



HAL
open science

Stein's method for extreme value distributions

Bruno Costacèque-Cecchi

► **To cite this version:**

Bruno Costacèque-Cecchi. Stein's method for extreme value distributions. Probability [math.PR]. Institut Polytechnique de Paris, 2024. English. NNT : 2024IPPAT045 . tel-04874813

HAL Id: tel-04874813

<https://theses.hal.science/tel-04874813v1>

Submitted on 8 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT
POLYTECHNIQUE
DE PARIS

NNT : 2024IPPAT045

Thèse de doctorat



Stein's method for extreme value distributions

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 12 décembre 2024, par

BRUNO COSTACÈQUE-CECCHI

Composition du Jury :

Laure COUTIN Professor, Université Toulouse III - Paul Sabatier	Présidente / Examinatrice
Giovanni PECCATI Professor, Luxembourg University	Rapporteur
Nicolas PRIVAULT Professor, Nanyang Technological University	Rapporteur
Caroline HILLAIRET Professor, ENSAE	Examinatrice
Philippe NAVEAU Senior Research Scientist, LSCE	Examineur
Laurent DECREUSEFOND Professor, Télécom Paris	Directeur de thèse

Contents

Remerciements	6
Notations	9
Introduction	11
Résumé	21
1 Preliminaries	31
1.1 Markov semi-groups	31
1.2 Poisson process	34
1.3 Extreme value theory	39
1.4 Distances between distributions	59
2 Stochastic quantization of max-stable distributions	63
2.1 Quantization of univariate extreme-value distributions	63
2.2 The Fréchet semi-group as a Markov process	88
2.3 Quantization of multivariate max-stable distributions	94
3 Stein's method for the Fréchet distribution	105
3.1 The univariate Fréchet distribution	105
3.2 Extension in higher dimension	121
4 Stein's method for the Gumbel and Weibull distributions	143
4.1 Properties of the Stein solution	143
4.2 Application to the coupon collector problem	153
5 Connections with Poisson stochastic analysis	171
5.1 Stochastic analysis for max-id distributions	171
5.2 Another quantization of the Poisson process	178
Bibliography	183

Acknowledgments

I would like first to thank the examiners for having accepted to be part of my jury. In particular, I am extremely grateful to Prof. Peccati and Prof. Privault for their review of my thesis.

Naturally I am heavily indebted to Prof. Decreusefond for supervising me during the last three years. He has always been there when I needed his help, a privilege many other PhD students would envy. And his patience with my cheeky comments and dumb jokes sure commands respect! Prof. Martins, a.k.a. "Papy", also deserves praise for his considerable tolerance to my humor, as well as the great assistance he has provided me regarding many things during those three years.

Special thanks are due to Prof. Naveau and Prof. Privault for their joint help to structure what will be my post-doctorate subject. Their invitations to respectively Aussois and Singapore also greatly honored me. Besides, I am very grateful to Y. Goude for having introduced me to Prof. Naveau, as well as for his great lecture about data mining at Orsay.

Those three years would have been far less easy without the good organization of TélécomParis. Be it for the courses, the doctoral formalities or the missions, I could always rely on Profs. Roueff, Nonnenmacher, Fontaine and Fercoq, as well as Mses. Kharmaze, Laouenan and Marin. Thank you all!

The other PhD students I have met also contributed greatly to the present work, either by giving me essential advice or by their constant support: thanks you so much Michael and Guillaume! Heartfelt thanks go to Christophe who gave me the instruction manual to work with Laurent, and has been a precious companion and friend during our time together as PhD students. Although we have met only recently, Ian has been of considerable assistance to me, and also completed my user guide of Laurent: for all of that, I am deeply thankful!

I am extremely thankful to the EDMH for having accepted my candidacy to the program "Junior Scientific Visibility", thus allowing me to visit Pr. Privault in Singapore in September. Those weeks spent at NTU were even more fruitful than expected, both for my PhD and my future post-doctorate!

There is a life beyond thesis, and it would be dull and worthless without friends. So thank you for having made it fun and exciting Thomas and Jean! A course on blockchain around a good dinner in a fancy restaurant and a passionate discussion about geopolitics while drinking beers are for sure memories I owe you and that I will cherish. Hugo, Jean-Baptiste, Alexandre, Amine, Alban, Louis and Maxime, a.k.a. "le Descamps": although you have annoyed me more than once (talk about an euphemism), I sure (somehow) enjoyed the time we have spent together... And, to answer your question, "oui, c'est bien une loi Gamma!".

My life would have been very different had I not met you, Quang and Wei Hung, so thank you so much for having been around all that time. Let's meet up soon!

"Last but not least" - how cliché that sentence is... - my family. I cannot thank enough my mother for all the love and support she has been providing me during those 28 years. You have been here whenever I needed it, and have always done your best for me. Even now, I barely realize how invaluable those things are... Special thanks to you Gérard, you have been the grandfather I never had. I sure will not forget the many important life lessons you taught me!

Notations

Abbreviations

a.s.	Almost surely
c.d.f.	Cumulative distribution function
p.d.f.	Probability density function
i.i.d.	Independent identically distributed
w.r.t.	With respect to
max-id	Max-infinitely divisible

Mathematical relations

x_+	$\max(0, x)$
x_-	$\max(0, -x)$
$x \oplus y$	Maximum of x and y ; also denoted by $\max(x, y)$
$x \odot y$	Minimum of x and y ; also denoted by $\min(x, y)$
$\mathbf{x} \oplus \mathbf{y}$	$(x^1 \oplus y^1, \dots, x^d \oplus y^d)$
$\mathbf{x} \odot \mathbf{y}$	$(x^1 \odot y^1, \dots, x^d \odot y^d)$
$\mathbf{x} < \mathbf{y}$	$x^j < y^j$, for all $j \in \llbracket 1, d \rrbracket$
$\mathbf{x} \not\leq \mathbf{y}$	$x^j > y^j$, for some $j \in \llbracket 1, d \rrbracket$
\lesssim	Lesser or equal up to a constant
\gtrsim	Greater or equal up to a constant

Set theory

\mathbb{N}	Set of integers including 0
\mathbb{N}^*	Set of integers excluding 0
\mathbb{R}	Set of real numbers
\mathbb{R}_+	Set of non-negative real numbers
\mathbb{R}^d	Set of real valued vectors of dimension d
\mathbb{Q}	Set of rational numbers
$\llbracket 1, d \rrbracket$	Set of integers between 1 and d

$\mathcal{B}(E)$ σ -field of the Borel sets of the topological space E

$A \uplus B$ Disjoint reunion of the sets A and B

Functional analysis

Id Identity on a given space

$\mathbf{L}^p(E, \mu)$ The set of p -integrable functions $\varphi : E \rightarrow \mathbb{R}$ w.r.t. a measure μ

$\mathbf{L}^\infty(E, \mu)$ The set of bounded functions $\varphi : E \rightarrow \mathbb{R}$ with respect to a measure μ

$\langle \cdot, \cdot \rangle_E$ Inner product on the vector space E

$\| \cdot \|_E$ Norm on the vector space E

$\| \cdot \|_p$ p -norm on \mathbb{R}^d , for $p \in [1, +\infty]$

φ^\leftarrow Pseudoinverse of a function φ

$d_{\text{TV}}(X, Y)$ Total variation distance between the distributions of X and Y

$d_{\text{K}}(X, Y)$ Kolmogorov distance between the distributions of X and Y

$d_{\text{W}}(X, Y)$ Wasserstein distance between the distributions of X and Y

$\mathbb{1}_A$ Indicator function of the set A

δ_x Dirac measure at x

λ_d Lebesgue measure on \mathbb{R}^d

$\lambda \otimes \mu$ Product measure of λ and μ

$\mathbf{0}$ Null vector of \mathbb{R}^d

$\mathbf{1}$ Unit vector of \mathbb{R}^d , whose coordinates are all equal to 1;
constant function equal to 1

∞ Vector of \mathbb{R}^d whose coordinates are all equal to $+\infty$

e_i i -th element of the canonical basis of \mathbb{R}^d , *i.e.* $e_i = (0, \dots, 0, 1, 0, \dots, 0)$

Probability

\mathbb{P}_X Law of X

$\mathbb{E}_\mu[\varphi]$ Integral of φ under the probability measure μ

$\mathbb{E}[X]$ Expectation of the random variable X

$\mathbb{E}[X|Y]$ Conditional expectation of X given the σ -algebra generated by the random variable Y

$\mathbb{V}(X)$ Variance of the random variable X

$\text{Cov}(X, Y)$ Covariance between the random variables X and Y

$X_n \xrightarrow{d} Y$ The sequence of random variables $(X_n)_{n \in \mathbb{N}}$ converges in distribution to the random variable Y

$X \stackrel{d}{=} Y$	The random variable X is equal in distribution to the random variable Y
$X \sim \mu$	X has the probability distribution μ
$\text{Geom}(p)$	Geometric distribution with parameter $p \in [0, 1]$
$\mathcal{VP}(\alpha, \sigma)$	Pareto distribution with shape parameter α and scale parameter σ
$\mathcal{E}(\lambda)$	Exponential distribution with scale parameter λ
$\mathcal{U}_{[a,b]}$	Uniform distribution on $[a, b]$
$\mathcal{F}(\alpha, \sigma)$	Fréchet distribution with shape parameter α and scale parameter σ
$\mathcal{G}(\mu, \sigma)$	Gumbel distribution with location parameter μ and scale parameter σ
$\mathcal{W}(\alpha, \sigma)$	Weibull distribution with shape parameter α and scale parameter σ
$\mathcal{MS}(\boldsymbol{\alpha}, \nu; \boldsymbol{\mu}, \boldsymbol{\sigma})$	Multivariate max-stable distribution with shape parameter $\boldsymbol{\alpha}$, angular measure ν , location parameter $\boldsymbol{\mu}$ and scale parameter $\boldsymbol{\sigma}$

Introduction

Extreme value theory

Catastrophes and extreme phenomena such as 2022 Hurricane Ian (\$113 billions of damage) 2021 European floods (\$11.8 billions) or 2023 Turkey–Syria earthquakes (\$148.8 billions) constitute an endless source of hazards, a reality made even more blatant nowadays due to the intensification of global warming and climate change. Human-related disasters such as 2008 financial crisis or the war in Ukraine started in 2014 have consequences whose range go beyond the financial field. To deal with such rare (?), but dramatic events, extreme value theory (EVT) constitutes a tool of choice. This branch of probability is precisely concerned with estimating the likelihood and the strength of extreme events, such as record rainfall, anomalous network traffic, massive losses in finance and insurance, or even the COVID-19 contagiousness ([Wong and Collins \[2020\]](#)), giving it a large range of applications. One of its precursors was M. Fréchet who showed in a paper published in [Fréchet \[1927\]](#) that, under certain assumptions on the base distribution, the only possible limiting distributions of a properly scaled maximum of n i.i.d. random variables when n goes to infinity are the Fréchet distribution and the Weibull distribution. His work was continued by two other important contributors to EVT: R. A. Fisher, the founding father of modern statistics, and L. H. C. Tippett, who was interested in the strength of textile fibers, which depends on the solidity of their weakest link. Their paper [Fisher and Tippett \[1928\]](#) generalized Fréchet’s result by allowing centering in the normalization of the maximum, thus managing to include the Gumbel distribution among the possible limits. R. von Mises and B. Gnedenko completed this work by characterizing the domain of attraction of univariate extreme value distributions. Together, those results are known as *Fisher-Tippett-Gnedenko’s theorem* (FTG theorem) and form the basis of EVT. In his monograph [Gumbel \[1958\]](#), E. J. Gumbel applied it to a wide range of problems, including hydrology, fatigue failures, aeronautics, geology, *etc.*

Today, EVT is an essential component of risk analysis in several domains, such as finance and insurance for instance, through the use of the concept of Value-at-Risk, the Peaks-over-threshold method (POT method), the Hill estimator, *etc.* A good account of those techniques can be found in *e.g.* [Embrechts et al. \[2013\]](#). An important issue with EVT is to know when the asymptotic results it relies on, such as the Fisher-Tippett-Gnedenko theorem, are reasonable approximations of the ‘true’, unknown model. Thus, it is essential to determine at which rate those convergences take place, in order to replace the unknown model behind data by the limit predicted by the theory while controlling the committed error. This problem is certainly not new: major contributions to it in the univariate setting include [Hall and Wellner \[1979\]](#), [Cohen \[1982\]](#), [Smith \[1982\]](#), [de Haan and Resnick \[1996\]](#), and many others. Results in higher dimension are much rarer, see for instance [Omey and Rachev \[1991\]](#) as well as A. Feidt’s PhD dissertation [Feidt \[2013\]](#). Most of the aforementioned works makes use of sophisticated analysis, like *regular variation with remainder* or *second-order regular variation*. They yield accurate results, at the price of increased complexity, and rely on assumptions which may be complicated to check in applications.

Modern EVT heavily relies on point processes techniques, as described in Leadbetter et al. [2012] for instance. More precisely, multivariate extreme value distributions can be seen as the maximum of a Poisson process whose intensity measure is homogeneous, a property which will play an essential role in this thesis. As such, Fisher-Tippett-Gnedenko's (FTG) theorem can be interpreted as the convergence of a certain point process to a Poisson process. Furthermore, the convergence of the point process of exceedances to a Poisson process under mild assumptions on the data constitutes another example of this strong connection between EVT and the point process theory. For all those reasons, the interplay between EVT and stochastic analysis for the Poisson process will be instrumental in the sequel.

Stein's method

Main concepts and approaches

A powerful technique to obtain rates of convergence in limit theorems in probability is the so-called Stein's method. It has been used for the first time by C. Stein himself in his seminal paper Stein [1972] to bound the Kolmogorov distance between a sum of dependent random variables and the normal distribution, in an effort to extend the Berry-Esseen theorem. Since then, his method has been greatly extended in many directions and has proved both flexible and effective. Its purpose is to bound certain types of distance between probability measures. More precisely, suppose μ and ν are two probability measures on some topological space E . The distances adapted to Stein's method are the ones of the form

$$d_{\mathcal{H}}(\mu, \nu) := \sup_{h \in \mathcal{H}} |\mathbb{E}_{\mu}[h] - \mathbb{E}_{\nu}[h]|,$$

where \mathcal{H} is a space of test-functions from E to \mathbb{R} . It is assumed to be large enough to characterize convergence in distribution w.r.t. the topology on E , and to contain only functions integrable w.r.t. both μ and ν . For example it can be the space of bounded continuous functions from E to \mathbb{R} , although less regular functions are also used (see *infra*). Such a metric is called an *integral probability metric*. Now assume we want to know the distance between μ and the target distribution ν , in the sense of a predefined metric. Stein's method relies on an operator \mathcal{L} which characterizes ν in a certain sense. In most, if not all, instances of Stein's method, the operator \mathcal{L} satisfies the following abstract property

\mathcal{L} is invertible on the space \mathcal{H}_0 of test-functions with zero mean w.r.t. the target distribution ν :

$$\mathcal{H}_0 := \{h \in \mathcal{H}, \mathbb{E}_{\nu}[h] = 0\}.$$

Such an operator is called a *Stein operator*. It yields a functional characterization of the target distribution ν known as *Stein's characterization*:

$$\mathbb{E}_{\mu}[\mathcal{L}g] = 0, \text{ for all } g \in \mathcal{G} \iff \mu = \nu. \quad (0.1)$$

The direct sense is the non-trivial part. It goes as follows. Let h be a test-function and set

$h^* := h - \mathbb{E}_\nu[h]$. We have:

$$\begin{aligned} \mathbb{E}_\mu[h] - \mathbb{E}_\nu[h] &= \mathbb{E}_\mu[h^*] \\ &= \mathbb{E}_\mu[\mathcal{L}\mathcal{L}^{-1}h^*] \\ &= \mathbb{E}_\mu[\mathcal{L}g_h] \\ &= 0, \end{aligned}$$

thanks to the assumptions we have made on \mathcal{L} . Since the class \mathcal{H} characterizes convergence in distribution, it characterizes in particular the equality in distribution. Thus $\mathbb{E}_\nu[h] = \mathbb{E}_\mu[h]$ for all h in \mathcal{H} implies that $\mu = \nu$. In many applications, the proof that \mathcal{L} characterizes ν uses more specific tools depending on the context, for it is not always possible to rely on some of the aforementioned assumptions. In any case, having a Stein operator for a given target distribution ν is the first step of the method. The historical example of such an operator is Stein's original characterization of the standard normal distribution:

$$\mathcal{L}h(x) := -xh(x) + h'(x), \quad x \in \mathbb{R}.$$

The fact that $\mathbb{E}[\mathcal{L}h(X)] = 0$ for all bounded Lipschitz functions h if and only if $X \sim \mathcal{N}(0, 1)$ is the so-called *Stein's lemma*, whose applications are numerous. One of its consequences is the fundamental inequality

$$d_{\mathcal{H}}(\mu, \nu) \leq \sup_{h \in \mathcal{H}} |\mathbb{E}_\mu[\mathcal{L}g_h]| = \sup_{f \in \mathcal{L}^{-1}\mathcal{H}_0} |\mathbb{E}_\mu[\mathcal{L}f]|. \quad (0.2)$$

Next, it is often very convenient to search and find bounds on $g_h := \mathcal{L}^{-1}h^*$ for all $h \in \mathcal{H}_0$, the solution to Stein's equation. This is because in (0.8) we work with functions of the form g_h for every h in \mathcal{H}_0 . It is not compulsory though, especially in the generator approach where one can work with the semi-group directly instead. Bounding g_h and its derivatives allows us to work with a more restricted class of functions than the test-functions associated with $d_{\mathcal{H}}$. For instance, one can prove (Nourdin and Peccati [2012]) that if one takes $h_z = \mathbf{1}_{(-\infty, z]}$ as the test-functions, so that the associated distance is the Kolmogorov distance, then one have

$$\|g_z\|_\infty \leq \frac{\sqrt{2\pi}}{4} \text{ and } \|g'_z\|_\infty \leq 1, \quad z \in \mathbb{R}.$$

Consequently, the Kolmogorov distance between a probability distribution μ and $\nu = \mathcal{N}(0, 1)$ satisfies

$$d_{\mathcal{H}}(\mu, \nu) \leq \sup_{g \in \mathcal{F}_{\text{Kol}}} |\mathbb{E}_\mu[\mathcal{L}g]| = \sup_{g \in \mathcal{F}_{\text{Kol}}} |\mathbb{E}[g'(X)] - \mathbb{E}[Xg(X)]|,$$

where X is a random variable with the distribution μ and

$$\mathcal{F}_K := \left\{ g : \mathbb{R} \rightarrow \mathbb{R}, \|g\|_\infty \leq \frac{\sqrt{2\pi}}{4} \text{ and } \|g'\|_\infty \leq 1 \right\}.$$

This set is sometimes called the *Stein class* associated with the normal distribution and Kolmogorov distance (see Mijoule et al. [2023] for instance).

The final step consists in proving that $\mathbb{E}_\mu[\mathcal{L}g_h]$ is 'small'. Surprisingly, it is often much easier than trying to control the difference $\mathbb{E}_\nu[h] - \mathbb{E}_\mu[h]$ directly. Many different tools have been developed for such a purpose, including exchangeable pairs, zero-bias couplings, size-bias couplings, *etc.*, see for instance Ross [2011], Ley et al. [2017-01], as well as Chen et al. [2010] for an exposition more

focused on the normal distribution. However, one must keep in mind that there is no general technique to deal with this part in a systematic way. Depending on how this step is performed, the resulting bound will be more or less accurate. An important matter is to find a Stein operator \mathcal{L} such that the final bound is sharp, or at least that this step is not too complicated to manage and still yields interesting results.

Applications of Stein’s method are numerous and concern various domains of probability, like random matrices, statistical mechanics, random graphs, *etc.* as well as statistics (see [Anastasiou et al. \[2021\]](#)), a testament of its great flexibility. The original method focused on the normal distribution in dimension 1 as the target distribution, and since that time has been extended in many directions: multivariate normal distribution ([Gotze \[1991\]](#)), Gaußian processes and diffusions ([Barbour \[1990\]](#)), Poisson distribution ([Chen \[1974\]](#)), Poisson point processes ([Barbour and Brown \[1992\]](#), [Chen and Xia \[2004\]](#)), univariate and multivariate stable distributions ([Chen et al. \[2022, 2024\]](#), [Coutin et al. \[2024\]](#)), Gamma distribution ([Arras and Swan \[2015\]](#)), *etc.*

Quantization and convex cones

When applying Stein’s method, the first problem one may meet is finding a ‘good’ Stein operator, even though it is not easy to come up with a general criterion to determine the quality of such an operator. Our line of thought will be that a Stein operator must have a ‘shape’ adapted to the application at hand. This point should be clarified in one of the applications given in this thesis: *the coupon collector problem*. Besides, the more intrinsic properties of the target distribution the Stein operator expresses, the better. We will elaborate more on this part later too. Another desirable property of a Stein operator is that it corresponds to the generator of some Markov semi-group. More precisely, the latter should admit the target distribution ν as an invariant measure. The act of building such a semi-group for a given probability measure is known as *quantization* in physics. This procedure has been established in [Parisi and Wu \[1981\]](#) and has been used, for instance, for numerical simulations of gauge theories with fermions ([Namiki \[1992\]](#)).

Our starting point will be the definition of the Markov semi-group $(P_t)_{t \geq 0}$ instead of its generator \mathcal{L} , the latter playing the role of the Stein operator. One can then prove under certain assumptions on the test function h that g_h exists and is equal to $-\int_0^\infty P_t h dt$. This avoids proving the often difficult equivalence (0.7) and makes the determination of the Stein’s solution g_h regularity easier. In certain situations, such as [Coutin and Decreusefond \[2020\]](#), it may be necessary to separate the integral from 0 to $+\infty$ into two parts: one from 0 to a variable threshold ε , and the other from ε to $+\infty$. Each term should be approached differently so that we do not end up with a simple upper-bound of the Lipschitz continuity, for instance, of the solution of the Stein’s equation. This difficulty will concern us.

We see the importance of being able to work with a Stein operator which is also the generator of some semi-group admitting the target distribution as an invariant measure. However constructing either the semi-group $(P_t)_{t \geq 0}$ or its generator \mathcal{L} first is never an easy task. This constitutes one of the main obstacles to the generalization of Stein’s method *via* the generator approach to distributions beyond the Gaußian and the Poisson measures. As of now, there is no known general recipe to achieve such a goal for an arbitrary probability measure, but see [Ley et al. \[2017-01\]](#) and references therein for a thorough survey of how to construct an operator satisfying (0.7) for a wide range of probability measures.

A powerful and generic way to quantize a given probability measure is to use an Itô process known

as *Langevin diffusion*. The application of this idea to Stein’s method has been introduced in [Gorham and Mackey \[2015\]](#) and [Gorham et al. \[2016\]](#). We do not delve further into this rich theory for, as powerful as it is, it may no be the most adapted tool to deal with non-Gaussian distributions depending on the context. For instance, the resulting semi-group and generator may not always be convenient to use depending on the situation, and the target distribution ν must be absolutely continuous with respect to Lebesgue measure. This will not be always the case when dealing with multivariate max-stable distributions. Let us just mention that, for the Gaussian distribution, it yields the same semi-group as the one obtained through the method we now expose.

The most common choice of Stein operator for the standard Gaussian measure on \mathbb{R}^d is

$$\mathcal{L}f(\mathbf{x}) = -\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \Delta f(\mathbf{x}),$$

for which the so-called Ornstein-Uhlenbeck semi-group is given by

$$P_t f(x) = \mathbb{E}\left[f\left(e^{-t}\mathbf{x} + \sqrt{1 - e^{-2t}}\mathbf{Z}\right)\right], \tag{0.3}$$

where $Z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and \mathbf{I} is the identity matrix. More generally, if the target distribution ν is the law of an α -stable distribution \mathbf{Z} , parametrized as in [Arras and Houdré \[2023\]](#), a semi-group which quantizes ν is

$$P_t f(x) = \mathbb{E}\left[f\left(e^{-t}\mathbf{x} + (1 - e^{-\alpha t})^{\frac{1}{\alpha}}\mathbf{Z}\right)\right]$$

where $\mathbf{Z} \sim \nu$ and $\alpha \in (0, 2]$. Both examples take place in an Euclidean setting but more exotic contexts are possible: define the *Heisenberg group* \mathbf{H} as the space matrices of the form:

$$\mathbf{H} := \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, (x, y, z) \in \mathbb{R}^3 \right\},$$

with the usual matrix product. Such a set endowed with that binary operation is indeed a group and it can be shown it is isomorphic to \mathbb{R}^3 equipped with the following group operation:

$$(x^1, y^1, z^1) \oplus (x^2, y^2, z^2) = \left(x^1 + x^2, y^1 + y^2, z^1 + z^2 + \frac{1}{2}(x^1 y^2 - x^2 y^1)\right).$$

This group possesses a notion of outer multiplication by non-negative scalars defined by:

$$d_a(x, y, z) = (ax, ay, a^2z), \quad a > 0$$

and which satisfies the axioms $d_a \circ d_b = d_{ab}$ and $d_a(\mathbf{x}_1 \oplus \mathbf{x}_2) = d_a \mathbf{x}_1 \oplus d_b \mathbf{x}_2$. Now, if we sum n i.i.d. , square-integrable random variables $\mathbf{X}_1, \dots, \mathbf{X}_n$ with values on \mathbf{H} and renormalize them appropriately, one gets when n goes to infinity a non-degenerate distribution which corresponds to the law of:

$$\mathbf{Z} = \left(B^1(1), B^2(1), \frac{1}{2}\left[\int_0^1 B^1 dB^2 - \int_0^1 B^2 dB^1\right]\right),$$

where B^1, B^2 are two independent standard Brownian motions on \mathbb{R} . The third term is also known as *Lévy’s stochastic area*. To be more precise, if $\mathbf{X}_1, \dots, \mathbf{X}_n$ are i.i.d. centered random variables with unit variance (componentwise speaking), then

$$\bigoplus_{i=1}^n d_{\frac{1}{\sqrt{n}}} \mathbf{X}_i \xrightarrow[n \rightarrow \infty]{} \mathbf{Z}.$$

Therefore, the distribution ν of \mathbf{Z} plays the same role on \mathbf{H} as the Gaussian distribution on \mathbb{R} : both are the limiting distributions of a version of the central limit theorem. A semi-group which quantizes ν has been introduced and studied thoroughly in [Lust-Piquard \[2010\]](#):

$$P_t f(\mathbf{x}) = \mathbb{E} \left[f \left(d_{e^{-t}} \mathbf{x} \oplus d_{\sqrt{1-e^{-2t}}} \mathbf{Z} \right) \right]$$

The crucial remark is that the semi-group property of the family of operators so defined is a mere rewriting of the stability satisfied by α -stable distributions and the Heisenberg Gaussian distribution: The standard Gaussian measure on \mathbb{R}^d is the unique probability law on \mathbb{R}^d such that

$$e^{-t} \mathbf{Z}_1 + (1 - e^{-\alpha t})^{\frac{1}{\alpha}} \mathbf{Z}_2 \stackrel{d}{=} \mathbf{Z} \quad (0.4)$$

where $\mathbf{Z}_1, \mathbf{Z}_2$ are two independent copies of \mathbf{Z} . And, although it is not obvious, one can also check that

$$d_{e^{-t}} \mathbf{Z}_1 \oplus d_{\sqrt{1-e^{-2t}}} \mathbf{Z}_2 \stackrel{d}{=} \mathbf{Z}_1, \quad (0.5)$$

Remark that at a formal level, equations (0.10) and (0.11) can be written as

$$d_a X_1 \oplus d_b X_2 \stackrel{d}{=} X_1 \quad (0.6)$$

where $a^\alpha + b^\alpha = 1$ for some $\alpha > 0$. In the first two cases the operator d_a is the multiplication of vectors by real numbers and the sign \oplus represents the addition of vectors in (0.10). This means that the algebraic structure at play here is that of a semi-group (the \oplus operation) with a group $(d_a)_{a \in \mathbb{R}_+}$ of automorphisms which satisfy the identity $d_a \circ d_b = d_{ab}$. Such a structure is called a *convex cone*. More precisely, an abelian topological semi-group (\mathbb{K}, \oplus) is a convex cone if there exists a continuous multiplication operator d_a of the elements x of \mathbb{K} by non-negative numbers a under which \mathbb{K} is stable and:

$$\begin{aligned} d_a(x \oplus y) &= d_a x \oplus d_a y, \quad a > 0, \quad x, y \in \mathbb{K} \\ d_a(d_b x) &= d_{ab} x, \quad a, b > 0, \quad x \in \mathbb{K} \\ d_1 x &= x, \quad x \in \mathbb{K} \\ d_a e &= e, \quad a > 0 \end{aligned}$$

where e is the neutral element of \mathbb{K} and must satisfy $x \oplus e = x$. In [Davydov et al. \[2008\]](#), the authors show that there exist numerous other examples of (strictly) stable distributions, *i.e.* distributions satisfying (0.12). These distributions are often interesting because their stability property causes them to appear in many limit theorems such as the central limit theorem or the law of rare events. This setting is also very broad and apply to Banach spaces as well as non-Euclidean contexts. Examples of convex cones are various and include the Heisenberg group (and more generally homogeneous groups), the set of compact convex sets, the set of upper semi-continuous functions, the set of finite measures, *etc.*) Each of those spaces is equipped with a binary operation \oplus and an outer multiplication $(d_a)_{a \geq 0}$ satisfying the previous properties. The most basic example of such a convex cone is \mathbb{R} with the addition and $d_a x = ax$, the usual multiplication. More sophisticated examples are the continuous positive functions on some cone E with the usual addition and

$$d_a f(x) = f(ax), \quad x \in E$$

or the set of compact convex sets with the Minkowski addition and the elementwise multiplication:

$$K_1 \oplus K_2 = \{x_1 + x_2, (x_1, x_2) \in K_1 \times K_2\},$$

and

$$d_a K = aK = \{ax, x \in K\}.$$

Other examples can be found in the paper by Davydov et al. Therein the existence of nontrivial stable random variables on convex cones is proved under certain general assumptions. Such distributions arise as limit candidates to renormalized 'sums' of i.i.d. random variables on \mathbb{K} *i.e.* a generalized central limit theorem. Being able to quantize them is therefore important to apply Stein's method and compute rates of convergence to those distributions. We present a general way to do so. It consists in defining a semi-group $(P_t)_{t \geq 0}$ directly through a Mehler formula:

$$P_t f(x) := \mathbb{E} \left[f \left(d_{e^{-t}} x \oplus d_{(1-e^{-t})^{1/\alpha}} Z \right) \right], \quad x \in \mathbb{K}$$

where Z is a *strictly stable random variable*, in the sense of (0.12). The stability property ensures that the family of operators $(P_t)_{t \geq 0}$ is indeed a semi-group which admits the distribution of Z as a stationary measure. Semi-groups admitting this kind of representation tend to have certain properties in common with the original Gaussian Ornstein-Uhlenbeck semi-group. However it does not seem possible at this point to establish a list of the properties they would all share without further assumptions on the convex cone \mathbb{K} . An important aspect of this thesis is to study the case where \oplus is the maximum operator and d_a is the usual multiplication. Some properties of the resulting semi-group will be proved by using rather generic arguments, while others will seem to be more specific to our special case.

Let us conclude this part by briefly mentioning another way to quantize certain distributions. It relies on the notion of *discrete stability*, see for instance Steutel and van Harn [1979], Davydov et al. [2011]. The goal was to extend the concept of stability to discrete spaces where rescaling by continuous arguments is not possible. Instead, a different, random rescaling is used. It must be formally analogous to a dilation, in the following sense: $p \circ (q \circ x)$ is equal in distribution to $(pq) \circ x$ and $p \circ (x \oplus y)$ has the same law as $(p \circ x) \oplus (p \circ y)$, for $x \in E$ and all p, q in a certain subset of \mathbb{R}_+ , *e.g.* $[0, 1]$. The most famous example of such a random rescaling is the *thinning* by $p \in [0, 1]$ operation on $E = \mathbb{N}$:

$$p \circ n = \sum_{i=1}^n B_i, \quad p \in [0, 1]$$

where the B_i are i.i.d. random variables with the Bernoulli distribution $\mathcal{B}(p)$. The Poisson distribution, which is not stable, is discrete stable however, in the sense that:

$$p \circ N_1 + (1 - p) \circ N_2 \stackrel{d}{=} N_1,$$

where N_1, N_2 are i.i.d. random variables with the Poisson distribution. Actually this result is but a reflect of the discrete stability of the Poisson process. If $\phi = \sum_{i=1}^n \delta_{x_i}$ is a finite configuration, then one defines the thinning of ϕ by p by removing points x_i at random, each with probability $1 - p$ and independently from the others. In other words

$$p \circ \phi = \sum_{i=1}^n B_i \delta_{x_i}.$$

Then one can prove that

$$p \circ \eta_1 \cup (1 - p) \circ \eta_2 \stackrel{d}{=} \eta_1,$$

where η_1, η_2 are two i.i.d. Poisson processes on some Polish space E , with the same finite intensity measure. This property of discrete stability allows us to construct a semi-group which quantizes the Poisson process *via* a Mehler formula by setting:

$$P_t f(\phi) = \mathbb{E}[f(e^{-t} \circ \phi + (1 - e^{-t}) \circ \eta)], \quad \phi \in \mathfrak{N}_E.$$

This is the so-called Glauber semi-group (see for instance [Decreusefond \[2022\]](#)). Those ideas have been successfully extended to the case of discrete α -stable processes by A. Vasseur in his PhD thesis [Vasseur \[2017\]](#), giving him the opportunity to apply Stein's method to limit theorems involving those processes. More examples of semi-groups should be constructable by using this principle. The present thesis will exemplify this fact.

Stochastic analysis for max-stable distributions

The ideas developed by Davydov et al. in [Davydov et al. \[2008\]](#) rely on the notion of *LePage series*. It allows one to represent stable (for some binary operation \oplus , like the addition or the maximum) random variable as the 'sum' (in the sense of \oplus) of a certain Poisson process whose intensity measure satisfies an homogeneity property. Such a representation has been known before in the field of α -stable random variables as well as in EVT, where it is referenced as the *de Haan decomposition*. This does not come as a surprise as $(\mathbb{R}, \oplus, \text{d})$ is indeed a convex cone if \oplus denotes the maximum operator and d the usual multiplication.

The de Haan decomposition makes a crucial connection between max-stable distributions and Poisson processes. It will appear under different appearances throughout this thesis, mainly *via* the semi-group $(\mathbf{P}_t^\alpha)_{t \geq 0}$ we will define to quantize a max-stable distribution ν . For instance, the generator \mathcal{L}_α of this semi-group will be expressed with the intensity measure of the same Poisson process behind the de Haan's decomposition of ν .

Another way of making use of this connection between Poisson processes and EVD is to leverage results from stochastic analysis for Poisson processes, such as the ones exposed in [Last and Penrose \[2011\]](#), [Last et al. \[2016\]](#), to obtain new results for EVD. This line of reasoning is extremely simple but brings results which would be difficult to prove directly.

Contributions

The contributions of this thesis are the followings: in chapter 2, we define several Markov semi-groups quantizing all the extreme value distributions, both in the univariate and the multivariate settings. We give several functional identities (commutation relations, Poincaré inequality, *etc.*) which illuminate the structure of those semi-groups and provide additional information about the extreme value distributions. We also exhibit a Markov process whose semi-group corresponds to the one quantizing the Fréchet distribution. Throughout the chapter, a parallel with the Gaussian Ornstein-Uhlenbeck semi-group is drawn.

Chapter 3 applies the previous results to Stein's method in the case where the target distribution is the Fréchet law (in any dimension). We state and solve Stein's equation, describing Stein's solution when the working distance is either the Wasserstein distance, the $\text{d}_{[2]}$ distance or the Kolmogorov distance. We then make use of those properties to find rates of convergence to the Fréchet distribution in several instances.

In chapter 4, we follow the same line of thought, but in the cases of the Gumbel and (positive) Weibull distributions. An application is given to the coupon collector problem, where we estimate the rate of convergence of the renormalized completion time of the collection to the Gumbel distribution in distance $d_{[2]}$.

Finally, chapter 5 focuses on the applications of stochastic analysis for the Poisson process to extreme value distributions, and more generally, to stable distributions and max-id distributions. In particular, we give a short proof of the multivariate version of Poincaré inequality for the family of semi-groups that we have introduced in chapter 2, as well as a covariance identity for max-id distributions. The last section of that chapter is more of an heuristic nature and defines a new semi-group quantizing the Poisson processes whose intensity measure satisfy a certain homogeneity property.

Résumé en français

La théorie des valeurs extrêmes

Les catastrophes naturelles et autres phénomènes extrêmes tels l'ouragan Ian de 2022 (dont le préjudice s'élève à 113 milliards \$), les inondations de 2021 en Europe (11.8 milliards \$), ou encore le séisme de 2023 en Turquie et en Syrie (148.8 milliards \$) représentent une source constante de risques. Le réchauffement climatique a conduit à une intensification marquée de ces aléas, tant en terme de fréquence que d'intensité. Les activités humaines sont elles aussi sources de contingence : la crise financière de 2008 ou encore la guerre en Ukraine, débutée en 2014, ont des conséquences de long-terme qui vont au-delà du "simple" préjudice financier. Pour faire face à ces événements rares (?) mais dramatiques, la théorie des valeurs extrêmes constitue un outil de choix. Cette branche de la théorie des probabilités s'attache précisément à estimer la fréquence et la violence d'événements extrêmes tels les inondations, une circulation anormale au sein d'un réseau, des pertes records en finance et en assurance ou encore le degré de contagiosité d'une épidémie comme celle du COVID-19 ([Wong and Collins \[2020\]](#)). L'un de ses plus célèbres précurseurs est M. Fréchet. Celui-ci a démontré dans [Fréchet \[1927\]](#) sous certaines hypothèses sur la loi des données que leur maximum renormalisé ne peut converger que vers une loi de Fréchet ou de Weibull. Ses travaux ont été poursuivis par deux autres grands auteurs : R. A. Fisher, le père fondateur des statistiques modernes, et L. H. C. Tippett, dont l'intérêt résidait dans la résistance des fibres textiles, elle-même dépendant de celle de leur lien le plus fragile. Leur article [Fisher and Tippett \[1928\]](#) généralise les résultats de Fréchet en autorisant la présence d'un paramètre de centrage dans la normalisation du maximum, avec pour conséquence de faire apparaître la loi de Gumbel comme distribution-limite possible. Par la suite, R. von Mises et B. Gnedenko ont parachevé ces travaux en caractérisant complètement le domaine d'attraction des loi d'extremum en dimension 1. Ces résultats sont aujourd'hui connus sous le nom de *théorème de Fisher-Tippett-Gnedenko* (théorème FTG) et sont la clef de voûte de la théorie des valeurs extrêmes. Celle-ci est appliquée par E. J. Gumbel dans son ouvrage [Gumbel \[1958\]](#) à une vaste étendue de problèmes, parmi lesquels l'hydrologie, la fatigue des matériaux, l'aéronautique, la géologie, *etc.*

Aujourd'hui, la théorie des valeurs extrêmes joue un rôle important en analyse de risque pour plusieurs domaines, tels la finance et l'assurance, pour ne citer qu'eux. Ces derniers font appel aux concepts de *value-at-risk*, de méthode POT, d'estimateur de Hill, *etc.*, qui sont tous issus de cette théorie. On pourra se référer à [Embrechts et al. \[2013\]](#) et les références citées pour d'avantage sur ce sujet. Les outils mentionnés reposent sur des théorèmes-limites dont la précision dépend du nombre d'observations disponibles. Ainsis, il est souvent nécessaire de savoir quand la distribution d'un maximum renormalisé de variables aléatoires peut être "convenablement" approchée par une loi d'extremum généralisée. Cela revient à demander à quelle vitesse a lieu la convergence du maximum vers sa limite, pour une certaine distance, l'objectif étant de substituer à la loi inconnue de la plus grande des observation la limite prédite par la théorie et sur laquelle plus d'informations sont disponibles. L'erreur due à cette substitution reste contrôlée dans le même temps. Ce problème n'est en rien nouveau : les principales contributions en dimension 1 incluent entre autres [Hall and](#)

Wellner [1979], Cohen [1982], Smith [1982], de Haan and Resnick [1996]. En revanche il est plus difficile de trouver des résultats similaires en dimension supérieure. On peut citer essentiellement Omev and Rachev [1991] ainsi que la thèse de A. Feidt's PhD Feidt [2013] et ses articles. La plupart des travaux mentionnés tendent à faire appel à des notions sophistiquées et peu maniables dans les applications, comme la *variation régulière avec reste* ou la *variation régulière d'ordre 2*. Bien qu'elles fournissent des bornes précises, cela se fait au prix d'une complexité importante et d'hypothèses délicates à vérifier en pratique.

La théorie moderne des valeurs extrêmes fait un usage important de celle des processus ponctuels. Plus précisément les loi d'extremum généralisées peuvent, quelque soit la dimension, être représentées comme le maximum d'un processus de Poisson dont la mesure intensité vérifie une certaine propriété d'homogénéité. En conséquence, le théorème de Fisher-Tippett-Gnedenko s'interprète comme la convergence d'un certain processus ponctuel fini vers un processus de Poisson. De même, le processus ponctuel des excès converge lui aussi vers un processus de Poisson lorsque le nombre d'observations tend vers l'infini, modulo des hypothèses raisonnables sur ces dernières. Ces exemples illustrent bien les fortes connexions qui existent entre théorie des valeurs extrêmes et théorie des processus ponctuels. Elle joueront un rôle important dans la suite de cette thèse.

Méthode de Stein

Principaux concepts et approches

Une technique flexible et puissante pour obtenir des taux de convergence dans des théorèmes-limites en probabilité est la célèbre méthode de Stein. Stein lui-même l'a utilisée pour la première fois dans son article pionnier Stein [1972] afin de borner la distance de Kolmogorov entre une somme de variables aléatoires dépendantes et la loi normale, ceci dans un effort de fournir une généralisation au théorème de Berry-Esseen. Depuis, cette méthode a été considérablement étendue à de nombreuses autres distributions. Elle vise à contrôler une certaine distance métrisant la convergence en loi entre une distribution-cible ν fixée, et une loi μ l'approchant. Plus précisément, on suppose que μ et ν sont deux lois de probabilité sur un même espace topologique E . Le type de distance utilisée pour quantifier la proximité entre μ et ν , et adaptée à la méthode de Stein, est de la forme suivante :

$$d_{\mathcal{H}}(\mu, \nu) := \sup_{h \in \mathcal{H}} |\mathbb{E}_{\mu}[h] - \mathbb{E}_{\nu}[h]|,$$

où \mathcal{H} est un espace de fonctions-tests de E dans \mathbb{R} . On suppose qu'il est suffisamment grand pour caractériser la convergence en loi engendrée par la topologie sur E , et aussi qu'il ne contient que des fonctions intégrables par rapport à μ et ν . Des exemples couramment rencontrés incluent l'espace des fonctions bornées continues de E dans \mathbb{R} , ou encore celui des fonctions 1-Lipschitz à dérivée 1-Lipschitz, si E est un espace métrique. Une telle métrique $d_{\mathcal{H}}$ sur les mesures de probabilité de E est appelée *integral probability metric* en anglais. Nous travaillerons toujours avec de telles métriques. Afin de majorer la distance $d_{\mathcal{H}}(\mu, \nu)$, où ν est la loi de probabilité cible, la méthode de Stein fait usage d'un certain opérateur linéaire qui satisfait la propriété suivante : \mathcal{L} est inversible sur l'espace \mathcal{H}_0 des fonctions-tests d'espérance nulle par rapport à ν :

$$\mathcal{H}_0 := \{h \in \mathcal{H}, \mathbb{E}_{\nu}[h] = 0\}.$$

Une application satisfaisant cette identité est appelée un *opérateur de Stein*. Elle fournit une caractérisation fonctionnelle de la loi-cible ν , couramment appelée *caractérisation de Stein*.

$$\mathbb{E}_{\mu}[\mathcal{L}g] = 0, \text{ pour tous } g \in \mathcal{G} \iff \mu = \nu. \quad (0.7)$$

L'implication directe constitue la partie non triviale et se démontre comme suit. Soit h une fonction-test. Posons $h^* := h - \mathbb{E}_\nu[h]$. Alors nous avons:

$$\begin{aligned} \mathbb{E}_\mu[h] - \mathbb{E}_\nu[h] &= \mathbb{E}_\mu[h^*] \\ &= \mathbb{E}_\mu[\mathcal{L}\mathcal{L}^{-1}h^*] \\ &= \mathbb{E}_\mu[\mathcal{L}g_h] \\ &= 0, \end{aligned}$$

grâce aux hypothèses que nous avons faites sur \mathcal{L} . Puisque la classe \mathcal{H} caractérise la convergence en loi, elle caractérise en particulier l'égalité en loi. Par conséquent $\mathbb{E}_\nu[h] = \mathbb{E}_\mu[h]$ quelque soit h dans \mathcal{H} implique que $\mu = \nu$. Dans de nombreuses applications, la preuve que \mathcal{L} caractérise ν repose sur des arguments plus spécifiques au contexte en jeu; en effet, certaines des hypothèses précédentes peuvent faire défaut. En tout cas, disposer d'un opérateur de Stein caractérisant la loi-cible ν représente la première étape de cette méthode. L'exemple historique d'un tel opérateur est celui donnée par Stein dans [Stein \[1972\]](#) pour la loi normale standard:

$$\mathcal{L}h(x) := -xh(x) + h'(x), \quad x \in \mathbb{R}.$$

Le fait que $\mathbb{E}[\mathcal{L}h(X)] = 0$ pour tout h bornée et Lipschitz si et seulement si $X \sim \mathcal{N}(0, 1)$ est encore appelé *lemme de Stein*. Ses applications sont nombreuses. L'une de ses conséquences est l'inégalité fondamentale :

$$d_{\mathcal{H}}(\mu, \nu) \leq \sup_{h \in \mathcal{H}} |\mathbb{E}_\mu[\mathcal{L}g_h]| = \sup_{f \in \mathcal{L}^{-1}\mathcal{H}_0} |\mathbb{E}_\mu[\mathcal{L}f]|. \quad (0.8)$$

Cela nous mène à la seconde étape de cette méthode : borner pour tout $h \in \mathcal{H}_0$ la solution de Stein $g_h := \mathcal{L}^{-1}h^*$. En effet, dans (0.8), on travaille avec des fonction de la forme g_h pour tout h dans \mathcal{H}_0 . Cette étape peut prendre des formes différentes selon la variante de la méthode de Stein employée, en particulier l'approche par générateur qui autorise à travailler avec le semi-groupe à la place. Un contrôle *a priori* sur g_h et ses dérivées permet de se restreindre à une classe plus restreinte de fonctions que les fonctions-tests associées à $d_{\mathcal{H}}$. Par exemple, on peut prouver ([Nourdin and Peccati \[2012\]](#)) que si l'on choisit comme fonction-test les $h_z = \mathbb{1}_{(-\infty, z]}$, c'est-à-dire qu'on utilise la distance de Kolmogorov, alors on a

$$\|g_z\|_\infty \leq \frac{\sqrt{2\pi}}{4} \text{ et } \|g'_z\|_\infty \leq 1, \quad z \in \mathbb{R}.$$

Par conséquent, la distance de Kolmogorov entre une loi de probabilité μ and $\nu = \mathcal{N}(0, 1)$ satisfait

$$d_{\mathcal{H}}(\mu, \nu) \leq \sup_{g \in \mathcal{F}_{\text{Kol}}} |\mathbb{E}_\mu[\mathcal{L}g]| = \sup_{g \in \mathcal{F}_{\text{Kol}}} |\mathbb{E}[g'(X)] - \mathbb{E}[Xg(X)]|,$$

où X est une variable aléatoire de loi μ et

$$\mathcal{F}_K := \left\{ g : \mathbb{R} \rightarrow \mathbb{R}, \|g\|_\infty \leq \frac{\sqrt{2\pi}}{4} \text{ et } \|g'\|_\infty \leq 1 \right\}.$$

Cet enesemble est parfois appelé la *classe de Stein* associée à la loi normale et à la distance de Kolmogorov (cf. [Mijoule et al. \[2023\]](#) par exemple).

Le but de la dernière étape est de prouver que $\mathbb{E}_\mu[\mathcal{L}g_h]$ est "petit". Surprenamment, cela est souvent bien plus simple que de chercher à contrôler la différence $\mathbb{E}_\nu[h] - \mathbb{E}_\mu[h]$ directement. De

nombreux outils ont été développés dans cette optique; citons entre autres les paires échangeables, les couplages zéro-biais, les couplages de biais par la taille, *etc.* On pourra se référer par exemple à [Ross \[2011\]](#), [Ley et al. \[2017-01\]](#), ainsi qu'à [Chen et al. \[2010\]](#) pour un exposé plus centré sur la loi normale. Quelque soit l'approche retenue, il est important de garder en tête qu'aucune technique ne fonctionne efficacement de manière systématique. Selon comment cette étape est menée, la borne obtenue sera plus ou moins fine. Par conséquent, il est utile de savoir construire des opérateurs de Stein \mathcal{L} tels que la borne finale soit la plus précise possible, ou tout du moins que cette étape ne soit pas trop compliquée à mener à son terme.

