + \cdots + \theta_d Z_d)$ lies in the univariate domain of V for every non-zero vector θ.*

Corollary 11.3 *The vertical component V of any high risk limit vector $W = (U, V)$ in (3.1) has a univariate GPD.*

Even if we do not assume that the limit vector has a multivariate GPD we can still make concrete statements about the domain of attraction. Suppose the random vector Z lies in the domain of attraction of a high risk limit distribution. If one of the components of the vector Z has a heavy tail in the sense that $E|Z_k|^m = \infty$ for some $k \in \{1, \ldots, d\}$ and some positive integer m then there is a minimal exponent $t \geq 0$ so that $E|Z_k|^s$ is finite for $s \in [0, t)$ and is infinite for $s > t$. By the remarks above it follows that t is positive and that the distributions F_θ of the random variables θZ, $\theta \neq 0$, all have the property that the tail function $s \mapsto 1 - F_\theta(s)$ varies regularly with exponent $-t$. Similarly if one of the components has a finite upper endpoint and the tail varies regularly in that endpoint then this holds for θZ for all $\theta \neq 0$ and the exponent of regular variation is the same in all directions.

In the introduction we imposed the condition that the distribution of θZ was continuous in its upper endpoint for each $\theta \neq 0$, see (3.2). We now see that for a vector Z in the domain of attraction of a high risk limit distribution all univariate distribution tails have to be asymptotically continuous:

Proposition 11.4 *If the vector Z lies in the domain of attraction of a high risk limit vector then*

$$P\{Z \in \partial H\}/P\{Z \in H\} \to 0 \qquad 0 < P\{Z \in H\} \to 0.$$

Proof If $W = (U, V)$ has a high risk limit distribution on H_+ then V has a GPD and hence $P\{V = 0\} = 0$. This implies the limit relation above. ¶

A simple calculation shows that in the univariate situation the high risk limit variables V all have the form $V = \gamma^T(0)$ where T is standard exponential and $\gamma : y \mapsto ay + b$ is an affine transformation on $\mathbf{R}$ with a and b positive. In a future publication we shall show that limit vectors for the relation (11.1) have the form

$$W = \alpha^T(S, 0)$$

where as above T is a standard exponential variable, S a random vector in $\mathbf{R}^{d-1}$, and T and S are independent. Here α^t, $t \in \mathbf{R}$, is a continuous one-parameter subgroup of $\mathcal{A}_+(d)$ such that $\tilde{\alpha}(0)$ is positive.

Finally we want to make a remark on independent components versus spherical symmetry. These classes reflect two ways to construct a multivariate distribution from a univariate distribution. The intersection of these two classes consists of the Gaussian distributions with i.i.d. centered components. In our theory spherical symmetry plays an essential role. The next result shows how small the role is of independence.

Proposition 11.5 *Suppose $d > 1$. Let the random vector Z in $\mathbf{R}^d$ lie in the domain of attraction of a GPD. If Z has independent components then Z is Gaussian.*

Proof Write $Z = (X, Y)$. Then the vertical coordinate Y lies in the domain of attraction of a univariate GPD, say $V_t = a_t Y^t \Rightarrow V$ as t tends to the upper endpoint of Y. Then $(X, V_t) \Rightarrow (X, V)$. By the convergence of types theorem $W = (X, V)$ has a multivariate GPD. Since the components are independent W is Gauss-exponential and X is Gaussian. Similarly one proves that the vector $(Z_2, \ldots, Z_d)$ is Gaussian by treating the first coordinate as the vertical. ¶

12 Alternative scenarios

We have defined a scenario for a random vector as a change in the underlying probability measure. We are interested in scenarios which will highlight extreme situations. In this paper a scenario simply was the vector conditioned on the event that it lies in a given half space H. If the probability $P\{X \in H\}$ is small, the event is extreme and one speaks of high risk. We shall now briefly sketch a number of alternative scenarios.