La méthode de Stein a un large champ d'applications et concerne de nombreux domaines majeurs des probabilités (matrices aléatoires, mécanique statistique, graphes aléatoires, *etc.*), ainsi que des statistiques (voir [Anastasiou et al. \[2021\]](#)). Cela constitue une preuve de sa grande souplesse d'utilisation. Cette méthode, telle que conçue par Stein originellement, est focalisée sur la loi normale univariée comme distribution-cible. Elle a depuis été développée dans de nombreuses directions : loi normale multivariée [Gotze \[1991\]](#), processus gaussiens et diffusions ([Barbour \[1990\]](#)), loi de Poisson ([Chen \[1974\]](#)), processus ponctuels de Poisson ([Barbour and Brown \[1992\]](#), [Chen and Xia \[2004\]](#)), lois stables en dimensions finie ([Chen et al. \[2022, 2024\]](#), [Coutin et al. \[2024\]](#)), loi gamma ([Arras and Swan \[2015\]](#)), *etc.*

Quantification et cônes convexes

Le premier problème susceptible d'apparaître en appliquant la méthode de Stein réside dans le fait de trouver un "bon" opérateur de Stein. Il est difficile de définir en général ce que devrait être un tel opérateur; notre point de vue est qu'il doit avoir une "forme" adaptée à l'application étudiée. Cet état d'esprit sera précisé et développé au cours de la présente thèse, en particulier dans l'une des applications proposées, le *problème du collectionneur*. En outre, par définition, l'opérateur de Stein encode complètement toute l'information disponible sur la loi-cible qu'il caractérise. Il est donc souhaitable que les propriétés fonctionnelles de l'opérateur de Stein révèle facilement ces informations. Enfin, il peut être utile que cet opérateur soit le générateur d'un certain *semi-groupe de Markov* de mesure stationnaire la loi-cible ν . Cette procédure a été utilisée en physique et porte en anglais le nom de *quantization*, voir [Parisi and Wu \[1981\]](#). Elle a servi entre autres à effectuer des simulation numériques en théorie de gauge pour les fermions ([Namiki \[1992\]](#)).

Notre point de départ sera la définition d'un tel semi-groupe markovien $(P_t)_{t \geq 0}$ de mesure stationnaire ν , puis le calcul de son générateur \mathcal{L} , qui sera alors automatiquement un opérateur de Stein pour ν . Ce détour par la théorie des semi-groupes est fertile car il fournit, sous certaines hypothèses sur les fonctions-tests h , une expression explicite de la solution de Stein g_h :

$$g_h = - \int_0^\infty P_t h \, dt.$$

On évite ainsi d'avoir à prouver la difficile équivalence (0.7) tout en ayant de plus grandes facilités à déterminer la régularité de g_h . Dans certains contextes (*e.g.* [Coutin and Decreusefond \[2020\]](#)), on peut néanmoins avoir à séparer l'intégrale de 0 à $+\infty$ en deux morceaux: l'un allant de 0 à un seuil ε arbitraire, et l'autre de ε à $+\infty$. Chaque terme doit être contrôlé séparément afin d'obtenir une borne la plus fine possible et qui n'explose pas au voisinage de 0 ou à l'infini. Nous aurons à faire face à cette difficulté. Lorsqu'on ne dispose pas d'un semi-groupe et qu'on travaille directement avec \mathcal{L} et g_h , ces problèmes apparaissent aussi, mais de façon plus cachée.

Pour toutes ces raisons, on voit l'importance de pouvoir travailler avec un semi-groupe $(P_t)_{t \geq 0}$ engendrant l'opérateur de Stein \mathcal{L} utilisé. Un problème de type "poule-œuf" se pose néanmoins : est-il plus simple de construire d'abord $(P_t)_{t \geq 0}$, ou de partir de \mathcal{L} et de prouver *via* un théorème à la Hille-Yoshida qu'il s'agit d'un générateur associé à un certain semi-groupe ? Cette difficulté a représenté un obstacle à la généralisation de l'approche par générateur de la méthode de Stein à d'autres lois que la gaussienne ou la loi de Poisson. Il n'existe pas de méthode universelle pour construire un opérateur de Stein qui satisfasse (0.7) pour une loi-cible arbitraire, mais on pourra consulter [Ley et al. \[2017-01\]](#) et les références citées pour un panorama des techniques actuelles.

Dans le cas particulier de l'approche par générateur, au mieux de nos connaissances, la seule technique générale permettant de construire un semi-groupe markovien de loi stationnaire fixée repose sur les *diffusions de Langevin*, voir par exemple [Gorham and Mackey \[2015\]](#) et [Gorham et al. \[2016\]](#). Bien que très souple, cette approche a ses limites : le semi-groupe et son générateur peuvent être compliqués à utiliser selon les applications, et de plus il est nécessaire que la loi-cible ν admette une densité pour la mesure de Lebesgue. Cela ne sera pas toujours vrai dans le cas des lois d'extremum en dimension supérieure. Mentionnons tout de même que cette méthode fournit, dans le cas de la loi normale, le même semi-groupe obtenu *via* l'approche que nous présentons maintenant.

Un choix commun d'opérateur de Stein pour la loi normale standard sur \mathbb{R}^d est

$$\mathcal{L}f(\mathbf{x}) = -\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \Delta f(\mathbf{x}).$$

Il s'agit du générateur du *semi-groupe d'Ornstein-Uhlenbeck* :

$$P_t f(x) = \mathbb{E}\left[f\left(e^{-t}\mathbf{x} + \sqrt{1 - e^{-2t}}\mathbf{Z}\right)\right], \quad (0.9)$$

où $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ et \mathbf{I} est la matrice identité. Plus généralement, si ν est une loi d'une variable aléatoire α -stable \mathbf{Z} , paramétrisée comme dans [Arras and Houdré \[2023\]](#), alors un semi-groupe quantifiant ν est donné par :

$$P_t f(x) = \mathbb{E}\left[f\left(e^{-t}\mathbf{x} + (1 - e^{-\alpha t})^{\frac{1}{\alpha}}\mathbf{Z}\right)\right]$$

où $\mathbf{Z} \sim \nu$ et $\alpha \in (0, 2]$. Ces deux exemples se situent dans un cadre euclidien, mais il est possible d'en donner dans des contextes plus exotiques : on définit le *groupe de Heisenberg* \mathbf{H} comme l'ensemble des matrices de la forme :

$$\mathbf{H} := \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, (x, y, z) \in \mathbb{R}^3 \right\},$$

avec le produit matriciel habituel. Un tel ensemble muni de cette opération forme bien un groupe, et on peut montrer qu'il est isomorphe à \mathbb{R}^3 doté de la loi de composition interne suivante :

$$(x^1, y^1, z^1) \oplus (x^2, y^2, z^2) = \left(x^1 + x^2, y^1 + y^2, z^1 + z^2 + \frac{1}{2}(x^1 y^2 - x^2 y^1)\right).$$

Ce groupe possède une notion de multiplication extérieure par les réels positifs donnée par :

$$d_a(x, y, z) = (ax, ay, a^2z), \quad a > 0$$

et qui satisfait les axiomes $d_a \circ d_b = d_{ab}$ et $d_a(\mathbf{x}_1 \oplus \mathbf{x}_2) = d_a \mathbf{x}_1 \oplus d_b \mathbf{x}_2$. Si maintenant l'on somme n variables aléatoires i.i.d. de carré intégrable $\mathbf{X}_1, \dots, \mathbf{X}_n$, à valeurs \mathbf{H} et qu'on les renormalise convenablement, une loi non dégénérée apparaît lorsque n tend vers l'infini :

$$\mathbf{Z} = \left(B^1(1), B^2(1), \frac{1}{2} \left[\int_0^1 B^1 dB^2 - \int_0^1 B^2 dB^1 \right] \right),$$

avec B^1, B^2 deux mouvements browniens standards sur \mathbb{R} . Le troisième terme est connu sous le nom d'*aire de Lévy stochastique*. Plus précisément, si $\mathbf{X}_1, \dots, \mathbf{X}_n$ sont des variables aléatoires i.i.d. centrées et de variance unité (coordonnée par coordonnée), alors

$$\bigoplus_{i=1}^n d_{\frac{1}{\sqrt{n}}} \mathbf{X}_i \xrightarrow{n \rightarrow \infty} \mathbf{Z}.$$

Par conséquent la loi ν de \mathbf{Z} joue le même rôle sur \mathbf{H} que la loi gaussienne sur \mathbb{R} : toutes les deux apparaissent comme lois-limites d'un certain théorème central limite. Un semi-groupe quantifiant ν a été introduit et étudié en détail dans [Lust-Piquard \[2010\]](#):

$$P_t f(\mathbf{x}) = \mathbb{E} \left[f \left(d_{e^{-t}} \mathbf{x} \oplus d_{\sqrt{1-e^{-2t}}} \mathbf{Z} \right) \right]$$

Il est important de remarquer que la propriété de semi-groupe de la famille d'opérateurs ainsi définie n'est qu'une réécriture de la stabilité des lois α -stables et la loi normale sur le groupe de Heisenberg : par exemple la mesure gaussienne sur \mathbb{R}^d est la seule loi sur \mathbb{R}^d telle que

$$e^{-t} \mathbf{Z}_1 + (1 - e^{-\alpha t})^{\frac{1}{\alpha}} \mathbf{Z}_2 \stackrel{d}{=} \mathbf{Z} \quad (0.10)$$

avec $\mathbf{Z}_1, \mathbf{Z}_2$ deux copies indépendantes de \mathbf{Z} . Et, bien que cela ne soit pas évident, on peut vérifier que

$$d_{e^{-t}} \mathbf{Z}_1 \oplus d_{\sqrt{1-e^{-2t}}} \mathbf{Z}_2 \stackrel{d}{=} \mathbf{Z}_1, \quad (0.11)$$

Observons qu'à un niveau plus formel, les équations (0.10) et (0.11) peuvent se réécrire sous la forme

$$d_a X_1 \oplus d_b X_2 \stackrel{d}{=} X_1 \quad (0.12)$$

avec $a^\alpha + b^\alpha = 1$ pour un certain $\alpha > 0$. Dans les deux premiers cas, l'opérateur d_a est la multiplication des vecteurs par les réels, et le signe \oplus représente l'addition des vecteurs dans (0.10). La structure algébrique sous-jacente est celle d'un semi-groupe, pour l'opération \oplus auquel on associe un groupe $(d_a)_{a \in \mathbb{R}_+}$ d'automorphismes satisfaisant l'identité $d_a \circ d_b = d_{ab}$. Une telle structure est appelée un *cône convexe*. Plus précisément, un semi-groupe topologique et abélien (\mathbb{K}, \oplus) est un cône convexe s'il existe un opérateur de multiplication continu d_a des éléments x de \mathbb{K} par les réels positifs a , laissant stable \mathbb{K} et satisfaisant:

$$\begin{aligned} d_a(x \oplus y) &= d_a x \oplus d_a y, \quad a > 0, \quad x, y \in \mathbb{K} \\ d_a(d_b x) &= d_{ab} x, \quad a, b > 0, \quad x \in \mathbb{K} \\ d_1 x &= x, \quad x \in \mathbb{K} \\ d_a e &= e, \quad a > 0 \end{aligned}$$

où e est l'élément neutre de \mathbb{K} et doit vérifier la relation $x \oplus e = x$. Dans [Davydov et al. \[2008\]](#), les auteurs montrent qu'il existe de nombreux exemples de lois (strictement) stables, c'est-à-dire satisfaisant (0.12). Ces distributions ont un intérêt particulier compte tenu de leur tendance à apparaître dans de nombreux résultats limites, comme la loi des événements rares ou le théorème central limite. Le cadre des cônes convexes est très large et inclut entre autres celui des espaces de Banach ainsi que de nombreux espaces non euclidiens, *e.g.* l'ensemble des compacts convexes, l'ensemble des fonction semi-continues supérieurement, l'ensemble des mesure finies, *etc.* Chacun de ces espaces est équipé d'une opération binaire \oplus et d'une multiplication extérieure $(d_a)_{a \geq 0}$ vérifiant les propriétés précédentes. Un exemple simple d'un tel cône convexe est \mathbb{R} doté de l'addition et

$d_a x = ax$, la multiplication habituelle. D'autres exemples plus sophistiqués incluent l'ensemble des fonctions continues positives définies sur un cône E , avec l'addition usuelle, et la multiplication définie par

$$d_a f(x) = f(ax), \quad x \in E$$

ou l'ensemble des compacts convexes muni de l'addition de Minkowski et de la multiplication élément par élément :

$$K_1 \oplus K_2 = \{x_1 + x_2, (x_1, x_2) \in K_1 \times K_2\},$$

et

$$d_a K = aK = \{ax, x \in K\}.$$

D'autres exemples sont présentés dans [Davydov et al. \[2008\]](#). L'existence de lois stables non triviales sur un cône convexe y est démontrée sous certaines hypothèses. Ces lois apparaissent comme limites possible pour une "somme" (toujours au sens de \oplus) renormalisée de variables aléatoires i.i.d. sur \mathbb{K} , c'est-à-dire, dans un théorème central limite généralisé. À ce titre il est donc important de pouvoir leur appliquer la méthode de Stein afin d'obtenir des vitesses de convergences pour ces théorèmes. Nous présentons une approche générale pour attaquer ce problème. Elle consiste à définir un semi-groupe $(P_t)_{t \geq 0}$ directement *via* une formule à la Mehler:

$$P_t f(x) := \mathbb{E} \left[f \left(d_{e^{-t}x} \oplus d_{(1-e^{-\alpha t})^{1/\alpha} Z} \right) \right], \quad x \in \mathbb{K}$$

où Z est une variable aléatoire *strictement stable*, au sens de (0.12). Cette propriété de stabilité garantit que la famille d'opérateurs $(P_t)_{t \geq 0}$ est bien un semi-groupe de mesure stationnaire la loi de Z . Les semi-groupes admettant ce type de représentation possèdent souvent certaines propriétés en commun avec le semi-groupe d'Ornstein-Uhlenbeck original. Il ne semble cependant pas possible à ce point d'établir une liste exhaustive des propriétés communes que satisfont ces différents semi-groupes, du moins pas sans faire des hypothèses supplémentaires sur le cône convexe \mathbb{K} . Dans cette thèse, nous nous concentrons principalement sur le cas où \oplus est le maximum et d_a la multiplication habituelle. Certaines propriétés du semi-groupe associé à ces opérations seront obtenus *via* des arguments génériques, tandis que d'autres seront plus spécifiques à notre cas particulier.

Nous concluons cette partie en mentionnant brièvement un autre moyen de quantifier certaines lois de probabilités. Il repose sur la notion de *stabilité discrète*, voir par exemple [Steutel and van Harn \[1979\]](#), [Davydov et al. \[2011\]](#). L'idée consiste à étendre le concept de stabilité aux lois discrètes. Pour ces dernières, une mise-à-l'échelle par un réel positif quelconque ne fait pas sens *a priori*. Pour pallier cette difficulté, une opération aléatoire remplace la mise-à-l'échelle déterministe. Elle continue à satisfaire, au moins formellement, les propriétés qu'on attend d'une multiplication, au sens suivant : $p \circ (q \circ x)$ est égal en loi à $(pq) \circ x$, et $p \circ (x \oplus y)$ a même loi que $(p \circ x) \oplus (p \circ y)$, pour x dans E et pour tout p, q in a certain subset of \mathbb{R}_+ , e.g $[0, 1]$. L'*amincissement* par $p \in [0, 1]$ est le principal exemple de ce type de mise à l'échelle sur $E = \mathbb{N}$:

$$p \circ n = \sum_{i=1}^n B_i, \quad p \in [0, 1]$$

où les B_i sont des variables aléatoires i.i.d. de loi $\mathcal{B}(p)$. La loi de Poisson n'est pas stable mais vérifie cependant la propriété suivante de stabilité discrète :

$$p \circ N_1 + (1 - p) \circ N_2 \stackrel{d}{=} N_1,$$

où N_1, N_2 sont des variables aléatoires i.i.d. de même loi de Poisson. En réalité, ce résultat n'est qu'un reflet de la propriété de stabilité discrète du processus de Poisson. Si $\phi = \sum_{i=1}^n \delta_{x_i}$ est une configuration finie, alors on définit l'amincissement de ϕ par p en retirant chaque point x_i de ϕ avec probabilité $1 - p$, indépendamment les uns des autres. Dit autrement :

$$p \circ \phi = \sum_{i=1}^n B_i \delta_{x_i}.$$

On peut alors prouver que :

$$p \circ \eta_1 \cup (1 - p) \circ \eta_2 \stackrel{d}{=} \eta_1,$$

où η_1, η_2 sont deux processus de Poisson i.i.d. sur un espace Polonais E , de même mesure intensité finie. Cette propriété de stabilité discrète nous permet de reproduire la construction de construire un semi-groupe de mesure stationnaire le processus de Poisson *via* une formule à la Mehler en posant :

$$P_t f(\phi) = \mathbb{E}[f(e^{-t} \circ \phi + (1 - e^{-t}) \circ \eta)], \phi \in \mathfrak{N}_E.$$

On retrouve ainsi le semi-groupe de Glauber (voir par exemple [Decreusefond \[2022\]](#)). Ces idées ont été adaptées avec succès au cas des processus discrets α -stables by A. Vasseur dans sa thèse de doctorat [Vasseur \[2017\]](#), lui donnant ainsi la possibilité d'appliquer la méthode de Stein pour quantifier des théorèmes-limites impliquant ces processus. Ce principe est général et peut être utilisé pour construire d'autres semi-groupes, *a priori* inédits. Cette thèse en est un exemple.

Analyse stochastique pour les lois max-stables

Les idées développées par Davydov et.al dans [Davydov et al. \[2008\]](#) reposent sur la notion de *série de LePage*. Ce concept permet de représenter une variable aléatoire stable pour une opération binaire \oplus donnée (telle l'addition, le maximum, *etc.*) comme une somme infinie (au sens de \oplus) d'un processus de Poisson dont la mesure intensité satisfait une certaine propriété d'homogénéité. Une telle représentation était déjà connue pour les variables aléatoires α -stables ainsi qu'en théorie des valeurs extrêmes, où elle porte le nom de *décomposition de Haan*. Cela ne constitue pas une surprise, dans la mesure où (\mathbb{R}, \oplus, d) est bien un cône convexe lorsque \oplus représente l'addition ou le maximum, et d la multiplication par un scalaire.

La décomposition de Haan permet d'établir un lien crucial entre les lois max-stables et les processus de Poisson. Elle apparaîtra sous différentes formes tout au long de cette thèse, principalement *via* le semi-groupe $(\mathbf{P}_t^\alpha)_{t \geq 0}$ que nous définirons plus tard et dont la mesure stationnaire sera une certaine loi max-stable ν (c'est-à-dire une loi d'extremum). Par exemple, le générateur \mathcal{L}_α de ce semi-group sera exprimé à l'aide de la mesure intensité du même processus de Poisson que celui qui intervient dans la décomposition de de Haan-LePage de ν .

Une autre façon d'employer cette connexion entre processus de Poisson et lois max-stables est de spécialiser des résultats classiques d'analyse stochastique pour les processus de Poisson, tels ceux exposés dans [Last and Penrose \[2011\]](#), [Last et al. \[2016\]](#). On obtient ainsi à peu de frais de nouveaux résultats pour les loi d'extremum, ou bien des preuves courtes et simples de théorèmes déjà connus.

Contributions

Les contributions de cette thèse sont les suivantes : au chapitre 2, nous définissons plusieurs semi-groupes de Markov admettant les lois d'extremum comme mesure stationnaire. Ces définitions sont données d'abord en dimension 1 puis étendues en dimension supérieure. Plusieurs identités fonctionnelles (relations de commutation, inégalité de Poincaré, *etc.*) sont présentées. Elles contribuent à éclairer la structure de ces semi-groupes et fournissent dans le même temps des informations supplémentaires sur les lois d'extremum. Nous exhibons aussi un processus de Markov dont le semi-groupe correspond à celui défini auparavant pour la loi de Fréchet. Tout au long de ce chapitre, un parallèle avec le semi-groupe d'Ornstein-Uhlenbeck est dressé.

Le chapitre 3 est centré sur les applications des résultats précédents à la méthode de Stein dans le cas où la loi-cible est la mesure de Fréchet, y compris sa généralisation multivariée. Nous donnons et résolvons l'équation de Stein, décrivons les propriétés de sa solution lorsque la distance utilisée est celle de Wasserstein, de Kolmogorov et la distance $d_{[2]}$. Nous faisons ensuite appel à ces propriétés pour obtenir des vitesses de convergence dans plusieurs théorèmes-limites impliquant la loi de Fréchet.

Nous suivons la même approche pour le chapitre 4, cette fois en nous concentrant sur la loi de Gumbel et celle de Weibull. Une application des résultats exposés concerne le problème du collectionneur. Nous y établissons un résultat de type vitesse de convergence du temps de complétion renormalisé vers la loi de Gumbel, et ce pour une distance similaire à celle de Wasserstein, la distance $d_{[2]}$.

Enfin, le chapitre 5 porte sur les connexions entre analyse stochastique pour les processus de Poisson et lois d'extremum généralisées, et plus généralement, avec les lois stables et les lois max-infiniment divisibles. En particulier, nous donnons une preuve courte d'une version multivariée de l'inégalité de Poincaré pour la famille de semi-groupes introduite au chapitre 2. Nous obtenons aussi une identité de covariance pour les lois max-infiniment divisibles. La dernière section de ce chapitre est de nature plus heuristique et vise à définir un nouveau semi-groupe sur l'espace des configurations, dont la mesure stationnaire est un processus de Poisson de mesure intensité vérifiant une certaine propriété d'homogénéité.

1 Preliminaries

Contents

1.1	Markov semi-groups	31
1.1.1	Functional inequalities	33
1.2	Poisson process	34
1.2.1	Definition and basic properties	34
1.2.2	Stochastic analysis for the Poisson process	36
1.2.3	The Fock space representation	37
1.2.4	The Glauber semi-group	38
1.3	Extreme value theory	39
1.3.1	The univariate case	39
1.3.2	The multivariate case	45
1.3.3	Max-infinitely divisible distributions	55
1.3.4	Construction and properties of the extremal integral	56
1.4	Distances between distributions	59

This chapter is dedicated to giving an overview of the notions essential to this work, as well to fixing some notations used throughout the next chapters.

1.1 Markov semi-groups

In this section we give the very basics of Markov semi-groups theory. Classic references on the subject include [Bakry et al. \[2014\]](#), [Ethier and Kurtz \[1986\]](#), [Rogers and D.Williams \[April 2000\]](#).

Definition 1.1.1 (Semi-group). Let (X, \mathcal{X}) be a measurable space. A *semi-group* on X is a family $(P_t)_{t \geq 0}$ of linear operators defined on a subspace V of real-valued, bounded measurable functions on X , such that the following properties are satisfied

1. For all $t \geq 0$ and all f in V , $P_t f$ is also measurable and bounded.
2. (**Conservation of mass**) $P_0 = \text{Id}$, the identity operator on V .
3. (**Semi-group property**) For all $t, s \geq 0$, we have $P_t \circ P_s = P_{t+s}$.

Definition 1.1.2 (Invariant measure). Let ν be a σ -finite measure on the measurable space

(X, \mathcal{X}) . A Markov semi-group $(P_t)_{t \geq 0}$ on (X, \mathcal{X}) admits ν as an *invariant measure* if

$$\int_X P_t f(x) \, d\nu(x) = \int_X f(x) \, d\nu(x), \quad t \geq 0, \quad f \in V.$$

Definition 1.1.3 (Markov semi-group). Let $(P_t)_{t \geq 0}$ be a semi-group on the measurable space (X, \mathcal{X}) with invariant measure ν . A semi-group which also satisfies the next two properties is called a *Markov semi-group*:

4. (**Markov property**) For all $t \geq 0$, $P_t f \geq 0$ if $f \in V$ takes only nonnegative values, and $P_t \mathbf{1} = \mathbf{1}$, the function constant equal to 1.
5. (**Right-continuity**) For all $f \in V$, the function $t \mapsto P_t f$ is right-continuous in $\mathbf{L}^2(X, \mathcal{X}, \nu)$.

Due to the semi-group property, it is clear one just needs to check the right-continuity property at $t = 0$. Furthermore, Markov semi-groups automatically satisfy the contraction property:

$$\|P_t f\|_{\mathbf{L}^2(X, \nu)} \leq \|f\|_{\mathbf{L}^2(X, \nu)}, \quad t \geq 0, \quad f \in V.$$

A semi-group $(P_t)_{t \geq 0}$ can be *symmetric* (or reversible) with respect to a measure ν , in which case the said measure is invariant under $(P_t)_{t \geq 0}$ (take $g = \mathbf{1}$ in the definition below):

Definition 1.1.4 (Symmetry). Let ν be a measure on (X, \mathcal{X}) . A semi-group $(P_t)_{t \geq 0}$ is *symmetric* with respect to a measure ν if P_t is a symmetric operator for the standard inner product on $\mathbf{L}^2(X, \nu)$ for any nonnegative t :

$$\langle P_t f, g \rangle_{\mathbf{L}^2(X, \nu)} = \langle f, P_t g \rangle_{\mathbf{L}^2(X, \nu)}, \quad t \geq 0, \quad f, g \in V.$$

Definition 1.1.5 (Domain and generator of a Markov semi-group). The *domain* of a Markov semi-group $(P_t)_{t \geq 0}$ on $\mathbf{L}^2(X, \nu)$ is defined as:

$$\text{Dom}(\mathcal{L}) := \left\{ f \in V, \lim_{t \rightarrow 0^+} \frac{P_t f - f}{t} \text{ exists a.s. and in } \mathbf{L}^2(X, \nu) \right\}.$$

In this case we denote by \mathcal{L} its *generator*:

$$\mathcal{L}f := \lim_{t \rightarrow 0^+} \frac{P_t f - f}{t}.$$

The defining assumptions of Markov semi-groups imply that their generators \mathcal{L} of $(P_t)_{t \geq 0}$ admit a pseudo-inverse.

Proposition 1.1.6 Let $(P_t)_{t \geq 0}$ be a Markov semi-group with invariant measure ν and generator \mathcal{L} . For all $f \in V$ such that $t \mapsto P_t f(x)$ is integrable for ν -almost all $x \in X$, we have

$$-\int_0^\infty \mathcal{L}P_t f(x) \, dt = f(x) \quad \text{a.e.} \tag{1.1}$$

if $\mathbb{E}_\nu[f] = 0$.

From now on, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 1.1.7 (Markov process). Let $(X_t)_{t \geq 0}$ be a measurable process on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $(\mathcal{F}_t)_{t \geq 0}$ be the filtration generated by $(X_t)_{t \geq 0}$. We say that $(X_t)_{t \geq 0}$ is a *Markov process* if it has the *Markov property*:

$$\mathcal{L}(X_{t+h} | \mathcal{F}_t, X_0 = x) = \mathcal{L}(X_h | X_0 = X_t), \quad t, h \geq 0, \quad x \in \mathbf{L}^2(\Omega, \mathbb{P}).$$

where $\mathcal{L}(X | \mathcal{F})$ denotes the distribution of X given the σ -field \mathcal{F} .

There exists a correspondence between a Markov process $(X_t)_{t \geq 0}$, its semi-group $(P_t)_{t \geq 0}$ and its generator \mathcal{L} : we have already seen how to derive the generator from the semi-group. The latter can be defined from the Markov process by setting:

$$P_t f(x) := \mathbb{E}[f(X_t) | X_0 = x].$$

By the Markov property, this expression indeed defines a Markov semi-group. Thus, knowing P_t is the same as knowing the distribution of X_t conditional on $X_0 = x$ for all t and x . On the other hand, P_t can be retrieved knowing \mathcal{L} via the informal equality $P_t = e^{t\mathcal{L}}$.

Through this connection between those three objects, it is possible to construct a martingale from a Markov process.

Theorem 1.1.8 Let $(X_t)_{t \geq 0}$ be a Markov process with generator \mathcal{L} , and f a function in the domain of \mathcal{L} such that $\mathbb{E}[\mathcal{L}f(X_t)]$ exists for all t . Then the process $(M_t)_{t \geq 0}$ defined by:

$$M_t := f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_u) \, du$$

is a martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$.

1.1.1 Functional inequalities

The main reference for this subsection is [Bakry et al. \[2014\]](#). Throughout this section we assume that a standard algebra exists for the considered Markov semi-groups:

Definition 1.1.9 (Standard algebra). Let \mathcal{A} be an algebra of functions included in every $\mathbf{L}^p(X, \nu)$, with $p \in (1, \infty)$. Assume it is dense in each of them. We will say that \mathcal{A} is a *standard algebra* if \mathcal{A} is stable under \mathcal{L} and all P_t . When ν is a probability measure, we will also require that \mathcal{A} contains every constant function.

This definition ensures us that all the objects defined in this section are well-defined. Typical examples for \mathcal{A} include the space of rapidly decreasing functions (Schwartz space) for the heat semi-group, or the space of slowly increasing functions for the Ornstein-Uhlenbeck semi-group.

Definition 1.1.10 (Carré du champ and energy). Let \mathcal{L} be the generator of some Markov semi-group $(P_t)_{t \geq 0}$ on (X, \mathcal{X}) , and μ a probability measure on (X, \mathcal{X}) .

- The *carré du champ operator* associated with \mathcal{L} is the symmetric bilinear form Γ defined as:

$$\Gamma(f, g) := \frac{1}{2}(\mathcal{L}(fg) - f\mathcal{L}(g) - g\mathcal{L}f), \quad f, g \in \mathcal{A}.$$

We shorten notations by writing $\Gamma(f)$ for $\Gamma(f, f)$.

- The *energy form* \mathcal{E}_μ is defined as:

$$\mathcal{E}_\mu(f) := -\mathbb{E}_\mu[f\mathcal{L}f], \quad f \in \mathcal{A}.$$

The next proposition gives the most basic properties of those two objects.

Proposition 1.1.11 *Let \mathcal{L} be a generator with invariant measure μ . For all $f \in \mathcal{A}$, we have:*

- $\Gamma(f) \geq 0$.
- $\mathcal{E}_\mu(f) = \mathbb{E}_\mu[\Gamma(f)]$.

Proposition 1.1.12 *Let $(P_t)_{t \geq 0}$ a semi-group with invariant measure μ and generator \mathcal{L} . The following assertions are equivalent:*

- The measure μ satisfies a Poincaré inequality with constant c :

$$\mathbb{V}_\mu(f) \leq c\mathcal{E}_\mu(f), \quad f \in \mathcal{A}.$$

- There exists a constant $\lambda > 0$ such that

$$\|P_t f - \mathbb{E}_\mu[f]\|_{\mathbf{L}^2(X, \mu)} \leq e^{-\lambda t} \|f\|_{\mathbf{L}^2(X, \mu)}, \quad t \in \mathbb{R}_+, \quad f \in \mathcal{A}$$

Furthermore the optimal constants are linked by the relation $c = \lambda^{-1}$.

1.2 Poisson process

The Poisson process plays a central role in extreme value theory and possesses many important properties. This section gathers those which will be the most relevant to us. Excellent references about the Poisson process are countless, but we will cite and use only a few of them: [Last and Penrose \[2011, 2017\]](#), [Decreusefond \[2022\]](#), [Privault \[2009\]](#). The reference textbooks by Resnick regarding extreme value theory [Resnick \[1987, 2006\]](#) also contain a crash course about point processes, including most, but not all, of the properties of the Poisson process we will make use of. In this section, unless specified otherwise, measures will always be σ -finite.

1.2.1 Definition and basic properties

Definition 1.2.1 (Configuration). Let E be a subset of \mathbb{R}^d . A *configuration* ϕ of E is a locally finite set of E , in the sense that every bounded subset of E contains only a finite number of points of ϕ , and where repetitions are allowed. We denote by \mathfrak{N}_E the set of configurations of E .

Remark 1.2.2. - The fact that repetitions are allowed means that $\{x, x\}$ is different from $\{x\}$.

- With this definition, a configuration is necessarily a countable subset of E . As a result, we will sometimes denote by $\phi = (x_i)_{i \in I}$ the points of a configuration, where I is a subset of \mathbb{N} . Another

way to denote a configuration is by seeing it as a random counting measure:

$$\phi_A := \phi(A) = \sum_{i \in I} \delta_{x_n}(A), \quad A \subseteq E$$

where δ_x is the Dirac measure at x .

- On the other hand, observe that a countable set is not always a configuration: the set $\phi = \{1/n, n \geq 1\}$ is not a configuration in $E = [0, 1]$ because it admits an accumulation point at 0: every interval $[0, \varepsilon]$ is bounded but contains an infinite number of points of ϕ , for $\varepsilon \in (0, 1]$.

■

Following the conventions of [Decreusefond \[2022\]](#) (page 121), we will indistinctly denote a point process as a random measure or a random set.

Several operations on configurations are possible.

Definition 1.2.3. - The *superposition* of two configurations ϕ_1, ϕ_2 is the reunion of ϕ_1 and ϕ_2 considered as sets.

- The *thinning* of a configuration $\phi = \sum_{i \in I} \delta_{x_i}$ by a scalar $p \in [0, 1]$ is defined as the random measure:

$$p \circ \phi := \sum_{i \in I} \varepsilon_i \delta_{x_i},$$

where the ε_i are i.i.d. random variables with the Bernoulli distribution $\mathcal{B}(p)$.

- The *dilatation* of a configuration $\phi = \sum_{i \in I} \delta_{x_i}$ by a scalar $\lambda \in \mathbb{R}_+^*$ is defined by:

$$\lambda \phi := \sum_{i \in I} \delta_{\lambda x_i},$$

while its *translation* by a vector $\tau \in \mathbb{R}^d$ is defined by

$$\phi + \tau := \sum_{i \in I} \delta_{x_i + \tau}.$$

Observe that the space of configurations \mathfrak{N}_E is closed under those operations.

Definition 1.2.4 (Point process). A *point process* η is an \mathfrak{N}_E -valued random variable.

The binomial process is one of the most basic instances of a point process and a prototype of the Poisson process.

Definition 1.2.5 (Binomial process). Let X_1, \dots, X_n be n i.i.d. random variables on \mathbb{R}^d with distribution μ . Then the configuration $\phi := \sum_{i=1}^n \delta_{X_i}$ is called a *binomial process* on \mathbb{R}^d with n points and base measure μ .

A (finite) Poisson process can be seen as a binomial process with a Poisson number N of points, the latter being independent of N . The general case builds on this intuition.

Definition 1.2.6 (Poisson process). Let λ be a σ -finite measure on \mathbb{R}^d . A *Poisson point process* η with intensity measure λ is a random counting measure which satisfies the two following properties:

1. η_A has the Poisson distribution $\mathcal{P}(\lambda(A))$ for all subset $A \subseteq E$.
2. For all $m \in \mathbb{N}^*$, the random variables $\eta_{A_1}, \dots, \eta_{A_m}$ are independent if the sets A_1, \dots, A_m are pairwise disjoint.

It is not obvious that only one point process, if any, satisfies all those properties. It is true however, even though we will not delve into the construction of the Poisson process. For this we refer to [Last and Penrose \[2017\]](#), among other classic references. The next result is used as the working definition of the Poisson process in [Privault \[2009\]](#).

Theorem 1.2.7 Let η be a Poisson process on \mathbb{R}^d with finite intensity measure λ . Then for all $f : \mathfrak{N}_E \rightarrow \mathbb{R}_+$,

$$\mathbb{E}[f(\eta)] = e^{-\lambda(E)} f(\emptyset) + e^{-\lambda(E)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_{E^n} f(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\lambda^n(\mathbf{x}_1, \dots, \mathbf{x}_n), \quad (1.2)$$

Poisson processes behave well with respect to superposition and thinning.

Proposition 1.2.8 Let η_1, η_2 two independent Poisson process with σ -finite intensity measures λ_1, λ_2 respectively, and $p \in [0, 1]$ Then

- $\eta_1 \cup \eta_2$ is a Poisson process with intensity measure $\lambda_1 + \lambda_2$.
- $p \circ \eta_1$ is a Poisson process with intensity measure $p\lambda_1$.

One of the most important properties of the Poisson process is the Campbell-Mecke formula. Actually, it characterizes the Poisson process.

Theorem 1.2.9 — Campbell-Mecke formula. Let η be a point process on E and λ a σ -finite measure on E . For any $f : E \times \mathfrak{N}_E \rightarrow \mathbb{R}$ such that

$$\mathbb{E} \left[\int_E |f(x, \eta)| d\lambda(x) \right] < +\infty,$$

we have

$$\mathbb{E} \left[\int_E f(x, \eta) d\eta(x) \right] = \int_E \mathbb{E}[f(x, \eta + \delta_x)] d\lambda(x) \quad (1.3)$$

if and only if η is a Poisson process with intensity measure λ .

1.2.2 Stochastic analysis for the Poisson process

Stochastic analysis for the Poisson process has its own notion of gradient, just like Malliavin calculus.

Definition 1.2.10 (Discrete gradient and its domain). We define the *discrete gradient* of $f : \mathfrak{N}_E \rightarrow \mathbb{R}$ as:

$$D_x^+ f(\phi) = f(\phi + \delta_x) - f(\phi), \quad \phi \in \mathfrak{N}_E$$

and

$$D_x^- f(\phi) = f(\phi - \delta_x) - f(\phi), \quad \phi \in \mathfrak{N}_E$$

where the notation $\phi - \delta_x$ should be understood as

$$\phi - \delta_x = \begin{cases} \phi \setminus \{x\} & \text{if } x \in \phi \\ \phi & \text{otherwise.} \end{cases}$$

Let η be a Poisson process on E with intensity measure λ . The *domain of the gradient* is defined as

$$\text{Dom}(D) := \left\{ f : \mathfrak{N}_E \rightarrow \mathbb{R} \mid \mathbb{E}[f(\eta)^2] < \infty \text{ and } \mathbb{E} \left[\int_E |f(\eta + \delta_x) - f(\eta)|^2 d\lambda(x) \right] < +\infty \right\}.$$

Definition 1.2.11 (Divergence and its domain). Let η be a Poisson process on E with intensity measure λ . The *domain of the divergence* is defined as

$$\text{Dom}_2(\delta) := \left\{ U : \mathfrak{N}_E \rightarrow \mathbb{R}, \mathbb{E} \left[\left(\int_E U(x, \eta - \delta_x) (d\eta(x) - d\lambda(x)) \right)^2 \right] < +\infty \right\}.$$

The divergence of U is

$$\delta U(\eta) := \int_E U(x, \eta - \delta_x) d\eta(x) - \int_E U(x, \eta) d\lambda(x).$$

The next result is a consequence of the Campbell-Mecke formula.

Theorem 1.2.12 — Integration-by-parts for the Poisson process. For any $f \in \text{Dom}(D)$ and any $U \in \text{Dom}_2(\delta)$, we have:

$$\mathbb{E} \left[\int_E D_x^+ f(\eta) U(x, \eta) d\lambda(x) \right] = \mathbb{E}[f(\eta) \delta U(\eta)]. \quad (1.4)$$

1.2.3 The Fock space representation

Denote by \mathbb{P}_η the distribution of a point process η . The inner product on $\mathbf{L}^2(E^n, \lambda^n)$ is defined as:

$$\langle f, g \rangle_n := \int_{E^n} f(x_1, \dots, x_n) g(x_1, \dots, x_n) d\lambda(x_1, \dots, x_n).$$

We also define for $n \geq 1$

$$T_n f(x_1, \dots, x_n) := \mathbb{E}[D_{x_1, \dots, x_n}^+ f(\eta)],$$

where $D_{x_1, \dots, x_n}^+ := D_{x_1}^+ \circ \dots \circ D_{x_n}^+$ (notice those operators commute). For $n = 0$, we set $T_0 f(x) := \mathbb{E}[f(\eta)]$.

Theorem 1.2.13 — Fock space representation. For all $f, g \in \mathbf{L}^2(\mathfrak{N}_E, \mathbb{P}_\eta)$,

$$\mathbb{C}\text{ov}(f(\eta), g(\eta)) = \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n f, T_n g \rangle_n. \quad (1.5)$$

A consequence of this result is the Poincaré inequality for the Poisson process.

Theorem 1.2.14 — Poincaré inequality. Assume that $f \in \mathbf{L}^2(\mathfrak{N}_E, \mathbb{P}_\eta)$. Then

$$\mathbb{V}(f(\eta)) \leq \int_E \mathbb{E}[(D_x^+ f(\eta))^2] d\lambda(x). \quad (1.6)$$

1.2.4 The Glauber semi-group

Definition 1.2.15 (Glauber semi-group). Let η be a Poisson process with σ -finite intensity measure λ . Set

$$\mathbf{P}_t^G f(\phi) := \mathbb{E}[f(e^{-t} \circ \phi \cup (1 - e^{-t}) \circ \eta)], \quad t \geq 0, \phi \in \mathfrak{N}_E, \quad (1.7)$$

for $f : \mathfrak{N}_E \rightarrow \mathbb{R}$.

Theorem 1.2.16 $(\mathbf{P}_t^G)_{t \geq 0}$ is an ergodic semi-group on \mathfrak{N}_E whose stationary measure is the distribution of the Poisson process with intensity measure λ . Its generator has its domain included in the set of bounded $f : \mathfrak{N}_E \rightarrow \mathbb{R}$ and is equal to:

$$\mathcal{L}_G f(\phi) = \int_E D_x^+ f(\phi) d\lambda(x) + \int_E D_x^- f(\phi) d\phi(x).$$

Furthermore it satisfies the commutation relation:

$$D_x^+ \mathbf{P}_t^G f(\phi) = e^{-t} \mathbf{P}_t^G D_x^+ f(\phi), \quad x \in E. \quad (1.8)$$

Finally, \mathcal{L}_G is invertible from $\mathbf{L}_0^2(\mathfrak{N}_E, \mathbb{P}_\eta)$ onto itself.

This semi-group yields two other covariance equalities which play a crucial role in the proof of a concentration inequality for functionals of the Poisson process.

Theorem 1.2.17 Let f, g be two functions belonging to $\text{Dom}(D)$. Then the two following identities hold:

$$\mathbb{C}\text{ov}(f(\eta), g(\eta)) = \int_0^\infty \int_E \mathbb{E}[D_x^+ \mathbf{P}_t^G f(\eta) D_x^+ g(\eta)] d\lambda(x) dt. \quad (1.9)$$

If furthermore f is centered, i.e. $\mathbb{E}[f(\eta)] = 0$,

$$\mathbb{C}\text{ov}(f(\eta), g(\eta)) = \int_E \mathbb{E}[D_x^+(\mathcal{L}_G^{-1} f)(\eta) D_x^+ g(\eta)] d\lambda(x).$$

Theorem 1.2.18 — Concentration inequality. Let η be a Poisson process with σ -finite intensity

measure λ and $f : \mathfrak{N}_E \rightarrow \mathbb{R}$. Assume that there exist non-negative α and β such that

$$D_x^+ f(\eta) \leq \alpha, \lambda \otimes \mathbb{P}_\eta\text{-a.s. and } \sup_{\phi \in \mathfrak{N}_E} \int_E D_x^+ f(\phi) d\lambda(x) \leq \beta, \mathbb{P}_\eta\text{-a.s.}$$

Then we have

$$\mathbb{P}(f(\eta) - \mathbb{E}[f(\eta)] \geq r) \leq e^{-\frac{r}{2\alpha} \log(1 + \frac{r\alpha}{\beta})}, r \geq 0. \tag{1.10}$$

1.3 Extreme value theory

Extreme value theory deals with records and their frequencies. It has been first applied to unidimensional observations, such as precipitation levels or loss amounts. The generalization of this theory to higher dimension has proved successful and allows for studying spatial extremes, that is, record observations over dependent sites. However it also comes with non-trivial difficulties, as the usual order on \mathbb{R} is no longer total on \mathbb{R}^d .

1.3.1 The univariate case

References for univariate EVT are numerous: let us mention among others [Resnick \[1987\]](#), [Leadbetter et al. \[2012\]](#), [Embrechts et al. \[2013\]](#), [Resnick \[2006\]](#), [de Haan and Ferreira \[2007\]](#).

The most basic record model in EVT consists in taking i.i.d. observations X_1, \dots, X_n and looking at the highest of them, *i.e.* $M_n = \max_{1 \leq i \leq n} X_i$. And just as in the study of sums of random variables, one is also interested with the asymptotic behavior of M_n . It will heavily depend on the support of the X_i : if it is unbounded on the right, then Borel-Cantelli's lemma tells us that M_n will a.s. go to infinity as the number of observations n increases. Otherwise, it is bounded on the right, say by x_1 , and then M_n will a.s. converge to x_1 . This result can be understood as an analog of the law of large numbers, where the sum $+$ operator has been replaced by the max operator \oplus . It is possible to quantify the speed of convergence of M_n to those two possible limits, *i.e.* an analog of the central limit theorem also exists. This is important in applications to know how many observations are necessary on average for M_n to go beyond a given threshold. Three possible distributions can play the role of the normal distribution (or the stable distributions when the X_i do not admit a variance). Just like the usual stable distributions, they appear naturally as the only distributions which are stable under the max operator.

Definition 1.3.1 (One-dimensional max-stability). Let μ be a probability distribution on \mathbb{R} . We say that μ is *max-stable* if for any $n \geq 1$, and if X_1, \dots, X_n are n i.i.d. random variables with common distribution μ , there exist a_n, b_n two deterministic constants in $\mathbb{R}_+ \times \mathbb{R}$ such that

$$\bigoplus_{i=1}^n X_i \stackrel{d}{=} a_n X_1 + b_n.$$

If $b_n = 0$, then we say that μ is *strictly max-stable*. A random vector \mathbf{X} is said to be *min-stable* if the distribution of $-\mathbf{X}$ is max-stable.

We have abused notations by having said that a random variable is max-stable if its distribution is max-stable. The next definition gives examples of max-stable distributions.

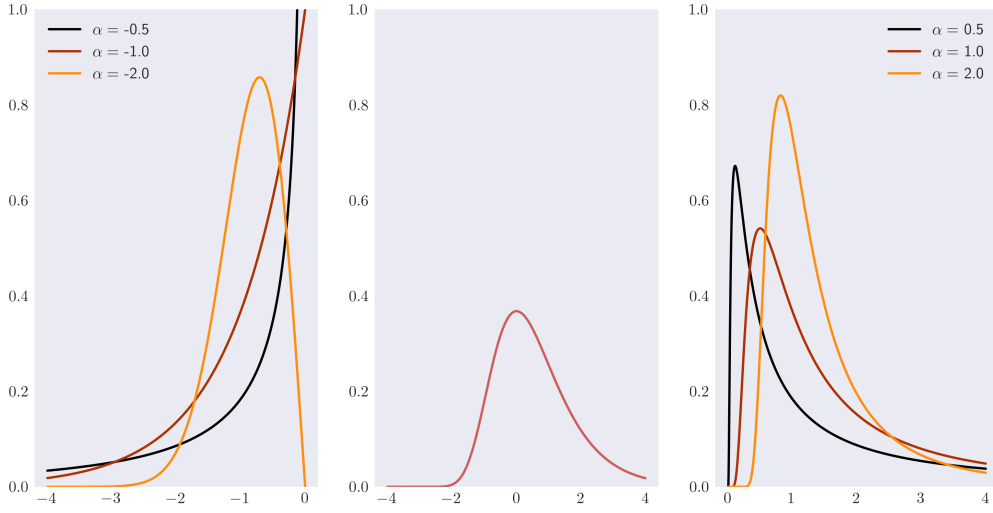


Figure 1.1: p.d.f of $\mathcal{W}(\alpha)$, $\mathcal{G}(0, 1)$ and $\mathcal{F}(\alpha)$ for $\alpha \in \{0.5, 1, 2\}$

Definition 1.3.2 (Fréchet, Gumbel, Weibull distributions). The next functions define probability distributions.

- The *Fréchet distribution* $\mathcal{F}(\alpha, \sigma)$ with shape parameter $\alpha > 0$ and scale parameter $\sigma > 0$ has c.d.f.

$$\Phi_{\alpha, \sigma}(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-\left(\frac{\sigma}{x}\right)^\alpha} & \text{otherwise.} \end{cases}$$

- The *Gumbel distribution* $\mathcal{G}(\mu, \sigma)$ with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ has c.d.f.

$$\Lambda_{\mu, \sigma}(x) := e^{-e^{-\frac{x-\mu}{\sigma}}}, \quad x \in \mathbb{R}.$$

- The *Weibull distribution* $\mathcal{W}(\alpha, \sigma)$ with shape parameter $\alpha > 0$ and scale parameter $\sigma > 0$ has c.d.f.

$$\Psi_{\alpha, \sigma}(x) := \begin{cases} e^{-\left(-\frac{x}{\sigma}\right)^\alpha} & \text{if } x \leq 0 \\ 1 & \text{otherwise.} \end{cases}$$

When $\sigma = 1$ (and $\mu = 0$ for the Gumbel distribution), we will drop the location and scale parameters in the notations above.

The probability density functions of those distributions have the following shapes:

Those distributions are indeed max-stable.

Proposition 1.3.3 — Max-stability. *Let X, Y be i.i.d. random variables and fix $\alpha > 0$.*

- Assume $X, Y \sim \mathcal{F}(\alpha, \sigma)$. Then:

$$[\lambda^{\frac{1}{\alpha}} X] \oplus [(1 - \lambda)^{\frac{1}{\alpha}} Y] \stackrel{d}{=} X, \lambda \in [0, 1], \quad (1.11)$$

- Assume $X, Y \sim \mathcal{G}(0, 1)$. Then:

$$[X + \log(\lambda)] \oplus [Y + \log(1 - \lambda)] \stackrel{d}{=} X, \lambda \in [0, 1]. \quad (1.12)$$

- Assume $X, Y \sim \mathcal{W}(\alpha, \sigma)$. Then:

$$[\lambda^{-\frac{1}{\alpha}} X] \oplus [(1 - \lambda)^{-\frac{1}{\alpha}} Y] \stackrel{d}{=} X, \lambda \in [0, 1]. \quad (1.13)$$

Remark 1.3.4. In the sequel we will work mainly with the usual Weibull distribution, which is the same as the negative Weibull distribution up to a change of sign. It will be usually called the *positive Weibull distribution* and its distribution will be denoted by $\Psi_{\alpha, \sigma}^+$. Therefore it has support in \mathbb{R}_+ and is min-stable. More precisely, denote by $x \odot y$ the minimum of x and y . The stability relation satisfied by the positive Weibull distribution writes as:

$$[\lambda^{-\frac{1}{\alpha}} X] \odot [(1 - \lambda)^{-\frac{1}{\alpha}} Y] \stackrel{d}{=} X, \lambda \in [0, 1], \quad (1.14)$$

where X and Y are independent and have the Weibull distribution $\mathcal{W}_+(\alpha, \sigma)$ with c.d.f.

$$\Psi_{\alpha, \sigma}^+(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-\left(\frac{x}{\sigma}\right)^\alpha} & \text{otherwise.} \end{cases}$$

The Weibull distribution generalizes classic probability laws such as the exponential distribution ($\alpha = 1$) or the Rayleigh distribution ($\alpha = 2$). ■

Actually, these are the only possible max-stable distributions in dimension 1. This point is proved in the proof of the next theorem (see for instance [Resnick \[1987\]](#)), which is a cornerstone of extreme value theory.

Theorem 1.3.5 — Fisher–Tippett–Gnedenko theorem. *Let $(X_i)_{i \geq 1}$ be a sequence of i.i.d. random variables with c.d.f. F and set $M_n := \max_{1 \leq i \leq n} X_i$. Assume that there exist two sequences of real numbers $(a_n)_{n \geq 1}, (b_n)_{n \geq 1}$, with $a_n > 0$ for all n , such that*

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \xrightarrow{n \rightarrow \infty} G(x),$$

where G is a non-degenerate c.d.f. and x is a point of continuity of G . Then G is the c.d.f. of either the Fréchet distribution, the Gumbel distribution or the Weibull distribution.

It is possible to standardize those three probability laws into one unique distribution. Although it allows for more elegant and compact formulations in certain contexts (see theorem [1.3.14](#)) below), we will not make much use of it in the sequel for it may prove unwieldy for our purposes.

Definition 1.3.6 (Generalized extreme value distribution). For $\xi \in \mathbb{R}$, denote by H_ξ the following c.d.f.

$$H_\xi(x) := \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0 \\ e^{-x} & \text{if } \xi = 0, \end{cases}$$

for the x such that $1 + \xi x > 0$.

Due to the importance of the three max-stable distributions, it is essential to know a few things about them, in particular their moments and their connections.

Proposition 1.3.7 - The Fréchet distribution $\mathcal{F}(\alpha, \sigma)$ admits moments of all order up to α excluded:

$$\mathbb{E}[Z^k] = \frac{1}{\sigma^k} \Gamma\left(1 - \frac{k}{\alpha}\right), \quad k < \alpha.$$

- The Gumbel distribution $\mathcal{G}(\mu, \sigma)$ admits moments of all order:

$$\mathbb{E}[(Z - \mu)^k] = \frac{1}{\sigma^k} \Gamma^{(k)}(1), \quad k \in \mathbb{N}.$$

- The negative Weibull distribution $\mathcal{W}(\alpha, \sigma)$ admits moments of all order:

$$\mathbb{E}[Z^k] = \left(-\frac{1}{\sigma}\right)^k \Gamma\left(1 + \frac{k}{\alpha}\right), \quad k \in \mathbb{N}.$$

Proposition 1.3.8 Let Z be a random variable with Fréchet distribution $\mathcal{F}(1)$. We have the following:

- $Z^{\frac{1}{\alpha}} \sim \mathcal{F}(\alpha)$.
- $-Z^{-\frac{1}{\alpha}} \sim \mathcal{W}(\alpha)$.
- $\log(Z) \sim \mathcal{G}(0, 1)$.

We will say that a random variable X belongs to the *domain of attraction* of the c.d.f. G (noted $F_X \in D(G)$) if one can find normalizing sequences $(a_n)_{n \geq 1}, (b_n)_{n \geq 1}$, with $a_n > 0$ for all n , such that $(M_n - b_n)/a_n$ converges weakly to the distribution with c.d.f. G . The convergence to types theorem (Resnick [1987] page 7) ensures that the domains of attractions associated to two c.d.f. do not intersect, unless those c.d.f. are equal up to an affine transformation of the argument.

Theorem 1.3.5 is essential in EVT but only gives the possible limits for a renormalized maximum of i.i.d. random variables; it does not state which limit (if any) appears for a given base distribution for the X_i . The next theorem clarifies this point and is also essential in EVT. A similar statement exists for the generalized central limit theorem (see [Durrett, 2019] as well as Embrechts et al. [2013] for a more detailed exposition). Before stating it, we need some further definitions.

Definition 1.3.9 (Regular variation). Let $h : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$. Such a function is said to be

regularly varying of index $\alpha \in \mathbb{R}$ at ∞ , noted $h \in \text{RV}_\alpha$, if

$$\lim_{t \rightarrow \infty} \frac{h(tx)}{h(t)} = x^\alpha, \quad x \in \mathbb{R}.$$

When $\alpha = 0$, we say that h is *slowly varying*.

A regularly varying function h with index α can always be written as

$$h(x) = \frac{L(x)}{x^\alpha}, \quad x > 0$$

where L is a slowly varying function.

Definition 1.3.10 (Von Mises function). A *Von Mises function* with auxiliary function f is a c.d.f. $F_\#$ with right endpoint x_1 which admits the representation

$$\bar{F}_\#(x) = ce^{-\int_{x_0}^x \frac{1}{f(u)} du}, \quad x \in (x_0, x_1)$$

for some scalar x_0 and where c is a positive constant, and f a positive, absolutely continuous function on (x_0, x_1) , whose derivative f' vanishes at x_1 .

Theorem 1.3.11 - X belongs to the domain of attraction of Φ_α if and only if its tail function $\bar{F}_X \in \text{RV}_{-\alpha}$. In this case $x_1 = +\infty$, and we can take $a_n = F^{\leftarrow}(1 - 1/n)$ and $b_n = 0$ as the normalizing sequences.

- X belongs to the domain of attraction of Λ if and only if there exist a Von Mises function $F_\#$ and a function c with finite limit at x_1 such that,

$$\bar{F}_X(x) = c(x)\bar{F}_\#(x), \quad x \in (z_0, x_1).$$

In this case, one may take $a_n = f(b_n)$ and $b_n = F^{\leftarrow}(1 - 1/n)$.

- X belongs to the domain of attraction of Ψ_α^- if and only if its right endpoint x_1 is finite and $x \mapsto \bar{F}_X(x_1 - x^{-1}) \in \text{RV}_\alpha$. In this case we can take $a_n = x_1 - F^{\leftarrow}(1 - 1/n)$ and $b_n = x_1$ as the normalizing sequences.

As one can see, the domain of attraction of the Gumbel distribution, as well as the normalizing sequences, are rather complicated to describe. On the other hand, it is usually much easier to work with the Fréchet distribution and the Weibull distribution. For instance, on many occasions, we will have

$$\bar{F}_X(x) \underset{x \rightarrow \infty}{\sim} \frac{c}{x^\alpha}$$

for some $c > 0$. This condition implies that \bar{F}_X is in $\text{RV}_{-\alpha}$ but also allows us to take $a_n = n^{1/\alpha}$, thus dispensing us of the need to know F_X and its pseudo-inverse. Many common distributions satisfy this stronger assumption, such as the Pareto distribution, the Fisher distribution, the Student distribution and so on. By contrast, much more complicated situations may arise in the case of the Gumbel distribution. For instance, both the exponential distribution and the normal distribution belong to $D(\Lambda)$. However, in the first case $a_n = 1$ and $b_n = \log n$, while in the second case we have the much more formidable constants

$$a_n = \frac{1}{\sqrt{2 \log n}} \quad \text{and} \quad b_n = \sqrt{2 \log n} - \frac{\log 4\pi + \log \log n}{2\sqrt{2 \log n}}.$$

This contrasts with the central limit theorem, where the normalizing constants depend on the base distribution only through its first two moments. However, similar problems arise too in the case of the generalized central limit theorem, where the base distribution does not have a second moment, and the limiting distribution is a non-Gaussian stable distribution. Notice that we always have an order on those constants though: $a_n < b_n$ for n large enough. To be more precise, if F belongs to the domain of attraction of the Gumbel distribution, we have (Resnick [1987]):

$$\lim_{n \rightarrow \infty} \frac{1 - F(a_n)}{1 - F(b_n)} = +\infty,$$

so in particular b_n reaches the right endpoint x_0 of the support of F faster than a_n could (assuming it does converge to it at all). Theorem 1.3.11 may be hard to apply in practice, so simpler conditions to check whether a distribution belongs to the domain of attraction of a max-stable distribution are welcome.

Proposition 1.3.12 *Let F be a distribution function with right endpoint x_1 and $\tau \in \mathbb{R}_+^*$. There exists a sequence $(u_n)_{n \geq 1}$ such that $n\bar{F}(u_n)$ converges to τ if and only if $F(x_1^-) = 1$ and*

$$\lim_{x \rightarrow x_1} \frac{\bar{F}(x)}{\bar{F}(x^-)} = 1.$$

An immediate consequence of this result is that no discrete distribution can be in the domain of attraction of the Weibull distribution, since then x_1 is finite while we cannot have $\bar{F}(x_1^-) = 1 = \bar{F}(x_1)$. It also implies that many common discrete distributions such as the Poisson distribution or the geometric distribution are not in the domain of attraction of the Gumbel distribution (see Embrechts et al. [2013] for more about this). This is a bit counter-intuitive since a geometric distribution can be seen as the discrete counterpart of the exponential distribution, which is a case in point of a distribution in the domain of attraction of the Gumbel measure.

This last result is used in conjunction with the next proposition, which plays an important part in the Poisson approximation approach to EVT and its connection with rare events theory, see for instance Feidt [2013].

Proposition 1.3.13 — Poisson approximation. *Let τ be a nonnegative scalar and $(u_n)_n$ a sequence of real numbers. We have the equivalence:*

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \tau \iff \lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_n) = e^{-\tau}. \quad (1.15)$$

That proposition is closely related to the next statement, taken from Embrechts et al. [2013]. It lies at the root of the most standard statistical techniques of extreme value theory.

Theorem 1.3.14 — Pickands-Balkema-de Haan Theorem. *Let F_X be the c.d.f. of some real random variable X with right endpoint x_1 . For any $\xi \in \mathbb{R}$, the following assertions are equivalent:*

1. F_X belongs to the domain of attraction of H_ξ .
2. There exists a positive, measurable function a such that for all x for which $1 + \xi x$ is positive,

we have

$$\lim_{u \rightarrow x_1} \frac{\overline{F}_X(u + xa(u))}{\overline{F}_X(u)} = \begin{cases} (1 + \xi x)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ e^{-x} & \text{if } \xi = 0. \end{cases} \quad (1.16)$$

3. Set $U_X(t) = F_X^{\leftarrow}(1 - t^{-1})$ for $t > 0$. For $x, y > 0$, with $y \neq 1$,

$$\lim_{u \rightarrow x_1} \frac{U_X(sx) - U_X(s)}{U_X(sy) - U_X(s)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & \text{if } \xi \neq 0 \\ \frac{\log x}{\log y} & \text{if } \xi = 0. \end{cases}$$

Remark 1.3.15. Assume $\xi = 1/\alpha > 0$. Then H_ξ is the c.d.f. of a translated Fréchet distribution $\alpha(Z - 1)$ with $Z \sim \mathcal{F}(\alpha)$.

- It is possible to show that $a(s)$ behaves like $\alpha^{-1}U_X(s)$ when s goes to ∞ . In particular, if $\overline{F}_X(s) \underset{s \rightarrow +\infty}{\sim} s^{-\alpha}$, then $a(s) \underset{s \rightarrow +\infty}{\sim} \alpha^{-1}s^{-1/\alpha}$. When the c.d.f. is differentiable, a possible choice for a is:

$$a(u) = \frac{\overline{F}_X(u)}{f_X(u)}.$$

See for instance [Papastathopoulos and Tawn \[2013\]](#) and the references therein.

- Since $a(u)$ is a positive function, we can then give a more probabilistic interpretation to equation (1.16):

$$\frac{\overline{F}_X(u + xa(u))}{\overline{F}_X(u)} = \frac{\mathbb{P}(X \geq u + xa(u))}{\mathbb{P}(X \geq u)} = \mathbb{P}(X \geq u + xa(u) \mid X \geq u),$$

meaning that the distribution of $(X - u)/a(u)$ conditional on $\{X \geq u\}$ converges weakly to a so-called generalized Pareto distribution (GPD). More precisely, given our assumption on α , it converges to a $\alpha(Y - 1)$, where $Y \sim \mathcal{VP}(\alpha, 1)$. In the same spirit, notice that the α^{-1} -regular variation hypothesis on \overline{F}_X can be interpreted as:

$$\mathcal{L}(u^{-1}X \mid X \geq u) \xrightarrow[u \rightarrow \infty]{d} \mathcal{VP}(\alpha, 1),$$

because, for $x \geq 1$, we have:

$$\mathbb{P}(u^{-1}X \geq x \mid X \geq u) = \frac{\mathbb{P}(X \geq ux)}{\mathbb{P}(X \geq u)} = \frac{\overline{F}_X(ux)}{\overline{F}_X(u)} \xrightarrow[u \rightarrow \infty]{} \frac{1}{x^\alpha} = \mathbb{P}(Y \geq x),$$

with $Y \sim \mathcal{VP}(\alpha, 1)$. ■

1.3.2 The multivariate case

Like Gaussian distributions and stable distributions, the notion of max-stable distribution can be extended into higher dimensions. Most, if not all results from this subsection, come from [Resnick \[1987\]](#) and [Resnick \[2006\]](#).

The maximum operator on \mathbb{R}^d is defined:

$$\mathbf{x} \oplus \mathbf{y} := (x^1 \oplus y^1, \dots, x^d \oplus y^d)$$

where x^j, y^j denote the coordinates of the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Although the following binary relation is not a true order, unlike its one-dimensional counterpart, it will be still useful in the sequel:

$$\mathbf{x} < \mathbf{y} \iff x^j < y^j, \text{ for all } j \in \llbracket 1, d \rrbracket. \quad (1.17)$$

By contrast, if at least one coordinate of \mathbf{x} is (strictly) greater than the corresponding coordinate of \mathbf{y} , we will say that $\mathbf{x} \not\leq \mathbf{y}$:

$$\mathbf{x} \not\leq \mathbf{y} \iff x^j > y^j, \text{ for some } j \in \llbracket 1, d \rrbracket.$$

We also denote hyperrectangles with the usual notations for segments:

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

Notice that $(-\infty, \mathbf{a}]^c = \{\mathbf{x} \in \mathbb{R}^d, \mathbf{x} \not\leq \mathbf{a}\}$, where $-\infty$ denotes the vector of \mathbb{R}^d whose coordinates are all equal to $-\infty$.

The c.d.f. $F_{\mathbf{X}}$ of a random vector \mathbf{X} in \mathbf{x} is defined as the probability that \mathbf{X} belongs to $(-\infty, \mathbf{x}]$:

$$F_{\mathbf{X}}(\mathbf{x}) := \mathbb{P}(\mathbf{X} \leq \mathbf{x}), \mathbf{x} \in \mathbb{R}^d.$$

Next we define what being max-stable in higher dimensions means.

Definition 1.3.16 (Multi-dimensional max-stability). Let μ be a probability distribution on \mathbb{R}^d . We say that μ is *max-stable* if for any $n \geq 1$ and any $\mathbf{X}_1, \dots, \mathbf{X}_n$ are n i.i.d. random vectors with common distribution μ , there exists $\mathbf{a}_n, \mathbf{b}_n$ two deterministic vectors in $(\mathbb{R}_+^*)^d \times \mathbb{R}^d$ such that

$$\bigoplus_{i=1}^n \mathbf{X}_i = \left(\bigoplus_{i=1}^n X_i^1, \dots, \bigoplus_{i=1}^n X_i^d \right) \stackrel{d}{=} \mathbf{a}_n \mathbf{X}_1 + \mathbf{b}_n.$$

where the multiplication between vectors is to be taken in a coordinate-wise manner. If $\mathbf{b}_n = \mathbf{0}$, then we say that μ is *strictly max-stable*. A random vector \mathbf{X} is said to be *min-stable* if $-\mathbf{X}$ is max-stable.

A basic property of max-stable random vectors is that, under an assumption on the normalizing vectors, they behave well under max-linear combinations, *i.e.* linear combinations where the sum operator $+$ has been replaced by the max operator \oplus .

Proposition 1.3.17 If $\mathbf{X} = (X^1, \dots, X^d)$ is a max-stable random vector such that there exist scalars a_n and b_n for which the normalizing vectors \mathbf{a}_n and \mathbf{b}_n satisfy

$$\mathbf{a}_n = a_n \mathbf{1} \text{ and } \mathbf{b}_n = b_n \mathbf{1}$$

for all $n \geq 1$, then any max-linear combination of its coordinates is a max-stable random variable.

Proof. Let \mathbf{v} be a vector in $(\mathbb{R}_+)^d$ and set

$$Z := \langle \mathbf{v}, \mathbf{X} \rangle_{\oplus} := \bigoplus_{j=1}^d v^j X^j.$$

Let Z_1, \dots, Z_n be n i.i.d. copies of Z . Then we have:

$$\begin{aligned} \mathbb{P}\left(\bigoplus_{i=1}^n Z_i \leq x\right) &= \mathbb{P}\left(\bigoplus_{i=1}^n \bigoplus_{j=1}^d v^j X_i^j \leq x\right) \\ &= \mathbb{P}\left(\bigoplus_{j=1}^d \bigoplus_{i=1}^n v^j X_i^j \leq x\right) \\ &= \mathbb{P}\left(\bigcap_{j=1}^d \left\{ \bigoplus_{i=1}^n X_i^j \leq \frac{x}{v_j} \right\}\right) \\ &= \mathbb{P}\left(\bigoplus_{j=1}^d v^j (a_n X_1^j + b_n) \leq x\right) \\ &= \mathbb{P}(a_n Z_1 + b_n \leq x). \end{aligned}$$

□

Before exposing the other principal properties of max-stable vectors, we give a few examples of such random vectors.

Example 1. - A random vector (X^1, \dots, X^d) with i.i.d. coordinates each having the Fréchet distribution $\mathcal{F}(1)$ is a max-stable random vector:

$$\mathbb{P}\left(\bigoplus_{i=1}^n \mathbf{X}_i \leq \mathbf{x}\right) = (\mathbb{P}(\mathbf{X}_1 \leq \mathbf{x}))^n = e^{-n \sum_{j=1}^d \frac{1}{x^j}}, \quad \mathbf{x} \in (\mathbb{R}_+)^d,$$

which is exactly the c.d.f. of $\mathbf{a}_n \mathbf{X}_1$, where $\mathbf{a}_n = n\mathbf{1} = n(1, \dots, 1)$ (and thus $\mathbf{b}_n = \mathbf{0}$).

- Let $X \sim \mathcal{F}(1)$. Then the random vector $\mathbf{X} = (X, \dots, X) = X\mathbf{1}$ is a max-stable random vector in \mathbb{R}^d . Here we have:

$$F_{\mathbf{X}}(\mathbf{x}) = e^{-\bigoplus_{j=1}^d \frac{1}{x^j}}, \quad \mathbf{x} \in (\mathbb{R}_+)^d.$$

The normalization vectors \mathbf{a}_n and \mathbf{b}_n are the same as in the previous examples, even though the dependency relations between the coordinates of each random vector are diametrically opposed (complete independence versus complete dependence).

- The random vector (X, Y) , where $X \sim \mathcal{F}(1)$, $Y \sim \mathcal{G}(0, 1)$ and X, Y are independent, is also max-stable. This time we have $\mathbf{a}_n = (n, 1)$ and $\mathbf{b}_n = (0, \log(n))$.

- The Weibull distribution $\mathcal{W}(1)$ of shape parameter 1 corresponds to the negative exponential distribution. Up to a change of sign, we recover the fact that the exponential distribution is min-stable. Set:

$$1 - F(x, y) := e^{-x} + e^{-y} - \frac{1}{e^x + e^y - 1}, \quad x, y > 0.$$

This defines a distribution function on \mathbb{R}^d with exponential margins known as the Marshall and Olkin bivariate exponential distribution (see [Marshall and Olkin \[1983\]](#)) Let (X, Y) be a random vector with this distribution function. One can easily prove that:

$$\mathbb{P}(X \geq x, Y \geq y) = \frac{1}{e^x + e^y - 1}, \quad x, y > 0.$$

Therefore, although the margins are min-stable, a random vector with this distribution is not min-stable, since we cannot find vectors $\mathbf{a}_n > \mathbf{0}$ and \mathbf{b}_n such that

$$(\mathbb{P}(X \geq x, Y \geq y))^n = \mathbb{P}(a_n^1 X + b_n^1 \geq x, a_n^2 Y + b_n^2 \geq y), \quad x, y > 0.$$

In all those examples, the marginal distributions had one of the one-dimensional max-stable distributions. This is a general fact: the marginal distributions of a max-stable random vector can only be Fréchet, Gumbel or Weibull. It is easy to prove by using the same arguments as in the proof of proposition 1.3.17 and taking $\mathbf{v} = \mathbf{e}_i$ (in this case the assumption on the normalizing vectors \mathbf{a}_n and \mathbf{b}_n is not necessary), thus showing that each coordinate X^j of a max-stable random vector \mathbf{X} is max-stable too. We then conclude by theorem 1.3.5. Alternatively, one can also write directly:

$$\begin{aligned} F_{\oplus_{i=1}^n X_i^j}(x) &= F_{\oplus_{i=1}^n \mathbf{X}_i}(+\infty, \dots, +\infty, x, +\infty, \dots, +\infty) \\ &= F_{\mathbf{a}_n \mathbf{X} + \mathbf{b}_n}(+\infty, \dots, +\infty, x, +\infty, \dots, +\infty) \\ &= F_{a_n^j X^j + b_n^j}(x). \end{aligned}$$

In the sequel, we will follow the custom and mainly focus on the *simple max-stable random vectors*, that is, max-stable random vectors whose marginals are Fréchet $\mathcal{F}(\alpha)$ with parameter $\alpha > 0$. Let us mention that the usual definition of simple max-stable random vectors [de Haan and Ferreira \[2007\]](#) (definition 6.1.13) assumes that $\alpha = 1$, but this will not fit into our framework, so we will drop this assumption.

Max-stable random vectors can have complicated dependency structures (and as such are not easy to simulate), but it is nonetheless possible to give a simple description of them thanks to the next result (see [Resnick \[1987\]](#)).

Theorem 1.3.18 — Spectral decomposition of max-stable random vectors. *Let \mathbf{Z} be a random vector in \mathbb{R}^d and define $E = [\mathbf{0}, +\infty]^d \setminus \{\mathbf{0}\}$. The following are equivalent:*

1. \mathbf{Z} is a simple max-stable random vector with $\mathcal{F}(\alpha)$ marginal distributions for some positive α .
2. There exists a finite measure ν on \mathbb{S}_+^{d-1} , called the *angular measure*, satisfying the moment constraints:

$$\int_{\mathbb{S}_+^{d-1}} u^j d\nu(\mathbf{u}) = 1, \quad j \in \llbracket 1, d \rrbracket \tag{1.18}$$

and a Poisson point process $\eta = \sum_{k=1}^{\infty} \delta_{(\rho_k, \mathbf{U}_k)}$ over $\mathbb{R}_+^* \times \mathbb{S}_+^{d-1}$ with intensity measure $\frac{\alpha}{r^{\alpha+1}} dr \otimes d\nu(\mathbf{u})$ such that:

$$\mathbf{Z} \stackrel{d}{=} \bigoplus_{k=1}^{\infty} \rho_k \mathbf{U}_k^{\frac{1}{\alpha}}. \tag{1.19}$$

3. There exists a finite measure ν on \mathbb{S}_+^{d-1} satisfying (1.18) such that for all $\mathbf{x} \in \mathbb{R}^d$

$$F_Z(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \notin E \\ \exp\left(-\int_{\mathbb{S}_+^{d-1}} \bigoplus_{j=1}^d \left(\frac{u^j}{x^j}\right)^\alpha d\nu(\mathbf{u})\right) & \text{if } \mathbf{x} \in E. \end{cases} \quad (1.20)$$

4. There exists a Radon measure μ , called the exponent measure, satisfying the two following conditions:

$$\mu(tA) = t^{-\alpha}\mu(A), \quad t \in \mathbb{R}_+^*, \quad A \in \mathcal{B}(E), \quad (1.21)$$

$$\mu\{\mathbf{x} \in E \mid x^j \geq 1\} = 1, \quad j \in \llbracket 1, d \rrbracket. \quad (1.22)$$

and such that for all $\mathbf{x} \in \mathbb{R}^d$

$$F_Z(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \notin E \\ \exp(-\mu[\mathbf{0}, \mathbf{x}]^c) & \text{if } \mathbf{x} \in E, \end{cases}$$

where $[\mathbf{0}, \mathbf{x}] := \prod_{j=1}^d [0, x^j]$ if $\mathbf{x} = (x^1, \dots, x^j)$.

This result will be central in what follows. In dimension 1, it simply states that a Fréchet $\mathcal{F}(1)$ random variable is the maximum of $(\Gamma_n^{-1})_{n \geq 1}$, where $(\Gamma_n)_{n \geq 1}$ is an homogeneous Poisson point process on \mathbb{R}_+ . Since the smallest point of the latter is distributed according to an exponential distribution $\mathcal{E}(1)$, this amounts to saying that if $Y \sim \mathcal{E}(1)$, then $Y^{-1} \sim \mathcal{F}(1)$, a standard result. Geometrically a typical realization of $(\Gamma_n^{-1})_{n \geq 1}$ looks like this



Figure 1.2: The 100 first points of $(\Gamma_n^{-1})_{n \geq 1}$

Notice the accumulation point around 0. In particular, this Poisson process has infinitely many points a.s. and its intensity measure $r^{-2} dr$ is σ -finite. This will stay true in higher dimension.

We will give later a more formal definition of the notion of exponent measure (see the subsection about max-infinitely divisible distributions). For now, observe that the exponent measure μ and the angular measure ν are related in the following way: first define a measure on \mathbb{R}_+^* by setting $\rho_\alpha[x, +\infty) := x^{-\alpha}$ for every positive x . Define also the product transformation by

$$\begin{aligned} \mathfrak{d} &: \mathbb{R}_+^* \times \mathbb{S}_+^{d-1} &\rightarrow E_{\mathbf{0}} \\ &(r, \mathbf{u}) &\mapsto r\mathbf{u} \end{aligned}$$

with $E_{\mathbf{0}} := (\overline{\mathbb{R}_+})^d \setminus \{\mathbf{0}\}$. Then we have:

$$\mu = \mathfrak{d}^*(\rho_\alpha \otimes \nu), \quad (1.23)$$

i.e. μ is the image-measure of $\rho_\alpha \otimes \nu$ by the product transformation d . This identity will be referred as the *polar decomposition* of μ (w.r.t. to the reference norm $\|\cdot\|$).

The equality in distribution (1.19) is also known as the representation in Lévy-LePage series in the more general context of Davydov et al. [2008] (dubbed here "de Haan-LePage series"). We will write in the more compact form:

$$\mathbf{Z} \stackrel{d}{=} \mathbf{m}(\eta), \quad (1.24)$$

where \mathbf{m} is defined as the coordinate-wise maximum of the elements $\rho_k \mathbf{U}_k^{1/\alpha}$. Now let us make a few additional remarks about this theorem: first we observe that the distribution of a simple max-stable random vector is completely determined by the knowledge of the angular measure. By putting such a vector at power $1/\alpha$ in a coordinate-wise manner (or applying log to it when $\alpha = 0$), translating it by $\mu \mathbf{1}$ and multiplying it by a scale parameter σ , we get three additional parameters to describe the family of max-stable random vectors whose coordinates share the same α , μ and σ . Thus we will denote any distribution belonging to this family by

$$\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu; \mu, \sigma).$$

When $\mu = 0$ and $\sigma = 1$, we will write $\mathcal{MS}(\alpha, \nu)$ for short. More generally, when the marginals of a max-stable random vectors have different max-stable distributions, and thus correspond to different α^j , μ^j and σ^j , we will denote this by $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu; \boldsymbol{\mu}, \boldsymbol{\sigma})$, where each vector contains the parameters of each marginal. Second, even though the moment constraints may seem mysterious, they are actually an easy consequence of the assumption that the marginal distributions of \mathbf{Z} are all unit Fréchet: by making x^1, \dots, x^{d-1} goes to infinity, and since Z^d has unit Fréchet distribution, we have

$$\exp\left(-\int_{\mathbb{S}_+^{d-1}} \frac{u^d}{x^d} d\nu(\mathbf{u})\right) = e^{-\frac{1}{x^d}}, \quad x^d \in \mathbb{R}_+^*$$

which yields the moment assumption for $j = d$. Applying the same argument to the $d - 1$ other variables, we recover (1.18). Furthermore the exponent measures μ in the second, third and fourth points are the same. Indeed, if \mathbf{Z} is defined by (1.19), then we have:

$$\begin{aligned} \mathbb{P}(\mathbf{Z} \leq \mathbf{x}) &= \mathbb{P}(\forall(\rho, \mathbf{U}) \in \eta, \rho \mathbf{U} \notin [\mathbf{0}, \mathbf{x}]^c) \\ &= \mathbb{P}(\eta[\mathbf{0}, \mathbf{x}]^c = 0) \\ &= e^{-\mu[\mathbf{0}, \mathbf{x}]^c}. \end{aligned}$$

It remains to prove that this is equal to (1.20). The quantity in exponent can be expressed as:

$$\begin{aligned} \mu[\mathbf{0}, \mathbf{x}]^c &= \int_{\mathbb{S}_+^{d-1}} \int_0^\infty \mathbf{1}_{[\mathbf{0}, \mathbf{x}]^c}(r\mathbf{u}) \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \int_{\mathbb{S}_+^{d-1}} \int_{\{r \in \mathbb{R}_+, r\mathbf{u} \in [\mathbf{0}, \mathbf{x}]^c\}} \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \int_{\mathbb{S}_+^{d-1}} \int_{\{r \in \mathbb{R}_+, r \geq \min \frac{x}{u}\}} \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \int_{\mathbb{S}_+^{d-1}} \bigoplus_{j=1}^d \frac{u^j}{x^j} d\nu(\mathbf{u}). \end{aligned}$$

Here \min denoted the usual minimum mapping (thus returning a scalar, not a vector). Now, to fix ideas, we give some examples of simple max-stable random vectors with explicit angular measures.

Example 2. - Take $\|\cdot\|_p$ as the reference norm for some $p \in [1, \infty]$. If $\nu = \sum_{j=1}^d \delta_{e_j}$ is the discrete measure concentrated on the canonical basis vectors, then it satisfies the moment constraints (1.18) and the associated exponent measure writes as:

$$\mu[\mathbf{0}, \mathbf{x}]^c = \int_{\mathbb{S}_+^{d-1}} \bigoplus_{j=1}^d \frac{w^j}{x^j} \, d\nu(\mathbf{u}) = \sum_{j=1}^d \frac{1}{x^j}, \quad \mathbf{x} \in E.$$

In other words, the max-stable distribution associated to this angular measure is the distribution of d i.i.d. random variables with unit Fréchet margins. Notice that in this case the angular measure is discrete and that the choice of the reference norm is immaterial here as long as $\|e_j\| = 1$. We illustrate the corresponding Poisson process, in the case of the Euclidean norm:

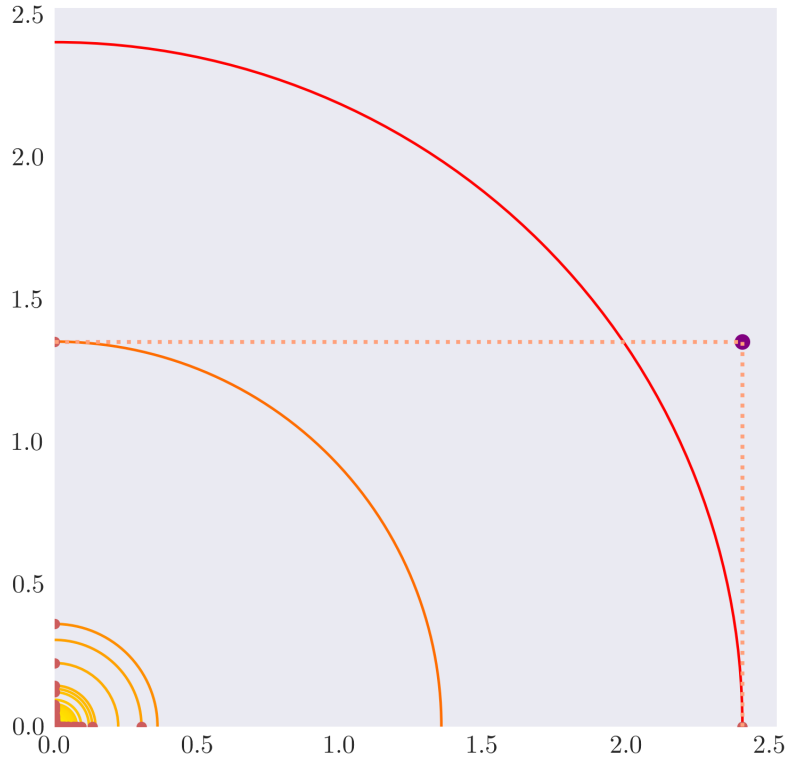


Figure 1.3: The 100 first points of a Poisson process with $\nu = \frac{1}{2}(\delta_{e_1} + \delta_{e_2})$

Each red circle has random radius $R_n = \Gamma_n^{-1}$, where $(\Gamma_n)_{n \geq 1}$ is an homogeneous Poisson point process on \mathbb{R}_+ . The red dots corresponds to $\Gamma_n \mathbf{U}_n$, where $\mathbf{U}_n = (B_n, 1 - B_n)$, and the B_n are i.i.d. Bernoulli $\mathcal{B}(1/2)$ random variables. Notice that the coordinate-wise maximum (the purple dot) never belongs to the Poisson process: at least two points of the process are needed for the maximum to be reached. It is distributed as a Fréchet random variable with shape parameter $\alpha = 1$.

- Now take $\|\cdot\|_2$ as the reference norm on \mathbb{R}^d and $\nu = \sqrt{d}\delta_{\frac{1}{\sqrt{d}}\mathbf{1}}$. Once more this measure satisfies the moment constraints (1.18) and we have that:

$$\mu[\mathbf{0}, \mathbf{x}]^c = \int_{\mathbb{S}_+^{d-1}} \bigoplus_{j=1}^d \frac{w^j}{x^j} d\nu(\mathbf{u}) = \bigoplus_{j=1}^d \frac{1}{x^j}, \quad \mathbf{x} \in E. \quad (1.25)$$

This is exactly the case of complete dependence between coordinates, i.e. the distribution of (X, \dots, X) where $X \sim \mathcal{F}(1)$. Visually, we get the following (the angular measure has not been renormalized by $\sqrt{2}$ here):

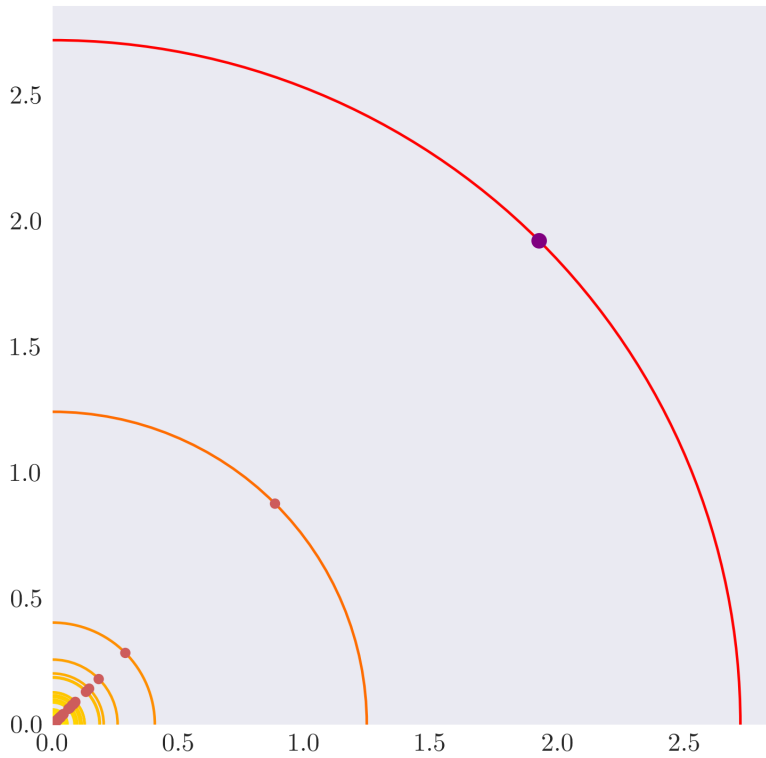


Figure 1.4: The 100 first points of a Poisson process with angular measure $\nu = \delta_{\frac{1}{\sqrt{2}}\mathbf{1}}$

This time $\mathbf{U}_n = 2^{-1/2}\mathbf{1}$, so the \mathbf{U}_n are deterministic. Besides, the coordinate-wise maximum of this Poisson point process always belongs to it.

- This time we work with a diffuse angular measure. We choose $\|\cdot\|_2$ on \mathbb{R}^2 as the reference measure and twice the uniform measure on the compact \mathbb{S}_+^1 for ν . This measure indeed satisfies (1.18):

$$\int_{\mathbb{S}_+^1} u^1 d\nu(\mathbf{u}) = 2 \int_0^1 r \int_0^{\pi/2} \cos(\theta) d\theta dr = 1 = 2 \int_0^1 r \int_0^{\pi/2} \sin(\theta) d\theta dr = \int_{\mathbb{S}_+^1} u^2 d\nu(\mathbf{u}).$$

As before, a simple computation gives the exponent measure. Let $\mathbf{x} = (x, y) \in E, .$ We have:

$$\begin{aligned} \mu[\mathbf{0}, \mathbf{x}]^c &= 2 \int_0^1 r \int_0^{\frac{\pi}{2}} \frac{\cos(\theta)}{x} \oplus \frac{\sin(\theta)}{y} d\theta dr \\ &= \frac{1}{x} \int_0^{\arctan(\frac{y}{x})} \cos(\theta) d\theta + \frac{1}{y} \int_{\arctan(\frac{y}{x})}^{\frac{\pi}{2}} \sin(\theta) d\theta \\ &= \frac{1}{x} \sin\left(\arctan\left(\frac{y}{x}\right)\right) + \frac{1}{y} \cos\left(\arctan\left(\frac{y}{x}\right)\right) \\ &= \frac{1}{\sqrt{x^2 + y^2}} \left(\frac{y}{x} + \frac{x}{y}\right), \end{aligned}$$

where we made use of the identities

$$\sin(\arctan(x)) = \frac{x}{\sqrt{1+x^2}} \text{ and } \cos(\arctan(x)) = \frac{1}{\sqrt{1+x^2}}, \quad x \in \mathbb{R}.$$

We illustrate the associated Poisson process, where the angular measure is renormalized to be a probability measure (the uniform distribution on \mathbb{S}_+^1):

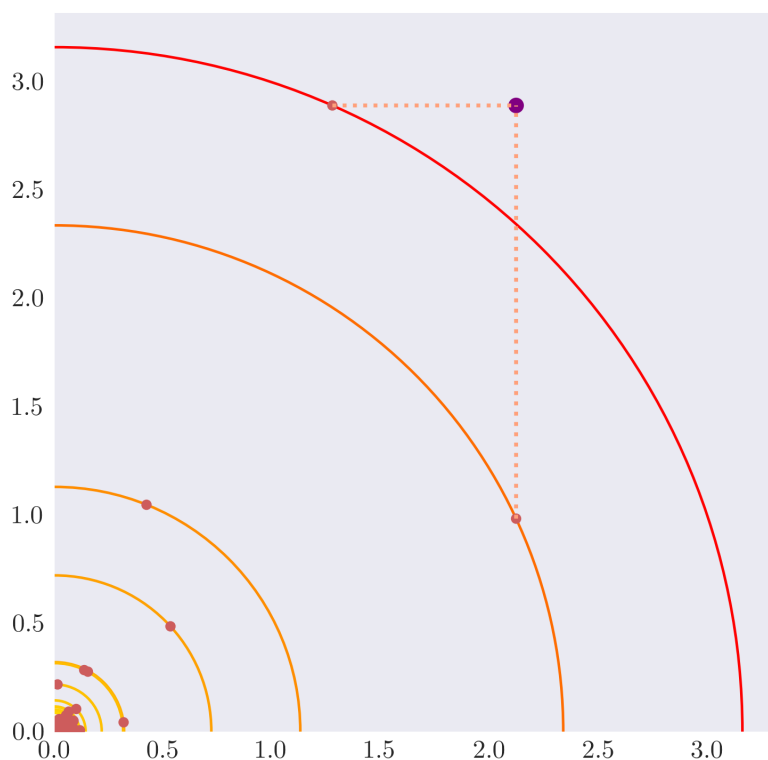


Figure 1.5: The 100 first points of a Poisson process with uniform angular measure ν for $\|\cdot\| = \|\cdot\|_2$

In this example, $\mathbf{U}_n = (\cos(\theta_n), \sin(\theta_n))$, where θ_n is random variable with uniform distribution on $[0, \pi/2]$. As such, \mathbf{U}_n is a uniform random variable on \mathbb{S}_+^1 . Depending on the realization of the Poisson point process, its component-wise maximum may or not belong to it.

- In the previous example, we chose the uniform measure (up to a constant) on the usual circle. Changing the reference norm (and thus the associated sphere) changes the dependency structure between the coordinates. For instance, take the norm $\|\cdot\|_\infty$ on \mathbb{R}^2 as the reference norm, and $\mathbb{S}_+^1 = ([0, 1] \times \{1\}) \uplus (\{1\} \times [0, 1])$ the associated unit circle. The uniform measure on this set becomes $\nu = 3(\lambda_1 \otimes \delta_1 + \delta_1 \otimes \lambda_1)/2$, where λ_1 is the Lebesgue measure on \mathbb{R} . But then the exponent measure writes as:

$$\mu[\mathbf{0}, \mathbf{x}]^c = \begin{cases} \frac{1}{x} + \frac{x}{3y^2} & \text{if } x \leq y \\ \frac{1}{y} + \frac{y}{3x^2} & \text{if } x > y, \end{cases} \quad \mathbf{x} = (x, y) \in E$$

which is very different from the previous example.

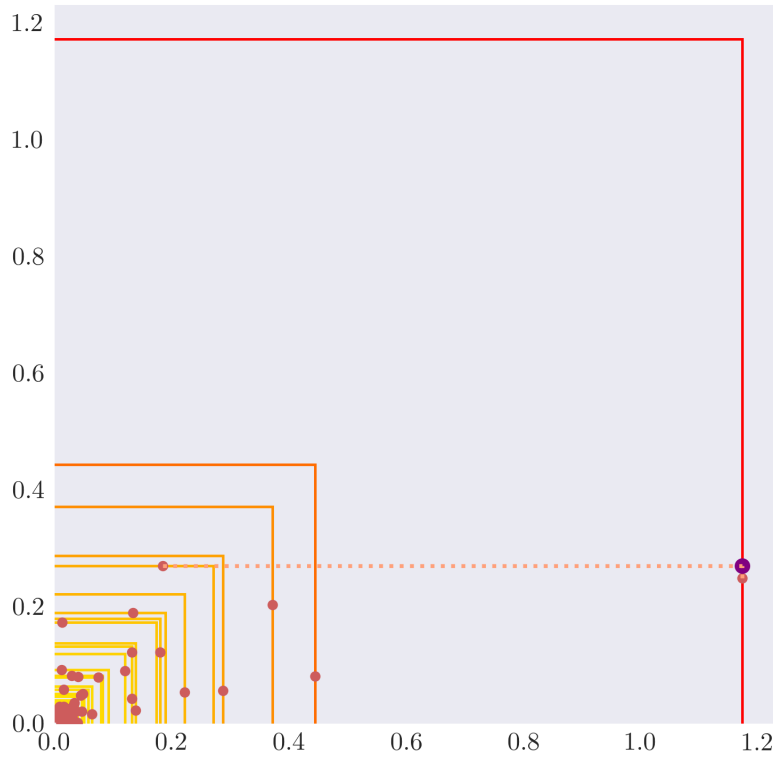


Figure 1.6: The 100 first points of a Poisson process with uniform angular measure ν for $\|\cdot\| = \|\cdot\|_\infty$

The random vectors \mathbf{U}_n are taken equal to $(B_n + (1 - B_n)X_n, (1 - B_n) + B_n X_n)$, where $B_n \sim \mathcal{B}(1/2)$ and $X_n \sim \mathcal{U}_{[0,1]}$. This indeed corresponds to the uniform distribution on the positive orthant \mathbb{S}_+^1 of the unit circle for the infinity norm $\|\cdot\|_\infty$.

The question of the existence of densities for max-stable random vectors is important in statistical applications. It will play a (minor) role later in this thesis to make use of a certain result from the theory of probability metrics. A paper by [Dombry et al. \[2016\]](#) gives necessary and sufficient conditions for a simple max-stable random vector (with $\alpha = 1$) to admit a density with respect to the Lebesgue measure λ_d on \mathbb{R}^d .

1.3.3 Max-infinitely divisible distributions

Max-infinitely divisible distributions (max-id distributions) are the counterparts of infinitely divisible distributions when the sum operator $+$ is replaced by the max operator \oplus (see for instance [Durrett \[2019\]](#) for more on infinitely divisible distributions). As such, they constitute a natural generalization of max-stable distributions. Just as infinitely divisible distributions are the only possible distributions of a Lévy process at time $t = 1$, max-id distributions are the only possible distributions of an extremal process at time $t = 1$. We do not delve into this theory and refer the interested reader to [Resnick \[1987\]](#). Let us finally mention that max-id distributions are the only possible limiting distributions for a maximum of a triangular array of n i.i.d. random vectors $(\mathbf{X}_{n,k})_{\substack{n \geq 1 \\ k \leq n}}$ when n goes to infinity.

Definition 1.3.19 (Max-id distributions). Let μ be a probability distribution on \mathbb{R}^d . We say that μ is *max-infinitely divisible distribution* if for any random vector \mathbf{X} with distribution μ , and any positive integer n , there exist n i.i.d. random vectors $\mathbf{X}_{n,1}, \dots, \mathbf{X}_{n,n}$ such that:

$$\mathbf{X} \stackrel{d}{=} \bigoplus_{i=1}^n \mathbf{X}_{n,i}.$$

A random vector \mathbf{X} is said to be *min-infinitely divisible* (min-id) if $-\mathbf{X}$ is max-id.

In particular, $F_{\mathbf{X}} = F_{\mathbf{X}_{1,n}}^n$, and thus it is equivalent to ask that $F_{\mathbf{X}}^{1/n}$ is a distribution function for all positive integers. Actually, an equivalent statement of the max-id property is that a random vector \mathbf{X} is max-id if and only if its distribution function $F_{\mathbf{X}}$ at any power $t > 0$ is still a distribution function. Before moving to the main result, we give a few examples of max-d distributions, as well as a counter-example.

- Every max-stable random vector \mathbf{X} with normalizing vectors \mathbf{a}_n and \mathbf{b}_n is max-id, with $\mathbf{X}_{n,i} := (\mathbf{X}_i - \mathbf{b}_n)/\mathbf{a}_n$, where $\mathbf{X}_1, \dots, \mathbf{X}_n$ are n i.i.d. copies of \mathbf{X} .
- Another important family of max-id distributions is given by the univariate probability distributions, since any distribution function on \mathbb{R} at any positive power remains a distribution function.
- Therefore, examples of distributions which are *not* max-id can only be found in higher dimensions: for instance, one can show that the Gaussian random vector $(X, X + \rho Y)$, where X and Y are i.i.d. with standard normal distribution $\mathcal{N}(0, 1)$ and $\rho \in \mathbb{R}$, is max-id if and only if ρ is nonnegative. Actually, a random vector \mathbf{X} can be max-id only if it is *associated*:

$$\text{Cov}(f_1(\mathbf{X}), f_2(\mathbf{X})) \geq 0,$$

for all f_1, f_2 which are real-valued, non-decreasing functions in each coordinate of their arguments on \mathbb{R}^d . Taking for f_1, f_2 the projections with respect to every pair of coordinate number

$j_1, j_2 \in \llbracket 1, d \rrbracket$, we see that a Gaussian vector can be max-id only if all the coefficients of its covariance matrix are nonnegative.

A generalization of theorem 1.3.21 to max-id distributions exists and will prove as useful as its max-stable counterpart, although less information can be available on the exponent measures at such a level of generality. We first define properly the notion of exponent measure:

Definition 1.3.20 (Exponent measure). We say that a Radon measure on \mathbb{R}^d is an *exponent measure* if there exists $\ell \in [-\infty, +\infty)^d$ such that by setting $E_\ell := [\ell, +\infty] \setminus \{\ell\}$, we have

$$\lim_{\mathbf{x} \rightarrow +\infty} \mu[-\infty, \mathbf{x}]^c = \mu\left(\bigcup_{j=1}^d \{\mathbf{x} \in E, x^j = +\infty\}\right) = 0$$

Either $(\ell > -\infty)$ or $(\mathbf{x} \geq \ell \text{ and } \exists j \in \llbracket 1, d \rrbracket, x^j = -\infty \implies \mu[-\infty, \mathbf{x}]^c = 0)$.