1) Instead of half spaces one may condition on Z belonging to a shifted orthant

$$[p, \infty) = [p_1, \infty) \times \cdots \times [p_d, \infty) \subset \mathbf{R}^d.$$

Note that $P\{Z \in [p, \infty)\} = P\{Z \geq p\}$ decreases as p increases coordinatewise. Let $Z^{[p,\infty)}$ denote the vector with the conditional distribution given $\{Z \geq p\}$. The vector $Z^{[p,\infty)} - p$ lives on the positive orthant $[0, \infty)^d$. Is it possible to scale these vectors by positive diagonal matrices A_p to yield a family of random vectors W_p which tend in distribution to a random limit vector W on $[0, \infty)^d$ with a non-degenerate distribution for $P\{Z \geq p\} \to 0$? It may be shown that the continuous limit distributions are multivariate generalized Pareto distributions but of a form different from those introduced in this paper. For $d = 2$ the heavy tailed limit distribution on $[0, \infty)^2$ either has i.i.d. components, or a density of the form $g(u, v) = c/(1 + u + v)^{t+2}$ with $t > 0$. See Balkema and Qi [3] for details.

In this approach one looks only at the behaviour of the distribution of the vector Z in certain directions. In addition the analysis relies heavily on the coordinate system, rather than geometric arguments.

2) Exponential tilting is a well known and elegant way to alter the underlying density f. Assume the density f has thin tails. Let $\Theta \subset \mathbf{R}^d$ be the set where the moment generating function K is finite. For $\theta \in \Theta$ define the probability density

$$f_\theta(x) = e^{\theta x} f(x)/K(\theta) \qquad K(\theta) = \int e^{\theta x} f(x) dx.$$

The factor $e^{\theta x}/K(\theta)$ pulls the mass of the distribution with density f in the direction θ. The underlying mathematical theory is elegant. In the mulltivariate case several classes of limit distributions are known. The normal distribution is a limit distribution, as are the multivariate Gamma distributions with independent components. There are others, like the density e^{-v}/C on the paraboloid $v > u^T u$, or $e^{-u^T u/2v+v}/C$ on $v > 0$, which are less obvious. See Nagaev and Zaigraev [30] and Barndorff-Nielsen and Klüppelberg [4] for details.

The theory of exponential tilting only works for light tails. The moment generating function has to exist on a non-empty region around the origin. The density $e^{-\|z\|^t/t}/C$ fails to satisfy this tail condition for $t \in (0,1)$. In financial mathematics we have to allow distributions whose tails are not light, and may even decrease like a power of $1/\|z\|$.

3) One may also condition on the extreme event $\{\|Z\| \geq r\}$ and let $r \to \infty$, see Section 8. The limit theory in this situation is closely related to multivariate extreme value theory, except that one uses polar coordinates rather than Cartesian coordinates. See Resnick [34]. Basrak et al. [5] show how the limit theory is related to regular variation of the distribution tails of linear functionals θZ, where θ varies over the unit sphere. From the mathematical point there is an important difference between this example and the other cases. Here we have a one-dimensional parameter r which tends to infinity, whereas in the previous three cases the parameters, H, p, θ, all were d-dimensional. As a result the limit distributions for $r \to \infty$ have a one-dimensional group of symmetries. The orbits of this group are rays from the origin rather than open sets in $\mathbf{R}^d$.

13 Conclusions

In how far have we succeeded in describing the high risk limit laws and their domains of attraction?

If the vector Z has a spherically symmetric density $f(z) = f_0(\|z\|)$, where f_0 is decreasing and satisfies a tail condition for the univariate domain of attraction for exceedances: f_0 varies regularly in its upper endpoint, or, for $\tau = 0$, satisfies the von Mises condition, $f_0 f_0''/f_0' f_0' \to 1$, then the high risk scenarios Z^H may be normalized to converge in distribution to a non-degenerate limit vector W on the upper half space H_+. We have introduced two generalizations by allowing the level sets of the density to be *rotund*, or egg-shaped, rather than concentric balls, and by multiplying the density by a *flat* function, a function which behaves like a positive constant in infinity. The class of limit distributions is the set of *Pareto-parabolic* or *multivariate generalized Pareto* laws listed in Section 10. We do not know whether there exist other limit laws.

Most of our results are strong limit theorems. This is a matter of convenience. One first proves pointwise convergence of the normalized densities, and then checks that convergence also holds in $\mathbf{L}^1$, see Proposition 3.7. Because our densities are continuous with convex level sets, convergence often holds uniformly on compact sets. Strong convergence opens up the possibility of limit laws for the conditional distributions on hyperplanes, half lines and convex sets. Weak convergence results may be obtained by roughening the underlying Lebesgue measure.

Let us now look at the domains of attraction of the three classes of multivariate GPDs introduced in this paper.