We are now in position to give the main result of this section.

Theorem 1.3.21 — Spectral decomposition of max-id random vectors. *Let \mathbf{Z} be a random vector in \mathbb{R}^d with distribution function $F_{\mathbf{Z}}$. The following are equivalent:*

1. \mathbf{Z} is a max-id random vector.
2. For some $\ell \in [-\infty, +\infty)^d$, there exists an exponent measure μ on $E_\ell = [\ell, +\infty] \setminus \{\ell\}$ such that

$$F_{\mathbf{Z}}(\mathbf{x}) = \begin{cases} e^{-\mu[-\infty, \mathbf{x}]^c} & \text{if } \mathbf{x} \geq \ell \\ 0 & \text{otherwise.} \end{cases}$$

3. For some $\ell \in [-\infty, +\infty)^d$, there exists an exponent measure μ on $E_\ell = [\ell, +\infty] \setminus \{\ell\}$ and a Poisson process $\eta = \sum_k \delta_{\mathbf{Z}_k}$ on E_ℓ with intensity measure μ such that

$$\mathbf{Z} \stackrel{d}{=} \bigoplus_{k=1}^{\infty} \mathbf{Z}_k. \tag{1.26}$$

If any of those conditions is satisfied, the exponent measure μ is characterized by:

$$\mu[-\infty, \mathbf{x}]^c = -\log F_{\mathbf{Z}}(\mathbf{x}), \quad \mathbf{x} > \ell. \tag{1.27}$$

1.3.4 Construction and properties of the extremal integral

The classic notion of stochastic integral has been extended in several ways. One of them, called *α -stable integral* is particularly suited to describe certain stable stochastic processes and is studied in e.g. Samorodnitsky and Taqqu [1994]. In the context of max-stable distributions, the corresponding notion is named *extremal integral*. Examples of papers relying on this notion include Vervaat [1988] and Resnick and Roy [1991]. Our presentation follows Stoev and Taqqu [2005].

We will denote the scale parameter of a Fréchet random variable Z by $\|Z\|_\alpha := \sigma$. The reason behind this notation will become clear in the sequel. Next we introduce the notions of α -Fréchet processes and max-linear combinations. Remember that Gaussian processes are defined as the

stochastic processes such that any linear combination of their marginals is still Gaussian. Here α -Fréchet processes are defined in a similar fashion.

Definition 1.3.22. Let \mathcal{X} be a subset of \mathbb{R}^d , for some $d \geq 1$. A real-valued stochastic process $Z = (Z(\mathbf{x}))_{\mathbf{x} \in \mathcal{X}}$ is α -Fréchet if any max-linear combination of its marginals has the Fréchet distribution $\mathcal{F}(\alpha, \sigma)$, for some σ and all $n \geq 1$:

$$\bigoplus_{j=1}^n a_j Z(\mathbf{x}_j) \stackrel{d}{=} \mathcal{F}(\alpha, \sigma), \quad (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathcal{X}^n, \quad (a_1, \dots, a_n) \in \mathbb{R}_+^n.$$

A result by de Haan states that for a process with α -Fréchet margins, this definition is equivalent to being max-stable. The next definition is the analogue of Gaussian spaces for α -Fréchet processes.

Definition 1.3.23. We say that a set \mathcal{M} of random variables is an α -Fréchet space if it closed under max-linear combinations and its elements are **jointly** α Fréchet random variables, that is:

1. (Max-linear space) For all $n \geq 1$, we have for $(Z_1, \dots, Z_n) \in \mathcal{M}$:

$$\bigoplus_{i=1}^n a_i Z_i \in \mathcal{M}, \quad (a_1, \dots, a_n) \in \mathbb{R}_+^n.$$

2. For all $Z \in \mathcal{M}$, Z is α -Fréchet. Furthermore \mathcal{M} is said to be *closed* if it is closed under taking limits in probability.

In the sequel we give a short account of the construction of the stochastic extremal integral as it is done in [Stoev and Taqqu \[2005\]](#). The reasoning is always the same: first construct the integral for simple functions, find some fundamental properties of this integral and then construct the integral for general integrands by approximating them with increasing sequences of simple functions. We will not detail the first and second parts, but will nonetheless explain the framework used to justify the approximation procedure since it will be useful for us later. Lastly we will enumerate some of the most important properties of the stochastic extremal integral and detail its links with max-stable processes.

Definition 1.3.24. Let (E, \mathcal{E}, μ) be a measure space and $\mathcal{E}_0 := \{A \in \mathcal{E}, \mu(A) < \infty\}$ and $\mathbf{L}^0(\Omega)$ the set of real random variables on Ω . Let $\alpha > 0$. We say that a function $M_\alpha : \mathcal{E}_0 \rightarrow \mathbf{L}^0(\Omega)$ is a *random sup-measure with control measure μ* if it satisfies the three following conditions:

1. (independently scattered) For any collection of disjoint sets $(A_j)_{1 \leq j \leq n}$ in \mathcal{E}_0 , the random variables $(M_\alpha(A_j))_{1 \leq j \leq n}$ are independent.
2. (α -Fréchet) For any $A \in \mathcal{E}_0$, $M_\alpha \sim \mathcal{F}(\alpha, \mu(A)^{1/\alpha})$, *i.e.*

$$\mathbb{P}(M_\alpha(A) \leq x) = \exp\left(-\frac{\mu(A)}{x^\alpha}\right) \mathbf{1}_{\{x \geq 0\}}, \quad x \in \mathbb{R}.$$

3. (σ -sup-additive) For any collection of disjoint sets $(A_i)_{i \geq 1}$ in \mathcal{E}_0 such that $\bigcup_i A_i \in \mathcal{E}_0$, we

have:

$$M_\alpha\left(\biguplus_{i=1}^{\infty} A_i\right) = \bigoplus_{i=1}^{\infty} M_\alpha(A_i) \text{ a.s..}$$

The existence of random sup-measures is not obvious but is guaranteed with no further assumption on μ . We will not dwell on the proof of this result and take it for granted (see the aforementioned paper by Stoev and Taqqu and the references therein). Now using this definition, we introduce the extremal integral for simple functions.

Definition 1.3.25. Let f be a simple function on E :

$$f(x) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(x), \quad x \in E$$

where a_i are non-negative numbers and the A_i are disjoint. The *extremal integral* of f with respect to the random sup-measure M_α is defined as:

$$\int_E^e f(x) \, dM_\alpha(x) := \bigoplus_{i=1}^n a_i M_\alpha(A_i).$$

To extend this definition to more general integrands, the next proposition is required. It is of interest on its own, and so is the lemma used to prove it.

Proposition 1.3.26 *Let \mathcal{M} be an α -Fréchet space and let $Z, Z_1, Z_2, \dots \in \mathcal{M}$. We have the equivalences:*

1. $Z_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} Z$,
2. $\rho_{\mathcal{M}, \alpha}(Z_n, Z) := 2\|Z_n \oplus Z\|_\alpha^\alpha - \|Z_n\|_\alpha^\alpha - \|Z\|_\alpha^\alpha \xrightarrow[n \rightarrow \infty]{} 0$,
3. $m_p(Z_n, Z) := \mathbb{E}[|Z_n^p - Z^p|] \xrightarrow[n \rightarrow \infty]{} 0$.

Moreover, the functionals $\rho_{\mathcal{M}, \alpha}$ and m_p are metrics which metrize the convergence in probability in \mathcal{M} . Finally, \mathcal{M} is a closed Fréchet space if and only if it is complete with either one or both of those metrics.

Lemma 1.3.27 *Let Z be a random variable with Fréchet distribution $\mathcal{F}(\alpha, \|Z\|_\alpha)$. If $p \in]0, \alpha[$, then $Z \in \mathbf{L}^p(\mathbb{R}_+^*)$ and*

$$\mathbb{E}[Z^p] = c_{p, \alpha} \|Z\|_\alpha^p := \Gamma\left(1 - \frac{p}{\alpha}\right) \|Z\|_\alpha^p,$$

where Γ denotes the Gamma function. Furthermore, if the sequence of random variables $(Z_n)_{n \geq 1}$ converges in distribution to some random variable Z , then $(\|Z_n\|_\alpha)_{n \geq 1}$ converges and

$$Z \sim \mathcal{F}\left(\alpha, \lim_{n \rightarrow \infty} \|Z_n\|_\alpha\right).$$

On the other hand, if there exists $\sigma \in \mathbb{R}_+$ such that $(\|Z_n\|_\alpha)_n$ converges to σ , then

$$Z_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{F}(\alpha, \sigma).$$

The previous results constitute the main ingredients to define the extremal integral for general integrands. It is constructed as the μ -almost sure limit of the sequence of extremal integrals of f_n , where $(f_n)_n$ is an increasing sequence of non-negative simple functions converging to f . Naturally this limit does not depend on the approximating sequence. Now we list the main properties of this integral:

Theorem 1.3.28 *Let f be a non-negative function such that $\int_E f(x)^\alpha d\mu(x)$ is finite. Then the extremal integral of f on E exists and is a random variable Z with Fréchet distribution $\mathcal{F}(\alpha, \|Z\|_\alpha)$, where*

$$\|Z\|_\alpha^\alpha = \left\| \int_E f(x) dM_\alpha(x) \right\|_\alpha^\alpha = \int_E f(x)^\alpha d\mu(x).$$

This theorem can be seen as an 'extremal' counterpart of the Itô isometry. But keep in mind that α -Fréchet random variables are *never* in $\mathbf{L}^\alpha(\mathbb{R}_+^*)$, so this analogy can be misleading. In the sequel we will use the notation

$$\mathbf{L}_+^\alpha(E, \mu) := \left\{ f : E \rightarrow \mathbb{R}_+, \int_E f(x)^\alpha d\mu(x) < +\infty \right\}.$$

The next proposition gathers several properties of the extremal integral, some of which being essential for the next subsection.

Proposition 1.3.29 1. (Max-linearity) *For all $f, g \in \mathbf{L}_+^\alpha(E, \mu)$, we have*

$$\left(\lambda \int_E f dM_\alpha \right) \oplus \left(\mu \int_E g dM_\alpha \right) = \int_E (\lambda f \oplus \mu g) dM_\alpha, \quad \lambda, \mu \geq 0.$$

2. (Independence) *The extremal integrals of f and g are independent if and only if f and g have disjoint supports, that is:*

$$\int_E f dM_\alpha \text{ and } \int_E g dM_\alpha \text{ are independent if and only if } fg = 0 \text{ } \mu\text{-a.e.}$$

3. (Monotonicity)

$$f \leq g \text{ } \mu\text{-a.e. if and only if } \int_E f dM_\alpha \leq \int_E g dM_\alpha \text{ } \mu\text{-a.e.}$$

In particular, $f = g$ μ -a.e. if and only if the associated extremal integrals are equal μ -a.e.

1.4 Distances between distributions

Every random element is defined on an appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

In this section we give a brief account of the main probability distances we will use throughout this dissertation. More on this topic can be found in the classic monographs [Nourdin and Peccati \[2012\]](#) and [Dudley \[2002\]](#), as well as in the reference papers [Zolotarev \[1984\]](#) and [Müller \[1997\]](#).

As stated in the introduction, Stein's method works for integral probability metrics between

probability measures μ, ν on \mathbb{R}^d of the form

$$d_{\mathcal{H}}(\mu, \nu) := \sup_{h \in \mathcal{H}} |\mathbb{E}_{\mu}[h] - \mathbb{E}_{\nu}[h]|, \quad (1.28)$$

where \mathcal{H} is a separating class of real-valued functions. We say that a class of functions \mathcal{H} is *separating* if all its elements h are Borel-measurable functions, belong to $\mathbf{L}^1(\Omega, \mathbb{P})$ and are such that

$$\mathbb{E}_{\mu}[h] = \mathbb{E}_{\nu}[h] \text{ for all } h \in \mathcal{H} \implies \mu = \nu.$$

The functions h of \mathcal{H} will often be called *test functions* while $d_{\mathcal{H}}$ will be referred as the distance associated to \mathcal{H} . An instance of a separating class of functions is the set of continuous, compactly-supported functions from \mathbb{R}^d to \mathbb{R} .

Equation (1.28) defines a *metric* on the subset of probability measures μ on \mathbb{R}^d such that h is integrable with respect to μ for all $h \in \mathcal{H}$: it is symmetric in its arguments μ and ν , it satisfies the triangle inequality, and it vanishes if and only if $\mu = \nu$. Notice however that the topology it induces has no reason to be the topology of the convergence in law in general.

Next we give a few examples of distances of the form (1.28). All of them will be used at some point in the sequel.

Definition 1.4.1. Let μ, ν be two probability measures on \mathbb{R}^d , and \mathbf{X}, \mathbf{Y} two random vectors with law μ and ν respectively. We do *not* need to assume they are defined on the same probability space.

- The *total variation distance* d_{TV} is the distance associated to

$$\mathcal{H} = \{h_A : \mathbf{x} \mapsto \mathbf{1}_A(\mathbf{x}), A \in \mathcal{B}(\mathbb{R}^d)\}.$$

Therefore, the total variation distance writes as

$$d_{\text{TV}}(\mathbf{X}, \mathbf{Y}) = \sup_{A \in \mathcal{B}(\mathbb{R}^d)} |\mathbb{P}(\mathbf{X} \in A) - \mathbb{P}(\mathbf{Y} \in A)|.$$

It can also be expressed as a \mathbf{L}^1 -distance between the densities $f_{\mathbf{X}}$ and $f_{\mathbf{Y}}$ of $\mathbb{P}_{\mathbf{X}}$ and $\mathbb{P}_{\mathbf{Y}}$ with respect to $(\mathbb{P}_{\mathbf{X}} + \mathbb{P}_{\mathbf{Y}})/2$:

$$d_{\text{TV}}(\mathbf{X}, \mathbf{Y}) = \frac{1}{2} \int_{\mathbb{R}^d} |f_{\mathbf{X}}(z) - f_{\mathbf{Y}}(z)| d(\mathbb{P}_{\mathbf{X}} + \mathbb{P}_{\mathbf{Y}})(z).$$

- The *Kolmogorov distance* d_{K} is the distance associated to

$$\mathcal{H} = \{h_z : \mathbf{x} \mapsto \mathbf{1}_{(-\infty, z]}(\mathbf{x}), z \in \mathbb{R}^d\}.$$

Therefore, the Kolmogorov distance writes as

$$d_{\text{K}}(\mathbf{X}, \mathbf{Y}) = \sup_{z \in \mathbb{R}^d} |F_{\mathbf{X}}(z) - F_{\mathbf{Y}}(z)| = \|F_{\mathbf{X}} - F_{\mathbf{Y}}\|_{\infty}.$$

- The *Wasserstein distance* d_W is the distance associated to the set of 1-Lipschitz functions

$$\mathcal{H} = \text{Lip}(\mathbb{R}^d, \mathbb{R}) = \{h : \mathbb{R}^d \rightarrow \mathbb{R}, |h(\mathbf{x}) - h(\mathbf{y})| \leq \|\mathbf{x} - \mathbf{y}\| \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^d\},$$

where $\|\cdot\|$ is some norm on \mathbb{R}^d .

- The *smooth Wasserstein distance* $d_{[2]}$ is the distance associated to the set of differentiable functions which are 1-Lipschitz, and whose derivative is also 1-Lipschitz

$$\mathcal{H} = \text{Lip}_{[2]}(\mathbb{R}^d, \mathbb{R}) = \{h : \mathbb{R}^d \rightarrow \mathbb{R}, h \text{ and } h' \text{ are 1-Lipschitz}\},$$

with respect to the same norm $\|\cdot\|$ on \mathbb{R}^d .

- The *stop-loss distance* d_{SL} can be considered a compromise between the Kolmogorov distance and the Wasserstein distance. The set of test-functions \mathcal{H} is:

$$\mathcal{H} = \{h_z : x \mapsto (x - z)_+, z \in \mathbb{R}\}.$$

Those functions are well-known in finance (to define European call options) and in deep learning (where they serve as activation functions for neural networks). Notice that h_z is 1-Lipschitz. In higher dimensions, a possible generalization would consist in taking functions of the form:

$$h_{\mathbf{z}}(\mathbf{x}) = \bigoplus_{j=1}^d h_{z_j}(x^j) = \bigoplus_{j=1}^d (x^j - z^j)_+, \mathbf{x} \in \mathbb{R}^d$$

where \mathbf{z} is some fixed vector in \mathbb{R}^d . Like their univariate counterparts, those functions are also 1-Lipschitz.

Many other choices exist and are commonly used in applications, such as the Fortet-Mourier distance (whose test-functions are the bounded Lipschitz functions) or the Radon distance (bounded continuous functions). Naturally, they are not equivalent in general, although it is possible to order them up to a certain extent. This is the purpose of the next result, which is stated and proved in greater generality in [Gaunt and Li \[2023\]](#).

Proposition 1.4.2 *Let μ, ν be two probability measures on \mathbb{R}^d .*

- *The Kolmogorov metric is bounded by the total variation metric:*

$$d_K(\mu, \nu) \leq d_{\text{TV}}(\mu, \nu).$$

- *The stop-loss distance is bounded by the Wasserstein distance:*

$$d_{\text{SL}}(\mu, \nu) \leq d_W(\mu, \nu).$$

- *If ν has a bounded density w.r.t. to Lebesgue measure on \mathbb{R}^d , then we also have:*

$$d_K(\mu, \nu) \leq C \sqrt{d_W(\mu, \nu)}, \tag{1.29}$$

for some constant C independent of μ .

- Under the same assumption on ν , we have:

$$d_K(\mu, \nu) \leq C \sqrt[3]{d_{[2]}(\mu, \nu)}, \quad (1.30)$$

for some constant C independent of μ .

Remark 1.4.3 (With which metric should one work?). As one can see, the choice of the working distance is not benign. It mainly depends on the application in mind, and more importantly, on the assumptions on the test-functions one needs. For instance, the total variation metric is well-suited to compare discrete distributions but cannot be used to determine the distance between a discrete distribution and a continuous one. Take for example X with an exponential distribution $\mathcal{E}(1)$, and X_n with a geometric distribution $\text{Geom}(1/n)$ divided by n . When n goes to infinity, $(X_n)_{n \geq 1}$ converges in distribution to X . However, X_n has its support in the discrete set $E_n = \{k/n, k \in \mathbb{N}^*\}$, which has Lebesgue measure equal to 0. In other words, we always have

$$1 = |\mathbb{P}(X_n \in E_n) - \mathbb{P}(X \in E_n)| \leq d_{\text{TV}}(X_n, X) \leq 1, \quad n \in \mathbb{N}^*,$$

and the convergence in distribution of $(X_n)_{n \geq 1}$ to X cannot be observed with this metric. On the other hand, the Kolmogorov distance is less demanding and allows for comparison between discrete and continuous distributions. But it is not always the most natural choice, especially in higher dimensions, and the indicator functions $\mathbb{1}_{(-\infty, z]}(\cdot)$ have little regularity. This may cause troubles when trying to apply Stein's method. The Wasserstein distance uses 1-Lipschitz functions as its test-functions. Because a norm $\|\cdot\|$ is always 1-Lipschitz with respect to itself, one must assume that both $\|\mathbf{X}\|$ et $\|\mathbf{Y}\|$ are integrable for $d_W(\mathbf{X}, \mathbf{Y})$ to exist. This condition is also sufficient:

$$|h(\mathbf{X}) - h(\mathbf{Y})| \leq \|\mathbf{X} - \mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|,$$

for any 1-Lipschitz function. Consequently, the Wasserstein distance is not suited to work with non-integrable random vectors, such as stable distributions, or Fréchet distributions, with shape parameter $\alpha \in (0, 1)$. On the other hand, a Lipschitz function h is absolutely continuous, so it is differentiable almost everywhere, and its derivative h' is bounded almost everywhere too (Rademacher's theorem, see Villani [2008]). Notice we can also express this distance as:

$$d_W(\mathbf{X}, \mathbf{Y}) = \int_{\mathbb{R}^d} |F_X(z) - F_Y(z)| dz = \inf_{\substack{\mathbf{X}' \sim \mathbf{X} \\ \mathbf{Y}' \sim \mathbf{Y}}} \mathbb{E}[\|\mathbf{X}' - \mathbf{Y}'\|], \quad (1.31)$$

where the infimum is taken on all coupling $(\mathbf{X}', \mathbf{Y}')$ such that \mathbf{X}' and \mathbf{Y}' have the same distributions as \mathbf{X} and \mathbf{Y} respectively. Those properties make many computations much more convenient when applying Stein's method to this distance. Besides, inequality (1.29) allows us to bound the Kolmogorov distance by using the Wasserstein distance. Thus, getting a rate of convergence for the former implies getting another one for the latter, although it may not be optimal due to the presence of the square root. Finally, the smooth Wasserstein distance imposes stronger assumptions on the test functions, which may prove necessary in certain applications. For instance, when one of the two distributions at least is not continuous, one cannot work anymore with h' if h is only Lipschitz, since the set of points at which it is not differentiable may intersect the set of atoms of the discontinuous distribution. Furthermore, better rates of convergence may be obtained thanks to those extra assumptions compared to the Wasserstein distance. An instance of this phenomenon can be found in Goldstein and Reinert [2001], and another one will appear later in this dissertation.

■

2 Stochastic quantization of max-stable distributions

Contents

2.1	Quantization of univariate extreme-value distributions	63
2.1.1	Quantization of the Fréchet distribution	63
2.1.2	Extension to the Gumbel and Weibull distributions	76
2.2	The Fréchet semi-group as a Markov process	88
2.3	Quantization of multivariate max-stable distributions	94
2.3.1	The case of the unit Fréchet max-stable random vectors	95
2.3.2	The general case	101

The goal of this chapter is to define the main tool of this thesis: a family of semi-groups which quantize the extreme value distributions. Since one-dimensional EVT has certain unique features compared to its multivariate counterpart, we start with the univariate case and then proceed to generalize our results to the multivariate case.

2.1 Quantization of univariate extreme-value distributions

As stated in the preliminary, extreme value distributions are parametrized by a shape parameter $\alpha \in \mathbb{R}$. The sign of α has a major impact on the behaviour of the extreme value distribution, and so does it on the semi-groups we introduce now.

2.1.1 Quantization of the Fréchet distribution

We fix $\alpha > 0$ in this subsection and let Z be a random variable with Fréchet distribution $\mathcal{F}(\alpha)$. Set $\gamma_t = e^t - 1$. We need to introduce the standard algebra (for a semi-group which has yet to be introduced) with which we will work in this subsection:

Definition 2.1.1 (Standard algebra - Fréchet case). We define the class of functions:

$$\mathcal{S}_\Phi := \left\{ f \in \mathcal{S}(\mathbb{R}_+^*), \sup_{x \in (0,1]} x^{-k} |f^{(n)}(x)| < \infty \text{ for all } (k, n) \in \mathbb{N} \times \mathbb{N}^* \right\},$$

where $\mathcal{S}(\mathbb{R}_+^*)$ is the Schwartz class of rapidly decreasing functions from \mathbb{R}_+^* to \mathbb{R} .

It is important to notice that we do *not* require f to vanish at 0 faster than any monomial x^k . This property only concerns f' and its derivatives. Observe that \mathcal{S}_Φ is included in $\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)$ for all

$p \geq 1$. Notice also that if f and g belong to \mathcal{S}_Φ , then so do $f + \lambda g$ and fg for any $\lambda \in \mathbb{R}$, thus \mathcal{S}_Φ is indeed an algebra of functions. Besides it is not empty.

Example 3. The space $\mathcal{C}_c^\infty(\mathbb{R}_+^*)$ of compactly supported, infinitely differentiable functions on \mathbb{R}_+^* is included in \mathcal{S}_Φ . Another example is the function $x \mapsto \exp(-x - x^{-1})$, which does not belong to $\mathcal{C}_c^\infty(\mathbb{R}_+^*)$.

Definition 2.1.2 (Fréchet semi-group). Let $\alpha \in \mathbb{R}_+^*$. Define the *Fréchet semi-group* $(\mathbf{P}_t^\alpha)_{t \geq 0}$ on the standard algebra $\mathcal{A} = \mathcal{S}_\Phi$ by

$$\mathbf{P}_t^\alpha f(x) := \mathbb{E}[f(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)], \quad x \geq 0, t \geq 0 \quad (2.1)$$

where $Z \sim \mathcal{F}(\alpha)$ and f belongs to \mathcal{S}_Φ .

Remark 2.1.3. This definition is formally similar to the Mehler's formula for the standard Ornstein-Uhlenbeck semi-group (0.9) as well as the definition of the stable semi-group introduced in Arras and Houdré [2023]:

$$P_t^\alpha f(x) := \mathbb{E}[f(e^{-t}x + (1 - e^{-\alpha t})^{\frac{1}{\alpha}}Z)],$$

where Z is an α -stable distribution on \mathbb{R} . Notice however that in our definition of the Fréchet semi-group, α also appears in the first exponential term. To have a definition closer to the original Mehler's formula, we could have set

$$\widetilde{\mathbf{P}}_t^\alpha f(x) := \mathbb{E}[f(e^{-t}x \oplus (1 - e^{-\alpha t})^{\frac{1}{\alpha}}Z)], \quad x \geq 0, t \geq 0.$$

With this definition, we have $\widetilde{\mathbf{P}}_t^\alpha f(x) = \mathbf{P}_{\alpha t}^\alpha f(x)$. But this choice would have led to more cumbersome formulae in the sequel so we prefer to stick with \mathbf{P}_t^α rather than $\widetilde{\mathbf{P}}_t^\alpha$. A connection between the semi-groups we will define for each EVD is possible too with this definition. ■

The goal of the next propositions is to check that $(\mathbf{P}_t^\alpha)_{t \geq 0}$ is indeed a Markov semi-group in the sense of definition 1.1.7. We will prove later that \mathcal{S}_Φ is indeed a standard algebra for $(\mathbf{P}_t^\alpha)_{t \geq 0}$. For now, simply keep in mind that because it contains $\mathcal{C}_c^\infty(\mathbb{R}_+^*)$, it is dense in every $\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)$.

Proposition 2.1.4 The family of operators $(\mathbf{P}_t^\alpha)_{t \geq 0}$ is a semi-group on \mathcal{S}_Φ .

Proof. It is clear that each \mathbf{P}_t^α is a well-defined linear operator for $f \in \mathcal{S}_\Phi$, we have:

$$\begin{aligned} \mathbb{E}[|f|(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)] &\leq C \mathbb{E}[(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)^{-1}] \\ &\leq C(1 - e^{-t})^{-\frac{1}{\alpha}} \mathbb{E}[Z^{-1}] \\ &< +\infty, \end{aligned}$$

where $C := \sup_{x \in \mathbb{R}_+^*} x f(x)$. Thanks to the previous inequality, one gets that $\mathbf{P}_t^\alpha f$ is measurable and bounded. We also see that $\mathbf{P}_0^\alpha = \text{Id}$ on \mathcal{S}_Φ . The semi-group property is a consequence of the

max-stability relation (1.11) satisfied by the Fréchet distribution:

$$\begin{aligned}
\mathbf{P}_t^\alpha(\mathbf{P}_s^\alpha f)(x) &= \mathbb{E}[(\mathbf{P}_s^\alpha f)(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)] \\
&= \mathbb{E}\left[\mathbb{E}\left[f\left((e^{-\frac{s}{\alpha}}(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z_1) \oplus (1 - e^{-s})^{\frac{1}{\alpha}}Z_2)\right)\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[f\left(e^{-\frac{1}{\alpha}(s+t)}x \oplus e^{-\frac{s}{\alpha}}(1 - e^{-t})^{\frac{1}{\alpha}}Z_1 \oplus (1 - e^{-s})^{\frac{1}{\alpha}}Z_2\right)\right]\right] \\
&= \mathbb{E}[f(e^{-\frac{1}{\alpha}(s+t)}x \oplus (1 - e^{-(s+t)})^{\frac{1}{\alpha}}Z)] \\
&= \mathbf{P}_{t+s}^\alpha f(x),
\end{aligned}$$

where Z_1 and Z_2 are i.i.d. random variables with the Fréchet distribution $\mathcal{F}(\alpha)$. The expectation of 1 w.r.t. any probability measure is still 1, so $\mathbf{P}_t^\alpha \mathbf{1} = \mathbf{1}$. Finally the positivity of the expectation operator implies that $(f \geq 0) \implies (\mathbf{P}_t^\alpha f \geq 0)$. This completes the proof. \square

We will need the following lemma, which gives an alternative expression of $\mathbf{P}_t^\alpha f$. It will prove very convenient throughout this section.

Lemma 2.1.5 *Let $f \in \mathcal{S}_\Phi$. The operator \mathbf{P}_t^α satisfies:*

$$\mathbf{P}_t^\alpha f(x) = f(e^{-\frac{t}{\alpha}}x)e^{-\frac{\gamma_t}{x^\alpha}} + \gamma_t \int_x^\infty f(e^{-\frac{t}{\alpha}}z)e^{-\frac{\gamma_t}{z^\alpha}} \frac{\alpha}{z^{\alpha+1}} dz. \quad (2.2)$$

Proof. It suffices to distinguish two cases, depending on whether $(1 - e^{-t})Z \leq e^{-t/\alpha}x$ or the opposite:

$$\begin{aligned}
\mathbf{P}_t^\alpha f(x) &= \mathbb{E}[f(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)] \\
&= \mathbb{E}\left[f(e^{-\frac{t}{\alpha}}x) \mathbb{1}_{\{e^{-t/\alpha}x \geq (1 - e^{-t})^{1/\alpha}Z\}}\right] + \mathbb{E}\left[f((1 - e^{-t})^{1/\alpha}Z) \mathbb{1}_{\{e^{-t/\alpha}x < (1 - e^{-t})^{1/\alpha}Z\}}\right] \\
&= f(e^{-\frac{t}{\alpha}}x) \mathbb{P}(\gamma_t^{1/\alpha}Z \leq x) + \mathbb{E}\left[f((1 - e^{-t})^{\frac{1}{\alpha}}Z) \mathbb{1}_{\{\gamma_t^{1/\alpha}Z > x\}}\right] \\
&= f(e^{-\frac{t}{\alpha}}x) \mathbb{P}(\gamma_t^{1/\alpha}Z \leq x) + \mathbb{E}\left[f(e^{-\frac{t}{\alpha}}\gamma_t^{1/\alpha}Z) \mathbb{1}_{\{\gamma_t^{1/\alpha}Z > x\}}\right].
\end{aligned}$$

Since the density of a Fréchet $\mathcal{F}(\alpha, \sigma)$ random variable Z and $\gamma_t^{1/\alpha}Z \sim \mathcal{F}(\alpha, \gamma_t^{1/\alpha})$ is

$$f_Z(x) = \sigma^\alpha e^{-\frac{\sigma^\alpha}{x^\alpha}} \frac{\alpha}{x^{\alpha+1}}, \quad x \geq 0,$$

and 0 for negative x , we get the desired result. \square

We next check the right-continuity property of $t \mapsto \mathbf{P}_t^\alpha$ in $\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)$.

Proposition 2.1.6 *The mapping $t \mapsto \mathbf{P}_t^\alpha$ is right-continuous at 0 in $\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)$.*

Proof. First assume that $\alpha > 2$ so that the integral $\int_{\mathbb{R}_+^*} r^2 \, d\Phi_\alpha(r)$ is finite. Let $f \in \mathcal{S}_\Phi$.

$$\begin{aligned}
& \|\mathbf{P}_t^\alpha f - f\|_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)} \\
&= \int_{\mathbb{R}_+^*} |\mathbf{P}_t^\alpha f(r) - f(r)|^2 \, d\Phi_\alpha(r) \\
&= \int_{\mathbb{R}_+^*} \left| f(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} + \gamma t \int_r^\infty f(e^{-\frac{t}{\alpha}} z) e^{-\frac{\gamma t}{z^\alpha}} \frac{\alpha}{z^{\alpha+1}} \, dz - f(r) \right|^2 \, d\Phi_\alpha(r) \\
&\leq 2 \int_{\mathbb{R}_+^*} |f(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} - f(r)|^2 \, d\Phi_\alpha(r) + 2\gamma t^2 \int_{\mathbb{R}_+^*} \left| \int_r^\infty f(e^{-\frac{t}{\alpha}} z) e^{-\frac{\gamma t}{z^\alpha}} \frac{\alpha}{z^{\alpha+1}} \, dz \right|^2 \, d\Phi_\alpha(r) \\
&\leq 2 \int_{\mathbb{R}_+^*} |f(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} - f(r)|^2 \, d\Phi_\alpha(r) + 2\|f\|_\infty^2 \int_{\mathbb{R}_+^*} (1 - e^{-\frac{\gamma t}{r^\alpha}})^2 \, d\Phi_\alpha(r).
\end{aligned}$$

By the dominated convergence theorem applied to the second integral (bound the integrand by 1), the second half of the previous expression vanishes when t goes to 0^+ . To deal with the first half, write

$$\begin{aligned}
|f(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} - f(r)| &\leq |f(e^{-\frac{t}{\alpha}} r) - f(r)| e^{-\frac{\gamma t}{r^\alpha}} + |f(r)| (1 - e^{-\frac{\gamma t}{r^\alpha}}) \\
&\leq \|f'\|_\infty (1 - e^{-\frac{t}{\alpha}}) r + \|f\|_\infty \frac{\gamma t}{r^\alpha},
\end{aligned}$$

since both f and $x \mapsto e^{-x}$ are Lipschitz on \mathbb{R}_+^* . Because we have assumed that $\alpha > 2$, we know that $r \mapsto r^2$ is integrable w.r.t. Φ_α . As a result, the square of each of the previous terms is integrable w.r.t. Φ_α and tends to 0 as t goes to 0^+ , which concludes the proof when α is greater than 2.

If $\alpha \in (0, 2]$, simply replace f by $g : r \mapsto f(r^k)$, where $k = \lfloor 3\alpha \rfloor$, and Z by $Z^{1/k}$. In this case we have that g is still in \mathcal{S}_Φ , and in particular remains bounded and Lipschitz, while $Z^{1/k}$ admits a variance. The previous arguments thus become valid again. \square

The next proposition is central for the applications and is another consequence of the max-stability argument used to prove that $(\mathbf{P}_t^\alpha)_{t \geq 0}$ has the semi-group property.

Proposition 2.1.7 *Let Φ_α denote the Fréchet probability measure $\mathcal{F}(\alpha)$. The semi-group $(\mathbf{P}_t^\alpha)_{t \geq 0}$ admits Φ_α has a stationary measure. As such, $(\mathbf{P}_t^\alpha)_{t \geq 0}$ is a Markov semi-group. Furthermore, $(\mathbf{P}_t^\alpha)_{t \geq 0}$ is ergodic:*

$$\mathbf{P}_t^\alpha f(x) \xrightarrow{t \rightarrow \infty} \mathbb{E}_{\Phi_\alpha}[f] = \int_{\mathbb{R}_+^*} f(r) \, d\Phi_\alpha(r), \quad x > 0, \quad f \in \mathcal{S}_\Phi.$$

Proof. Let $Z_1, Z_2 \sim \mathcal{F}(\alpha)$ be two i.i.d. random variables with the Fréchet distribution. We have

$$\int_{\mathbb{R}_+^*} \mathbf{P}_t^\alpha f(r) \, d\Phi_\alpha(r) = \mathbb{E}[f(e^{-\frac{t}{\alpha}} Z_1 \oplus (1 - e^{-t})^{\frac{1}{\alpha}} Z_2)] = \mathbb{E}[f(Z)] = \int_{\mathbb{R}_+^*} f(r) \, d\Phi_\alpha(r).$$

The proof of the ergodicity is a consequence of the dominated convergence theorem since

$$e^{-\frac{t}{\alpha}} r \oplus (1 - e^{-t})^{\frac{1}{\alpha}} Z \xrightarrow{t \rightarrow \infty} Z \text{ a.s.}$$

and $f \in \mathcal{S}_\Phi$. \square

Proposition 2.1.8 *For every $t \geq 0$ and every $p \in [1, \infty)$, the operator \mathbf{P}_t^α extends to a linear contraction operator on $\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)$.*

Proof. The argument is generic and is the same as the one used in [Nourdin and Peccati \[2012\]](#) to prove a similar property for the Ornstein-Uhlenbeck semi-group. By Jensen's inequality applied to the probability measure Φ_α and the convex function $x \mapsto x^p$, we have

$$\begin{aligned} \|\mathbf{P}_t^\alpha f\|_{\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)}^p &= \int_{\mathbb{R}_+^*} |\mathbb{E}[f(e^{-\frac{t}{\alpha}}r \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)]|^p d\Phi_\alpha(r) \\ &\leq \int_{\mathbb{R}_+^*} \mathbb{E}[|f|^p(e^{-\frac{t}{\alpha}}r \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)] d\Phi_\alpha(r) \\ &= \mathbb{E}[|f|^p(e^{-\frac{t}{\alpha}}Z_1 \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z_2)] \\ &= \mathbb{E}[|f|^p(Z)] \\ &= \|f\|_{\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)}^p, \end{aligned}$$

where Z_1 and Z_2 are i.i.d. random variables with the Fréchet distribution $\mathcal{F}(\alpha)$. We have once again used relation (1.11), or equivalently, that Φ_α is a (the) stationary measure for $(\mathbf{P}_t^\alpha)_{t \geq 0}$. □

We give other useful alternative expressions of \mathbf{P}_t^α , its derivative as well as of its generator, all based on (2.2).

Proposition 2.1.9 *Let $f \in \mathcal{S}_\Phi$ and $x \in \mathbb{R}_+^*$. We have the following:*

1. *The derivative of $\mathbf{P}_t^\alpha f$ satisfies:*

$$(\mathbf{P}_t^\alpha f)'(x) = e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} f'(e^{-\frac{t}{\alpha}}x) \quad (2.3)$$

2. *$\mathbf{P}_t^\alpha f$ can be rewritten as:*

$$\mathbf{P}_t^\alpha f(x) = -e^{-\frac{t}{\alpha}} \int_x^\infty e^{-\frac{\gamma t}{r^\alpha}} f'(e^{-\frac{t}{\alpha}}r) dr. \quad (2.4)$$

Proof. 1. To prove equation (2.3), we simply differentiate (2.2) w.r.t. x

$$\begin{aligned} \frac{d}{dx} \left(f(e^{-\frac{t}{\alpha}}x) e^{-\frac{\gamma t}{x^\alpha}} + \gamma t \int_x^\infty f(e^{-\frac{t}{\alpha}}z) e^{-\frac{\gamma t}{z^\alpha}} \frac{\alpha}{z^{\alpha+1}} dz \right) \\ = e^{-\frac{t}{\alpha}} f'(e^{-\frac{t}{\alpha}}x) e^{-\frac{\gamma t}{x^\alpha}} + \gamma t f(e^{-\frac{t}{\alpha}}x) e^{-\frac{\gamma t}{x^\alpha}} \frac{\alpha}{x^{\alpha+1}} - \gamma t f(e^{-\frac{t}{\alpha}}x) e^{-\frac{\gamma t}{x^\alpha}} \frac{\alpha}{x^{\alpha+1}} \\ = e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} f'(e^{-\frac{t}{\alpha}}x). \end{aligned}$$

2. This is a direct consequence of the previous point and of the fundamental theorem of calculus. □

Now that we have proved that $(\mathbf{P}_t^\alpha)_{t \geq 0}$ is a Markov semi-group, we can focus on computing its generator.

Proposition 2.1.10 *The Markov semi-group $(\mathbf{P}_t^\alpha)_{t \geq 0}$ has generator \mathcal{L}_α :*

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha} x f'(x) + \int_0^\infty (f(x \oplus r) - f(x)) \, d\rho_\alpha(r), \quad x > 0$$

where $d\rho_\alpha(r) := \alpha r^{-(\alpha+1)} \, dr$ on \mathbb{R}_+^* .

Proof. For starters, observe that

$$\int_0^\infty (f(x \oplus r) - f(x)) \, d\rho_\alpha(r) = -\frac{f(x)}{x^\alpha} + \int_x^\infty f(r) \, d\rho_\alpha(r). \quad (2.5)$$

We compute the limit of $(\mathbf{P}_t^\alpha f(x) - f(x))/t$ for all positive x and f in the Schwartz class $f \in \mathcal{S}_\Phi$ by relying once again on (2.2).

$$\begin{aligned} \frac{1}{t}(\mathbf{P}_t^\alpha f(x) - f(x)) - \mathcal{L}_\alpha f(x) &= \frac{1}{t} \left(f(e^{-\frac{t}{\alpha}} x) e^{-\frac{\gamma t}{x^\alpha}} + \gamma t \int_x^\infty f(e^{-\frac{t}{\alpha}} z) e^{-\frac{\gamma t}{r^\alpha}} \, d\rho_\alpha(r) - f(x) \right) - \mathcal{L}_\alpha f(x) \\ &= \frac{1}{t} \left(f(e^{-\frac{t}{\alpha}} x) e^{-\frac{\gamma t}{x^\alpha}} - f(x) \right) + \frac{1}{\alpha} x f'(x) + \frac{f(x)}{x^\alpha} \\ &\quad + \frac{\gamma t}{t} \int_x^\infty f(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} \, d\rho_\alpha(r) - \int_x^\infty f(r) \, d\rho_\alpha(r) \\ &= (1) + (2). \end{aligned}$$

To deal with (1), assume as before $\alpha > 2$ and use the same decomposition as in the proof of proposition 2.1.6:

$$\frac{1}{t} \left(f(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} - f(r) \right) = \frac{1}{t} \left(f(e^{-\frac{t}{\alpha}} r) - f(r) \right) + \frac{1}{t} f(r) (e^{-\frac{\gamma t}{r^\alpha}} - 1).$$

The second part converges to $r^{-\alpha} f(r)$ a.s. and in $\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)$ thanks to a dominated convergence argument. To deal with the first part, use a Taylor formula with remainder: there exists a $\xi_{r,t} \in]r e^{-t/\alpha}, r[$ such that

$$f(e^{-\frac{t}{\alpha}} r) = f(r) + (e^{-\frac{t}{\alpha}} - 1) r f'(r) + \frac{1}{2} (e^{-\frac{t}{\alpha}} - 1)^2 r^2 f''(\xi_{r,t}).$$

Since f is in \mathcal{S}_Φ and α is greater than 2, f'' is bounded and the bounded convergence theorem applies, so we get that (1) goes to 0 a.s. and in $\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)$. Using the same argument as in the proof of proposition 2.1.6 with $k = \lfloor 4\alpha \rfloor$, we can extend this proof to the general case.

As for (2), since the Schwartz class $\mathcal{S}(\mathbb{R}_+^*)$ is stable under integration, similar arguments as the ones described above apply and give that (2) goes to 0 too, both a.s. and in $\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)$. \square

An important observation about \mathcal{L}_α is that it can be decomposed into two distinct parts, each with its own role:

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha} x f'(x) + \int_0^\infty (f(x \oplus r) - f(x)) \, d\rho_\alpha(r) = \mathfrak{D}_\alpha f(x) + \mathbf{D}_\alpha f(x), \quad (2.6)$$

where \mathfrak{D}_α is the generator of the dilatations at power α , and \mathbf{D}_α is defined by:

$$\mathbf{D}_\alpha f(x) := \int_0^\infty (f(x \oplus r) - f(x)) \, d\rho_\alpha(r).$$

We will show later in this chapter that \mathbf{D}_α is also the generator of a certain (Markov) semi-group.

It is well-known that the Fréchet distribution and the Pareto distribution are intimately connected. This link manifests in the next proposition too which gives two different formulae for the generator \mathcal{L}_α .

Proposition 2.1.11 *Let $f \in \mathcal{S}_\Phi$ and $x \in \mathbb{R}_+^*$ and $Z \sim \mathcal{F}(\alpha)$. Let $Y \sim \mathcal{VP}(\alpha, 1)$. We have the following:*

1. *The generator of the Fréchet semi-group is connected to the Pareto distribution with shape parameter α and scale parameter 1:*

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha} x f'(x) + \frac{1}{x^\alpha} \int_1^\infty (f(xr) - f(x)) \, d\rho_\alpha(r) \quad (2.7)$$

$$= -\frac{1}{\alpha} x f'(x) + \frac{1}{x^\alpha} \mathbb{E}[f(xY) - f(x)]. \quad (2.8)$$

2. *$\mathcal{L}_\alpha f$ can be rewritten as*

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha} x f'(x) + \frac{1}{\alpha} \int_x^\infty r f'(r) \, d\rho_\alpha(r) \quad (2.9)$$

$$= -\frac{1}{\alpha} x f'(x) + \frac{1}{\alpha} \frac{1}{x^{\alpha-1}} \mathbb{E}[Y f'(xY)]. \quad (2.10)$$

Proof. 1. The change of variable $u = r/x$ gives immediately the result:

$$\mathbf{D}_\alpha f(x) = \int_x^\infty (f(r) - f(x)) \frac{\alpha}{r^{\alpha+1}} \, dr = \frac{1}{x^\alpha} \int_1^\infty (f(xu) - f(x)) \, d\rho_\alpha(u).$$

Because the Pareto distribution $\mathcal{VP}(\alpha, 1)$ has density $r \mapsto \alpha r^{-(\alpha+1)} \mathbb{1}_{\{r \geq 1\}}$, we obtain the second identity just as easily.

2. An integration-by-parts (differentiate $r \mapsto f(r) - f(x)$ and integrate $r \mapsto \alpha r^{-(\alpha+1)}$) yields the desired formula:

$$\mathbf{D}_\alpha f(x) = \int_x^\infty (f(r) - f(x)) \frac{\alpha}{r^{\alpha+1}} \, dr = \int_x^\infty f'(r) \frac{1}{r^\alpha} \, dr = \frac{1}{\alpha} \int_x^\infty r f'(r) \, d\rho_\alpha(r).$$

□

Although \mathbf{D}_α is not a true derivation, it satisfies a pseudo Leibniz rule.

Proposition 2.1.12 — Pseudo Leibniz rule - Fréchet case. *For every $\alpha > 0$ and $f, g \in \mathcal{S}_\Phi$, we have*

$$\begin{aligned} \mathbf{D}_\alpha(fg)(x) &= \mathbf{D}_\alpha f(x)g(x) + f(x)\mathbf{D}_\alpha g(x) \\ &\quad + \int_x^\infty (f(r) - f(x))(g(r) - g(x)) \, d\rho_\alpha(r), \quad x > 0. \end{aligned}$$

This is also equal to:

$$\begin{aligned} \mathbf{D}_\alpha(fg)(x) &= \mathbf{D}_\alpha f(x)g(x) + f(x)\mathbf{D}_\alpha g(x) \\ &\quad + \int_x^\infty \mathbf{D}_\alpha f(y)g'(y) \, dy + \int_x^\infty f'(y)\mathbf{D}_\alpha g(y) \, dy, \quad x > 0. \end{aligned}$$

Proof. The first identity results from the definition of \mathbf{D}_α and the relation

$$f(r)g(r) - f(x)g(x) = f(x)(g(r) - g(x)) + g(x)(f(r) - f(x)) + (f(r) - f(x))(g(r) - g(x)).$$

This result is a consequence of the integration-by-parts formula. \square

Now we have the tools to prove that \mathcal{S}_Φ is a standard algebra for $(\mathbf{P}_t^\alpha)_{t \geq 0}$.

Proposition 2.1.13 *The class of functions \mathcal{S}_Φ is stable under both \mathcal{L}_α and \mathbf{P}_t^α for all nonnegative t : if f belongs to \mathcal{S}_Φ then so do $\mathcal{L}_\alpha f$ and $\mathbf{P}_t^\alpha f$.*

Proof. - First we prove that $\mathbf{P}_t^\alpha f$ belongs to \mathcal{S}_Φ if f belongs to \mathcal{S}_Φ . For every positive x and $k \in \mathbb{N}$,

$$\begin{aligned} x^k |\mathbf{P}_t^\alpha f(x)| &\leq x^k \mathbb{E}[|f|(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)] \\ &= e^{\frac{kt}{\alpha}} (e^{-\frac{t}{\alpha}}x)^k \mathbb{E}[|f|(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)] \\ &\leq e^{\frac{kt}{\alpha}} \mathbb{E}\left[(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)^k |f|(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)\right] \\ &\leq e^{\frac{kt}{\alpha}} \sup_{y \in \mathbb{R}_+^*} y^k |f|(y) \\ &< +\infty. \end{aligned}$$

As for the derivatives of $\mathbf{P}_t^\alpha f$, it suffices to use (2.3) to see that $(\mathbf{P}_t^\alpha f)^{(n)}$ behaves as required in 0 and at $+\infty$ when $n \in \mathbb{N}^*$.

- The term $xf'(x)$ does not pose any problem and is clearly in \mathcal{S}_Φ . The troublesome part is $\mathbf{D}_\alpha f$. First we have for $x > 0$ and $k \in \mathbb{N}$

$$\begin{aligned} x^k |\mathbf{D}_\alpha f(x)| &\leq x^k \int_x^\infty r^{-\alpha} |f'(r)| \, dr \\ &\leq \int_x^\infty r^{k-\alpha} |f'(r)| \, dr \\ &\leq \int_0^\infty r^{k-\alpha} |f'(r)| \, dr \\ &= \int_0^1 r^{k-\alpha} |f'(r)| \, dr + \int_1^\infty r^{k-\alpha} |f'(r)| \, dr \\ &= \sup_{y \in (0,1]} y^{k-\alpha} |f'(y)| + \sup_{y \in \mathbb{R}_+^*} y^{k+1} |f'(y)| \int_1^\infty \frac{1}{r^{\alpha+1}} \, dr \\ &< +\infty. \end{aligned}$$

And since $(\mathbf{D}_\alpha f)'(x) = x^{-\alpha} f'(x)$, we get without difficulty that the derivatives of $\mathbf{D}_\alpha f$ go to 0 more quickly than any x^k for all $k \in \mathbb{Z}$. \square

A second consequence of (2.9) is that \mathbf{D}_α satisfies a commutation relation with \mathbf{P}_t^α akin to the commutation rule of the Gaussian OU semi-group, *i.e.* $\nabla P_t = e^{-t} P_t \nabla$, where ∇ is the usual derivative.

Proposition 2.1.14 — **Commutation relation for \mathbf{D}_α - Fréchet case.** *We have for all $f \in \mathcal{S}_\Phi$*

$$\mathbf{D}_\alpha \mathbf{P}_t^\alpha f(x) = e^{-t} \mathbf{P}_t^\alpha \mathbf{D}_\alpha f(x), \quad x \in \mathbb{R}_+^*, \quad t \geq 0. \quad (2.11)$$

Proof. We use the relation (2.4) and the alternative expression of \mathbf{D}_α :

$$\begin{aligned} \mathbf{D}_\alpha \mathbf{P}_t^\alpha f(x) &= \int_x^\infty (\mathbf{P}_t^\alpha f)'(r) \frac{1}{r^\alpha} dr \\ &= -e^{-\frac{t}{\alpha}} \int_x^\infty f'(e^{-\frac{t}{\alpha}} u) e^{-\frac{\gamma t}{r^\alpha}} \frac{1}{r^\alpha} dr. \end{aligned}$$

On the other hand, computing $\mathbf{P}_t^\alpha \mathbf{D}_\alpha f(x)$, we get:

$$\begin{aligned} \mathbf{P}_t^\alpha \mathbf{D}_\alpha f(x) &= -e^{-\frac{t}{\alpha}} \int_x^\infty (\mathbf{D}_\alpha f)'(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} dr \\ &= -e^{-\frac{t}{\alpha}} \int_x^\infty f'(e^{-\frac{t}{\alpha}} r) \frac{1}{e^{-t} r^\alpha} e^{-\frac{\gamma t}{r^\alpha}} dr \\ &= -e^t e^{-\frac{t}{\alpha}} \int_x^\infty f'(e^{-\frac{t}{\alpha}} r) e^{-\frac{\gamma t}{r^\alpha}} \frac{1}{r^\alpha} dr \\ &= e^t \mathbf{D}_\alpha \mathbf{P}_t^\alpha f(x). \end{aligned}$$

□

Another consequence of (2.3) is the exponential decaying in time of the derivative of $\mathbf{P}_t^\alpha f$ in \mathbf{L}^p -norm .

Proposition 2.1.15 *Assume f is in \mathcal{S}_Φ . Then for every $p \geq 1$,*

$$\|\nabla \mathbf{P}_t^\alpha f\|_{\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)} \leq e^{-(\frac{1}{p} + \frac{1}{\alpha})t} \|\nabla f\|_{\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)}, \quad t \geq 0.$$

Proof. The change of variable $u = e^{-t/\alpha} r$ yields:

$$\begin{aligned} \|\nabla \mathbf{P}_t^\alpha f\|_{\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)}^p &= e^{-\frac{pt}{\alpha}} \int_{\mathbb{R}_+} |f'|^p(e^{-\frac{t}{\alpha}} r) r e^{-p\frac{\gamma t}{r^\alpha}} e^{-\frac{1}{r^\alpha}} \frac{\alpha}{r^{\alpha+1}} dx \\ &= e^{-\frac{pt}{\alpha}} \int_{\mathbb{R}_+} |f'|^p(e^{-\frac{t}{\alpha}} r) e^{-\frac{pe^t - (p-1)}{r^\alpha}} \frac{\alpha}{r^{\alpha+1}} dr \\ &= e^{-(1+\frac{p}{\alpha})t} \int_{\mathbb{R}_+} |f'(u)|^p e^{-\frac{p-(p-1)e^{-t}}{u^\alpha}} \frac{\alpha}{u^{\alpha+1}} e^{-\alpha t} du \\ &= e^{-(1+\frac{p}{\alpha})t} \int_{\mathbb{R}_+} |f'(u)|^p e^{-(p-1)(1-e^{-t})\frac{1}{u^\alpha}} e^{-\frac{1}{u^\alpha}} \frac{\alpha}{u^{\alpha+1}} du \\ &\leq e^{-(1+\frac{p}{\alpha})t} \int_{\mathbb{R}_+} |f'(u)|^p e^{-\frac{1}{u^\alpha}} \frac{\alpha}{u^{\alpha+1}} du \\ &= e^{-(1+\frac{p}{\alpha})t} \|\nabla f\|_{\mathbf{L}^p(\mathbb{R}_+^*, \Phi_\alpha)}^p \end{aligned}$$

since $(p-1)(1-e^{-t})$ is non-negative for $t \geq 0$ and $p \geq 1$.

□

It is well-known that the generator \mathcal{L}_γ of the OU semi-group satisfies

$$\mathcal{L}_\gamma f(x) = -xf'(x) + f''(x) = (\delta \circ \nabla)f(x), \quad (2.12)$$

where $\delta := -x + \nabla$ is known as the divergence operator. It is equal to the adjoint of the usual derivative operator ∇ with respect to the scalar product $\langle f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \gamma)}$ and γ denotes the standard normal distribution $\mathcal{N}(0, 1)$. In particular \mathcal{L}_γ is self-adjoint. We will see that our generator \mathcal{L}_α does not share this property, but it nonetheless satisfies a similar relation:

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha}xf'(x) + \frac{1}{\alpha} \int_x^\infty rf'(r) \, d\rho_\alpha(r) = (\delta_\alpha \circ \mathbf{D}_\alpha)f(x), \quad (2.13)$$

where the operator δ_α is equal to

$$\delta_\alpha f(x) := \alpha^{-1}(x^{\alpha+1}\nabla + \alpha\text{Id})f(x) = \alpha^{-1}x^{\alpha+1}f'(x) + f(x), \quad f \in \mathcal{S}_\Phi.$$

This operator is actually a Stein operator, as proved in [Bartholome and Swan \[2013\]](#), [Kusumoto and Takeuchi \[2020\]](#), where it is denoted by \mathcal{T}_α and does not have the α^{-1} term. Equality (2.13) makes a connection between their operator and \mathcal{L}_α which thus enjoy similar properties. However, we will see in the next chapter that being connected to a semi-group makes \mathcal{L}_α more convenient to work with, in our opinion at least. We now give three interesting properties of δ_α .

Proposition 2.1.16 — Integration-by-parts formula - Fréchet case. *Let f, g belong to \mathcal{S}_Φ . Then we have*

$$\langle \delta_\alpha f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)} = -\langle f, \alpha^{-1}r^{\alpha+1}\nabla g \rangle_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)} \quad (2.14)$$

Proof. The proof of this result stated by [Bartholome and Swan \[2013\]](#) follows from the usual integration-by-parts formula. \square

Just like \mathbf{D}_α , the operator δ_α "almost" commutes with \mathbf{P}_t^α .

Proposition 2.1.17 — Commutation relation for δ_α - Fréchet case. *We have for all $f \in \mathcal{S}_\Phi$*

$$\delta_\alpha \mathbf{P}_t^\alpha f(x) = e^t \mathbf{P}_t^\alpha \delta_\alpha f(x), \quad x \in \mathbb{R}_+^*, \quad t \geq 0. \quad (2.15)$$

Proof. As before, we compute each side of the equality. We start with $\delta_\alpha \mathbf{P}_t^\alpha f$:

$$\delta_\alpha \mathbf{P}_t^\alpha f(x) = \alpha^{-1}x^{\alpha+1}f'(e^{-\frac{t}{\alpha}}x)e^{-\frac{t}{\alpha}}e^{-\frac{\gamma t}{x^\alpha}} + \mathbf{P}_t^\alpha f(x).$$

The second part is equal to:

$$\begin{aligned} \mathbf{P}_t^\alpha \delta_\alpha f(x) &= -e^{-\frac{t}{\alpha}} \int_x^\infty (\delta_\alpha f)'(e^{-\frac{t}{\alpha}}r) e^{-\frac{\gamma t}{r^\alpha}} \, dr \\ &= -e^{-\frac{t}{\alpha}} \int_x^\infty (\alpha^{-1}x^{\alpha+1}f' + f)'(e^{-\frac{t}{\alpha}}r) e^{-\frac{\gamma t}{r^\alpha}} \, dr \\ &= -e^{-\frac{t}{\alpha}} \int_x^\infty ((1 + \alpha^{-1})x^\alpha f' + \alpha^{-1}x^{\alpha+1}f'')(e^{-\frac{t}{\alpha}}r) e^{-\frac{\gamma t}{r^\alpha}} \, dr + \mathbf{P}_t^\alpha f(x) \\ &= e^{-t} \alpha^{-1}x^{\alpha+1}f'(e^{-\frac{t}{\alpha}}x) e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} + e^{-t} \mathbf{P}_t^\alpha f(x), \end{aligned}$$

by integrating by parts the integral with the $\alpha^{-1}x^{\alpha+1}f''$ term (integrating $e^{-t/\alpha}f''(e^{-t/\alpha}r)$ and differentiating $\alpha^{-1}e^{-t}r^{\alpha+1}e^{-\gamma t/r^\alpha}$). \square

Notice that term e^{-t} in the right-hand side of (2.15) has been replaced by e^t . A shorter, but formal, proof of that result consists in using the commutation relation for \mathbf{D}_α together with the well-known fact that \mathcal{L}_α and \mathbf{P}_t^α commute:

$$\begin{aligned}\mathcal{L}_\alpha \circ \mathbf{P}_t^\alpha &= \mathbf{P}_t^\alpha \circ \mathcal{L}_\alpha \iff \delta_\alpha \circ \mathbf{D}_\alpha \circ \mathbf{P}_t^\alpha = (\mathbf{P}_t^\alpha \circ \delta_\alpha) \circ \mathbf{D}_\alpha \\ &\iff (e^{-t}\delta_\alpha \circ \mathbf{P}_t^\alpha) \circ \mathbf{D}_\alpha = (\mathbf{P}_t^\alpha \circ \delta_\alpha) \circ \mathbf{D}_\alpha.\end{aligned}$$

By "simplifying" by \mathbf{D}_α , one recovers the previous result.

Denote by $[A, B] := A \circ B - B \circ A$ the *commutator* between two endomorphisms of \mathcal{S}_Φ . It serves as a tool to measure the lack of commutativity between A and B since $[A, B] = 0$ (the null operator) if and only if A and B commute. The commutator plays a fundamental role in quantum mechanics, see for instance Hall [2013]. The following identities reveals an interesting structure behind \mathcal{L}_α , \mathbf{D}_α and δ_α .

Proposition 2.1.18 — Commutator identities - Fréchet case. *We have the following relations:*

$$[\delta_\alpha, \mathbf{D}_\alpha] = \text{Id}. \quad (2.16)$$

$$[\mathcal{L}_\alpha, \mathbf{D}_\alpha] = \mathbf{D}_\alpha. \quad (2.17)$$

$$[\delta_\alpha, \mathcal{L}_\alpha] = \delta_\alpha. \quad (2.18)$$

Furthermore, \mathcal{L}_α , \mathbf{D}_α and δ_α satisfy the Jacobi identity:

$$[\mathcal{L}_\alpha, [\mathbf{D}_\alpha, \delta_\alpha]] + [\mathbf{D}_\alpha, [\delta_\alpha, \mathcal{L}_\alpha]] + [\delta_\alpha, [\mathcal{L}_\alpha, \mathbf{D}_\alpha]] = 0. \quad (2.19)$$

Proof. Once more we rely on the simpler expression of the generator (2.9). Notice we can ignore the identity part in $\delta_\alpha = \alpha^{-1}x^{\alpha+1}\nabla + \text{Id}$ since it commutes with \mathbf{D}_α . Let $f \in \mathcal{S}_\Phi$ and $x \in \mathbb{R}_+^*$.

$$\begin{aligned}\alpha[\delta_\alpha, \mathbf{D}_\alpha]f(x) &= [x^{\alpha+1}\nabla, \mathbf{D}_\alpha]f(x) \\ &= x^{\alpha+1}\nabla(\mathbf{D}_\alpha f)(x) - \mathbf{D}_\alpha(x^{\alpha+1}f')(x) \\ &= x^{\alpha+1}(-x^{-\alpha}f'(x)) - \int_x^\infty (r^{\alpha+1}f'(r))' \frac{dr}{r^\alpha} \\ &= -xf'(x) + xf'(x) - \alpha \int_x^\infty f'(r) dr \\ &= \alpha f(x).\end{aligned}$$

The proof of the second identity is rather similar:

$$\begin{aligned}[\mathcal{L}_\alpha, \mathbf{D}_\alpha]f(x) &= -\frac{1}{\alpha}[x\nabla, \mathbf{D}_\alpha]f(x) \\ &= -\frac{1}{\alpha}x(\mathbf{D}_\alpha f)'(x) + \frac{1}{\alpha}\mathbf{D}_\alpha(xf')(x) \\ &= -\frac{1}{\alpha}x^{-\alpha} + \frac{1}{\alpha} \int_x^\infty (rf'(r))' \frac{dr}{r^\alpha} \\ &= \int_x^\infty f'(r) \frac{dr}{r^\alpha}.\end{aligned}$$

The final identity is not much harder to prove thanks to the first relation:

$$\begin{aligned}
\alpha[\delta_\alpha, \mathcal{L}_\alpha]f(x) &= -\frac{1}{\alpha}[x^{\alpha+1}\nabla, x\nabla]f(x) + [x^{\alpha+1}\nabla, \mathbf{D}_\alpha]f(x) \\
&= \alpha f(x) + \frac{1}{\alpha}[x\nabla, x^{\alpha+1}\nabla]f(x) \\
&= \alpha f(x) + \frac{1}{\alpha}\left(x(x^{\alpha+1}f'(x))' - x^{\alpha+1}(xf'(x))'\right) \\
&= \alpha f(x) + x^{\alpha+1}f'(x).
\end{aligned}$$

Finally, we have

$$[\mathcal{L}_\alpha, [\mathbf{D}_\alpha, \delta_\alpha]] + [\mathbf{D}_\alpha, [\delta_\alpha, \mathcal{L}_\alpha]] + [\delta_\alpha, [\mathcal{L}_\alpha, \mathbf{D}_\alpha]] = [\mathbf{D}_\alpha, \delta_\alpha] + [\delta_\alpha, \mathbf{D}_\alpha] = 0.$$

□

The Jacobi identity above is part of the definition of a *Lie algebra*. The next definitions are taken from Hall [2013] as well as from Faraut [2008]. We say that a \mathbb{R} -vector space \mathfrak{g} with a bilinear map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ is a (real) *Lie algebra* if it satisfies the following properties:

1. (*Anti-symmetry*) : $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$
2. (*Jacobi identity*) : $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in \mathfrak{g}$.

In that case, the set $[\mathfrak{g}, \mathfrak{g}] := \{[x, y], x, y \in \mathfrak{g}\}$ equipped with $[\cdot, \cdot]$ is a Lie algebra as well. Set $\mathfrak{g}_0 := \mathfrak{g}$ and

$$\mathcal{D}^{k+1}(\mathfrak{g}) := [\mathcal{D}^k(\mathfrak{g}), \mathcal{D}^k(\mathfrak{g})], \quad k \in \mathbb{N}.$$

We call a Lie algebra is *solvable* if there exists some $k \in \mathbb{N}$ such that $\mathcal{D}^k(\mathfrak{g}) = \{0\}$. It is easy to check that the vector space spanned by \mathcal{L}_α , \mathbf{D}_α and δ_α with respect to linear combinations and equipped with the commutator is a Lie algebra. Actually, it is even solvable.

Proposition 2.1.19 *The vector space $\mathfrak{g}_\alpha := \text{span}(\mathcal{L}_\alpha, \mathbf{D}_\alpha, \delta_\alpha, \text{Id})$ equipped with the commutator $[\cdot, \cdot]$ is a solvable Lie algebra.*

Proof. We have already proved that \mathfrak{g}_α is a Lie algebra thanks to equality (2.19). The fact it is solvable comes from noticing that $\mathcal{D}^1(\mathfrak{g}_\alpha) = [\mathfrak{g}_\alpha, \mathfrak{g}_\alpha] = \text{span}(\mathbf{D}_\alpha, \delta_\alpha)$, so that $\mathcal{D}^3(\mathfrak{g}_\alpha) = \{0\}$, thanks to identities (2.16), (2.17) and (2.18). □

Lie algebras have been thoroughly classified. Where does \mathfrak{g}_α sit in this classification? In MacCallum [1999], a complete classification of 4 dimensional Lie algebras is exposed, with applications to cosmology in view. However \mathfrak{g}_α does not appear explicitly in it, because the commutation relations we have just given do not match with any of the ones presented in that book. Nonetheless, if we set:

$$\begin{aligned}
X_1 &:= \mathcal{L}_\alpha \\
X_2 &:= -\mathbf{D}_\alpha + \delta_\alpha \\
X_3 &:= \mathbf{D}_\alpha + \delta_\alpha \\
X_4 &:= -2\text{Id},
\end{aligned}$$

then the previous commutations become

$$\begin{aligned} [X_1, X_2] &= -X_3 \\ [X_1, X_3] &= -X_2 \\ [X_2, X_3] &= X_4 \\ [X_i, X_4] &= 0, \quad i \in \llbracket 1, 3 \rrbracket. \end{aligned}$$

This matches the class U310 defined in MacCallum [1999] (p. 307), implying that \mathfrak{g}_α is isomorphic to that Lie algebra. Notice also that if we restrict ourselves to X_2 , X_3 and X_4 , we get the commutation relations characteristic of the *Heisenberg algebra*, so that \mathfrak{g}_α contains a subalgebra isomorphic to it. The implications of those results, if any, remain to study.

Now, let us sum up the parallels we have drawn between the Ornstein-Uhlenbeck semi-group and its max-stable counterpart: the usual derivative ∇ is replaced by \mathbf{D}_α , while δ changes for δ_α . The generator of the OU semi-group \mathcal{L}_γ satisfies $\mathcal{L}_\gamma = \delta \circ \nabla$ which turns into (2.13) in the Fréchet case. The commutation relation $\nabla P_t = e^{-t} P_t \nabla$ becomes the equality (2.11). The divergence δ is the adjoint of the derivative operator w.r.t. the scalar product induced by the normal distribution, and δ_α is the adjoint of $-x^{\alpha+1} \nabla$ w.r.t. the scalar product associated with the Fréchet distribution. Finally, it is well-known that $[\delta, \nabla] = \text{Id}$, an equality satisfied by δ_α and \mathbf{D}_α as well. The relation $[\mathcal{L}_\gamma, \nabla] = \nabla$ also holds true in the Fréchet case, this time as $[\mathcal{L}_\alpha, \mathbf{D}_\alpha] = \mathbf{D}_\alpha$. Notice that the vector space spanned by \mathcal{L}_γ , ∇ , δ and Id is also a solvable Lie algebra, thanks to the exact same arguments we used for \mathfrak{g}_α .

We continue our tour of the properties of $(\mathbf{P}_t^\alpha)_{t \geq 0}$ with several functional inequalities. Before stating them, we give two covariance identity which will play an important role in the sequel. The proofs will be given in the next subsection in the Gumbel case, where they take a simpler form.

Proposition 2.1.20 — Covariance identities - Fréchet case. *Let $f, g \in \mathcal{S}_\Phi$ and $Z \sim \mathcal{F}(\alpha)$.*

- *Let $Y \sim \mathcal{VP}(\alpha, 1)$ be a random variable with Pareto distribution, independent of Z . Then:*

$$\langle \mathcal{L}_\alpha f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)} = \mathbb{E}[\mathcal{L}_\alpha f(Z)g(Z)] = -\frac{1}{\alpha^2} \mathbb{E}[YZ^2 f'(YZ)g'(Z)]. \quad (2.20)$$

- *Assume further that f has zero mean: $\mathbb{E}[f(Z)] = 0$. Then:*

$$\langle f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)} = -\frac{1}{\alpha^2} \mathbb{E}[YZ^2 (\mathcal{L}_\alpha^{-1} f)'(YZ)g'(Z)]. \quad (2.21)$$

Recall the energy form $\mathcal{E}_{\Phi_\alpha}$ defined as

$$\mathcal{E}_{\Phi_\alpha}(f) = -\mathbb{E}_{\Phi_\alpha}[f \mathcal{L}_\alpha f] = -\mathbb{E}[f(Z) \mathcal{L}_\alpha f(Z)].$$

Theorem 2.1.21 — Poincaré inequality - Fréchet case. *Let $f \in \mathcal{S}_\Phi$. Let $Z \sim \mathcal{F}(\alpha)$ and $Y \sim \mathcal{VP}(\alpha, 1)$ be a random variable with Pareto distribution, independent of Z .*

$$\mathbb{V}(f(Z)) \leq 2\mathcal{E}_{\Phi_\alpha}(f) = \frac{2}{\alpha^2} \mathbb{E}[YZ^2 f'(YZ)g'(Z)]. \quad (2.22)$$

Thanks to proposition 1.1.12, this inequality implies the exponential convergence of $\mathbf{P}_t^\alpha f$ to $\mathbb{E}_{\Phi_\alpha}[f]$ in \mathbf{L}^2 -norm at exponential speed.

Corollary 2.1.22 Let $f \in \mathcal{S}_\Phi$ and Φ_α denote the probability measure of $\mathcal{F}(\alpha)$. We have:

$$\|\mathbf{P}_t^\alpha f - \mathbb{E}_{\Phi_\alpha}[f]\|_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)} \leq e^{-\frac{t}{2}} \|f\|_{\mathbf{L}^2(\mathbb{R}_+^*, \Phi_\alpha)}.$$

2.1.2 Extension to the Gumbel and Weibull distributions

The definition of $(\mathbf{P}_t^\alpha)_{t \geq 0}$ can be adapted easily to the two other extreme value distributions, *i.e.* $\alpha \in (-\infty, 0]$. We will see we can connect them through simple transformations, allowing us to transfer all the previously proved properties in the Fréchet case to the two remaining ones.

The Gumbel case

A standard algebra adapted to the Gumbel distribution is the following.

Definition 2.1.23 (Standard algebra - Gumbel case). We define the class of functions:

$$\mathcal{S}_\Lambda := \{f(e^x), f \in \mathcal{S}_\Phi\} = \{f \in \mathcal{S}(\mathbb{R}), \sup_{x \in \mathbb{R}_-} e^{-kx} |f^{(n)}(e^x)| < \infty \text{ for all } (k, n) \in \mathbb{N} \times \mathbb{N}^*\}.$$

Just as in the previous subsection, the max-stability relation satisfied by the Gumbel distribution yields another semi-group.

Definition 2.1.24 (Gumbel semi-group). Define the *Gumbel semi-group* $(\mathbf{P}_t^0)_{t \geq 0}$ on the standard algebra $\mathcal{A} = \mathcal{S}_\Lambda$.

$$\mathbf{P}_t^0 f(x) := \mathbb{E}\left[f\left((x-t) \oplus (Z + \log(1 - e^{-t}))\right)\right], \quad x \in [-\infty, +\infty), \quad t \geq 0 \quad (2.23)$$

where $Z \sim \mathcal{G}(0, 1)$ and f belongs to \mathcal{S}_Λ .

Instead of proving again that (2.23) defines a semi-group, we instead show it is actually the same semi-group as the one defined previously, up to a transformation of the argument.

Lemma 2.1.25 For $f \in \mathcal{S}_\Phi$, we have

$$\mathbf{P}_t^0 f(x) = \mathbf{P}_t^\alpha (f \circ \log x^\alpha)(e^{\frac{x}{\alpha}}), \quad x \in \mathbb{R}, \quad (2.24)$$

which can also be written as the following intertwining relation

$$\mathbf{P}_t^0 f(\log x^\alpha) = \mathbf{P}_t^\alpha (f \circ \log x^\alpha)(x), \quad x \in \mathbb{R}_+^*.$$

Proof. Because $Z' = \alpha \log Z$ has the Gumbel $\mathcal{G}(0, 1)$ distribution if Z has the Fréchet distribution $\mathcal{F}(\alpha)$, and since $z \mapsto \alpha \log z$ is non-decreasing on \mathbb{R}_+^* , we have:

$$\begin{aligned} \mathbf{P}_t^\alpha (f \circ \log x^\alpha)(e^{\frac{x}{\alpha}}) &= \mathbb{E}\left[(f \circ \log x^\alpha)\left(e^{-\frac{t}{\alpha}} e^{\frac{x}{\alpha}} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} Z\right)\right] \\ &= \mathbb{E}\left[f\left((x-t) \oplus (Z' + \log(1 - e^{-t}))\right)\right]. \end{aligned}$$

The last equality stems from the fact that $g(x \oplus y) = g(x) \oplus g(y)$ if g is a non-decreasing function. \square

Due to this result, by taking $\alpha = 1$, we see why \mathcal{S}_Λ is an appropriate candidate to be a standard algebra for $(\mathbf{P}_t^0)_{t \geq 0}$. Besides, it is not hard to see that this class is an algebra of functions which contains compactly supported, infinitely differentiable functions on \mathbb{R} , as well as functions such as $x \mapsto \exp(\cosh(x))$ defined on \mathbb{R} . Finally, because the functions $z \mapsto \alpha \log z$ and $z \mapsto \exp(z/\alpha)$ are inverse to each other, it is clear from (2.24) that $(\mathbf{P}_t^0)_{t \geq 0}$ is a semi-group. We also get the other defining properties of a Markov semi-group, such as right-continuity, by using this relation and the definition of \mathcal{S}_Φ . This gives us the next results.

Proposition 2.1.26 *Let Λ denote the Gumbel probability measure $\mathcal{G}(0, 1)$. The semi-group $(\mathbf{P}_t^0)_{t \geq 0}$ admits Λ as a stationary measure. As such, $(\mathbf{P}_t^0)_{t \geq 0}$ is a Markov semi-group. Furthermore, $(\mathbf{P}_t^0)_{t \geq 0}$ is ergodic:*

$$\mathbf{P}_t^0 f(x) \xrightarrow{t \rightarrow \infty} \mathbb{E}_\Lambda[f] = \int_{\mathbb{R}} f(r) d\Lambda(r), \quad x \in \mathbb{R}, \quad f \in \mathcal{S}_\Lambda.$$

Proposition 2.1.27 *For every $t \geq 0$ and every $p \in [1, \infty)$, the operator \mathbf{P}_t^0 extends to a linear contraction operator on $\mathbf{L}^p(\mathbb{R}, \Lambda)$.*

The generator of $(\mathbf{P}_t^0)_{t \geq 0}$ can be deduced from the generator \mathcal{L}_1 of $(\mathbf{P}_t^1)_{t \geq 0}$ by using relation (2.24). The computation itself is given in the Weibull case for ease of notations.

Proposition 2.1.28 *The Markov semi-group $(\mathbf{P}_t^0)_{t \geq 0}$ has generator \mathcal{L}_0 :*

$$\mathcal{L}_0 f(x) = -f'(x) + \int_{\mathbb{R}} (f(x \oplus r) - f(x)) d\rho_0(r),$$

where $d\rho_0(r) := e^{-r} dr$ on \mathbb{R} .

The pseudo Leibniz rule takes formally the same form as in the Fréchet case.

Proposition 2.1.29 — Pseudo Leibniz rule - Gumbel case. *We have for $f, g \in \mathcal{S}_\Lambda$:*

$$\begin{aligned} \mathbf{D}_0(fg)(x) &= \mathbf{D}_0 f(x)g(x) + f(x)\mathbf{D}_0 g(x) \\ &\quad + \int_x^\infty (f(r) - f(x))(g(r) - g(x)) d\rho_0(r), \quad x \in \mathbb{R}. \end{aligned}$$

This is also equal to:

$$\begin{aligned} \mathbf{D}_0(fg)(x) &= \mathbf{D}_0 f(x)g(x) + f(x)\mathbf{D}_0 g(x) \\ &\quad + \int_x^\infty \mathbf{D}_0 f(y)g'(y) dy + \int_x^\infty f'(y)\mathbf{D}_0 g(y) dy, \quad x \in \mathbb{R}. \end{aligned}$$

We summarize the counterparts of proposition 2.1.9 and 2.1.11 for the Gumbel case in the next proposition.

Proposition 2.1.30 *Let $f \in \mathcal{S}_\Lambda$ and $x \in \mathbb{R}$. Let $Y \sim \mathcal{E}(1)$.*

1. *The derivative of $\mathbf{P}_t^0 f$ satisfies:*

$$(\mathbf{P}_t^0 f)'(x) = e^{-\gamma t e^{-x}} f'(x - t). \quad (2.25)$$

2. $\mathbf{P}_t^0 f$ can be rewritten as:

$$\mathbf{P}_t^0 f(x) = - \int_x^\infty e^{-\gamma t e^{-r}} f'(r-t) \, dr. \quad (2.26)$$

3. The generator of the Gumbel semi-group is connected to the exponential distribution with scale parameter 1:

$$\mathcal{L}_0 f(x) = -f'(x) + e^{-x} \int_0^\infty (f(x+r) - f(x)) \, d\rho_0(r) \quad (2.27)$$

$$= -f'(x) + e^{-x} \mathbb{E}[f(x+Y) - f(x)]. \quad (2.28)$$

4. $\mathcal{L}_0 f$ can be rewritten as:

$$\mathcal{L}_0 f(x) = -f'(x) + \int_x^\infty f'(r) \, d\rho_0(r) \quad (2.29)$$

$$= -f'(x) + e^{-x} \mathbb{E}[f'(x+Y)] \quad (2.30)$$

As before, we obtain that \mathcal{S}_Λ is a standard algebra for $(\mathbf{P}_t^0)_t$. Furthermore, the commutation rule between \mathbf{D}_0 and \mathbf{P}_t^0 holds here too.

Proposition 2.1.31 — **Commutation relation for \mathbf{D}_0 - Gumbel case.** *We have for all $f \in \mathcal{S}_\Lambda$*

$$\mathbf{D}_\alpha \mathbf{P}_t^0 f(x) = e^{-t} \mathbf{P}_t^0 \mathbf{D}_0 f(x), \quad x \in \mathbb{R}, \quad t \geq 0. \quad (2.31)$$

We also have an inequality for the \mathbf{L}^p -norm of the gradient of $\mathbf{P}_t^0 f$.

Proposition 2.1.32 *Assume f is in \mathcal{S}_Λ . Then, for every $p \geq 1$,*

$$\|\nabla \mathbf{P}_t^0 f\|_{\mathbf{L}^p(\mathbb{R}, \Lambda)} \leq e^{-\frac{t}{p}} \|\nabla f\|_{\mathbf{L}^p(\mathbb{R}, \Lambda)}, \quad t \geq 0.$$

The equivalent of (2.12) and of (2.13) in the Gumbel case is

$$\mathcal{L}_0 f(x) = -f'(x) + \int_x^\infty f'(r) \, d\rho_0(r) = (\delta_0 \circ \mathbf{D}_0) f(x), \quad (2.32)$$

where the operator δ_0 is equal to

$$\delta_0 f(x) := (e^x \nabla + \text{Id}) f(x) = e^x f'(x) + f(x), \quad f \in \mathcal{S}_\Lambda.$$

Once more, this defines a Stein operator and it is the one used in [Kusumoto and Takeuchi \[2020\]](#) to treat limit theorems involving the Gumbel case, where it is denoted by \mathcal{S}_0 .

Proposition 2.1.33 — **Integration-by-parts formula - Gumbel case.** *Let f, g belong to \mathcal{S}_Λ . Then we have*

$$\langle \delta_0 f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} = - \langle f, e^r \nabla g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} \quad (2.33)$$

Proof. The proof follows from the usual integration-by-parts formula. □

Just like \mathbf{D}_0 , this divergence operator satisfies a commutation rule with \mathbf{P}_t^0 :

Proposition 2.1.34 — Commutation relation for δ_0 - Gumbel case. *We have for all $f \in \mathcal{S}_\Lambda$*

$$\delta_0 \mathbf{P}_t^0 f(x) = e^t \mathbf{P}_t^0 \delta_0 f(x), \quad x \in \mathbb{R}, \quad t \geq 0. \quad (2.34)$$

The commutation identities between δ_0 and \mathbf{D}_0 , and \mathcal{L}_0 and \mathbf{D}_0 , are formally the same as in the Gumbel case and we recover the solvable Lie algebra structure we found in the previous section. This suggests us that some connection exists between \mathcal{L}_α , \mathbf{D}_α and δ_α on the one hand, and \mathcal{L}_0 , \mathbf{D}_0 and δ_0 on the other. The last subsection of this chapter give a clear statement of this intuition.

Proposition 2.1.35 — Commutator identities - Gumbel case. *We have the following relations:*

$$[\delta_0, \mathbf{D}_0] = \text{Id}. \quad (2.35)$$

$$[\mathcal{L}_0, \mathbf{D}_0] = \mathbf{D}_0. \quad (2.36)$$

$$[\delta_0, \mathcal{L}_0] = \delta_0. \quad (2.37)$$

Furthermore, \mathcal{L}_0 , \mathbf{D}_0 and δ_0 satisfy the Jacobi identity:

$$[\mathcal{L}_0, [\mathbf{D}_0, \delta_0]] + [\mathbf{D}_0, [\delta_0, \mathcal{L}_0]] + [\delta_0, [\mathcal{L}_0, \mathbf{D}_0]] = 0. \quad (2.38)$$

In particular, the vector space \mathfrak{g}_0 spanned by those three operators and the identity with respect to linear combinations is a solvable Lie algebra.

Now we give the proofs of the covariance identities and functional inequalities stated at the end of the section about the quantization of the Fréchet distribution in the Gumbel case.

Proposition 2.1.36 — Covariance identities - Gumbel. *Let $f, g \in \mathcal{S}_\Lambda$ and $Z \sim \mathcal{G}(0, 1)$.*

- *Let $Y \sim \mathcal{E}(1)$ be a random variable with exponential distribution, independent of Z . Then:*

$$\langle \mathcal{L}_0 f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} = -\mathbb{E}[f'(Z + Y)g'(Z)]. \quad (2.39)$$

- *Assume further that f has zero mean: $\mathbb{E}[f(Z)] = 0$. Then:*

$$\langle f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} = -\mathbb{E}[(\mathcal{L}_0^{-1} f)'(Z + Y)g'(Z)]. \quad (2.40)$$

Proof. Integrating the density of the Gumbel distribution and differentiating the rest in the second

term below, one finds:

$$\begin{aligned}
\langle \mathcal{L}_0 f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} &= -\langle \nabla f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} + \langle \mathbf{D}_0 f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} \\
&= -\langle \nabla f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} + \int_{\mathbb{R}} \left(\int_x^\infty e^{-y} f'(y) \, dy \right) g(x) e^{-(x+e^{-x})} \, dx \\
&= -\langle \nabla f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} + \langle \nabla f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \Lambda)} - \int_{\mathbb{R}} \left(\int_x^\infty e^{-y} f'(y) \, dy \right) g'(x) e^{-e^{-x}} \, dx \\
&= - \int_{\mathbb{R}} \left(\int_0^\infty e^{-(y+x)} f'(y+x) \, dy \right) g'(x) e^{-e^{-x}} \, dx \\
&= - \int_{\mathbb{R}} \left(\int_0^\infty e^{-y} f'(y+x) \, dy \right) g'(x) e^{-(x+e^{-x})} \, dx \\
&= -\mathbb{E}[f'(Z+Y)g'(X)].
\end{aligned}$$

The second equality is a direct consequence of the first, by replacing f by $\mathcal{L}_0^{-1}f$. \square

Remark 2.1.37. In the case of the usual Ornstein-Uhlenbeck semi-group, (2.39) instead becomes:

$$\langle \mathcal{L}_\gamma f, g \rangle_{\mathbf{L}^2(\mathbb{R}, \gamma)} = -\mathbb{E}[f'(Z)g'(Z)], \quad f, g \in \mathcal{S}(\mathbb{R})$$

where $Z \sim \mathcal{N}(0, 1)$. Besides, equation (2.39) makes obvious that \mathcal{L}_0 is not symmetric. This is true as well for the other max-stable semi-groups \blacksquare

The energy form \mathcal{E}_Λ equals

$$\mathcal{E}_\Lambda(f) = -\mathbb{E}_\Lambda[f \mathcal{L}_0 f] = -\mathbb{E}[f(Z) \mathcal{L}_0 f(Z)].$$

Theorem 2.1.38 — Poincaré inequality - Gumbel case. Let $f \in \mathcal{S}_\Lambda$. Let $Z \sim \mathcal{G}(0, 1)$ and $Y \sim \mathcal{E}(1)$ be a random variable with exponential distribution, independent of Z .

$$\mathbb{V}(f(Z)) \leq 2\mathbb{E}[f'(Z)f'(Z+Y)]. \quad (2.41)$$

Proof. Define

$$\beta(s) := \mathbf{P}_s^0(\mathbf{P}_{t-s}^0 f)^2, \quad s \in [0, t]$$

and notice that $\beta(t) - \beta(0) = \mathbf{P}_t^0(f^2) - (\mathbf{P}_t^0(f))^2$. Furthermore, set $g := \mathbf{P}_{t-s}^0 f$. Then we have

$$\beta'(s) = \mathbf{P}_s^0(\mathcal{L}_0(g^2) - 2g\mathcal{L}_0 g) = \mathbf{P}_s^0(\Gamma(g)).$$

Set $\Gamma_\Lambda(g) := (\mathcal{L}_0(g^2) - 2g\mathcal{L}_0 g)/2$, the carré du champ operator, and define $G(x) := \mathbf{D}_0 g(x)$. Let $x \in \mathbb{R}$. The carré du champ operator equals:

$$2\Gamma_\Lambda(g)(x) = -2g(x)g'(x) + 2 \int_x^\infty e^{-y} g'(y)g(y) \, dy + 2g(x)g'(x) - 2g(x) \int_x^\infty e^{-y} g'(y) \, dy.$$

Then we get

$$\begin{aligned}
2\Gamma_\Lambda(g)(x) &= 2 \int_x^\infty e^{-y} g'(y) \int_x^y g'(r) \, dr \, dy \\
&= 2 \int_x^\infty g'(r) \int_r^\infty e^{-y} g'(y) \, dy \, dr \\
&= -2 \int_x^\infty e^r G'(r)G(r) \, dr \\
&= e^x(G(x))^2 + \int_x^\infty e^r (G(r))^2 \, dr.
\end{aligned}$$

Since $G(x) = \mathbf{D}_0 g(x)$, and since $g = \mathbf{P}_{t-s}^0 f$, by the commutation rule the previous equality becomes:

$$\begin{aligned} 2\Gamma_\Lambda(g)(x) &= e^x (G(x))^2 + \int_x^\infty e^r (G(r))^2 dr \\ &= e^x (\mathbf{D}_0 \mathbf{P}_{t-s}^0 f(x))^2 + \int_x^\infty e^r (\mathbf{D}_0 \mathbf{P}_{t-s}^0 f(r))^2 dr \\ &= e^{-2(t-s)} [e^x (\mathbf{P}_{t-s}^0 \mathbf{D}_0 f(x))^2 + \int_x^\infty e^r (\mathbf{P}_{t-s}^0 \mathbf{D}_0 f(r))^2 dr] \\ &\leq e^{-2(t-s)} [e^x \mathbf{P}_{t-s}^0 (\mathbf{D}_0 f)^2(x) + \int_x^\infty e^r \mathbf{P}_{t-s}^0 (\mathbf{D}_0 f)^2(r) dr]. \end{aligned}$$

We have used Cauchy-Schwartz inequality to get the last inequality. Once more integrating by parts (integrating the exponential and differentiating the semi-group), we find:

$$\begin{aligned} 2\Gamma_\Lambda(g)(x) &\leq e^{-2(t-s)} [e^x \mathbf{P}_{t-s}^0 (\mathbf{D}_0 f)^2(x) + \int_x^\infty e^r \mathbf{P}_{t-s}^0 (\mathbf{D}_0 f)^2(r) dr] \\ &= -e^{-2(t-s)} \int_x^\infty e^r (\mathbf{P}_{t-s}^0 (\mathbf{D}_0 f)^2)'(r) dr \end{aligned}$$

Recall that we have:

$$\mathbf{P}_s^0 f(x) = - \int_x^\infty e^{-\gamma_s e^{-r}} f'(r-s) dr.$$

Evaluating the first equality in $\mathcal{L}_0(g^2) - 2g\mathcal{L}_0(g)$, we get the following expression for $\beta'(s)$:

$$\begin{aligned} \beta'(s) &= 2\mathbf{P}_s^0 \Gamma_\Lambda(g)(x) \\ &\leq -e^{-2(t-s)} \int_x^\infty e^{-\gamma_s e^{-r}} e^{r-s} (\mathbf{P}_{t-s}^0 (\mathbf{D}_0 f)^2)'(r-s) dr \\ &= -e^{-2(t-s)} \int_x^\infty e^{-\gamma_s e^{-r}} e^{r-s} e^{-\gamma_{t-s} e^{-(r-s)}} ((\mathbf{D}_0 f)^2)'(r-s-t+s) dr \\ &= -e^{-2(t-s)} \int_x^\infty e^{-\gamma_s e^{-r}} e^{r-s} e^{-(e^t - e^s) e^{-r}} ((\mathbf{D}_0 f)^2)'(r-t) dr \\ &= -e^{-2(t-s)} \int_x^\infty e^{-\gamma_t e^{-r}} e^{r-s} ((\mathbf{D}_0 f)^2)'(r-t) dr \\ &= -e^{-2(t-s)} \int_{x-t}^\infty e^{-(1-e^{-t})e^{-r}} e^{r+t-s} ((\mathbf{D}_0 f)^2)'(r) dr \\ &= -e^{-(t-s)} \int_{x-t}^\infty e^r e^{-(1-e^{-t})e^{-r}} ((\mathbf{D}_0 f)^2)'(r) dr. \end{aligned}$$

Therefore, we have the inequality:

$$\mathbf{P}_t^0(f^2) - (\mathbf{P}_t^0(f))^2 = \beta(t) - \beta(0) \leq -(1-e^{-t}) \int_{-t}^\infty e^r e^{-(1-e^{-t})e^{-r}} ((\mathbf{D}_0 f)^2)'(r) dr, \quad t \geq 0. \quad (2.42)$$

We can dominate the expression inside the integral by $e^r |((\mathbf{D}_0 f)^2)'(r)|$ when (say) r is positive (recall that $f \in \mathcal{S}_\Phi$), while we bound it by $r \mapsto C e^r$ for negative r . Here C is equal to:

$$C = \|e^{-r} \nabla f\|_\infty \int_{\mathbb{R}} e^{-y} |f'(y)| dy.$$

Finally, by the dominating convergence theorem, we can let t goes to infinity in (2.42) to get the announced inequality. □

Corollary 2.1.39 *Let $f \in \mathcal{S}_\Lambda$ and Λ the probability distribution $\mathcal{G}(0, 1)$. We have:*

$$\|\mathbf{P}_t^0 f - \mathbb{E}_\Lambda[f]\|_{\mathbf{L}^2(\mathbb{R}, \Lambda)} \leq e^{-\frac{t}{2}} \|f\|_{\mathbf{L}^2(\mathbb{R}, \Lambda)}.$$

Thanks to $(\mathbf{P}_t^0)_{t \geq 0}$, we can prove a concentration inequality for functionals of a Gumbel random variable.

Proposition 2.1.40 *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded \mathcal{C}^1 -function with bounded derivative, and such that f' is integrable with respect to Lebesgue measure on \mathbb{R} . Let Z be a random variable with Gumbel distribution. Then we have:*

$$\mathbb{P}(f(Z) - \mathbb{E}[f(Z)] \geq x) \leq e^{-\frac{x^2}{4\|f\|_\infty \|f'\|_\infty}}, x > 0.$$

Proof. Replacing f by $f - \mathbb{E}[f(Z)]$, we can assume that f is centered with respect to Λ . We use the so-called *Herbst argument* (see for instance [Decreusefond \[2022\]](#)), that is, we bound the logarithmic derivative of $\varphi(\theta) := \mathbb{E}[e^{\theta f(Z)}]$. We start from

$$\varphi'(\theta) = \mathbb{E}[f(Z)e^{\theta f(Z)}]$$

and apply the covariance identity (2.40) to f and $e^{\theta f}$. This yields:

$$\varphi'(\theta) = -\mathbb{E}[(\mathcal{L}_0^{-1} f)'(Z + Y)\theta e^{\theta f(Z)}] = -\theta \mathbb{E}[(\mathcal{L}_0^{-1} f)'(Z + Y)f'(Z)e^{\theta f(Z)}].$$

Next we need to control the terms before $e^{\theta f(Z)}$. We cannot use Hölder's inequality and the likes, because we would then be unable to make $\mathbb{E}[e^{\theta f(Z)}] = \varphi(\theta)$ appear on the right-hand side. This is why we need such stringent assumptions on f . Using (2.25), we have to bound:

$$\begin{aligned} (\mathcal{L}_0^{-1} f)'(x) &= \left(\int_0^\infty \int_x^\infty f'(y-t)e^{-\gamma_t e^{-y}} dy dt \right)' \\ &= - \int_0^\infty f'(x-t)e^{-\gamma_t e^{-x}} dt \\ &= -f(x) + \int_0^\infty f(x-t)e^t e^{-(x+\gamma_t e^{-x})} dt \\ &= -f(x) + \int_0^\infty f(x - \log(1+s))e^{-(x+se^{-x})} ds \\ &= -f(x) + \int_0^\infty f(x - \log(1+te^x))e^{-t} dt \\ &\leq 2\|f\|_\infty. \end{aligned}$$

The third equality stems from an integration by parts (integrating $t \mapsto -u'(y-t)$) while the fourth comes from the change of variable $s = \gamma_t = e^t - 1$. The fifth is just the change of variable $t = e^{-x}s$. Since f' is also assumed to be bounded, we find:

$$\varphi'(\theta) \leq 2\|f\|_\infty \|f'\|_\infty \theta \varphi(\theta),$$

and by integrating, we get:

$$\mathbb{E}[e^{\theta f(Z)}] = \varphi(\theta) \leq e^{\|f\|_\infty \|f'\|_\infty \theta^2}.$$

Applying Markov's inequality, this yields:

$$\mathbb{P}(f(Z) - \mathbb{E}[f(Z)] \geq x) \leq e^{\|f\|_\infty \|f'\|_\infty \theta^2 - x\theta}, x > 0$$

and by optimizing in θ , we find the announced result. \square

The Weibull case

We make the choice to deal with the positive Weibull distribution instead of its negative, max-stable counterpart. This brings several advantages such as simpler notations and results expressed directly for the standard exponential distribution. It also allows us to show concretely what changes when working with min-stable distributions rather than max-stable distributions.

As before, the natural choice of standard algebra for the Weibull distribution is dictated by the transformation which sends the Fréchet distribution to the Weibull distribution, *i.e.* $x \mapsto -x^{-1}$ for the negative max-stable Weibull distribution and $x \mapsto x^{-1}$ for the positive, min-stable Weibull distribution.

Definition 2.1.41 (Standard algebra - Weibull case). We define the classes of functions:

$$\mathcal{S}_\Psi := \{f(-x^{-1}), f \in \mathcal{S}_\Phi\} = \{f \in \mathcal{S}(\mathbb{R}_+^*), \sup_{x \in [1, +\infty)} x^k |f^{(n)}(-x^{-1})| < \infty \text{ for all } (k, n) \in \mathbb{N} \times \mathbb{N}^*\}.$$

$$\mathcal{S}_{\Psi,+} := \{f(x^{-1}), f \in \mathcal{S}_\Phi\} = \{f \in \mathcal{S}(\mathbb{R}_+^*), \sup_{x \in [1, +\infty)} x^k |f^{(n)}(x^{-1})| < \infty \text{ for all } (k, n) \in \mathbb{N} \times \mathbb{N}^*\}.$$

The max and min-stability relations satisfied by the negative and positive Weibull distribution gives two Mehler's formulas.

Definition 2.1.42 (Positive and negative Weibull semi-groups). Let $\alpha \in \mathbb{R}_+^*$.

- Define the *Weibull semi-group* $(\mathbf{P}_t^{-\alpha})_{t \geq 0}$ on the standard algebra $\mathcal{A} = \mathcal{S}_\Psi$ by

$$\mathbf{P}_t^{-\alpha} f(x) := \mathbb{E}[f(e^{\frac{t}{\alpha}} x \oplus (1 - e^{-t})^{-\frac{1}{\alpha}} Z)], \quad x \leq 0, \quad t \geq 0 \quad (2.43)$$

where $Z \sim \mathcal{W}(\alpha)$ and f belongs to \mathcal{S}_Ψ .

- Define the *positive Weibull semi-group* $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$ on the standard algebra $\mathcal{A} = \mathcal{S}_{\Psi,+}$ by

$$\mathbf{P}_t^{-\alpha,+} f(x) := \mathbb{E}[f(e^{\frac{t}{\alpha}} x \odot (1 - e^{-t})^{-\frac{1}{\alpha}} Z)], \quad x \geq 0, \quad t \geq 0 \quad (2.44)$$

where $Z \sim \mathcal{W}_+(\alpha)$ and f belongs to $\mathcal{S}_{\Psi,+}$.

Since we have the equality

$$\mathbf{P}_t^{-\alpha} f(-x) = \mathbf{P}_t^{-\alpha,+} (f \circ (-\text{Id}))(x), \quad x \geq 0$$

we will focus on the positive Weibull semi-group only in this subsection.