For limit distributions with bounded support, Section 9 gives a fairly complete description of the domain of (strong) attraction. Normalization is determined by the boundary of the domain of the density of Z. It should be pointed out however that rotundity of the domain is a sufficient, but not a necessary condition. Just as there exist functions which are pointwise differentiable, but whose derivative is not continuous, so there exist bounded convex open sets D whose boundary is not C^2, and whose curvature may vanish or be infinite in certain points, but for which the uniform distribution on D lies in the domain of attraction of the uniform distribution on a parabolic cap.

For the heavy tailed limit laws the domain of attraction is smaller; see Section 8. If Z has density $f(z) \sim c/n(z)^{t+d}$ where $n = n_D$ is the gauge function of a rotund set D then D has to be a centered ellipsoid. Vectors in the spherical Pareto domain may have densities f with elliptic level sets $\{f > c\}$ where the shape of the ellipsoid varies continuously for $c \downarrow 0$ without tending to a limit. The rate at which the shape changes goes to zero. In applications, after a suitable choice of coordinates, one may therefore use a model with a density whose level sets are concentric balls to describe the distribution above the halo.

For the Gauss-exponential domain we have exhibited in Section 5 a large class of interesting densities, $f_0 = e^{-\psi \circ n}/C$. These may be used as a first order approximation to model the behaviour of the distribution over the halo of a sample cloud. The class has been extended somewhat in Sections 6 and 7. However we are far from having a good description of the whole domain, even for dimension $d = 2$. Many questions of theoretical interest are still open. Do there exist densities f in the

Gauss-exponential domain such that $\{f > 0\}$ is an open half space, a paraboloid, or a cube? Do there exist densities which decrease like e^{-r} in one direction and like e^{-r^2} in another? Given a sequence of rotund sets D_n of volume one, does there exist a density f in the Gauss-exponential domain, and sequences $c_n \downarrow 0$ and $r_n \uparrow \infty$ so that $\{f > c_n\} = r_n D_n$?

Our aim was to open up a field of investigation, motivated by questions in risk management, insurance and finance, and present some non–trivial mathematical results. We have included some theoretical results without proof so as to give a representative overview of the theory, and sketched possible further lines of research, such as limit theorems for the sample point processes, in order to clarify the relation between high risk limit theory and multivariate extremal theory. Other topics such as rates of convergence, and projection theorems, will be treated in future papers. The transition from the domain of the heavy tailed distributions to the domain of the Gauss–exponential law is not well understood. For applications, this parameter region is of considerable interest, and needs to be investigated.

Acknowledgements

The first author takes pleasure in thanking the Forschungsinstitut für Mathematik (FIM) of the ETH Zurich for its kind hospitality during several visits. Both authors acknowledge financial support from the Isaac Newton Institute. The results of this paper were first presented at the meeting “Managing uncertainty - New tools for insurance, economics and finance” at the Isaac Newton Institute for Mathematical Sciences, Cambridge, 23/7 - 10/8, 2001.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath (1999) Coherent measures of risk. *Math. Finance* **9**, 203–228.

[2] A.A. Balkema and L. de Haan (1974) Residual lifetime at great age. *Ann. Probab.* **2**, 792–804.

[3] A.A. Balkema and Y.–C. Qi (1998) Strictly stable laws for multivariate residual lifetimes. *J. Math. Sci.* **92**, 3873–3880.

[4] O.E. Barndorff–Nielsen and C. Klüppelberg (1999) Tail exactness of multivariate saddlepoint approximations. *Scand. J. Statist.* **26**, 253–264.

[5] B. Basrak, R.A. Davis and T. Mikosch (2002) A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

[6] N.H. Bingham, C.M. Goldie and J.L. Teugels (1989) *Regular Variation*, Cambridge UP.

[7] N.H. Bingham and R. Kiesel (2001) Semi–parametric modelling in finance: theoretical foundations. *Quantitative Finance* **2**, 241–250.

[8] S. Boyd and L. Vandenberghe (2004) *Convex Optimization*, Cambridge University Press.

[9] W. Breymann, A. Dias and P. Embrechts (2003) Dependence structures for multivariate high–frequency data in finance. *Quantitative Finance* **3**, 1–16.

[10] B.M. Brown and S.I. Resnick (1977) Extreme values of independent stochastic processes. *J. Appl. Probab.* **14**, 732-739.