Lemma 2.1.43 For $f \in \mathcal{S}_\Phi$, we have

$$\mathbf{P}_t^{-\alpha,+} f(x) = \mathbf{P}_t^\alpha (f \circ x^{-1})(x^{-1}), \quad x \in \mathbb{R}_+^*, \quad (2.45)$$

which can also be written as the following intertwining relation

$$\mathbf{P}_t^{-\alpha,+} f(x^{-1}) = \mathbf{P}_t^\alpha (f \circ x^{-1})(x), \quad x \in \mathbb{R}_+^*.$$

The same line of reasoning as in the Gumbel case allows us to transfer the main properties of $(\mathbf{P}_t^\alpha)_{t \geq 0}$ to $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$. We list them now without proof, except when a result is specific to the Weibull case.

Proposition 2.1.44 Let Ψ_α^+ denote the Weibull probability measure $\mathcal{W}_+(\alpha)$. The semi-group $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$ admits Ψ_α^+ as a stationary measure. As such, $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$ is a Markov semi-group. Furthermore, $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$ is ergodic:

$$\mathbf{P}_t^{-\alpha,+} f(x) \xrightarrow{t \rightarrow \infty} \mathbb{E}_{\Psi_\alpha^+}[f] = \int_{\mathbb{R}_+} f(r) \, d\Psi_\alpha^+(r), \quad x \in \mathbb{R}_+, \quad f \in \mathcal{S}_{\Psi,+}.$$

Proposition 2.1.45 For every $t \geq 0$ and every $p \in [1, \infty)$, the operator $\mathbf{P}_t^{-\alpha,+}$ extends to a linear contraction operator on $\mathbf{L}^p(\mathbb{R}_+, \Psi_\alpha^+)$.

Formally speaking the generator takes the same form as in the Fréchet case, except that the dilation semi-group takes a minus sign, the maximum operation \oplus is replaced by the minimum operation \odot and the measure ρ_α becomes $\rho_{-\alpha,+}$, the latter being also defined on \mathbb{R}_+^* .

Proposition 2.1.46 The Markov semi-group $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$ has generator $\mathcal{L}_{-\alpha,+}$:

$$\mathcal{L}_{-\alpha,+} f(x) = \frac{1}{\alpha} x f'(x) + \int_{\mathbb{R}_+} (f(x \odot r) - f(x)) \, d\rho_{-\alpha,+}(r),$$

where $d\rho_{-\alpha,+}(r) := \alpha r^{\alpha-1} \, dr$ on \mathbb{R}_+ .

Proof. Let $f \in \mathcal{S}_{\Psi,+}$ and set $g : x \mapsto f(x^{-1})$. Observe that f belongs to \mathcal{S}_Φ . Set also $y = x^{-1}$. With those notations, (2.45) becomes $\mathbf{P}_t^{-\alpha,+} f(x) = \mathbf{P}_t^\alpha g(y)$. By using the expression of the generator \mathcal{L}_α of $(\mathbf{P}_t^\alpha)_{t \geq 0}$ and applying carefully the chain rule, we find:

$$\begin{aligned} \mathcal{L}_{-\alpha,+} f(x) &= \left. \frac{d}{dt} \right|_{t=0} \mathbf{P}_t^{-\alpha,+} f(x) \\ &= (\mathcal{L}_\alpha g)(y) \\ &= -\frac{1}{\alpha} y g'(y) + \int_0^\infty (g(y \oplus r) - g(y)) \, d\rho_\alpha(r) \\ &= -\frac{1}{\alpha} y (f(y^{-1}))' + \int_0^\infty (f(x \odot r^{-1}) - f(x)) \, d\rho_\alpha(r) \\ &= \frac{1}{\alpha} x f'(x) + \int_0^\infty (f(x \odot r^{-1}) - f(x)) \frac{\alpha}{r^{\alpha+1}} \\ &= \frac{1}{\alpha} x f'(x) + \int_0^\infty (f(x \odot u) - f(x)) \alpha u^{\alpha-1} \, du. \end{aligned}$$

We have used the change of variable $u = r^{-1}$ to get the last line. We recognize the measure $\rho_{-\alpha,+}$ in the integral, which concludes the proof. □

Like before, the pseudo Leibniz rule has the same shape as in the Fréchet and Gumbel cases, although with an appropriate $\mathbf{D}_{-\alpha,+}$.

Proposition 2.1.47 — Pseudo Leibniz rule - Weibull case. For every $\alpha > 0$ and $f, g \in \mathcal{S}_\Phi$, we

have

$$\begin{aligned} \mathbf{D}_\alpha(fg)(x) &= \mathbf{D}_\alpha f(x)g(x) + f(x)\mathbf{D}_\alpha g(x) \\ &\quad + \int_x^\infty (f(r) - f(x))(g(r) - g(x)) \, d\rho_\alpha(r), \quad x > 0. \end{aligned}$$

This is also equal to:

$$\begin{aligned} \mathbf{D}_\alpha(fg)(x) &= \mathbf{D}_\alpha f(x)g(x) + f(x)\mathbf{D}_\alpha g(x) \\ &\quad + \int_x^\infty \mathbf{D}_\alpha f(y)g'(y) \, dy + \int_x^\infty f'(y)\mathbf{D}_\alpha g(y) \, dy, \quad x > 0. \end{aligned}$$

The version of proposition 2.1.9 and 2.1.11 for the Weibull case is given in the next proposition. Before this, let us introduce the beta distribution $B(\alpha, \beta)$ with positive shape parameters α, β . It has density

$$f_{\alpha, \beta}(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbf{1}_{(0,1)}(x),$$

where B is the beta function. The special case $\beta = 1$ gives us the distribution $B(\alpha, 1)$, with density $\alpha x^{\alpha-1} \mathbf{1}_{(0,1)}(x)$. It plays in the Weibull case the same role as the Pareto distribution and exponential distribution for the Fréchet distribution and Gumbel distribution respectively.

Proposition 2.1.48 *Let $f \in \mathcal{S}_{\Psi,+}$ and $x \in \mathbb{R}_+$. Let $Y \sim B(1, \alpha)$.*

1. *The derivative of $\mathbf{P}_t^{-\alpha,+} f$ satisfies:*

$$(\mathbf{P}_t^{-\alpha,+} f)'(x) = e^{\frac{t}{\alpha}} e^{-\gamma t x^\alpha} f'(e^{\frac{t}{\alpha}} x) \quad (2.46)$$

2. *$\mathbf{P}_t^{-\alpha,+} f$ can be rewritten as:*

$$\mathbf{P}_t^{-\alpha,+} f(x) = e^{\frac{t}{\alpha}} \int_0^x e^{-\gamma t r^\alpha} f'(e^{\frac{t}{\alpha}} r) \, dr. \quad (2.47)$$

3. *The generator of the Weibull semi-group is connected to the Beta distribution with parameter α and 1:*

$$\mathcal{L}_{-\alpha,+} f(x) = \frac{1}{\alpha} x f'(x) + x^\alpha \int_0^\infty (f(xr) - f(x)) \, d\rho_{-\alpha,+}(r) \quad (2.48)$$

$$= -\frac{1}{\alpha} x f'(x) + x^\alpha \mathbb{E}[f(xY) - f(x)]. \quad (2.49)$$

4. *$\mathcal{L}_{-\alpha,+} f$ can be rewritten as:*

$$\mathcal{L}_{-\alpha,+} f(x) = \frac{1}{\alpha} x f'(x) - \frac{1}{\alpha} \int_0^x r f'(r) \, d\rho_{-\alpha,+}(r) \quad (2.50)$$

$$= \frac{1}{\alpha} x f'(x) - \frac{1}{\alpha} x^{\alpha+1} \mathbb{E}[Y f'(xY)]. \quad (2.51)$$

Just as in the Fréchet and Gumbel cases, we get that $\mathcal{S}_{\Psi,+}$ is a standard algebra for $(\mathbf{P}_t^{-\alpha,+})_t$. The commutation rule between $\mathbf{D}_{-\alpha,+}$ and $\mathbf{P}_t^{-\alpha,+}$ works as well.

Proposition 2.1.49 — **Commutation relation for $\mathbf{D}_{-\alpha,+}$ - Weibull case.** *We have for all $f \in \mathcal{S}_{\Psi,+}$*

$$\mathbf{D}_{-\alpha,+} \mathbf{P}_t^{-\alpha,+} f(x) = e^{-t} \mathbf{P}_t^{-\alpha,+} \mathbf{D}_{-\alpha,+} f(x), \quad x \in \mathbb{R}_+, \quad t \geq 0. \quad (2.52)$$

The inequality for the \mathbf{L}^p -norm of the gradient of $\mathbf{P}_t^{-\alpha,+} f$ now writes as:

Proposition 2.1.50 *Assume f is in $\mathcal{S}_{\Psi,+}$. Then for every $p \geq 1$,*

$$\|\nabla \mathbf{P}_t^{-\alpha,+} f\|_{\mathbf{L}^p(\mathbb{R}_+, \Psi_\alpha^+)} \leq e^{-\frac{t}{p}} \|\nabla f\|_{\mathbf{L}^p(\mathbb{R}_+, \Psi_\alpha^+)}, \quad t \geq 0.$$

The equivalent of (2.12) and of (2.13) in the Weibull case is

$$\mathcal{L}_{-\alpha,+} f(x) = \frac{1}{\alpha} x f'(x) - \int_0^x r f'(r) \, d\rho_{-\alpha,+}(r) = (\delta_{-\alpha,+} \circ \mathbf{D}_{-\alpha,+}) f(x), \quad (2.53)$$

where the operator $\delta_{-\alpha,+}$ is equal to

$$\delta_{-\alpha,+} f(x) := \alpha^{-1} (r^{-\alpha+1} \nabla - \alpha \text{Id}) f(x) = \alpha^{-1} r^{-\alpha+1} f'(x) - f(x), \quad f \in \mathcal{S}_{\Psi,+}.$$

Once more, this defines a Stein operator and it is the one used in [Kusumoto and Takeuchi \[2020\]](#) to treat limit theorems involving the Weibull case, where it is denoted by \mathcal{S}_α .

Proposition 2.1.51 — **Integration-by-parts formula - Weibull case.** *Let f, g belong to $\mathcal{S}_{\Psi,+}$. Then we have*

$$\langle \delta_{-\alpha,+} f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+, \Psi_\alpha^+)} = -\langle f, \alpha r^{-\alpha+1} \nabla g \rangle_{\mathbf{L}^2(\mathbb{R}_+, \Psi_\alpha^+)} \quad (2.54)$$

This operator satisfies a commutation rule with $\mathbf{P}_t^{-\alpha,+}$:

Proposition 2.1.52 — **Commutation relation for $\delta_{-\alpha,+}$ - Weibull case.** *We have for all $f \in \mathcal{S}_{\Psi,+}$*

$$\delta_{-\alpha,+} \mathbf{P}_t^{-\alpha,+} f(x) = e^t \mathbf{P}_t^{-\alpha,+} \delta_{-\alpha,+} f(x), \quad x \in \mathbb{R}_+, \quad t \geq 0. \quad (2.55)$$

As in the Fréchet case, the commutator of $\mathbf{D}_{-\alpha,+}$ and $\delta_{-\alpha,+}$, or $\mathbf{D}_{-\alpha,+}$ and $\mathcal{L}_{-\alpha,+}$, has formally the same shape.

Proposition 2.1.53 — **Commutator identities - Weibull case.** *We have the following relations:*

$$[\delta_{-\alpha,+}, \mathbf{D}_{-\alpha,+}] = \text{Id}. \quad (2.56)$$

$$[\mathcal{L}_{-\alpha,+}, \mathbf{D}_{-\alpha,+}] = \mathbf{D}_{-\alpha,+}. \quad (2.57)$$

$$[\delta_{-\alpha,+}, \mathcal{L}_{-\alpha,+}] = \delta_{-\alpha,+}. \quad (2.58)$$

Furthermore, $\mathcal{L}_{-\alpha,+}$, $\mathbf{D}_{-\alpha,+}$ and $\delta_{-\alpha,+}$ satisfy the Jacobi identity:

$$[\mathcal{L}_{-\alpha,+}, [\mathbf{D}_{-\alpha,+}, \delta_{-\alpha,+}]] + [\mathbf{D}_{-\alpha,+}, [\delta_{-\alpha,+}, \mathcal{L}_{-\alpha,+}]] + [\delta_{-\alpha,+}, [\mathcal{L}_{-\alpha,+}, \mathbf{D}_{-\alpha,+}]] = 0. \quad (2.59)$$

In particular, the vector space $\mathfrak{g}_{-\alpha,+}$ spanned by those three operators and the identity with respect to linear combinations and $[\cdot, \cdot]$ is a solvable Lie algebra.

Proposition 2.1.54 — Covariance identities - Weibull case. Let $f, g \in \mathcal{S}_{\Psi,+}$ and $Z \sim \mathcal{W}_+(\alpha)$.

- Let Y be a random variable with Beta distribution $B(\alpha, 1)$, independent of Z . Then:

$$\langle \mathcal{L}_{-\alpha,+} f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+, \Psi_\alpha^+)} = -\frac{1}{\alpha^2} \mathbb{E}[Y Z^2 f'(YZ) g'(Z)]. \quad (2.60)$$

- Assume further that f has zero mean: $\mathbb{E}[f(Z)] = 0$. Then:

$$\langle f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+, \Psi_\alpha^+)} = -\frac{1}{\alpha^2} \mathbb{E}[Y Z^2 (\mathcal{L}_{-\alpha,+}^{-1} f)'(YZ) g'(Z)]. \quad (2.61)$$

Proof. We give a short proof of (2.60) which relies on (2.39). Let $\varphi = x \mapsto -\alpha \log x$ and define $T_\varphi f = f \circ \varphi$ whenever this composition is well-defined. We have $\varphi^{-1}(x) = \exp(-x/\alpha)$. Notice that if $X \sim \mathcal{G}(0, 1)$ then $Z = \varphi^{-1}(X) \sim \mathcal{W}_+(\alpha)$ and also that if $W \sim \mathcal{E}(1)$, then $Y = \varphi^{-1}(W) \sim B(\alpha, 1)$. Therefore, one has

$$f(Z) \stackrel{d}{=} (T_{\varphi^{-1}} f)(X) \text{ and } f(ZY) \stackrel{d}{=} (T_{\varphi^{-1}} f)(X + W),$$

as well as

$$\mathbf{P}_t^{-\alpha,+} f(x) = (T_\varphi \mathbf{P}_t^0 T_{\varphi^{-1}}) f(x), x \geq 0, f \in \mathcal{S}_{\Psi,+}.$$

Therefore one can write $\mathcal{L}_{-\alpha,+} = T_\varphi \mathcal{L}_0 T_{\varphi^{-1}}$. Consequently, thanks to (2.39),

$$\begin{aligned} \langle \mathcal{L}_{-\alpha,+} f, g \rangle_{\mathbf{L}^2(\mathbb{R}_+, \Psi_\alpha^+)} &= \mathbb{E}[(\mathcal{L}_{-\alpha,+} f)(Z) g(Z)] \\ &= \mathbb{E}[(T_{\varphi^{-1}} T_\varphi \mathcal{L}_0 T_{\varphi^{-1}}) f(X) (T_{\varphi^{-1}} g)(X)] \\ &= \mathbb{E}[\mathcal{L}_0 (T_{\varphi^{-1}} f)(X) (T_{\varphi^{-1}} g)(X)] \\ &= -\mathbb{E}[(T_{\varphi^{-1}} f)'(X + W) (T_{\varphi^{-1}} g)'(X)] \\ &= -\frac{1}{\alpha^2} \mathbb{E}[Y Z^2 f'(YZ) g'(Z)]. \end{aligned}$$

□

Proposition 2.1.55 — Poincaré inequality - Weibull case. Let $f \in \mathcal{S}_{\Psi,+}$. Let $Z \sim \mathcal{W}_+(\alpha)$ and $Y \sim B(\alpha, 1)$ be a random variable with beta distribution, independent of Z .

$$\mathbb{V}(f(Z)) \leq \mathbb{E}[\mathcal{L}_\alpha f(Z) g(Z)] = -\frac{2}{\alpha^2} \mathbb{E}[Y Z^2 f'(YZ) g'(Z)]. \quad (2.62)$$

The generator of the Weibull semi-group with an integer parameter α satisfies a specific property which is once again proved by simple (iterated) integration-by-parts.

Proposition 2.1.56 Take $\alpha = m \in \mathbb{N}$. Denote by $f^{(1)}(x) := f'(x)$ and $f^{(k-1)}(x) := \int_0^x f^{(k)}(r) dr$

for all integers k in \mathbb{Z} which are less or equal than 1. Then we have

$$\mathcal{L}_{-m}f(x) = \frac{1}{m}xf'(x) + \sum_{k=1}^m (-1)^k k! x^{m-k+1} f^{(-k+1)}(x) + (-1)^{m+1} m! f^{(-m+1)}(x), \quad f \in \mathcal{S}_{\Psi}.$$

Proof. For $\alpha = m$, iterated integration-by-parts in the integral

$$\frac{1}{\alpha} \mathbf{D}_{-\alpha,+} f(x) = \frac{1}{\alpha} \int_0^x r f'(r) \, d\rho_{-\alpha,+}(r) = \int_0^x r^m f'(r) \, dr$$

yield the desired result. \square

Example 4. As stated before, the exponential and Rayleigh distributions are special cases of the Weibull distribution. Since they correspond to $\alpha = 1$ and $\alpha = 2$ respectively, the previous result is applicable. For the exponential case, it writes as

$$\mathcal{L}_{-1,+}f(x) = xf'(x) - (x-1)f(x), \quad f \in \mathcal{S}_{\Psi,+},$$

while in the Rayleigh case we obtain

$$\mathcal{L}_{-2,+}f(x) = \frac{1}{2}xf'(x) - x^2f(x) + 2(x-1)f^{(-1)}(x), \quad f \in \mathcal{S}_{\Psi,+}.$$

Remark 2.1.57. A comparative look at our three max-stable semi-groups shows that their generators can be expressed in an unified way:

$$\mathcal{L}_{\alpha}f(x) = \mathfrak{d}_{\alpha}f(x) + \int_x^{\infty} \mathfrak{d}_{\alpha}f(r) \, d\rho_{\alpha}(r), \quad f \in \mathcal{S}_{\Phi} \quad (2.63)$$

for every $\alpha \in \mathbb{R}$, with

$$\mathfrak{d}_{\alpha}f(x) := \begin{cases} -\frac{1}{\alpha}xf'(x) & \text{if } \alpha \neq 0 \\ -f'(x) & \text{if } \alpha = 0. \end{cases}$$

The operator \mathfrak{d}_{α} corresponds to the generator of the dilation semi-group $(\mathfrak{p}_t^{\alpha})_{t \geq 0}$ when α is different of 0 while we get the translation semi-group $(\mathfrak{p}_t^0)_{t \geq 0}$ when $\alpha = 0$:

$$\mathfrak{p}_t^{\alpha}f(x) := \begin{cases} f(e^{-\frac{t}{\alpha}}x) & \text{if } \alpha \neq 0 \\ f(x-t) & \text{if } \alpha = 0. \end{cases}$$

■

2.2 The Fréchet semi-group as a Markov process

The goal of this section is to build a Markov process whose semi-group is $(\mathbf{P}_t^{\alpha})_{t \geq 0}$. It will be expressed in terms of extremal integrals, as defined in section 1.3.4. The main motivation of this section is to continue the analogy with the usual Ornstein-Uhlenbeck process, which can be described as a stochastic integral depending on a time parameter.

Definition 2.2.1. The *max-stable Ornstein-Uhlenbeck process* is defined as :

$$X_t := e^{-\frac{t}{\alpha}} X_0 \oplus \int_0^t e^{-\frac{1}{\alpha}(t-s)} dM_\alpha(s).$$

where M_α is a α -Fréchet random sup-measure with Lebesgue control measure.

Formally this process is the exact counterpart of the standard Ornstein-Uhlenbeck semi-group, except that the sum is replaced by the maximum, and the stochastic integral by the extremal integral. It remains to prove that this process is indeed Markovian and has the right semi-group:

Proposition 2.2.2 *The process $(X_t)_{t \geq 0}$ is a Markov process and*

$$\mathbb{E}[f(X_t) | X_0 = x] = \mathbf{P}_t^\alpha f(x), \quad x \in \mathbb{R}_+, \quad f \in \mathcal{S}_\Phi.$$

Proof. Let us note provisionally $\widetilde{\mathbf{P}}_t^\alpha f(x) := \mathbb{E}[f(X_t) | X_0 = x]$. It is clear that for all nonnegative t , $\widetilde{\mathbf{P}}_t^\alpha$ is a linear operator and that $\widetilde{\mathbf{P}}_0^\alpha = \text{Id}$. Now we need to check that $\widetilde{\mathbf{P}}_t^\alpha \circ \widetilde{\mathbf{P}}_s^\alpha = \widetilde{\mathbf{P}}_{t+s}^\alpha$:

$$\begin{aligned} (\widetilde{\mathbf{P}}_t^\alpha \circ \widetilde{\mathbf{P}}_s^\alpha) f(x) &= \mathbb{E} \left[f \left(e^{-\frac{t}{\alpha}} X_s \oplus \int_0^t e^{-\frac{1}{\alpha}(t-u)} dM_\alpha(u) \right) \mid X_0 = x \right] \\ &= \mathbb{E} \left[f \left(e^{-\frac{1}{\alpha}(t+s)} x \oplus \int_0^t e^{-\frac{1}{\alpha}(t+s-u)} dM_\alpha(u) \oplus \int_0^s e^{-\frac{1}{\alpha}(s-u)} dM'_\alpha(u) \right) \right], \end{aligned}$$

where M'_α denotes an independent copy of M_α . Furthermore, by the isometry property, we have that:

$$\int_0^s e^{-(s-u)} dM_\alpha(u) \stackrel{d}{=} \mathcal{F} \left(1, \left(\int_0^s e^{-(s-u)} du \right)^{1/\alpha} \right) = \mathcal{F} \left(\alpha, \left(\int_t^{t+s} e^{-(t+s-u)} du \right)^{1/\alpha} \right).$$

Consequently, injecting this result in the previous computation:

$$\begin{aligned} (\widetilde{\mathbf{P}}_t^\alpha \circ \widetilde{\mathbf{P}}_s^\alpha) f(x) &= \mathbb{E} \left[f \left(e^{-\frac{1}{\alpha}(t+s)} x \oplus \int_0^t e^{-\frac{1}{\alpha}(t+s-u)} dM_\alpha(u) \oplus \int_0^s e^{-\frac{1}{\alpha}(s-u)} dM'_\alpha(u) \right) \right] \\ &= \mathbb{E} \left[f \left(e^{-\frac{1}{\alpha}(t+s)} x \oplus \int_0^t e^{-(t+s-u)} dM_\alpha(u) \oplus \int_t^{t+s} e^{-\frac{1}{\alpha}(t+s-u)} dM'_\alpha(u) \right) \right] \\ &= \mathbb{E} \left[f \left(e^{-\frac{1}{\alpha}(t+s)} x \oplus \int_0^{t+s} e^{-\frac{1}{\alpha}(t+s-u)} dM_\alpha(u) \right) \right] \\ &= \widetilde{\mathbf{P}}_{t+s}^\alpha f(x). \end{aligned}$$

Furthermore, since $\int_0^t e^{-(t-s)} ds = 1 - e^{-t}$, we have that $\int_0^t e^{-\frac{1}{\alpha}(t-u)} dM_\alpha(u) \stackrel{d}{=} \mathcal{F}(\alpha, (1 - e^{-t})^{1/\alpha})$, so

$$\int_0^t e^{-\frac{1}{\alpha}(t-s)} dM_\alpha(s) \stackrel{d}{=} (1 - e^{-t})^{\frac{1}{\alpha}} Z,$$

where Z is a random variable with Fréchet distribution $\mathcal{F}(\alpha)$. Therefore $(\widetilde{\mathbf{P}}_t^\alpha)_{t \geq 0} = (\mathbf{P}_t^\alpha)_{t \geq 0}$.

□

Remark 2.2.3. Using the terminology of [Stoev and Taqqu \[2005\]](#), this process is an instance of an *integral moving maximum process*:

$$X_t = \int_{\mathbb{R}_+}^e f(t-u) \, dM_\alpha(u),$$

where $f \in \mathbf{L}_+^\alpha(\mathbb{R}_+, \lambda)$ and λ is the Lebesgue measure. Here $f = u \mapsto e^{-\frac{1}{\alpha}u} \mathbf{1}_{\mathbb{R}_+}(u)$. ■

The notion of extremal integral allows us to define another stochastic process of interest here.

Definition 2.2.4. Let Z_0 be a positive random variable and α a positive number. An *α -max-stable motion* $(Z_t)_{t \geq 0}$ is a stochastic process such that there exists an α -Fréchet random measure satisfying

$$Z_t = Z_0 \oplus \int_0^e dM_\alpha(s), \quad t \geq 0.$$

The next proposition links this stochastic process to the max-stable Ornstein-Uhlenbeck process.

Proposition 2.2.5 $(Z_t)_{t \geq 0}$ is a Markov process whose generator \mathcal{K}_α is:

$$\mathcal{K}_\alpha f(x) = \mathbf{D}_\alpha f(x) = \int_x^\infty (f(x \oplus r) - f(x)) \frac{\alpha}{r^{\alpha+1}} \, dr, \quad x \in \mathbb{R}_+^*, \quad f \in \mathcal{S}_\Phi.$$

Proof. Let f be in \mathcal{S}_Φ . Furthermore, define $Q_t f(x) := \mathbb{E}[f(Z_t) | X_0 = x]$. By definition of Z_t , we have:

$$\begin{aligned} Z_{t+s} &= Z_0 \oplus \int_0^e dM_\alpha(u) \\ &= Z_0 \oplus \int_0^e dM_\alpha(u) \oplus \int_s^e dM_\alpha(u) \\ &\stackrel{d}{=} Z_s \oplus tZ, \end{aligned}$$

where Z is a random variable with Fréchet distribution $\mathcal{F}(\alpha)$ independent of $\sigma(Z_u, u \leq s)$. This proves that $(Z_t)_{t \geq 0}$ is a Markov process with semi-group

$$Q_t f(x) = \mathbb{E}[f(x \oplus tZ)] = f(x)e^{-\frac{t}{x^\alpha}} + t \int_x^\infty f(r)e^{-\frac{t}{r^\alpha}} \frac{\alpha}{r^{\alpha+1}} \, dr, \quad x > 0.$$

An easy calculation yields the generator of the proposition. □

In other words, the generator of the max-stable Ornstein-Uhlenbeck process writes as the generator of the dilation semi-group plus the generator of an α -max-stable motion. The former "counteracts" the "diffusion" effect of \mathbf{D}_α , the generator of the max-stable motion. This phenomenon is similar to what is observed in the standard Ornstein-Uhlenbeck process (where the max-stable motion is replaced by the Brownian motion), or more generally with α -stable Ornstein-Uhlenbeck processes (see [Samorodnitsky and Taqqu \[1994\]](#)).

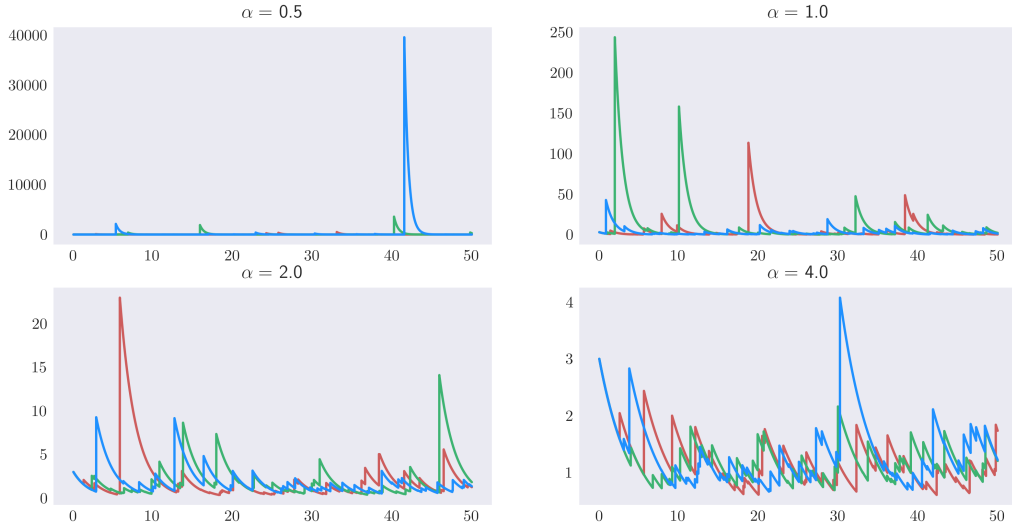


Figure 2.1: Three paths of X_t for $\alpha \in \{\frac{1}{3}, 1, 3\}$

It is well known that one can construct martingales from Markov processes by using their generators (see for instance [Rogers and D.Williams \[April 2000\]](#), or theorem 6.14 in [Gall \[2016\]](#)). This is the content of proposition 1.1.8. We apply it to $(X_t)_{t \geq 0}$, as defined in the previous remark. For f a function in the domain of \mathcal{L}_α such that $\mathbb{E}[\mathcal{L}_\alpha f(X_t)]$ exists for all t , the process defined by

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}_\alpha f(X_u) \, du$$

is a martingale. This result takes an especially simple form for $f = \log$:

Proposition 2.2.6 *Let X_0 be a positive random variable and define*

$$M_t := \log X_t - \log X_0 + \frac{t}{\alpha} - \frac{1}{\alpha} \int_0^t \frac{1}{X_u^\alpha} \, du.$$

Then $(M_t)_{t \geq 0}$ is a martingale for the filtration $(\mathcal{F}_t)_{t \geq 0} = \sigma(X_t, t \geq 0)$.

Proof. We could apply directly the general theorem to $f = \log$, as stated in [Rogers and D.Williams \[April 2000\]](#), even though $(X_t)_{t \geq 0}$ is not Feller (the proof of the needed implication does not require the Feller property). In that case, we have

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha} x f'(x) + \int_x^\infty \frac{f'(y)}{y^\alpha} \, dy = -\frac{1}{\alpha} + \int_x^\infty \frac{1}{y^{\alpha+1}} \, dy = -\frac{1}{\alpha} + \frac{1}{\alpha} \frac{1}{x^\alpha},$$

which explains the form of $(M_t)_{t \geq 0}$. Nonetheless we give a direct proof in order to make things more enlightening. We will need the identity:

$$\mathbb{E}[\log (e^{-\frac{t}{\alpha}} x \oplus (1 - e^{-t})^{\frac{1}{\alpha}} Z_\alpha)] - \frac{1}{\alpha} \int_0^t \mathbb{E}\left[\frac{1}{e^{-u} x^\alpha \oplus (1 - e^{-u}) Z_\alpha^\alpha}\right] \, du = \log x - \frac{t}{\alpha}, \quad (2.64)$$

where $Z_\alpha \sim \mathcal{F}(\alpha)$. Actually, we just need to prove it for $\alpha = 1$: by setting $y = x^\alpha$ and using that $Z := Z_\alpha^\alpha$ has the Fréchet distribution $\mathcal{F}(1)$, the previous equality becomes:

$$\mathbb{E}[\log(e^{-t}y \oplus (1 - e^{-t})Z)] - \int_0^t \mathbb{E}\left[\frac{1}{e^{-u}y \oplus (1 - e^{-u})Z}\right] du = \log y - t, \quad (2.65)$$

$$\begin{aligned} \mathbb{E}[\log(e^{-t}y \oplus (1 - e^{-t})Z)] &= \log(e^{-t}y)e^{-\frac{\gamma t}{y}} + \int_{\frac{y}{\gamma t}}^\infty \log((1 - e^{-t})z) \frac{e^{-\frac{1}{z}}}{z^2} dz \\ &= \log(e^{-t}y)e^{-\frac{\gamma t}{y}} + \log(1 - e^{-t})(1 - e^{-\frac{\gamma t}{y}}) + \int_{\frac{y}{\gamma t}}^\infty \log z \frac{e^{-\frac{1}{z}}}{z^2} dz \end{aligned} \quad (2.66)$$

On the other hand

$$\begin{aligned} \int_0^t \mathbb{E}\left[\frac{1}{e^{-u}y \oplus (1 - e^{-u})Z}\right] du &= \int_0^t e^u \frac{e^{-\frac{\gamma u}{y}}}{y} + \frac{1}{1 - e^{-u}} \int_{\frac{y}{\gamma u}}^\infty \frac{e^{-\frac{1}{z}}}{z^3} dz du \\ &= 1 - e^{-\frac{\gamma t}{y}} + \int_0^t \int_{\frac{y}{\gamma u}}^\infty \frac{1}{1 - e^{-u}} \frac{e^{-\frac{1}{z}}}{z^3} dz du \\ &= 1 - e^{-\frac{\gamma t}{y}} + \int_{\frac{y}{\gamma t}}^\infty \int_{\log(1 + \frac{y}{z})}^t \frac{1}{1 - e^{-u}} \frac{e^{-\frac{1}{z}}}{z^3} du dz \\ &= 1 - e^{-\frac{\gamma t}{y}} + \int_{\frac{y}{\gamma t}}^\infty \frac{e^{-\frac{1}{z}}}{z^3} \left[\log \gamma t - \log\left(\frac{y}{z}\right) \right] dz \\ &= 1 - e^{-\frac{\gamma t}{y}} + \log\left(\frac{\gamma t}{y}\right) \int_{\frac{y}{\gamma t}}^\infty \frac{e^{-\frac{1}{z}}}{z^3} dz + \int_{\frac{y}{\gamma t}}^\infty \log z \frac{e^{-\frac{1}{z}}}{z^3} dz \\ &= 1 - e^{-\frac{\gamma t}{y}} + \log\left(\frac{\gamma t}{y}\right) \int_{\frac{y}{\gamma t}}^\infty \frac{e^{-\frac{1}{z}}}{z^3} dz \\ &\quad + \frac{\gamma t}{y} \log\left(\frac{\gamma t}{y}\right) e^{-\frac{\gamma t}{y}} + \int_{\frac{y}{\gamma t}}^\infty \log z \frac{e^{-\frac{1}{z}}}{z^2} dz - \int_{\frac{y}{\gamma t}}^\infty \frac{e^{-\frac{1}{z}}}{z^2} dz \\ &= \frac{\gamma t}{y} \log\left(\frac{\gamma t}{y}\right) e^{-\frac{\gamma t}{y}} + \log\left(\frac{\gamma t}{y}\right) \int_{\frac{y}{\gamma t}}^\infty \frac{e^{-\frac{1}{z}}}{z^3} dz + \int_{\frac{y}{\gamma t}}^\infty \log z \frac{e^{-\frac{1}{z}}}{z^2} dz, \end{aligned}$$

the penultimate equality stemming from an integration by parts (differentiate $z \mapsto z^{-1} \log z$). The first integral in the last equality is just an avatar of the incomplete Gamma function with an integer argument. As such it can be computed explicitly (once more by integration by parts) and thus we find:

$$\begin{aligned} \int_0^t \mathbb{E}\left[\frac{1}{e^{-u}y \oplus (1 - e^{-u})Z}\right] du &= \frac{\gamma t}{y} \log\left(\frac{\gamma t}{y}\right) e^{-\frac{\gamma t}{y}} + \log\left(\frac{\gamma t}{y}\right) (1 - e^{-\frac{\gamma t}{y}} - \frac{\gamma t}{y} e^{-\frac{\gamma t}{y}}) \\ &\quad + \int_{\frac{y}{\gamma t}}^\infty \log z \frac{e^{-\frac{1}{z}}}{z^2} dz \\ &= \log\left(\frac{\gamma t}{y}\right) (1 - e^{-\frac{\gamma t}{y}}) + \int_{\frac{y}{\gamma t}}^\infty \log z \frac{e^{-\frac{1}{z}}}{z^2} dz. \end{aligned} \quad (2.67)$$

Subtracting (2.67) to (2.66), we finally obtain:

$$\begin{aligned} \mathbb{E}[\log(e^{-t}y \oplus (1 - e^{-t})Z)] - \int_0^t \mathbb{E}\left[\frac{1}{e^{-u}y \oplus (1 - e^{-u})Z}\right] du &= \log(e^{-t}y)e^{-\frac{\gamma t}{y}} + \log(1 - e^{-t})(1 - e^{-\frac{\gamma t}{y}}) \\ &\quad - \log\left(\frac{\gamma t}{y}\right)(1 - e^{-\frac{\gamma t}{y}}) \\ &= \log y - t, \end{aligned}$$

as announced. From this equality and the Markov property of $(X_t)_{t \geq 0}$, it is easy to obtain the martingale property of $(M_t)_{t \geq 0}$. □

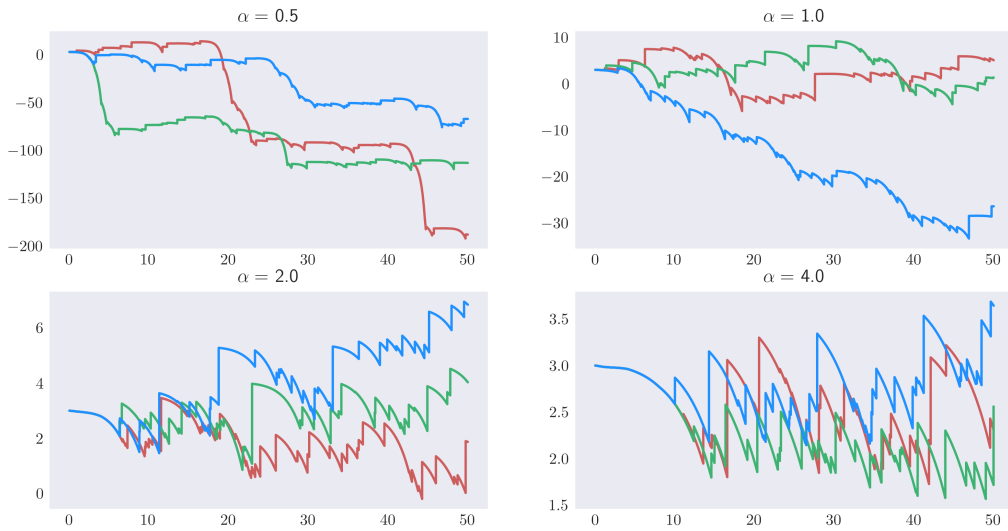


Figure 2.2: Three paths of M_t for $\alpha \in \{\frac{1}{3}, 1, 3\}$

Remark 2.2.7. Set $Y_t = \alpha \log X_t$ and $y = \alpha \log X_0$. Then it is easy to see that $(Y_t)_{t \geq 0}$ is a Markov process with semi-group $(\mathbf{P}_t^0)_{t \geq 0}$. With those notations, after being multiplied by α , the previous martingale becomes:

$$M_t = Y_t - Y_0 + t - \int_0^t e^{-Y_u} du.$$

This is exactly the same martingale which would have appeared had we applied theorem 1.1.8 of Gall [2016] to the Markov process associated to the Gumbel distribution, and $f = \text{Id}$. The same line of reasoning could be applied to find the martingale connected to the Weibull distribution. In other words, up to an obvious change of variable, the martingale $(M_t)_{t \geq 0}$ is the same for the three max-stable distributions. ■

To end that section, we give a table comparing the properties satisfied by the Gaussian Ornstein-Uhlenbeck semi-group, and the Fréchet semi-group. There, G will denote a random variable with the standard normal distribution $\mathcal{N}(0, 1)$, $(B_t)_{t \geq 0}$ a standard Brownian motion, Z a random variable with the Fréchet distribution $\mathcal{F}(\alpha)$, and Y a random variable with the Pareto distribution $\mathcal{VP}(\alpha, 1)$.

Property	Ornstein-Uhlenbeck semi-group	Fréchet semi-group
Mehler's formula	$\mathbb{E}\left[f(e^{-t}x + \sqrt{1 - e^{-2t}}G)\right]$	$\mathbb{E}\left[f(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{\frac{1}{\alpha}}Z)\right]$
Generator	$-xf'(x) + f''(x)$	$-\frac{1}{\alpha}xf'(x) + \mathbf{D}_\alpha f(x)$
Derivative	∇	\mathbf{D}_α
Divergence	$-xf(x) + f'(x)$	$\frac{1}{\alpha}x^{\alpha+1}f'(x) + f(x)$
First commutation rule	$[P_t, \nabla] = (1 - e^{-t})P_t\nabla$	$[\mathbf{P}_t^\alpha, \mathbf{D}_\alpha] = (1 - e^{-t})\mathbf{P}_t^\alpha\mathbf{D}_\alpha$
Second commutation rule	$[P_t, \delta] = (1 - e^{-t})P_t\delta$	$[\mathbf{P}_t^\alpha, \delta_\alpha] = (1 - e^{-t})\mathbf{P}_t^\alpha\delta_\alpha$
Creation and annihilation rule	$[\delta, \nabla] = \text{Id}$	$[\delta_\alpha, \mathbf{D}_\alpha] = \text{Id}$
First infinitesimal commutation rule	$[\mathcal{L}_\gamma, \nabla] = \nabla$	$[\mathcal{L}_\alpha, \mathbf{D}_\alpha] = \mathbf{D}_\alpha$
Second infinitesimal commutation rule	$[\delta, \mathcal{L}_\gamma] = \delta$	$[\delta_\alpha, \mathcal{L}_\alpha] = \delta_\alpha$
Energy form	$-\mathbb{E}[f'(G)g'(G)]$	$-\frac{1}{\alpha^2}\mathbb{E}[YZ^2f'(YZ)g'(Z)]$
Stochastic integral representation	$X_t = e^{-t}X_0 + \sqrt{2} \int_0^t e^{-(t-s)} dB(s)$	$X_t = e^{-\frac{t}{\alpha}}X_0 \oplus \int_0^t e^{-\frac{1}{\alpha}(t-s)} dM_\alpha(s)$

2.3 Quantization of multivariate max-stable distributions

The last section of this chapter is dedicated to the quantization of multivariate max stable distributions. We will rely heavily on the properties of the Poisson process recalled in section 1.2 as well as on the notations introduced in section 1.3.2.

We need to keep in mind that unlike the unidimensional case, the marginals of a max-stable random vector may be quite cumbersome to describe. Indeed, recall the example given in the preliminaries: the vector (X, Y) , with $X \sim \mathcal{F}(1)$ and $Y \sim \mathcal{G}(0, 1)$, and X, Y independent, is max-stable but has different marginal distributions. However, up to a non-decreasing transformation, the marginal distributions of a max-stable random vector can be normalized to be *e.g.* unit Fréchet $\mathcal{F}(1)$. This does not change the dependency structure of the vector, which is given by the angular measure. For this reason, it is easier to work with max-stable random vectors whose marginals have the same Fréchet $\mathcal{F}(1)$. Since the definition of our semi-group is 'compatible' with such transformations, we will not lose in generality by doing so. Therefore, the philosophy of this section is to first define the semi-group associated to the multivariate max-stable random vectors with unit Fréchet marginals, and then to apply a certain transformation to it to deduce a quantization of max-stable distributions in general.

2.3.1 The case of the unit Fréchet max-stable random vectors

In our context, the right generalization of \mathbb{R}_+^* in higher dimensions is not $(\mathbb{R}_+^*)^d$ but E_0 defined as follows

$$E_0 := \{\mathbf{x} \in \overline{\mathbb{R}}_+^d, \exists j \in \llbracket 1, d \rrbracket, x^j > 0\} = \{\mathbf{x} \in \overline{\mathbb{R}}_+^d, \mathbf{x} \not\leq \mathbf{0}\} = [0, +\infty]^d \setminus \{\mathbf{0}\}.$$

The Schwartz class on E_0 is defined as

$$\mathcal{S}(E_0) := \{f \in \mathcal{C}^\infty(E_0), \mathbf{x}^k \partial_j^n |f(\mathbf{x})| < \infty, \text{ for all } (\mathbf{k}, n) \in \mathbb{N}^d \times \mathbb{N}, \mathbf{j} \in \mathbb{N}^n\},$$

where $\mathbf{x}^k = \prod_{j=1}^d (x^j)^{k_j}$ and $\partial_j^n f(\mathbf{x}) := \partial_{j_1} \dots \partial_{j_n} f(\mathbf{x})$.

Definition 2.3.1 (Standard algebra). We define the class of functions:

$$\mathcal{S}_{\Phi, d} := \{f \in \mathcal{S}(E_0), \sup_{\mathbf{x} \in (0, 1]^d} \mathbf{x}^{-k} |\partial_j^n f(\mathbf{x})| < \infty \text{ for all } (\mathbf{k}, n) \in \mathbb{N}^d \times \mathbb{N}^*, \mathbf{j} \in \mathbb{N}^n\}.$$

The multivariate extension of the Fréchet semi-group takes formally the same shape as in the univariate case.

Definition 2.3.2 (Multivariate unit Fréchet semi-group). Fix a reference norm $\|\cdot\|$ on \mathbb{R}^d . Define the *multivariate unit Fréchet semi-group* $(\mathbf{P}_t^{1, \nu})_{t \geq 0}$ on the standard algebra $\mathcal{A} = \mathcal{S}_{\Phi, d}$ by

$$\mathbf{P}_t^{1, \nu} f(\mathbf{x}) := \mathbb{E}[f(e^{-t} \mathbf{x} \oplus (1 - e^{-t}) \mathbf{Z})], \quad \mathbf{x} \geq \mathbf{0}, t \geq 0 \quad (2.68)$$

where $\mathbf{Z} \sim \mathcal{MS}(1, \nu)$, with ν a finite measure on \mathbb{S}_+^{d-1} satisfying the moment constraints (1.18).

Without further assumptions on the angular measure ν , the set $\mathcal{S}_{\Phi, d}$ is *not* a standard algebra for $(\mathbf{P}_t^{1, \nu})_{t \geq 0}$ in general. This difficulty is caused by the lack of differentiability of $\mathbf{x} \mapsto \mathbf{x} \oplus \mathbf{y}$ (for some fixed \mathbf{y}) in the \mathbf{x} which share a coordinate with \mathbf{y} . In dimension 1, this problem did not exist because we worked with the measure ρ_α , which is absolutely continuous w.r.t. Lebesgue's measure. In higher dimensions, we will need to impose that the angular measure puts zero mass on sets included in

$$\Delta := \{\mathbf{x} \in E_0, x^i = x^j \text{ for some } i \neq j\}.$$

However we will usually not need to work with a true standard algebra, except when iterating the generator of $(\mathbf{P}_t^{1, \nu})_{t \geq 0}$. For most of our needs, working on $\mathcal{S}_{\Phi, d}$ with no additional assumptions on ν will be enough to ensure the objects we want to define are well-posed.

Proposition 2.3.3 *The family of operators $(\mathbf{P}_t^{1, \nu})_{t \geq 0}$ is a semi-group on $\mathcal{S}_{\Phi, d}$.*

Proof. Clearly $\mathbf{P}_t^{1, \nu}$ is a linear operator on $\mathcal{S}_{\Phi, d}$. Furthermore, if $f \in \mathcal{S}_{\Phi, d}$, then f is bounded. Thus $\mathbf{P}_s^{1, \nu} f$ is well-defined and bounded too, as well as $\mathbf{P}_t^{1, \nu}(\mathbf{P}_s^{1, \nu} f)$, for all nonnegative t, s . The

semi-group relation is satisfied:

$$\begin{aligned}
(\mathbf{P}_t^{1,\nu} \circ \mathbf{P}_s^{1,\nu})f(\mathbf{x}) &= \mathbb{E}[(\mathbf{P}_s^{1,\nu})f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z})] \\
&= \mathbb{E}\left[f\left(e^{-s}(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z}) \oplus (1 - e^{-s})\mathbf{Z}'\right)\right] \\
&= \mathbb{E}\left[f\left(e^{-(t+s)}\mathbf{x} \oplus (e^{-s}(1 - e^{-t})\mathbf{Z} \oplus (1 - e^{-s})\mathbf{Z}')\right)\right] \\
&= \mathbb{E}[f(e^{-(t+s)}\mathbf{x} \oplus (1 - e^{-(t+s)})\mathbf{Z})]
\end{aligned}$$

where \mathbf{Z}' is an independent copy of \mathbf{Z} . Using representation (1.20), it is clear that $e^{-s}(1 - e^{-t})\mathbf{Z} \oplus (1 - e^{-s})\mathbf{Z}' \stackrel{d}{=} (1 - e^{-(t+s)})\mathbf{Z}$. \square

The max-stability relation upon which we have built our semi-group implies that it quantizes the multivariate unit Fréchet distribution.

Proposition 2.3.4 $(\mathbf{P}_t^{1,\nu})_{t \geq 0}$ is ergodic and admits $\Phi_{1,\nu}$ as a stationary distribution, where $\Phi_{1,\nu}$ denotes the distribution of $\mathbf{Z} \sim \mathcal{M}(1, \nu)$.

Proof. By the definition of $\mathbf{P}_t^{1,\nu}f$ and a dominated convergence argument, we get

$$\mathbf{P}_t^{1,\nu}f(\mathbf{x}) \xrightarrow{t \rightarrow \infty} \mathbb{E}[f(\mathbf{Z})],$$

which means that $(\mathbf{P}_t^{1,\nu})_{t \geq 0}$ is ergodic. Furthermore, if $\mathbf{Z}' \sim \mathcal{MS}(1, \nu)$, then it is clear that:

$$\mathbb{E}[\mathbf{P}_t^{1,\nu}f(\mathbf{Z}')] = \mathbb{E}[f(e^{-t}\mathbf{Z}' \oplus (1 - e^{-t})\mathbf{Z})] = \mathbb{E}[f(\mathbf{Z}')],$$

due to the max-stability property. \square

Proving that $t \mapsto \mathbf{P}_t^{1,\nu}$ is right-continuous at 0 and computing the generator of $(\mathbf{P}_t^{1,\nu})_{t \geq 0}$ are a bit trickier than in dimension 1 because decomposition (2.2) is not as obvious to write. Actually, it can be seen as a very particular case of Mecke's formula. This heuristic still holds in higher dimension and gives us the counterpart of (2.2) we need.

Lemma 2.3.5 Denote by $d\mu = T^*(d\rho_1 \otimes d\nu)$ the exponent measure of the max-stable distribution $\mathcal{MS}(1, \nu)$ (see (1.23)). Let \mathbf{Z} be a random vector with this distribution. We have for $f \in \mathcal{S}_{\Phi,d}$ and $x \in E_0$:

$$\begin{aligned}
\mathbf{P}_t^{1,\nu}f(\mathbf{x}) &= F_{\gamma_t \mathbf{Z}}(\mathbf{x})f(e^{-t}\mathbf{x}) \\
&\quad + F_{\gamma_t \mathbf{Z}}(\mathbf{x})\gamma_t \int_{[0,\mathbf{x}]^c} f(e^{-t}(\mathbf{x} \oplus r\mathbf{u})) \frac{1}{r^2} dr d\nu(\mathbf{u}) \\
&\quad + F_{\gamma_t \mathbf{Z}}(\mathbf{x}) \sum_{n=2}^{\infty} \frac{\gamma_t^n}{n!} \int_{([0,\mathbf{x}]^c)^n} f(e^{-t}(\mathbf{x} \oplus r_1\mathbf{u}_1 \oplus \cdots \oplus r_n\mathbf{u}_n)) \prod_{i=1}^n \frac{1}{r_i^2} dr_i d\nu(\mathbf{u}_i),
\end{aligned} \tag{2.69}$$

where we abused notations by integrating over $[0, \mathbf{x}]^c$, which denotes here the set $\{(r, \mathbf{u}) \in E_{\text{pol}}, r\mathbf{u} \not\leq \mathbf{x}\}$, where

$$E_{\text{pol}} := \overline{\mathbb{R}}_+^* \times \mathbb{S}_+^{d-1}.$$

Proof. Remember that \mathbf{Z} can be seen as the coordinate-wise maximum of a certain Poisson point process thanks to (1.19). As a result we can rewrite $\mathbf{P}_t^{1,\nu} f(\mathbf{x})$ as an expectation with respect to a Poisson process η with intensity measure $d\mu = r^{-2} dr \otimes d\nu(\mathbf{u})$ on $E_{\mathbf{0}}$ and apply (1.2) to expand it. The problem lies in the fact that the intensity measure of η is not finite because it has an accumulation point at $\mathbf{0}$. But observe that we do not need all the points (r, \mathbf{u}) of η , only those such that $r\mathbf{u}$ is *not less* than $\gamma_t^{-1}\mathbf{x}$, i.e. those which belong to

$$A := \{(r, \mathbf{u}) \in \overline{\mathbb{R}}_+^* \times \mathbb{S}_+^{d-1}, r\mathbf{u} \notin [\mathbf{0}, \gamma_t^{-1}\mathbf{x}]\}.$$

This random set $\eta \cap A$ is still a Poisson process, this time with intensity measure μ restricted to A . Thus it has finite intensity measure since it avoids the accumulation point at $\mathbf{0}$. We can apply (1.2) to the function $\varphi(\cdot) = f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{m}(\cdot))$, where f belongs to $\mathcal{S}_{\Phi,d}$, to get:

$$\begin{aligned} \mathbf{P}_t^{1,\nu} f(\mathbf{x}) &= e^{-\mu[\mathbf{0}, \gamma_t^{-1}\mathbf{x}]^c} f(e^{-t}\mathbf{x}) \\ &\quad + e^{-\mu[\mathbf{0}, \gamma_t^{-1}\mathbf{x}]^c} \sum_{n=1}^{\infty} \frac{1}{n!} \int_{A^n} f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{m}\{r_1\mathbf{u}_1, \dots, r_n\mathbf{u}_n\}) \prod_{i=1}^n \frac{1}{r_i^2} dr_i d\nu(\mathbf{u}_i) \\ &= e^{-\gamma_t\mu[\mathbf{0}, \mathbf{x}]^c} f(e^{-t}\mathbf{x}) \\ &\quad + e^{-\gamma_t\mu[\mathbf{0}, \mathbf{x}]^c} \sum_{n=1}^{\infty} \frac{1}{n!} \int_{A^n} f(e^{-t}(\mathbf{x} \oplus \gamma_t r_1\mathbf{u}_1 \oplus \dots \oplus r_n\mathbf{u}_n)) \prod_{i=1}^n \frac{1}{r_i^2} dr_i d\nu(\mathbf{u}_i) \\ &= e^{-\gamma_t\mu[\mathbf{0}, \mathbf{x}]^c} f(e^{-t}\mathbf{x}) \\ &\quad + e^{-\gamma_t\mu[\mathbf{0}, \mathbf{x}]^c} \sum_{n=1}^{\infty} \frac{\gamma_t^n}{n!} \int_{([\mathbf{0}, \mathbf{x}]^c)^n} f(e^{-t}(\mathbf{x} \oplus r_1\mathbf{u}_1 \oplus \dots \oplus r_n\mathbf{u}_n)) \prod_{i=1}^n \frac{1}{r_i^2} dr_i d\nu(\mathbf{u}_i), \end{aligned}$$

since $\gamma_t A = [\mathbf{0}, \mathbf{x}]^c$. By using that $[\mathbf{0}, \gamma_t^{-1}\mathbf{x}]^c = \gamma_t^{-1}[\mathbf{0}, \mathbf{x}]^c$ and equation (1.21), we found that $\mu[\mathbf{0}, \gamma_t^{-1}\mathbf{x}]^c = \gamma_t\mu[\mathbf{0}, \mathbf{x}]^c$. Finally, notice that

$$e^{-\gamma_t\mu[\mathbf{0}, \mathbf{x}]^c} = \mathbb{P}(\mathbf{Z} \leq \gamma_t^{-1}\mathbf{x}) = F_{\gamma_t\mathbf{Z}}(\mathbf{x}).$$

□

Proposition 2.3.6 *The family of operators $(\mathbf{P}_t^\alpha)_{t \geq 0}$ satisfies the following properties:*

- The application $t \mapsto \mathbf{P}_t^{1,\nu}$ is right-continuous at 0, and so $(\mathbf{P}_t^{1,\nu})_{t \geq 0}$ is a Markov semi-group.
- The Markov semi-group $(\mathbf{P}_t^{1,\nu})_{t \geq 0}$ has generator $\mathcal{L}_{1,\nu}$, given by:

$$\mathcal{L}_{1,\nu} f(\mathbf{x}) = -\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \int_{E_{\text{pol}}} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}), \quad \mathbf{x} \in E_{\mathbf{0}}, f \in \mathcal{S}_{\Phi,d}$$

and where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product on \mathbb{R}^d .

Proof. - We prove that $\mathbf{P}_t^{1,\nu} f$ converges to f in norm $\mathbf{L}^2(E_{\mathbf{0}}, \Phi_{1,\nu})$. Clearly

$$F_{\gamma_t\mathbf{Z}}(\mathbf{x}) f(e^{-t}\mathbf{x}) = e^{-\gamma_t\mu[\mathbf{0}, \mathbf{x}]^c} f(e^{-t}\mathbf{x}) \xrightarrow[t \rightarrow 0^+]{} f(\mathbf{x}) \text{ a.e.}$$

and thus $\Phi_{1,\nu}$ -a.s. too. It is dominated by $C(1 + \|\sqrt[4]{\mathbf{x}}\|)$ for some constant $C > 0$ since $f(\mathbf{x}) = \tilde{f}(\mathbf{x})$, where $\tilde{f}(\mathbf{x}) = f(\mathbf{x}^4)$, and \tilde{f} still belongs to $\mathcal{S}_{\Phi,d}$. In particular, \tilde{f} is Lipschitz on

E_0 . Moreover, because all norms are equivalent in finite dimension, one can find another constant C such that $\|\sqrt[4]{\mathbf{x}}\| \leq C\|\sqrt[4]{\mathbf{x}}\|_1$, making it clear that $\mathbf{x} \mapsto \|\sqrt[4]{\mathbf{x}}\|$ is $\Phi_{1,\nu}$ -square integrable. By bounded convergence theorem, we conclude that $\exp(-\gamma_t \mu[\mathbf{0}, \mathbf{x}]^c) f(e^{-t} \mathbf{x})$ converges to 0 in norm $\mathbf{L}^2(E_0, \Phi_{1,\nu})$. The two other terms of (2.69) likewise converge to 0 in norm $\mathbf{L}^2(E_0, \Phi_{1,\nu})$ because $\gamma_t = e^t - 1$ vanishes when t goes to 0 and the multiple integrals are easily dominated by $(\mu[\mathbf{0}, \mathbf{x}]^c)^n$.

- We differentiate equality (2.69) with respect to t at $t = 0$. The first term gives two parts of the generator:

$$\begin{aligned} \frac{d}{dt} \Big|_{t=0} e^{-\gamma_t \mu[\mathbf{0}, \mathbf{x}]^c} f(e^{-t} \mathbf{x}) &= -\mu[\mathbf{0}, \mathbf{x}]^c f(\mathbf{x}) - \langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle \\ &= -\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle - \int_{[0, \mathbf{x}]^c} f(\mathbf{x}) \frac{1}{r^2} dr d\nu(\mathbf{u}). \end{aligned}$$

As for the first integral term,

$$\begin{aligned} \frac{d}{dt} \Big|_{t=0} \gamma_t e^{-\gamma_t \mu[\mathbf{0}, \mathbf{x}]^c} \int_{[0, \mathbf{x}]^c} f(e^{-t}(\mathbf{x} \oplus r\mathbf{u})) \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \frac{d}{dt} \Big|_{t=0} \left\{ e^{-\gamma_t \mu[\mathbf{0}, \mathbf{x}]^c} \gamma_t \int_{[0, \mathbf{x}]^c} f(e^{-t}(\mathbf{x} \oplus r\mathbf{u})) \frac{1}{r^2} dr d\nu(\mathbf{u}) \right\} \\ &= \int_{[0, \mathbf{x}]^c} f(\mathbf{x} \oplus r\mathbf{u}) \frac{1}{r^2} dr d\nu(\mathbf{u}). \end{aligned}$$

The series converges to 0 at speed $O(t^2)$, so we can neglect it. Putting all the parts together and noticing that $(r, \mathbf{u}) \mapsto f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})$ vanishes on $[\mathbf{0}, \mathbf{x}]$, we find the announced result:

$$\int_{[0, \mathbf{x}]^c} f(\mathbf{x} \oplus r\mathbf{u}) \frac{1}{r^2} dr d\nu(\mathbf{u}) - \int_{[0, \mathbf{x}]^c} f(\mathbf{x}) \frac{1}{r^2} dr d\nu(\mathbf{u}) = \int_{E_{\text{pol}}} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}).$$

The only thing left to do is proving that the convergence takes place in norm $\mathbf{L}^2(E_0, \Phi_{1,\nu})$. This is done in a similar manner as the proof of right-continuity, but with much heavier notations, so we skip this part.

□

As in the 1-dimensional case, the max-stable Ornstein-Uhlenbeck semi-group enjoys several useful properties which we sum up in the next proposition:

Proposition 2.3.7 *The multivariate Fréchet semi-group satisfies the following.*

1. (Commutation rule) Set:

$$\mathbf{D}_{1,\nu} f(\mathbf{x}) := \int_{E_{\text{pol}}} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}), \quad \mathbf{x} \in E_0.$$

Then for all $f \in \mathcal{S}_{\Phi,d}$ we have:

$$\mathbf{D}_{1,\nu} \mathbf{P}_t^{1,\nu} f(\mathbf{x}) = e^{-t} \mathbf{P}_t^{1,\nu} \mathbf{D}_{1,\nu} f(\mathbf{x}), \quad t \geq 0.$$

2. The operator $\mathbf{D}_{1,\nu}$ can be rewritten as:

$$\mathbf{D}_{1,\nu}f(\mathbf{x}) = \int_{E_{\text{pol}}} \langle r\mathbf{u}\mathbf{1}_{r\mathbf{u},\mathbf{x}}, \nabla f(\mathbf{x} \oplus r\mathbf{u}) \rangle \frac{1}{r^2} dr d\nu(\mathbf{u}), \quad f \in \mathcal{S}_{\Phi,d} \quad (2.70)$$

where $\mathbf{1}_{\mathbf{y},\mathbf{x}}^j$ equals 1 if $y^j \geq x^j$ and 0 otherwise, i.e. $\mathbf{1}_{\mathbf{y},\mathbf{x}}^j = \mathbb{1}_{\{y^j \geq x^j\}}$.

3. This is also equal to:

$$\mathbf{D}_{1,\nu}f(\mathbf{x}) = \sum_{j=1}^d \int_{\{r\mathbf{u}^j \geq x^j\}} u^j \partial_j f(\mathbf{x} \oplus r\mathbf{u}) \frac{1}{r} dr d\nu(\mathbf{u}), \quad f \in \mathcal{S}_{\Phi,d}, \quad (2.71)$$

with $\{r\mathbf{u}^j \geq x^j\}$ the subset of E_{pol} of (r, \mathbf{u}) such that $r u^j \geq x^j$.

Proof. 1. First we compute $\mathbf{P}_t^{1,\nu} \mathbf{D}_{1,\nu}f(\mathbf{x})$:

$$\begin{aligned} \mathbf{P}_t^{1,\nu} \mathbf{D}_{1,\nu}f(\mathbf{x}) &= \mathbb{E}[(\mathbf{D}_{1,\nu}f)(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z})] \\ &= \mathbb{E}\left[\int_{E_{\text{pol}}} \left(f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z} \oplus r\mathbf{u}) - f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z})\right) \frac{1}{r^2} dr d\nu(\mathbf{u})\right] \\ &= \int_{E_{\text{pol}}} \mathbb{E}\left[f(e^{-t}\mathbf{x} \oplus r\mathbf{u} \oplus (1 - e^{-t})\mathbf{Z}) - f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z})\right] \frac{1}{r^2} dr d\nu(\mathbf{u}). \end{aligned}$$

On the other hand, a change of variable yields:

$$\begin{aligned} \mathbf{D}_{1,\nu} \mathbf{P}_t^{1,\nu} f(\mathbf{x}) &= \int_{E_{\text{pol}}} \left(\mathbf{P}_t^{1,\nu} f(\mathbf{x} \oplus r\mathbf{u}) - \mathbf{P}_t^{1,\nu} f(\mathbf{x})\right) \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \int_{E_{\text{pol}}} \mathbb{E}\left[\left(f(e^{-t}(\mathbf{x} \oplus r\mathbf{u}) \oplus (1 - e^{-t})\mathbf{Z}) - f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z})\right) \frac{1}{r^2} dr d\nu(\mathbf{u})\right] \\ &= e^{-t} \int_{E_{\text{pol}}} \mathbb{E}\left[\left(f(e^{-t}\mathbf{x} \oplus r\mathbf{u} \oplus (1 - e^{-t})\mathbf{Z}) - f(e^{-t}\mathbf{x} \oplus (1 - e^{-t})\mathbf{Z})\right) \frac{1}{r^2} dr d\nu(\mathbf{u})\right] \\ &= e^{-t} \mathbf{P}_t^{1,\nu} \mathbf{D}_{1,\nu}f(\mathbf{x}). \end{aligned}$$

2. To prove the alternate expression of $\mathbf{D}_{1,\nu}$, we write:

$$\begin{aligned} \mathbf{D}_{1,\nu}f(\mathbf{x}) &= \int_{\mathbb{S}_+^{d-1}} \int_{r\mathbf{u} \in [\mathbf{0}, \mathbf{x}]^c} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \int_{\mathbb{S}_+^{d-1}} \int_{\min \frac{\mathbf{u}}{\mathbf{x}}}^{\infty} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}). \end{aligned}$$

Next, we perform an integration-by-parts in the inner integral, integrating $r \mapsto r^{-2}$ and differentiating $r \mapsto f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})$. Notice that this last function has at least one argument which is not constant with respect to r by definition of the integration domain. This accounts for the presence of the vector $\mathbf{1}_{r\mathbf{u},\mathbf{x}}$. Since the function $r \mapsto \mathbf{x} \oplus r\mathbf{u}$ is piecewise \mathcal{C}^1 , the

integration-by-parts makes sense and we get:

$$\begin{aligned} \int_{\mathbb{S}_+^{d-1}} \int_{\min \frac{\mathbf{u}}{\mathbf{x}}}^{\infty} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}) &= \int_{\mathbb{S}_+^{d-1}} \int_{\min \frac{\mathbf{u}}{\mathbf{x}}}^{\infty} \langle \mathbf{u} \mathbf{1}_{r\mathbf{u}, \mathbf{x}}, \nabla f(\mathbf{x} \oplus r\mathbf{u}) \rangle \frac{dr}{r} d\nu(\mathbf{u}) \\ &= \int_{\{r\mathbf{u} \in [0, \mathbf{x}]^c\}} \langle r\mathbf{u} \mathbf{1}_{r\mathbf{u}, \mathbf{x}}, \nabla f(\mathbf{x} \oplus r\mathbf{u}) \rangle \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \int_{E_{\text{pol}}} \langle r\mathbf{u} \mathbf{1}_{r\mathbf{u}, \mathbf{x}}, \nabla f(\mathbf{x} \oplus r\mathbf{u}) \rangle \frac{1}{r^2} dr d\nu(\mathbf{u}). \end{aligned}$$

□

Example 5. Depending on the angular measure, the form of $\mathbf{D}_{1, \nu}$ changes drastically. For the sake of clarity, assume that the reference norm is the infinity norm $\|\cdot\|_{\infty}$ on \mathbb{R}^d .

- In the case of complete independence, $\nu = \sum_{j=1}^d \delta_{\mathbf{e}_j}$, where \mathbf{e}_j is the j -th vector of the canonical basis of \mathbb{R}^d , so that:

$$\begin{aligned} \mathbf{D}_{1, \nu} f(\mathbf{x}) &= \sum_{j=1}^d \int_0^{\infty} (f(\mathbf{x} \oplus r\mathbf{e}_j) - f(\mathbf{x})) \frac{1}{r^2} dr \\ &= \sum_{j=1}^d \int_{x^j}^{\infty} r \partial_j f(\mathbf{x} \oplus r\mathbf{e}_j) \frac{1}{r^2} dr, \quad \mathbf{x} \in E_0. \end{aligned}$$

Notice that $\mathbf{D}_{1, \nu} f(\mathbf{x})$ is still infinitely differentiable with respect to each x^j , because the \oplus operator has actually vanished already. This is explained by the very particular shape of the angular measure.

- On the other hand, in the case of complete dependence, i.e. $\nu = \delta_{\mathbf{1}}$, we get:

$$\begin{aligned} \mathbf{D}_{1, \nu} f(\mathbf{x}) &= \int_{\min \mathbf{x}}^{\infty} (f(\mathbf{x} \oplus r\mathbf{1}) - f(\mathbf{x})) \frac{1}{r^2} dr \\ &= \sum_{j=1}^d \int_{x^j}^{\infty} r \partial_j f(\mathbf{x} \oplus r\mathbf{1}) \frac{1}{r^2} dr, \quad \mathbf{x} \in E_0, \end{aligned}$$

where $\min \mathbf{x} = x^1 \odot \dots \odot x^d$. Notice that $\mathbf{D}_{1, \nu} f(\mathbf{x})$ is still differentiable with respect to each x^j once but not more in general.

A pseudo Leibniz rule still holds in higher dimensions for $\mathbf{D}_{1, \nu}$:

Proposition 2.3.8 — Pseudo Leibniz rule. For every $f, g \in \mathcal{S}_{\Phi, d}$, we have for $\mathbf{x} \in E_0$:

$$\begin{aligned} \mathbf{D}_{1, \nu}(fg)(\mathbf{x}) &= \mathbf{D}_{1, \nu} f(\mathbf{x})g(\mathbf{x}) + f(\mathbf{x})\mathbf{D}_{1, \nu} g(\mathbf{x}) \\ &\quad + \int_{E_{\text{pol}}} (f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x}))(g(\mathbf{x} \oplus r\mathbf{u}) - g(\mathbf{x})) \frac{1}{r^2} dr d\nu(\mathbf{u}). \end{aligned}$$

Proof. The proof relies on the same kind of decomposition we have used to prove proposition 2.1.12, with $r\mathbf{u}^{1/\alpha}$ replacing r . □

The commutation rule between $\mathbf{P}_t^{1,\nu}$ and $\mathbf{D}_{1,\nu}$ admits an "infinitesimal version", to quote the expression of Chafaï [2006], page 6. It corresponds to the commutator relation (2.17) we have proved in the univariate case through a direct computation. Things are less obvious in higher dimensions, but we give a general argument which works for other semi-groups, like the α -stable semi-groups.

Proposition 2.3.9 *We have the commutation relation:*

$$[\mathcal{L}_{1,\nu}, \mathbf{D}_{1,\nu}] = \mathbf{D}_{1,\nu}. \quad (2.72)$$

Proof. Equation (2.71) makes it clear that $\mathcal{L}_{1,\nu}\mathbf{D}_{1,\nu}$ and $\mathbf{D}_{1,\nu}\mathcal{L}_{1,\nu}$ are well-defined for every choice of angular measure. The same goes for $\mathbf{P}_t^{1,\nu}\mathbf{D}_{1,\nu}$ and $\mathbf{D}_{1,\nu}\mathbf{P}_t^{1,\nu}$ thanks to the commutation rule. We write

$$\begin{aligned} [\mathcal{L}_{1,\nu}, \mathbf{D}_{1,\nu}] &= \lim_{t \rightarrow 0^+} \frac{1}{t} [\mathbf{P}_t^{1,\nu} - \text{Id}, \mathbf{D}_{1,\nu}] \\ &= \lim_{t \rightarrow 0^+} \frac{1}{t} [\mathbf{P}_t^{1,\nu}, \mathbf{D}_{1,\nu}] \\ &= \lim_{t \rightarrow 0^+} \frac{1}{t} (\mathbf{P}_t^{1,\nu} \mathbf{D}_{1,\nu} - \mathbf{D}_{1,\nu} \mathbf{P}_t^{1,\nu}) \\ &= \lim_{t \rightarrow 0^+} \frac{1 - e^{-t}}{t} \mathbf{P}_t^{1,\nu} \mathbf{D}_{1,\nu} \\ &= \mathbf{D}_{1,\nu}. \end{aligned}$$

□

Remark 2.3.10. In this subsection we have made great use of the polar decomposition to express $\mathbf{D}_{1,\nu}$ and $\mathcal{L}_{1,\nu}$. The advantage of this formulation is that it fits our choice of parametrization of max-stable distributions, which relies on the stability index α and the angular measure ν . Another choice of parametrization would have consisted in using only the exponent measure μ instead of decomposing it as the product of the measure ρ_α and the angular measure ν . This yields more compact formulae, *e.g.*

$$\mathbf{D}_\mu f(\mathbf{x}) := \int_{E_0} (f(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x})) \, d\mu(\mathbf{y}) = \mathbf{D}_{1,\nu} f(\mathbf{x}),$$

if μ is the image measure of $r^{-2} \, dr \otimes d\nu(\mathbf{u})$ by $(r, \mathbf{u}) \mapsto r\mathbf{u}$. However, one drawback of that parametrization is that the homogeneity and diffusivity properties of μ become less easy to exploit.

■

2.3.2 The general case

To extend the results of the previous section to max-stable distributions, we will need a few notations and objects. First, for $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$, denote by T_φ the mapping defined by $T_\varphi f = f \circ \varphi$ for all $f : \mathbb{R}^d \rightarrow \mathbb{R}$ such that this composition is well-defined. It is clear that if φ is invertible, then $T_\varphi^{-1} = T_{\varphi^{-1}}$.

Second, recall the inverse probability integral transform: if X is a real random variable with c.d.f. F_X , then $F_X(X)$ has the uniform distribution $\mathcal{U}_{[0,1]}$. Equivalently, $-1/\log F_X(X)$ has the unit

Fréchet distribution $\mathcal{F}(1)$, and thus

$$X \stackrel{d}{=} F_X^{\leftarrow}(\exp(-Z^{-1}))$$

where $Z \sim \mathcal{F}(1)$ and F_X^{\leftarrow} denotes the pseudo-inverse of F_X . Such a transformation can be applied to any simple max-stable random vector *i.e.* max-stable random vector with unit Fréchet margins, to give it any desired margins. In the sequel X will often be a max-stable distribution, so will mainly work with invertible functions such as $x \mapsto x^\alpha$ or $x \mapsto \log x$, so we are not losing much in generality by assuming that F_X is invertible. The next lemma is basic in multivariate extreme value theory.

Lemma 2.3.11 *Let \mathbf{X} be a max-stable random vector in \mathbb{R}^d . Denote by F_{X_j} the c.d.f. of its j -th marginal and set $\varphi := (-1/\log F_{X_1}, \dots, -1/\log F_{X_d})$. Assume that each F_{X_j} is invertible. Then $\varphi(\mathbf{X})$ is a simple max-stable random vector.*

Remark 2.3.12. In the previous lemma, is not possible to assume that \mathbf{X} is only max-id instead of max-stable in general. Otherwise it would mean that all max-id random vectors could be normalized into simple max-stable random vectors through a non-decreasing transformation. This is not true, as the following counterexample shows:

$$F(x, y) := \begin{cases} (e^{\frac{1}{x}} + e^{\frac{1}{y}} - 1)^{-1} & \text{if } x, y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

This expression defines a bivariate c.d.f. with unit Fréchet $\mathcal{F}(1)$ margins. Furthermore the function $Q(x, y) := -\log F(x, y)$ satisfies $\partial_{x,y}^2 Q(x, y) \leq 0$ for all $x, y \in \mathbb{R}_+$. By a criterion given in Resnick [1987] (proposition 5.3 page 254), this implies that F is the c.d.f. of a max-id random vector. However, it is clear that F is not max-stable. This implies that max-id random vectors may have a more complex dependency structure than max-stable random vectors, even when they have the same margins. On the other hand, lemma 2.3.11 shows why max-stable random vectors can be parametrized just by using a vector of shape parameters α and a finite measure on \mathbb{S}_+^{d-1} . ■

Let $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu)$ and denote by φ_α the non-decreasing ν transformation which turns the distribution of \mathbf{Z} into $\mathcal{MS}(\mathbf{1}, \nu)$, that is $\varphi_\alpha := (-1/\log F_{Z_1}, \dots, -1/\log F_{Z_d})$ For ease of notation, we will write $T_{\varphi_\alpha} = T_\alpha$. The standard algebra is defined by applying φ_α to $\mathcal{S}_{\Phi, d}$:

Definition 2.3.13 (Standard algebra). We define the class of functions:

$$\mathcal{S}_{\alpha, d} := \{f \circ \varphi, f \in \mathcal{S}_{\Phi, d}\} = T_\alpha \mathcal{S}_{\Phi, d}.$$

The general definition of our semi-group for max-stable random vectors then goes as follows:

Definition 2.3.14 (Multivariate max-stable semi-group). Define the *multivariate unit Fréchet semi-group* $(\mathbf{P}_t^{\alpha, \nu})_{t \geq 0}$ on the standard algebra $\mathcal{A} = \mathcal{S}_{\alpha, d}$ by setting

$$\mathbf{P}_t^{\alpha, \nu} f(\mathbf{x}) := (T_\alpha \mathbf{P}_t^{1, \nu} T_\alpha^{-1}) f(\mathbf{x}), \quad t \geq 0, \quad \mathbf{x} \in \varphi(E_0) \quad (2.73)$$

where $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu)$, with ν a finite measure on \mathbb{S}_+^{d-1} satisfying the moment constraints (1.18).

This definition makes immediate several properties of $(\mathbf{P}_t^{\alpha, \nu})_{t \geq 0}$, such as the semi-group property

$$\mathbf{P}_t^{\alpha, \nu} \mathbf{P}_s^{\alpha, \nu} = (T_\alpha \mathbf{P}_t^{1, \nu} T_\alpha^{-1})(T_\alpha \mathbf{P}_s^{1, \nu} T_\alpha^{-1}) = T_\alpha \mathbf{P}_{t+s}^{1, \nu} T_\alpha^{-1} = \mathbf{P}_{t+s}^{\alpha, \nu}.$$

The fact that $\mathcal{MS}(\boldsymbol{\alpha}, \nu)$ is an invariant measure of $(\mathbf{P}_t^{\boldsymbol{\alpha}, \nu})_{t \geq 0}$, as well as the ergodicity and right-continuity of this semi-group are other easy consequences of its definition and the properties of $(\mathbf{P}_t^{1, \nu})_{t \geq 0}$. Its generator can be expressed as:

$$\mathcal{L}_{\boldsymbol{\alpha}, \nu} = T_{\boldsymbol{\alpha}} \mathcal{L}_{1, \nu} T_{\boldsymbol{\alpha}}^{-1} = T_{\boldsymbol{\alpha}} (\mathfrak{d}_1 + \mathbf{D}_{1, \nu}) T_{\boldsymbol{\alpha}}^{-1} = T_{\boldsymbol{\alpha}} \mathfrak{d}_1 T_{\boldsymbol{\alpha}}^{-1} + \mathbf{D}_{\boldsymbol{\alpha}, \nu},$$

where $\mathfrak{d}_1 f(\mathbf{x}) = -\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle$ is the generator of the dilatation semi-group in dimension d , and $\mathbf{D}_{\boldsymbol{\alpha}, \nu}$ is defined as $\mathbf{D}_{\boldsymbol{\alpha}, \nu} := T_{\boldsymbol{\alpha}} \mathbf{D}_{1, \nu} T_{\boldsymbol{\alpha}}^{-1}$. This operator also satisfies the commutation rule:

$$\begin{aligned} \mathbf{D}_{\boldsymbol{\alpha}, \nu} \mathbf{P}^{\boldsymbol{\alpha}, \nu} &= (T_{\boldsymbol{\alpha}} \mathbf{D}_{1, \nu} T_{\boldsymbol{\alpha}}^{-1}) (T_{\boldsymbol{\alpha}} \mathbf{P}_t^{1, \nu} T_{\boldsymbol{\alpha}}^{-1}) \\ &= T_{\boldsymbol{\alpha}} \mathbf{D}_{1, \nu} \mathbf{P}_t^{1, \nu} T_{\boldsymbol{\alpha}}^{-1} \\ &= e^{-t} T_{\boldsymbol{\alpha}} \mathbf{P}_t^{1, \nu} \mathbf{D}_{1, \nu} T_{\boldsymbol{\alpha}}^{-1} \\ &= e^{-t} \mathbf{P}_t^{\boldsymbol{\alpha}, \nu} \mathbf{D}_{\boldsymbol{\alpha}, \nu}. \end{aligned}$$

For instance, when $\alpha > 0$ and $\boldsymbol{\alpha} = \alpha \mathbf{1}$, $T_{\boldsymbol{\alpha}}(\mathbf{x}) = \mathbf{x}^{\alpha}$, and the generator of the multivariate Fréchet semi-group with parameters $\boldsymbol{\alpha}$ and ν is

$$\begin{aligned} \mathcal{L}_{\boldsymbol{\alpha}, \nu} f(\mathbf{x}) &= -\frac{1}{\alpha} \langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \int_{E_{\text{pol}}} (f(\mathbf{x} \oplus r \mathbf{u}^{1/\alpha}) - f(\mathbf{x})) \frac{\alpha}{r^{\alpha+1}} \, dr \, d\nu(\mathbf{u}) \\ &= -\frac{1}{\alpha} \langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle + \frac{1}{\alpha} \int_{E_{\text{pol}}} \langle r \mathbf{u}^{1/\alpha} \mathbf{1}_{r\mathbf{u}, \mathbf{x}}, \nabla f(\mathbf{x} \oplus r \mathbf{u}^{1/\alpha}) \rangle \frac{\alpha}{r^{\alpha+1}} \, dr \, d\nu(\mathbf{u}). \end{aligned} \quad (2.74)$$

In dimension 1, the application T_{α} is actually a *Lie algebra isomorphism* on the Lie algebra spanned by $\text{span}(\mathcal{L}_1, \mathbf{D}_1, \delta_1)$, in the sense that:

$$T_{\alpha}[\phi_1, \phi_2] = [T_{\alpha} \phi_1, T_{\alpha} \phi_2], \quad \forall \phi_1, \phi_2 \in \text{span}(\mathcal{L}_1, \mathbf{D}_1, \delta_1).$$

We thus see that δ_{α} satisfies $T_{\alpha} \delta_1$, just like \mathcal{L}_{α} and \mathbf{D}_{α} . This gives a short proof of proposition 2.1.31 and 2.1.53.

We conclude this section by noticing that since all univariate distributions are max-id and have the same dependency structure (!), one can quantize them by resorting to the same method we have used to deal with max-stable distributions. For instance, if $Z \sim \mathcal{F}(1)$, we know that $\exp(-1/Z) \sim \mathcal{U}_{[0,1]}$, so a semi-group quantifying the uniform distribution is given by

$$P_t f(x) = \mathbb{E} \left[f(x^{e^t} \oplus U^{\frac{1}{1-e^{-t}}}) \right], \quad x \in [0, 1]$$

where $U \sim \mathcal{U}[0, 1]$, while for the logistic distribution with c.d.f. $(1 + e^{-x})^{-1}$ on \mathbb{R} , we get

$$P_t f(x) = \mathbb{E} \left[f \left(-\log \left((1 + e^{-x})^t - 1 \right) \oplus -\log \left(e^{\frac{1}{(1-e^{-t})Z}} - 1 \right) \right) \right], \quad x \in \mathbb{R}$$

since $-\log(e^{1/Z} - 1)$ has the logistic distribution if Z has the unit Fréchet distribution. This time we cannot give a Mehler formula for $P_t f(x)$ by using a random variable having the target logistic distribution. This stems from the fact that the distribution of a maximum of two i.i.d. logistic random variables is not easily expressed in terms of one logistic distribution. However, the maximum of two i.i.d. uniform random variables U_1, U_2 has the same distribution as $\sqrt{U_1}$, which explains the simpler expression of the former semi-group. In the same way, the *minimum* of two

i.i.d. Pareto $\mathcal{VP}(1)$ random variables X_1, X_2 has the same distribution as $\sqrt{X_1}$, so a semi-group quantifying this distribution would be:

$$P_t f(x) = \mathbb{E} \left[f(x^{e^t} \odot X^{\frac{1}{1-e^{-t}}}) \right], \quad x \in [1, +\infty),$$

where $X \sim \mathcal{VP}(1)$. Here we have used the *nonincreasing* transformation $z \mapsto \exp(1/Z)$ to change from the Fréchet distribution to the Pareto one, hence the presence of the minimum operator \odot instead of \oplus .

3 Stein’s method for the Fréchet distribution

Contents

3.1	The univariate Fréchet distribution	105
3.1.1	Stein’s equation and its solution	105
3.1.2	Rates of convergence to the Fréchet distribution	111
3.2	Extension in higher dimension	121
3.2.1	Regularity of Stein’s solutions in higher dimension	122
3.2.2	Applications	128

In this chapter we show how the Fréchet semi-group we have introduced may prove useful to apply Stein’s method to extreme value distributions. We start with the univariate case, where we give a few results regarding the Stein’s solution associated to \mathcal{L}_α , and then apply the obtained results to prove a bound on the distance between renormalized maxima of random variables and the Fréchet distribution. We then proceed to show how those results extend in higher dimension and give some further applications.

3.1 The univariate Fréchet distribution

3.1.1 Stein’s equation and its solution

We first check that the generator \mathcal{L}_α is indeed a characterizing operator of the Fréchet distribution Φ_α . The proof relies on the Mellin transform and is not given. Instead we will prove this result in the case of the Gumbel distribution; the proof will make use of the more familiar Laplace transform. The space of test-functions is chosen to be as large as possible while ensuring that each part of $\mathcal{L}_\alpha f$ is well-defined and integrable.

Proposition 3.1.1 *Let X be a random variable on \mathbb{R}_+ such that $\mathbb{E}[X^{-1}]$ is finite. Then X has the Fréchet distribution $\mathcal{F}(\alpha)$ if and only if*

$$\mathbb{E}[X f'(X)] = \alpha \mathbb{E}[\mathbf{D}_\alpha f(X)],$$

for all $f \in \mathcal{C}^1(\mathbb{R}_+^)$ such that there exists a constant C for which $|f'(z)| \leq Cz^{-1}$ for z large enough. If $\alpha > 1$ and $X \sim \mathcal{F}(\alpha)$, the identity holds true if f is assumed to be Lipschitz instead.*

The generator \mathcal{L}_α of $(\mathbf{P}_t^\alpha)_{t \geq 0}$ can be written as

$$\mathcal{L}_\alpha f(x) = -\frac{1}{\alpha} x f'(x) + \int_x^\infty f'(r) \frac{dr}{r^\alpha}, \quad x > 0, \quad f \in \mathcal{S}_\Phi.$$

In the sequel, we will need to work with test functions which do not belong to this space, such as 1-Lipschitz functions. Therefore we must make sure $\mathbf{P}_t^\alpha f$ and $\mathcal{L}_\alpha f$ have a sense. Furthermore, $\mathbf{P}_t^\alpha f$ must be nice enough so that $-\int_0^\infty \mathbf{P}_t^\alpha f dt = \mathcal{L}_\alpha^{-1} f$ is defined for centered f . Answering those questions is the content of the next proposition.

Proposition 3.1.2 *Let $\alpha > 0$ and $Z \sim \mathcal{F}(\alpha)$. Set $h^* = h - \mathbb{E}[h(Z)]$ if h is \mathbb{P}_Z -integrable. We have the following:*

- $\mathbf{P}_t^\alpha h$ and $\mathcal{L}_\alpha h$ exist for all $h \in \text{Lip}_1(\mathbb{R}_+, \mathbb{R})$ if and only if $\alpha > 1$.
- For all $z \in \mathbb{R}_+$, set $h_z = \mathbf{1}_{(-\infty, z]}$. Then $\mathbf{P}_t^\alpha h_z$ is well defined.

In both cases, $t \mapsto \mathbf{P}_t^\alpha h^*(x)$ is integrable for all $x \in \mathbb{R}_+^*$, so that $\mathcal{L}_\alpha^{-1} h^*$ exists. In the second case, the pseudo-inverse of h_z^* equals:

$$\mathcal{L}_\alpha^{-1} h_z^*(x) = \alpha(\log xz^{-1})_+ e^{-\frac{1}{z^\alpha}} - \int_{\alpha(\log xz^{-1})_+}^\infty \left(e^{-\frac{1-e^{-t}}{z^\alpha}} - e^{-\frac{1}{z^\alpha}} \right) dt. \quad (3.1)$$

Proof. - If $\alpha \leq 1$, take $h = \text{Id}$. This function is 1-Lipschitz, but if $Z \sim \mathcal{F}(\alpha)$, then $e^{-t/\alpha}x \oplus (1 - e^{-t})^{1/\alpha}Z \geq (1 - e^{-t})^{1/\alpha}Z$ is not integrable for all $t > 0$. On the other hand, if $\alpha > 1$, then we have

$$|h(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{1/\alpha}Z)| \leq h(0) + e^{-\frac{t}{\alpha}}x + Z,$$

because f is 1-Lipschitz on \mathbb{R}_+ . As Z is integrable, this proves that $\mathbf{P}_t^\alpha f$ is defined on \mathbb{R}_+ . Now, recall that a 1-Lipschitz function is differentiable a.e. and that its derivative h' is bounded by 1 a.e. (Villani [2008]). Thus we find

$$\int_x^\infty |h'(r)| \frac{dr}{r^\alpha} \leq \int_x^\infty \frac{dr}{r^\alpha} = \frac{1}{\alpha - 1} \frac{1}{x^{\alpha-1}}, \quad x > 0,$$

meaning that $\mathcal{L}_\alpha h(x)$ is well-defined for all positive x .

- Things are even simpler in this case, since we have for all nonnegative z

$$\mathbf{P}_t^\alpha h_z(x) = \mathbb{P}(e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{1/\alpha}Z \leq z) = e^{-\frac{1-e^{-t}}{z^\alpha}} \mathbf{1}_{\{e^{-t/\alpha}x \leq z\}}, \quad x \geq 0. \quad (3.2)$$

If h is 1-Lipschitz, then

$$\begin{aligned} |\mathbf{P}_t^\alpha h^*(x)| &= |\mathbf{P}_t^\alpha h(x) - \mathbb{E}[h(Z)]| \\ &\leq \mathbb{E}[|e^{-\frac{t}{\alpha}}x \oplus (1 - e^{-t})^{1/\alpha}Z - Z|] \\ &\leq \mathbb{E}[|e^{-\frac{t}{\alpha}}x - Z| \mathbf{1}_{\{Z \leq \gamma_t^{-1/\alpha}x\}}] + (1 - (1 - e^{-t})^{1/\alpha}) \\ &\leq \mathbb{E}[|e^{-\frac{t}{\alpha}}x - Z|^\beta]^{\frac{1}{\beta}} \mathbb{P}(Z \leq \gamma_t^{-1/\alpha}x)^{\frac{1}{\beta}} + (1 - (1 - e^{-t})^{1/\alpha}) \\ &= \mathbb{E}[|e^{-\frac{t}{\alpha}}x - Z|^\beta]^{\frac{1}{\beta}} e^{-\left(\frac{\gamma_t}{x^\alpha}\right)^{\frac{1}{\beta}}} + (1 - (1 - e^{-t})^{1/\alpha}), \end{aligned}$$

by Hölder's inequality, with $\beta = (\alpha + 1)/2$, so that $\beta \in (1, \alpha)$. The first term is clearly integrable with respect to t , while the second one is equivalent to e^{-t}/α when t goes to ∞ . Thus $t \mapsto \mathbf{P}_t^\alpha h(x) - \mathbb{E}[h(Z)]$ is integrable w.r.t. Lebesgue measure, and by proposition 1.1.6, $\mathcal{L}_\alpha^{-1} h^*(x)$ exists and is given by:

$$\mathcal{L}_\alpha^{-1} h^*(x) = - \int_0^\infty \mathbf{P}_t^\alpha h^*(x) dt, \quad x > 0.$$

In the case of the indicator functions h_z , expression (3.2) tells us that

$$\begin{aligned} \mathbf{P}_t^\alpha h_z^*(x) &= -e^{-\frac{1}{z^\alpha}} + e^{-\frac{1-e^{-t}}{z^\alpha}} \mathbf{1}_{\{e^{-t/\alpha} x \leq z\}} \\ &= -e^{-\frac{1}{z^\alpha}} + e^{-\frac{1-e^{-t}}{z^\alpha}} \mathbf{1}_{\{t \geq \alpha(\log xz^{-1})_+\}} \\ &= -e^{-\frac{1}{z^\alpha}} \mathbf{1}_{\{t \leq \alpha(\log xz^{-1})_+\}} + (e^{-\frac{1-e^{-t}}{z^\alpha}} - e^{-\frac{1}{z^\alpha}}) \mathbf{1}_{\{t \geq \alpha(\log xz^{-1})_+\}}. \end{aligned}$$

The presence of the positive part is due to the constraint that t is always non-negative. The first term is equivalent to e^{-t}/z^α , and the second is null outside the bounded interval $[0, \alpha(\log xz^{-1})_+]$, and continuous w.r.t. t inside, hence the existence of $\mathcal{L}_\alpha^{-1} h_z^*(x)$. Equality (3.1) immediately follows. \square

We next list some properties of those Stein solutions, starting with the Lipschitz case.

Proposition 3.1.3 *Let $\alpha > 1$.*

1. *Let h be a 1-Lipschitz function on \mathbb{R}_+^* . The associated Stein solution $g_h := \mathcal{L}_\alpha^{-1} h^*$ is α -Lipschitz and its derivative satisfies:*

$$g'_h(x) = - \int_0^\infty e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} h'(e^{-\frac{t}{\alpha}} x) dt, \quad x > 0. \quad (3.3)$$

As a result, we have

$$|g'_h(x)| \leq \min(\alpha, x^\alpha), \quad x > 0. \quad (3.4)$$

2. *Consequently, $\mathbf{D}_\alpha g_h(0)$ is well-defined and finite.*
3. *Furthermore assume that $h \in \text{Lip}_{[2]}(\mathbb{R}_+^*, \mathbb{R})$, i.e. h and h' are Lipschitz on \mathbb{R}_+^* . Then g'_h is C_α -Lipschitz, with C_α a positive constant depending only on α .*

Proof. 1. We already know from (2.3) that for $f \in \mathcal{S}_\Phi$,

$$(\mathbf{P}_t^\alpha f)'(x) = e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} f'(e^{-\frac{t}{\alpha}} x), \quad x > 0.$$

It is not hard to check that $\mathbf{P}_t^\alpha f$ is still absolutely continuous if f is 1-Lipschitz, and that the previous identity keeps holding. By dominating this function by $e^{-t/\alpha}$, we can apply the dominated convergence theorem and differentiate g_h under the integral sign:

$$g'_h(x) = - \frac{d}{dx} \left(\int_0^\infty (\mathbf{P}_t^\alpha h(x) - \mathbb{E}[h(Z)]) dt \right) = \int_0^\infty e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} h'(e^{-\frac{t}{\alpha}} x) dt.$$

Applying the triangle inequality and using that $|h'(x)| \leq 1$ a.e., we get $|g'_h(x)| \leq \alpha$. Actually we can be more precise:

$$\begin{aligned} |g'_h(x)| &\leq \int_0^\infty e^{-\frac{t}{\alpha}} e^{-\frac{\gamma t}{x^\alpha}} dt \\ &= \int_0^\infty \frac{x^\alpha}{(x^\alpha u + 1)^{1+\frac{1}{\alpha}}} e^{-u} du. \end{aligned}$$

The second expression gives $|g'_h(x)| \leq x^\alpha$ since the denominator is greater than 1 for all u and x in \mathbb{R}_+^* .

2. This is a direct consequence of the previous item:

$$\int_0^\infty |g'_h(r)| \frac{1}{r^\alpha} dr \leq \int_0^\infty \min(\alpha, r^{-\alpha}) dr < +\infty.$$

3. Representation (3.3) tells us that g'_h is the integral of a product of two *bounded* Lipschitz function, since h is also supposed to be Lipschitz. The first one is $r \mapsto \exp(-\gamma_t r^{-\alpha})$, whose derivative $\alpha\gamma_t r^{-(\alpha+1)} \exp(-\gamma_t r^{-\alpha})$ is positive and whose maximum is

$$\left(1 + \frac{1}{\alpha}\right)^{1+\frac{1}{\alpha}} \gamma_t^{-\frac{1}{\alpha}} \leq 4\gamma_t^{-\frac{1}{\alpha}}$$

because $\alpha > 1$. Notice this bound is integrable w.r.t. t on \mathbb{R}_+ . The second Lipschitz function is $r \mapsto h(e^{-t/\alpha} r)$, which is $e^{-t/\alpha}$ -Lipschitz. Therefore the integrand is $4\gamma_t^{-1/\alpha} + e^{-t/\alpha}$ Lipschitz and g'_h is C_α -Lipschitz, with

$$C_\alpha = \int_0^\infty e^{-\frac{t}{\alpha}} (4\gamma_t^{-1/\alpha} + e^{-t/\alpha}) dt = \frac{2}{\alpha} + 4 \int_0^\infty e^{-\frac{t}{\alpha}} \gamma_t^{-1/\alpha} dt.$$

□

Remark 3.1.4. The interest of working with doubly Lipschitz functions, *i.e.* $h \in \text{Lip}_{[2]}(\mathbb{R}_+^*, \mathbb{R})$, lies in the lack of regularity of g_h when h is a Lipschitz function or some indicator function $\mathbb{1}_{(-\infty, z]}$. In any of those cases, g_h is Lipschitz but its weak derivative has no reason to be defined everywhere. This is not a cause of problems when working with continuous random variables, but it will be when it comes to comparing a discrete random variable to an extreme value distribution, as we shall do later (see section 4.2). ■

The Stein's solution for test functions equal to h_z being more explicit, it can be described more thoroughly.

Proposition 3.1.5 *Let $z \in \mathbb{R}_+^*$. Denote by g_z the pseudo-inverse of $h_z^* = h_z - \mathbb{P}(Z \leq z)$. We have the following:*

1. g_z is non-decreasing and constant over $[0, z]$:

$$g_z(x) = - \int_0^\infty \left(e^{-\frac{1-e^{-t}}{z^\alpha}} - e^{-\frac{1}{z^\alpha}} \right) dt, \quad x \in [0, z].$$

Furthermore we have the equivalent

$$g_z(x) \underset{x \rightarrow +\infty}{\sim} \alpha e^{-\frac{1}{z^\alpha}} \log x.$$

2. g_z is continuously differentiable over $\mathbb{R}_+^* \setminus \{z\}$ and its derivative equals:

$$g'_z(x) = \begin{cases} \frac{\alpha}{x} e^{-\left(\frac{1}{z^\alpha} - \frac{1}{x^\alpha}\right)} & \text{if } x > z \\ 0 & \text{otherwise.} \end{cases} \quad (3.5)$$

3. g'_z satisfies the inequality

$$|g'_z(x)| \leq \frac{\alpha}{x} \mathbf{1}_{\{x > z\}}, \quad x > 0. \quad (3.6)$$

Thus it also satisfies for all positive x, y :

$$\alpha^{-1} |g_z(x) - g_z(y)| \leq \frac{1}{z} |x \oplus z - y \oplus z| \odot |\log(x \oplus z) - \log(y \oplus z)|. \quad (3.7)$$

In particular, g_z is αz^{-1} -Lipschitz.