[11] H.A. Brozius and L. de Haan (1987) On limiting laws for the convex hull of a sample. *J. Appl. Probab.***24**, 852-862.

[12] A.J. Cabo and P. Groeneboom (1994) Limit theorems for functionals of convex hulls. *Probab. Theory Related Fields***100**, 31–55.

[13] P. Carr, H. Geman, D. Madan and M. Yor (2000) The fine structure of asset returns: an empirical investigation. Preprint, University of Maryland, College Park.

[14] M. Crouhy, D. Galai and R. Mark (2001) *Risk Management.* McGraw–Hill, New York.

[15] E. Eberlein (2001) Application of generalized hyperbolic Lévy motions in finance. In: *Lévy Processes: Theory and Applications*, O.E. Barndorff–Nielsen, T. Mikosch and S. Resnick (Eds.), Birkhäuser, pp. 319–337.

[16] W.F. Eddy (1980) The distribution of the convex hull of a Gaussian sample. *J. Appl. Prob.* **17**, 686-695.

[17] P. Embrechts (Ed.) (2000) *Extremes and Integrated Risk Management*, Risk Waters Group Ltd., Risk Books, London.

[18] P. Embrechts, C. Klüppelberg and T. Mikosch (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[19] P. Embrechts, A. McNeil and D. Straumann (2002) Correlation and dependence in risk management: properties and pitfalls. In: *Risk Management: Value at Risk and Beyond.* M.A.H. Dempster, Cambridge University Press, Cambridge, pp. 176–223.

[20] J. Faraut and A. Korànyi (1994) *Analysis on symmetric cones.* Oxford University Press, New York

[21] H. Föllmer and M. Schweizer (1993) A microeconomic approach to diffusion models for stock prices. *Math. Finance* **3**, 1–23; Erratum (1994), *Math. Finance* **4**, 285.

[22] E. Hult and F. Lindskog (2001) Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv. Appl. Probab.* **34**, 587–608.

[23] P. Jorion (2002) *Value at Risk: The New Benchmark for Managing Financial Risk*, 2nd Edition, McGraw–Hill.

[24] A. Juri and M.V. Wüthrich (2002) Copula convergence theorems for tail events. *Insurance Math. Econom.* **30**, 405–420.

[25] S. Kotz, T.J. Kozubowski and K. Podgorski (2001) *The Laplace Distribution and Generalizations: A Revisit With Appications to Communications, Economics, Engineering and Finance.* Birkhäuser, Basel.

[26] A. Levin and A. Tchernitser (2001) Multifactor stochastic variance models in risk management: maximum entropy approach and Lévy processes. Preprint, Bank of Montreal.

[27] H.-J. Luethi and G. Studer (1997) Maximum loss for risk measurement of portfolio. Operations Research Proceedings 1997, Springer, Berlin, pp. 386–391.

[28] H.-J. Luethi and G. Studer (1997) Quadratic maximum loss. In: *VaR: Understanding and Applying Value–at–Risk*. KPMG and Risk Publications, London, pp. 307–316.

[29] R.M. Mazo (2002) *Brownian Motion: Fluctuations, Dynamics and Applications*. Clarendon Press, Oxford.

[30] A. Nagaev and A. Zaigraev (2000) Abelian theorems for a class of probability distributions in $\mathbf{R}^d$ and their applications. *J. Math. Sci.* **99**, 1454–1462.

[31] M.H. Pesaran, T. Schuermann and S. Weiner (2004) Modelling regional interdependencies using a global error–correcting macroeconometric model. *J. Business Econ. and Statist.* **22**, 129–162.

[32] J. Pickands III (1975) Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119–131.

[33] A. Rényi and R. Sulanka (1963) Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahrscheinlichkeitsth. verw. Gebiete* **2**, 75-84.

[34] S. Resnick (2004) On the foundations of multivariate heavy tail analysis. In *Stochastic Methods and their Applications*, Publisher: Applied Probability Trust, J. Applied Probability Special Volume 41A; C.C. Heyde Festschrift.

[35] J. Rosenberg and T. Schuermann (2004) A general approach to integrated risk management with skewed, fat–tailed risks. Federal Reserve Bank of New York, Staff Report 185, May 2004.

[36] M.V. Wüthrich (2003) Asymptotic Value–at–Risk estimates for sums of dependent random variables. *Astin Bulletin*, **33**, 75–92.