4. We have the inequality:

$$0 \leq \mathbf{D}_\alpha g_z(x) \leq \frac{1}{x^\alpha}, \quad x > 0.$$

More precisely

$$\mathbf{D}_\alpha g_z(x) = e^{-\frac{1}{z^\alpha}} \left(e^{\frac{1}{x^\alpha} \odot \frac{1}{z^\alpha}} - 1 \right), \quad x > 0. \quad (3.8)$$

Consequently, $\sup_{z \in \mathbb{R}_+} \mathbf{D}_\alpha g_z(0)$ is finite.

Proof. 1. The function $h_z = \mathbf{1}_{(-\infty, z]}$ is non-increasing and $(\mathbf{P}_t^\alpha)_{t \geq 0}$ is a Markov semi-group, thus $-\mathbf{P}_t^\alpha h_z^*$ is also non-decreasing. Integrating over $[0, +\infty)$ w.r.t. t , we get the monotony of g_z over $[0, z]$. In the same way, $x \mapsto h_z(e^{-t/\alpha}x)$ is constant on $[0, z]$ for $e^{-t/\alpha}x$ is always less than z for all $t \geq 0$. When x goes to $+\infty$, the integral converges to 0, leaving the logarithm term exploding to $+\infty$, hence the announced asymptotic equivalent. In particular g_z is not bounded.

2. The function $x \mapsto (\log xz^{-1})_+$ is clearly piecewise continuously differentiable over \mathbb{R}_+ . It is continuous everywhere, but its derivative is not in z . Consequently, g_z is the composition of a piecewise continuously differentiable function by the smooth function:

$$y \mapsto ye^{-\frac{1}{z^\alpha}} - \int_y^\infty \left(e^{-\frac{1-e^{-t}}{z^\alpha}} - e^{-\frac{1}{z^\alpha}} \right) dt.$$

As such, it is also piecewise continuously differentiable. An application of the chain rule yields the announced result:

$$g'_z(x) = \frac{\alpha}{x} e^{-\frac{1}{z^\alpha}} + \frac{\alpha}{x} \left(e^{-\frac{1-(zx^{-1})^\alpha}{z^\alpha}} - e^{-\frac{1}{z^\alpha}} \right) = \frac{\alpha}{x} e^{-\left(\frac{1}{z^\alpha} - \frac{1}{x^\alpha}\right)}.$$

for $x > z$. If x is less than z , we already know that g_z is constant and thus its derivative is null.

3. Assume $x \leq y$. By integrating g'_z over $[x, y]$ and using inequality (3.6), one finds

$$\begin{aligned} |g_z(x) - g_z(y)| &\leq \int_x^y \frac{\alpha}{r} \mathbf{1}_{\{r > z\}} dr \\ &= \alpha (\log y - \log(x \oplus z))_+ \\ &= \alpha |\log(y \oplus z) - \log(x \oplus z)|. \end{aligned}$$

Alternatively, one could have bounded $1/r$ by $1/z$ to obtain the $|x \oplus z - y \oplus z|$ term. Because $x \mapsto x \oplus z$ is 1-Lipschitz, one deduces from this last inequality that g_z is αz^{-1} -Lipschitz.

4. The fact that $\mathbf{D}_\alpha g_z$ is non-negative is a by-product of the monotony of g_z , which gives $g_z(x \oplus r) - g_z(x) \geq 0$. As for the second part of the inequality, one can use the bound obtained in the previous point to find

$$\mathbf{D}_\alpha g_z(x) = \frac{1}{\alpha} \int_x^\infty r g'_z(r) \, d\rho_\alpha \leq \int_x^\infty d\rho_\alpha(r) = \frac{1}{x^\alpha}.$$

A more precise result can be proved by using (3.5):

$$\begin{aligned} \mathbf{D}_\alpha g_z(x) &= \frac{1}{\alpha} \int_x^\infty r g'_z(r) \, d\rho_\alpha(r) \\ &= \int_{x \oplus z}^\infty e^{-\left(\frac{1}{z^\alpha} - \frac{1}{r^\alpha}\right)} \, d\rho_\alpha(r) \\ &= e^{-\frac{1}{z^\alpha}} \int_{x \oplus z}^\infty e^{\frac{1}{r^\alpha}} \, d\rho_\alpha(r) \\ &= e^{-\frac{1}{z^\alpha}} \int_0^{(x \oplus z)^{-\alpha}} e^r \, dr \\ &= e^{-\frac{1}{z^\alpha}} \left(e^{\frac{1}{(x \oplus z)^\alpha}} - 1 \right). \end{aligned}$$

Notice that we can check that indeed $\mathcal{L}_\alpha g_z(x) = h_z^*$ by using (3.5) and (3.8) together. Finally,

$$\mathbf{D}_\alpha g_z(0) = e^{-\frac{1}{z^\alpha}} \left(e^{\frac{1}{z^\alpha}} - 1 \right) = 1 - e^{-\frac{1}{z^\alpha}},$$

a function which is bounded w.r.t. $z \in \mathbb{R}_+$.

□

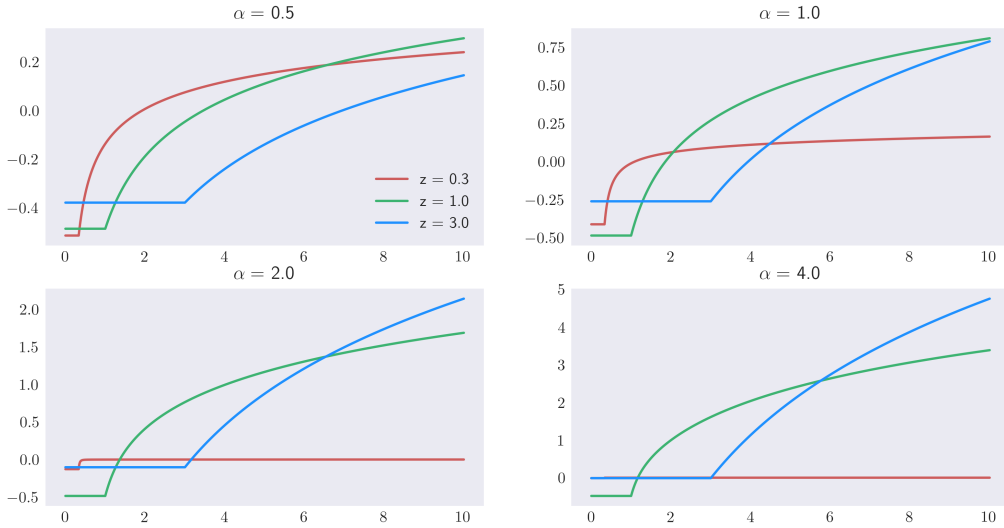


Figure 3.1: Graphs of g_z over $[1, 10]$ for $z \in \{\frac{1}{3}, 1, 3\}$ and $\alpha \in \{\frac{1}{2}, 1, 2, 4\}$

Notice that the order of magnitude of g_z decreases with respect to α . This is the opposite of what happens with Fréchet $\mathcal{F}(\alpha)$ random variables: the lower α is, the more explosive their values tend

to be. The constancy of g_z on $[0, z]$, its non-differentiability in z , as well as its logarithmic growth are easily observed too.

3.1.2 Rates of convergence to the Fréchet distribution

Before turning to applications, we must raise a technical difficulty specific to the Fréchet distribution (and to a lesser extent, to the Weibull distribution): it is possible that a sequence of random variables with support in a space larger than \mathbb{R}_+^* converges in distribution to a Fréchet distribution. This is the case of a renormalized maximum of n i.i.d. Cauchy random variables for instance. But the generator of the Fréchet semi-group is defined only on \mathbb{R}_+^* , so we must deal with the negative values of the approaching random variable before applying Stein's method. This is an instance of the so-called *problem of support* (see for example [Döbler \[2012\]](#) page 23).

We give a general argument in the case of continuous W . Assume that W is a real continuous random variable with density f_W , at least beyond a certain positive threshold K :

$$F_W(x) = F_W(K) + \int_K^x f_W(r) \, dr, \quad x \geq K.$$

Assume also that h is some nice test-function. Then we have

$$\begin{aligned} |\mathbb{E}[h(W)] - \mathbb{E}[h(Z)]| &= |\mathbb{E}[h^*(W)]| \\ &\leq |\mathbb{E}[h^*(W)\mathbb{1}_{\{W \leq K\}}]| + |\mathbb{E}[h^*(W)\mathbb{1}_{\{W > K\}}]| \end{aligned} \quad (3.9)$$

In applications, the first term usually goes to 0 fast (at exponential rate for instance), so it is the second one which will give the rate of convergence. To deal with it, we write

$$\begin{aligned} -\mathbb{E}[h^*(W)\mathbb{1}_{\{W > K\}}] &= \mathbb{E}[\mathcal{L}_\alpha \mathcal{L}_\alpha^{-1} h^*(W)\mathbb{1}_{\{W > K\}}] \\ &= \mathbb{E}[\mathcal{L}_\alpha g_h(W)\mathbb{1}_{\{W > K\}}] \\ &= -\frac{1}{\alpha} \mathbb{E}[W g'_h(W)\mathbb{1}_{\{W > K\}}] + \mathbb{E}[\mathbf{D}_\alpha g_h(W)\mathbb{1}_{\{W > K\}}]. \end{aligned}$$

Rewriting each term by using the density f_W , one gets

$$\begin{aligned} \alpha \mathbb{E}[h^*(W)\mathbb{1}_{\{W > K\}}] &= \int_K^\infty r g'_h(r) f_W(r) \, dr - \int_K^\infty r g'_h(r) \mathbb{P}(K < W \leq r) \, d\rho_\alpha(r) \\ &= \int_K^\infty r g'_h(r) f_W(r) \, dr - \int_K^\infty r g'_h(r) F_W(r) \, d\rho_\alpha(r) + F_W(K) \int_K^\infty r g'_h(r) \, d\rho_\alpha(r) \\ &= \int_K^\infty r g'_h(r) F_W(r) \left[\frac{f_W(r)}{F_W(r)} \, dr - d\rho_\alpha(r) \right] + F_W(K) \int_K^\infty r g'_h(r) \, d\rho_\alpha(r). \end{aligned}$$

We see that this error term is controlled by the distance between the measure ρ_α and the logarithmic derivative of F_W , plus a remainder which corresponds to the mass put by \mathbb{P}_W on \mathbb{R}_- . This gives us the following abstract bound on the Wasserstein distance and Kolmogorov distance between W and Z :

Proposition 3.1.6 *Let W be a random variable such that there exists some positive K for which F_W is absolutely continuous on $[K, +\infty)$. Assume that Z is a random variable with the Fréchet distribution $\mathcal{F}(\alpha)$ for some $\alpha > 0$. Let also \mathcal{H} be a space of test-functions.*

- *If $\mathcal{H} = \text{Lip}_1(\mathbb{R}, \mathbb{R})$, take $\alpha > 1$ and assume there exists $p \in (1, \alpha)$ such that $|W|^p$ is integrable. Let q be the conjugate of p i.e. $p^{-1} + q^{-1} = 1$. Then there exists a positive*

constant C_α such that we have

$$\begin{aligned} d_W(W, Z) &\leq \left(C_\alpha + \mathbb{E}[|W|^p]^{\frac{1}{p}} \right) F_W(K)^{\frac{1}{q}} \\ &\quad + \frac{1}{\alpha} \int_K^\infty r \min(\alpha, r^\alpha) F_W(r) \left| \frac{f_W(r)}{F_W(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr. \end{aligned} \quad (3.10)$$

- If $\mathcal{H} = \{h_z = x \mapsto \mathbb{1}_{(-\infty, z]}(x), z \in \mathbb{R}\}$, then

$$d_K(W, Z) \leq 2F_W(K) + \sup_{z > K} \int_z^\infty e^{-\left(\frac{1}{z^\alpha} - \frac{1}{r^\alpha}\right)} F_W(r) \left| \frac{f_W(r)}{F_W(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr. \quad (3.11)$$

Proof. - Since \mathbb{P}_W admits a first moment, and $\alpha > 1$, the Wasserstein distance between \mathbb{P}_W and \mathbb{P}_Z is finite. We start from (3.9). Consider the trivial coupling between \mathbb{P}_W and \mathbb{P}_Z : let W' and Z' be independent random variables defined on the same probability space, with distribution \mathbb{P}_W and \mathbb{P}_Z respectively. Then, for all 1-Lipschitz functions h , we see that

$$\begin{aligned} |\mathbb{E}[h^*(W)\mathbb{1}_{\{W \leq K\}}]| &\leq \mathbb{E}[|h(W') - h(Z')|\mathbb{1}_{\{W' \leq K\}}] \\ &\leq \mathbb{E}[|W' - Z'|\mathbb{1}_{\{W' \leq K\}}] \\ &\leq \mathbb{E}[|W|\mathbb{1}_{\{W \leq K\}}] + \mathbb{E}[Z]F_W(K) \\ &\leq \mathbb{E}[|W|^p]^{\frac{1}{p}} F_W(K)^{\frac{1}{q}} + \mathbb{E}[Z]F_W(K) \\ &\leq \left(C + \mathbb{E}[|W|^p]^{\frac{1}{p}} \right) F_W(K)^{\frac{1}{q}}, \end{aligned}$$

with $C = \mathbb{E}[Z] = \Gamma(1 - \alpha^{-1})$. We have applied Hölder's inequality to find the penultimate inequality. Next we deal with the second part of (3.9). Thanks to inequality (3.4), we can write

$$\begin{aligned} |\mathbb{E}[h^*(W)\mathbb{1}_{\{W > K\}}]| &\leq \frac{1}{\alpha} \int_K^\infty r |g'_h(r)| F_W(r) \left| \frac{f_W(r)}{F_W(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr + \frac{1}{\alpha} F_W(K) \int_K^\infty r |g'_h(r)| d\rho_\alpha(r) \\ &\leq \frac{1}{\alpha} \int_K^\infty r \min(\alpha, r^\alpha) F_W(r) \left| \frac{f_W(r)}{F_W(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr + F_W(K) \int_K^\infty \min(\alpha, r^\alpha) \frac{dr}{r^\alpha} \\ &\leq \frac{1}{\alpha} \int_K^\infty r \min(\alpha, r^\alpha) F_W(r) \left| \frac{f_W(r)}{F_W(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr + \left(1 + \frac{\alpha}{\alpha - 1} \right) F_W(K). \end{aligned}$$

Because $q \geq 1$ and $F_W(K) \in [0, 1]$, we have $F_W(K) \leq F_W(K)^{1/q}$, giving us the announced bound.

- Once more we start from (3.9). For all $z \in \mathbb{R}$, we have

$$\begin{aligned} |\mathbb{E}[h_z^*(W)\mathbb{1}_{\{W \leq K\}}]| &= |\mathbb{P}(W \leq z, W \leq K) - \mathbb{P}(Z \leq z)\mathbb{P}(W \leq K)| \\ &\leq \mathbb{P}(W \leq z, W \leq K) + \mathbb{P}(W \leq K) \\ &\leq 2\mathbb{P}(W \leq K). \end{aligned}$$

so that we have

$$|\mathbb{E}[h_z^*(W)\mathbb{1}_{\{W \leq K\}}]| \leq 2F_W(K), \quad z \in \mathbb{R}.$$

□

Remark 3.1.7. The assumption that F_W is absolutely continuous on $[K, +\infty)$ means our result covers the case of random variables which may not be continuous over the whole real line, but ultimately admit a density. ■

We now apply this line of reasoning to the classical question of computing rates of convergence in extreme value theory, in the case where the limiting distribution is Fréchet.

Assume that $(X_i)_{i \geq 1}$ is a sequence of i.i.d. random variables and set

$$M_n := \max_{1 \leq i \leq n} X_i \text{ and } Z_n := \frac{M_n - b_n}{a_n}, \quad n \geq 1$$

where $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$ are sequences of real numbers, with $a_n > 0$ for all n . Under reasonable assumptions on the common c.d.f. F_X of the X_i , we want to get an estimate of the rate at which the convergence of $(Z_n)_{n \geq 1}$ to the Fréchet distribution occurs. We already know that for $(Z_n)_{n \geq 1}$ to converge to the Fréchet distribution, \bar{F}_X must have infinite right end-point, *i.e.* $x_1 = +\infty$, and be regularly varying at $+\infty$ with order $\alpha > 0$, *i.e.* it must be of the form

$$\bar{F}_X(x) = \frac{L(x)}{x^\alpha}$$

with L a slowly-varying function at $+\infty$, for some $\alpha > 0$ and x large enough. Any function with a finite limit at $+\infty$ is slowly varying, but so is the logarithm. The asymptotic behaviour of L at $+\infty$ plays an important role in the determination of rates of convergence, so we need to make further assumptions on L . We follow the example of [Chen et al. \[2022\]](#) by restricting ourselves to the case where L converges to some limit ℓ when x goes to $+\infty$. Before stating and proving our result, we give a useful lemma.

Lemma 3.1.8 *Let $\alpha, \beta, \gamma \in \mathbb{R}_+^*$, with $\alpha + \beta > 1$. Then we have:*

$$\int_0^\infty \frac{e^{-\frac{\gamma}{r^\alpha}}}{r^{\alpha+\beta}} dr = \frac{1}{\alpha} \frac{1}{\gamma^{1+\frac{\beta-1}{\alpha}}} \Gamma\left(1 + \frac{\beta-1}{\alpha}\right).$$

Proof. The changes of variable $s = r^{-\alpha}$ and $t = \gamma s$ yield:

$$\begin{aligned} \int_0^\infty \frac{e^{-\frac{\gamma}{r^\alpha}}}{r^{\alpha+\beta}} dr &= \frac{1}{\alpha} \int_0^\infty s^{\frac{\beta-1}{\alpha}} e^{-\gamma s} ds \\ &= \frac{1}{\alpha} \frac{1}{\gamma^{1+\frac{\beta-1}{\alpha}}} \int_0^\infty t^{\frac{\beta-1}{\alpha}} e^{-t} dt \\ &= \frac{1}{\alpha} \frac{1}{\gamma^{1+\frac{\beta-1}{\alpha}}} \Gamma\left(1 + \frac{\beta-1}{\alpha}\right). \end{aligned}$$

□

Proposition 3.1.9 *Let $(X_i)_{i \geq 1}$ be a sequence of i.i.d. random variables with common distribution F_X . Suppose also there exists $\alpha > 0$ for which we have*

$$\bar{F}_X(x) \underset{x \rightarrow +\infty}{\sim} \frac{1}{x^\alpha}.$$

We further assume that there is a positive K such that \bar{F}_X is differentiable over $[K, +\infty)$ and

which satisfies:

$$\frac{\delta}{r^\alpha} \leq \bar{F}_X(r) \leq \frac{\delta'}{r^\alpha}, \quad r \geq K,$$

for some $\delta, \delta' > 0$. Take $a_n = n^{1/\alpha}$ and $b_n = 0$, so that $Z_n = n^{-1/\alpha} M_n$. Then:

- Assume that $\alpha > 1$ and that X_- , the negative part of X , is integrable. Then there exists a constant C_{α, F_X}^W depending only on α and F_X such that:

$$d_W(Z_n, Z) \leq C_{\alpha, F_X}^W \left[\frac{1}{n} + n^{1-\frac{1}{\alpha}} \int_K^\infty e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} |r^{\alpha+1} f_X(r) - \alpha| \frac{dr}{r^\alpha} \right], \quad n \geq 2. \quad (3.12)$$

- There exists another constant C_{α, F_X}^K depending only on α and F_X such that:

$$d_K(Z_n, Z) \leq C_{\alpha, F_X}^K \left[\frac{1}{n} + n \int_K^\infty e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} |r^{\alpha+1} f_X(r) - \alpha| \frac{dr}{r^{\alpha+1}} \right], \quad n \geq 2. \quad (3.13)$$

Proof. The assumption on the c.d.f. of X implies that it belongs to the domain of attraction of the Fréchet distribution $\mathcal{F}(\alpha)$ and also that X_+ is integrable. Since we have assumed its negative part to be integrable, this also implies that X is integrable, and thus Z_n too.

- We take care first of the mass that Z_n puts on $(-\infty, K]$. Because X belongs to the domain of attraction of $\mathcal{F}(\alpha)$, a minor modification of the proof of a result found in Resnick [1987] at page 77 yields that

$$\mathbb{E}[|Z_n|^p] \xrightarrow{n \rightarrow \infty} \mathbb{E}[Z^p] = \Gamma\left(1 - \frac{p}{\alpha}\right), \quad p \in (1, \alpha).$$

In particular this sequence of expectations is bounded. On the other hand,

$$F_{Z_n}(K)^{\frac{1}{q}} = F_X(K)^{\frac{n}{q}},$$

and this quantity decreases to 0 exponentially fast since $F_X(K)$ is strictly less than 1. Notice that in our setting, $(Z_n)_{n \geq 1}$ cannot converge in distribution to Z faster than at exponential speed, except in trivial cases. This is due to a result given in Rootzén [1984] which states that a faster than exponential rate of convergence implies that Z_n actually is a max-stable distribution for n large enough.

To deal with the values of z_n greater than K , notice first we can reformulate the assumption on F_X as

$$\bar{F}_X(r) = \frac{1 + \varepsilon(r)}{r^\alpha}, \quad r \geq K$$

where ε is a differentiable function vanishing at $+\infty$ and such that $1 + \varepsilon(r) \geq \delta$ if $r \geq K$. Since we have already dealt with the values of Z_n lower than K , we can make use of this assumption on X , and thus say that Z_n has a density $f_{Z_n}(r) = n a_n F_X(a_n r)^{n-1} f_X(a_n r)$. This implies in particular that:

$$\frac{f_{Z_n}(r)}{F_{Z_n}(r)} = n a_n \frac{f_X(a_n r)}{F_X(a_n r)}, \quad r \geq a_n^{-1} K.$$

Therefore we can continue the proof of proposition 3.1.6 to bound $|\mathbb{E}[h^*(W) \mathbb{1}_{\{W > K\}}]|$ (with

$K' = a_n^{-1}K$:

$$\begin{aligned}
|\mathbb{E}[h^*(W)\mathbb{1}_{\{W>K\}}]| &\leq \int_{\frac{K}{a_n}}^{\infty} r \min(\alpha, r^\alpha) F_X(a_n r)^n \left| n a_n \frac{f_X(a_n r)}{F_X(a_n r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&\leq \alpha \int_{\frac{K}{a_n}}^{\infty} F_X(a_n r)^n \left| n a_n \frac{f_X(a_n r)}{F_X(a_n r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&= \alpha a_n^{-2} \int_K^{\infty} r F_X(r)^n \left| n a_n \frac{f_X(r)}{F_X(r)} - n a_n \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&= \alpha n^{1-\frac{1}{\alpha}} \int_K^{\infty} r F_X(r)^n \left| \frac{f_X(r)}{F_X(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr.
\end{aligned}$$

Notice we have used that $a_n^{\alpha+1} = n a_n$. Now we need to investigate why $f_X/F_X = (\log F_X)'$ should be close to $r \mapsto \alpha r^{-(\alpha+1)}$. We re-write the expression in the absolute value as:

$$\frac{f_X(r)}{F_X(r)} - \frac{\alpha}{r^{\alpha+1}} = \frac{\alpha}{r^{\alpha+1}} \frac{\bar{F}_X(r)}{F_X(r)} + \frac{1}{r^{\alpha+1}} \frac{r^{\alpha+1} f_X(r) - \alpha}{F_X(r)}.$$

Plugging this expression in the previous integral and applying the triangle inequality, we obtain

$$\begin{aligned}
|\mathbb{E}[h^*(W)\mathbb{1}_{\{W>K\}}]| &\leq \alpha n^{1-\frac{1}{\alpha}} \int_K^{\infty} r F_X(r)^n \left| \frac{f_X(r)}{F_X(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&= \alpha n^{1-\frac{1}{\alpha}} \int_K^{\infty} F_X(r)^{n-1} |\alpha \bar{F}_X(r) + r^{\alpha+1} f_X(r) - \alpha| \frac{dr}{r^\alpha} \\
&\leq \alpha n^{1-\frac{1}{\alpha}} \int_K^{\infty} e^{-\frac{(n-1)\delta}{r^\alpha}} |\alpha \bar{F}_X(r) + r^{\alpha+1} f_X(r) - \alpha| \frac{dr}{r^\alpha} \\
&\leq \alpha n^{1-\frac{1}{\alpha}} \left(\delta' \alpha \int_K^{\infty} e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} \frac{dr}{r^{2\alpha}} + \int_K^{\infty} e^{-\delta \frac{n-1}{r^\alpha}} |r^{\alpha+1} f_X(r) - \alpha| \frac{dr}{r^\alpha} \right).
\end{aligned}$$

We have used that $n-1 \geq n/2$, as soon as $n \geq 2$. The penultimate bound stems from the assumption that $\bar{F}_X(r)$ is more than $\delta r^{-\alpha}$ if $r \geq K$. This implies in particular $1 + \varepsilon(r) \geq \delta$. The standard inequality $1-x \leq e^{-x}$ with $x = \bar{F}_X(r)$ then gives the exponential bound. We have also used that there exists some positive C such that $\bar{F}_X(r) \leq \delta' r^{-\alpha}$. The first integral is bounded *via* lemma 3.1.8:

$$\begin{aligned}
n^{1-\frac{1}{\alpha}} \int_K^{\infty} e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} \frac{dr}{r^{2\alpha}} &\leq n^{1-\frac{1}{\alpha}} \int_0^{\infty} e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} \frac{dr}{r^{2\alpha}} \\
&= \frac{2^{2-\frac{1}{\alpha}} n^{1-\frac{1}{\alpha}}}{\alpha \delta^{2-\frac{1}{\alpha}} n^{2-\frac{1}{\alpha}}} \Gamma\left(2 - \frac{1}{\alpha}\right) \\
&\lesssim \frac{1}{n}.
\end{aligned}$$

This concludes for the case of the Wasserstein distance.

- The case of the Kolmogorov distance is similar and relies on the same arguments. One must be careful that $r g'_z(r)$ is bounded by 1, so the change of variable $s = a_n r$ will make appear a_n^{-1}

before the integral, instead of a_n^{-2} as in the Wasserstein case:

$$\begin{aligned}
|\mathbb{E}[h^*(W)\mathbb{1}_{\{W>K\}}]| &\leq \sup_{z \geq a_n^{-1}K} \int_z^\infty e^{-(\frac{1}{z^\alpha} - \frac{1}{r^\alpha})} F_X(a_nr)^n \left| na_n \frac{f_X(a_nr)}{F_X(a_nr)} - \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&\leq \int_{\frac{K}{a_n}}^\infty F_X(a_nr)^n \left| na_n \frac{f_X(a_nr)}{F_X(a_nr)} - \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&= a_n^{-1} \int_K^\infty F_X(r)^n \left| na_n \frac{f_X(r)}{F_X(r)} - na_n \frac{\alpha}{r^{\alpha+1}} \right| dr \\
&= n \int_K^\infty F_X(r)^n \left| \frac{f_X(r)}{F_X(r)} - \frac{\alpha}{r^{\alpha+1}} \right| dr.
\end{aligned}$$

In particular we recover the n^{-1} term in the final bound:

$$n \int_K^\infty e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} \frac{dr}{r^{2\alpha+1}} \lesssim \frac{n}{n^2} = \frac{1}{n}.$$

□

Remark 3.1.10. - As mentioned in the introduction, quantifying the speed of convergence in univariate extreme-value theory is a problem which has been thoroughly studied before. A reference paper regarding this question is [Smith \[1982\]](#). By making use of the notion of regular variation with a remainder ([Goldie and Smith \[1987\]](#)), Smith was able to obtain sharp rates of convergence in various settings. One downside of this method is that the remainder may be difficult to compute for an arbitrary distribution function. By contrast, our approach gives rates of convergence expressed only with the density function, the main information one has about probability distributions in many situations. At the cost of a loss of generality and minimal assumptions on the c.d.f. we obtain a bound which is more convenient to use, as the next examples will show. Other approaches to derive rates of convergence in EVT can be found in [Resnick \[1987\]](#) and make use of different tools, such as von Mises functions. Once more, they come at the price of generality, but yield easier-to-use bounds.

- Our bounds yield speeds of convergence which cannot be faster than n^{-1} . In particular we will see in the examples that because of this term, we will not always get optimal rates of convergence. This is because we have bounded

$$|r^{\alpha+1}f_X(r) - \alpha F_X(r)| = |\alpha \bar{F}_X(r) + r^{\alpha+1}f_X(r) - \alpha|$$

by $\delta' \alpha r^{-\alpha} + |r^{\alpha+1}f_X(r) - \alpha|$. A closer look tells us that when r goes to $+\infty$, the term $r^{\alpha+1}f_X(r)$ converges to α and $\bar{F}_X(r)$ goes to 0, while $\alpha \bar{F}_X(r)$ tends to α and $r^{\alpha+1}f_X(r)$ vanishes when r goes to the left of the support of X (maybe below K). This heuristic suggests us to bound the previous expression by

$$|\alpha \bar{F}_X(r) + r^{\alpha+1}f_X(r) - \alpha| \leq |\alpha \bar{F}_X(r) - \alpha p(r)| + |r^{\alpha+1}f_X(r) - \alpha q(r)|,$$

where p, q are positive functions such that $p(r) + q(r) = 1$. Taking $p(r) = \bar{F}_X(r)$ brings us back to the initial inequality. One could also take $p(r) = \mathbb{1}_{\{r \leq c\}}$ and then try to optimize on c to get a better rate of convergence. What we have done amounts to choose $p(r) = 0$. We still get an interesting rate this way, and we do not need to use more information on \bar{F}_X to conclude.

- To the best of our knowledge, those results constitute one of the few applications of Stein's method to EVT so far. The other two we are aware of are [Bartholome and Swan \[2013\]](#) and [Kusumoto and Takeuchi \[2020\]](#)), as well as Feidt's doctoral thesis ([Feidt \[2013\]](#)). The first two papers define directly Stein's operators (the same one in the Fréchet case) through integration-by-parts, although only the former actually makes use of those results to get rates of convergence, by relying on certain arguments of Smith. In the second work, the author employs completely different arguments and instead relies on Stein's method for Poisson process to assess the distance between the point process of exceedances and the limiting Poisson process. She then deduces rates of convergence for the maximum of n random variables to one of the extreme value distributions *via* results such as [proposition 1.15](#). The connection between EVD and Poisson processes is well-known now, and we will also exploit it, although our strategy will be different from Feidt's.
- The approach we have made use of is quite different from what has been done in similar contexts, even in the framework of Stein's method. In [Xu \[2019\]](#), the author bounds the Wasserstein distance between a renormalized sum of i.i.d. random variables with regularly varying tails and a stable distribution by using a Taylor formula and a truncation argument. The Stein's operator is also given by a semi-group $(Q_t^\alpha)_{t \geq 0}$ which quantizes stable distributions. In [Chen et al. \[2022\]](#), the authors rely on the leave-one-out approach and use the regularity properties of the semi-group $(Q_t^\alpha)_{t \geq 0}$ to circumvent the non-integrability of their random variables, as $\alpha \in (0, 1]$. Those arguments do not seem to work with max-stable distributions, unlike our method. Its main drawback is that it relies heavily on the existence of a density for X_1, \dots, X_n (an assumption which also appears in [Chen et al. \[2022\]](#)). We will later a way to go beyond this limitation, at the cost of practicability.
- Those bonds hold for the Wasserstein distance, so they also control the stop-loss distance (see [1.4.1](#)). In particular, we have a rate of convergence for the sequence $(\mathbb{E}[(Z_n - z)_+])_{n \geq 1}$ to its limit $\mathbb{E}[(Z - z)_+]$ for any z in \mathbb{R} whenever α is greater than 1. ■

Example 3.1.11. As per the custom, we apply our result first to the case where $X \sim \mathcal{VP}(\alpha)$. This example is especially simple since in that case $K = 1$ and $r^{\alpha+1} f_X(r) - \alpha = 0$ on $[1, +\infty)$. Thus, only the n^{-1} error term remains for both metrics. In the case of the Kolmogorov distance, it is well-known since [Hall and Wellner \[1979\]](#) that a maximum of n i.i.d. Pareto random variables renormalized by $n^{1/\alpha}$ converges in the Kolmogorov distance to the Fréchet distribution $\mathcal{F}(\alpha)$ at rate n^{-1} . That the same result holds in the Wasserstein distance is new, to the best of our knowledge.

The next statement is a corollary of [proposition 3.1.9](#) and will be very convenient in the applications:

Corollary 3.1.12 *Assume that F_X is a distribution function satisfying the assumptions of [proposition 3.1.9](#). If its derivative also satisfies*

$$|r^{\alpha+1} f_X(r) - \alpha| \lesssim \frac{1}{r^\gamma}, \quad r \geq K \tag{3.14}$$

for some $\gamma > 0$, then

- in the Wasserstein case, if $\alpha > 1$, there exists C_{α, F_X}^W such that:

$$d_W(Z_n, Z) \leq C_{\alpha, F_X}^W \left[\frac{1}{n} + \frac{1}{n^{\frac{\alpha}{2}}} \right], \quad n \geq 1, \quad (3.15)$$

- in the Kolmogorov case, for all $\alpha > 0$, there exists C_{α, F_X}^K such that:

$$d_K(Z_n, Z) \leq C_{\alpha, F_X}^K \left[\frac{1}{n} + \frac{1}{n^{\frac{\alpha}{2}}} \right], \quad n \geq 1. \quad (3.16)$$

Proof. We treat only the Kolmogorov case (just keep in mind the constant C_{α, F_X} is different in the Wasserstein case and explodes when α tends to 1^+); by plugging the assumption into inequality (3.13) and using lemma 3.1.8 with $\beta = \gamma + 1$, we get

$$\begin{aligned} n \int_K^\infty e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} |r^{\alpha+1} f(r) - \alpha| \frac{dr}{r^{\alpha+1}} &\leq n \int_0^\infty e^{-\frac{\delta}{2} \frac{n}{r^\alpha}} \frac{dr}{r^{\alpha+\gamma+1}} \\ &\lesssim \frac{n}{n^{1+\frac{\alpha}{2}}} \\ &= \frac{1}{n^{\frac{\alpha}{2}}}. \end{aligned}$$

□

Example 3.1.13. We give various examples to illustrate the usefulness and simplicity of use of proposition 3.1.9 and corollary 3.1.12.

1. **Log-logistic distribution** - In this case $\bar{F}_X(r) = (1 + r^\alpha)^{-1}$ for $r \geq 0$. By taking $K = 1$, we find

$$\bar{F}_X(r) = \frac{1}{r^\alpha} \frac{1}{1 + r^{-\alpha}} \geq \frac{1}{2r^\alpha}, \quad r \geq 1$$

so that $\delta = 1/2$. We need to bound $r \mapsto r^\alpha f(r) - \alpha$ on $[1, \infty)$:

$$\begin{aligned} |r^{\alpha+1} f_X(r) - \alpha| &= \left| \alpha \frac{r^{2\alpha}}{(1 + r^\alpha)^2} - \alpha \right| \\ &= \alpha \left(1 - \frac{1}{(1 + r^{-\alpha})^2} \right) \\ &\leq \alpha ((1 + r^{-\alpha})^2 - 1) \\ &\leq \frac{2\alpha}{r^\alpha}, \end{aligned}$$

since the function $r \mapsto (1 + r)^2$ is 2-Lipschitz on $[0, 1]$. Therefore the assumption (3.14) is satisfied with $\gamma = \alpha$, and so inequality (3.15) holds, yielding a rate of convergence of n^{-1} in both the Wasserstein metric (if $\alpha > 1$) and the Kolmogorov metric.

2. **Student distribution** - This is our first example of distribution whose c.d.f. is not available in a closed form, and also whose support is \mathbb{R} . The density of the Student

distribution with ν degrees of freedom is explicit though:

$$f_X(r) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} \left(1 + \frac{r^2}{\nu}\right)^{-\frac{\nu+1}{2}} =: \frac{1}{C} \left(1 + \frac{r^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

This function is regularly varying of order $\nu + 1$, so we take $\alpha = \nu$. It is clear that $r^{\nu+1}f(r) - \nu$ will not vanish when r goes to ∞ . Therefore we will work with $\sigma^{-1}X$, where X has the Student distribution with parameter ν , and σ is a positive scalar to be determined. Its density $f_{\sigma^{-1}X}$ is:

$$f_{\sigma^{-1}X}(r) = \sigma f_X(\sigma r) = \frac{\sigma}{C} \left(1 + \frac{\sigma^2 r^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

Now we want to find σ such that $r^{\nu+1}f_{\sigma^{-1}X}(r) - \nu$ goes to 0, and to determine at which speed it does. As in the previous examples, we will take $K = 1$.

$$\begin{aligned} |r^{\nu+1}f_{\sigma^{-1}X}(r) - \nu| &= \left| r^{\nu+1} \frac{\sigma}{C} \left(1 + \frac{\sigma^2 r^2}{\nu}\right)^{-\frac{\nu+1}{2}} - \nu \right| \\ &= \left| \frac{\sigma}{C} (\nu^{-1}\sigma^2 + r^{-2})^{-\frac{\nu+1}{2}} - \nu \right| \\ &= \left| \frac{\nu^{\frac{\nu+1}{2}}}{C\sigma^\nu} (1 + \nu\sigma^{-2}r^{-2})^{-\frac{\nu+1}{2}} - \nu \right|. \end{aligned}$$

We see that σ must be equal to $\nu C^{-2/(\nu-1)}$ for the last term to vanish at infinity. This implies that $\sigma = C^{-1/\nu} \nu^{(\nu-1)/2\nu}$. Taking this value of σ , we can factorize to obtain

$$\begin{aligned} |r^{\nu+1}f_{\sigma^{-1}X}(r) - \nu| &= \nu \left| \frac{\nu^{\frac{\nu-1}{2}}}{C\sigma^\nu} (1 + \nu\sigma^{-2}r^{-2})^{-\frac{\nu+1}{2}} - 1 \right| \\ &= \nu \left| (1 + \nu\sigma^{-2}r^{-2})^{-\frac{\nu+1}{2}} - 1 \right| \\ &\lesssim \frac{1}{r^2}, \end{aligned}$$

as the function $r \mapsto (1+r)^{(\nu+1)/2}$ is Lipschitz on $[0, 1]$. Consequently assumption (3.14) is satisfied with $\delta = 2$ and we get

$$d(\sigma^{-1}Z_n, Z) \lesssim \frac{1}{n} + \frac{1}{n^{\frac{2}{\nu}}},$$

where d can be either the Wasserstein (for $\alpha > 1$) or Kolmogorov distance. Numerical estimations suggest that the true rate of convergence is always $n^{-2/\alpha}$, even when $\nu \in (0, 2)$.

3. **Generalized Beta prime distribution** - This distribution has four positive parameters α, β, p, q , the first three being shape parameters while q is a scale parameter. Its p.d.f. is

$$f_X(r) = \frac{p}{qB(\alpha, \beta)} \frac{(r/q)^{\alpha p - 1}}{(1 + (r/q)^p)^{\alpha + \beta}} = \frac{q^{p\beta}}{\beta B(\alpha, \beta)} \frac{p\beta}{r^{p\beta + 1}} \frac{1}{(1 + (r/q)^{-p})^{\alpha + \beta}}, \quad r > 0$$

where $B(\alpha, \beta)$ is the beta function in α and β . By choosing $q = B(\alpha, \beta)^{-1/p\beta}$, we obtain

for $r \geq K = 1$

$$\begin{aligned} |r^{p\beta+1}f(r) - p\beta| &= p\beta \left| \frac{1}{(1 + (r/q)^{-p})^{\alpha+\beta}} - 1 \right| \\ &= \frac{p\beta}{(1 + (r/q)^{-p})^{\alpha+\beta}} \left[\left(1 + \left(\frac{q}{r}\right)^p\right)^{\alpha+\beta} - 1 \right] \\ &\lesssim \frac{1}{r^p}. \end{aligned}$$

By the same arguments as in the previous examples, we find that:

$$d(\sigma^{-1}Z_n, Z) \lesssim \frac{1}{n} + \frac{1}{n^{\frac{p}{\beta}}}.$$

This distribution admits several interesting particular cases, like the beta prime distribution, the Fisher distribution, the Pareto distribution, the log-logistic distribution, as well as more exotic laws (Dagum, Singh–Maddala, *etc.*).

4. **Stable distributions** - A central family of distributions belonging to the domain of attraction of Φ_α are the stable distributions with index $\alpha \in (0, 2)$. Except when $\alpha = 1/2$ (Lévy distribution) or 1 (Cauchy distribution), they do not admit explicit density functions, but we still have the following result from Bergström [1953] for one-sided α -stable distributions:

$$f_X(x) = \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k-1} \sin(\alpha\pi k) \frac{\Gamma(\alpha k + 1)}{\Gamma(k + 1)} \frac{1}{x^{k\alpha+1}}, \quad x > 0.$$

The first term equals $\Gamma(\alpha + 1) \sin(\alpha\pi) x^{-(\alpha+1)}$, which means that we need to renormalize by $\sigma = (\Gamma(\alpha) \sin(\alpha\pi))^{1/\alpha}$. The remainder is then:

$$\begin{aligned} |r^{\alpha+1}f_{\sigma^{-1}X}(r) - \alpha| &= \frac{\alpha}{\pi} \left| \sum_{k=2}^{\infty} (-1)^{k-1} \sin(\alpha\pi k) \frac{\Gamma(\alpha k + 1)}{\Gamma(k + 1)} \frac{1}{r^{(k-1)\alpha}} \right| \\ &= \frac{\alpha}{\pi} \left| \sum_{k=1}^{\infty} (-1)^k \sin(\alpha\pi(k + 1)) \frac{\Gamma(\alpha(k + 1) + 1)}{\Gamma(k + 2)} \frac{1}{r^{k\alpha}} \right|. \end{aligned}$$

Furthermore, the density of stable distributions is bounded (see Zolotarev [1986] page 87), so the previous series is bounded too. The order of convergence is given by the first term of the series which is different of 0; let us call its index k_0 . Then the same method as before gives us a rate of convergence of $n^{-k_0\alpha/\alpha} + n^{-1} = n^{-k_0} + n^{-1}$. For instance, in the case of the Lévy distribution, which is a one-sided 1/2-stable distribution, the term of index 1 in the above series vanishes but not the second, so $k_0 = 2$, hence a rate of convergence of n^{-1} in Kolmogorov distance.

To conclude this section, we present another method to obtain rates of convergence in extreme value theory. It is inspired by the classic leave-one-out approach and relies on different arguments. We find it less practical to yield though, hence our preference for the above approach.

Remark 3.1.14. For ease of notations, set:

$$X_{n,i} := a_n^{-1}X_i \text{ and } M_{n \setminus i} := \max_{k \neq i} X_k.$$

In particular, $M_{n \setminus n} = M_{n-1}$. Set also $Z_{n,n-1} := a_n^{-1} M_{n-1}$. As before, one must show that

$$\mathbb{E}[Z_n g'(Z_n)] \simeq \alpha \mathbb{E}[\mathbf{D}_\alpha g(Z_n)] = \mathbb{E}[Z_n^{-(\alpha-1)} Y g'(Y Z_n)],$$

where $g = g_h$ is some Stein's solution and we have used the alternative expression of \mathbf{D}_α given in (2.9), with $Y \sim \mathcal{VP}(\alpha, 1)$. Take $\psi = x \mapsto x g'(x)$ and proceed as follows:

$$\begin{aligned} \mathbb{E}[\psi(Z_n)] &= \sum_{i=1}^n \mathbb{E}[\psi(X_{n,i}) \mathbf{1}_{\{X_i \geq M_{n \setminus i}\}}] \\ &= n \mathbb{E}[\psi(X_{n,n}) \mathbf{1}_{\{X_n \geq M_{n-1}\}}]. \end{aligned}$$

Conditioning on $z = M_{n-1}$ in the previous inequality, which is independent of $X_{n,n}$, we find:

$$\mathbb{E}[X_{n,n} g'(X_{n,n}) \mathbf{1}_{\{X_n \geq z\}}] = \bar{F}_X(z) \mathbb{E}[X_{n,n} g'(X_{n,n}) | X_n \geq z].$$

From this we deduce that:

$$\begin{aligned} \mathbb{E}[\psi(Z_n)] &= n \mathbb{E}[\psi(X_{n,n}) \mathbf{1}_{\{X_n \geq M_{n-1}\}}] \\ &= n \mathbb{E}\left[\mathbb{E}[\psi(X_{n,n}) \mathbf{1}_{\{X_n \geq M_{n-1}\}} | M_{n-1}]\right] \\ &= n \mathbb{E}\left[\bar{F}_X(M_{n-1}) \mathbb{E}[\psi(X_{n,n}) | X_n \geq M_{n-1}]\right] \\ &= n \mathbb{E}\left[\bar{F}_{a_n^{-1} X}(Z_{n,n-1}) \mathbb{E}[\psi(Z_{n,n-1} M_{n-1}^{-1} X_n) | X_n \geq M_{n-1}]\right], \end{aligned}$$

By using that $n \bar{F}_{a_n^{-1} X}(r) \xrightarrow[n \rightarrow \infty]{} r^{-\alpha}$ if \bar{F}_X is α^{-1} -varying, we find:

$$\begin{aligned} \mathbb{E}[\psi(Z_n)] &\simeq \mathbb{E}\left[Z_{n,n-1}^{-\alpha} \mathbb{E}[\psi(Z_{n,n-1} M_{n-1}^{-1} X_n) | X_n \geq M_{n-1}]\right] \\ &\simeq \mathbb{E}[Z_{n,n-1}^{-\alpha} \psi(Y Z_{n,n-1})] \\ &= \mathbb{E}[Z_{n,n-1}^{-(\alpha-1)} Y g'(Y Z_{n,n-1})]. \end{aligned}$$

The second approximation results from theorem 1.3.14 and the remark following it: conditional to M_{n-1} , for n large enough, the distribution of $M_{n-1}^{-1} X_n$ when $X_n \geq M_{n-1}$ is approximately a Pareto distribution $\mathcal{VP}(\alpha, 1)$ because M_{n-1} almost surely goes to infinity. One advantage of that method is that we do not need to assume that X admits a density for large enough values. On the other hand it relies on approximations which may be difficult to quantify in general. ■

3.2 Extension in higher dimension

The purpose of this section is to go beyond the univariate setting and to establish tools to quantify the speed of convergence to max-stable random vectors. To keep notations simple, we will focus on the multivariate max-stable distributions whose marginals are all Fréchet $\mathcal{F}(\alpha)$ for some $\alpha > 0$. When $\alpha = \alpha \mathbf{1}$ for some α , and ν is some angular measure on \mathbb{S}_+^{d-1} w.r.t. some norm, we will denote the associated semi-group by $\mathbf{P}_t^{\alpha, \nu} := \mathbf{P}_t^{\alpha, \nu}$.

3.2.1 Regularity of Stein's solutions in higher dimension

Here the benefits of working with the generator approach will be made obvious; although it is not too difficult to find a Stein operator for the univariate extreme-value distributions, as they have explicit p.d.f. and c.d.f., the same cannot be said about their multivariate counterparts. On the other hand, the Mehler's formula (2.68) through which we have defined the Fréchet semi-group $(\mathbf{P}_t^\alpha)_{t \geq 0}$ readily extends to higher dimensions, as we have seen in the previous chapter. This multivariate semi-group $(\mathbf{P}_t^{\alpha, \nu})_{t \geq 0}$ admits $\mathcal{MS}(\alpha, \nu)$ as its stationary distribution, and thus its generator $\mathcal{L}_{\alpha, \nu}$ is a Stein operator for this distribution:

$$\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu) \implies \mathbb{E}[\mathcal{L}_{\alpha, \nu} f(\mathbf{Z})] = 0, \quad f \in \mathcal{S}_{\alpha, d}. \quad (3.17)$$

We aim at using those tools to apply Stein's method to multivariate extreme value distribution. To do so, it is in our opinion instructive to see first how equality naturally appears. Its proof is the consequence of a general argument concerning semi-groups with invariant measure, but we want to understand why it works in this specific case through direct arguments. This will be of great help in the applications. For ease of notations, we assume provisionally $\alpha = 1$.

We must show that:

$$\mathbb{E}[\langle \mathbf{Z}, \nabla f(\mathbf{Z}) \rangle] = \mathbb{E}[\mathbf{D}_{1, \nu} f(\mathbf{Z})], \quad f \in \mathcal{S}_{\Phi, d}.$$

By Mecke's formula applied to the functional

$$\varphi : (\mathbf{x}, \eta) \mapsto \langle r\mathbf{u}\mathbf{1}_{r\mathbf{u}, \mathbf{m}(\eta)}, \nabla f(\mathbf{m}(\eta) \oplus r\mathbf{u}) \rangle.$$

we see that

$$\begin{aligned} \mathbb{E}[\mathbf{D}_{1, \nu} f(\mathbf{Z})] &= \mathbb{E}[\mathbf{D}_{1, \nu} f(\mathbf{m}(\eta))] \\ &= \int_{E_{\text{pol}}} \mathbb{E}[\langle r\mathbf{u}\mathbf{1}_{r\mathbf{u}, \mathbf{m}(\eta)}, \nabla f(\mathbf{m}(\eta) \oplus r\mathbf{u}) \rangle] \frac{1}{r^2} dr d\nu(\mathbf{u}) \\ &= \mathbb{E}\left[\int_{E_{\text{pol}}} \langle r\mathbf{u}\mathbf{1}_{r\mathbf{u}, \mathbf{m}(\eta - \delta_{(r, \mathbf{u})})}, \nabla f(\mathbf{m}(\eta) \oplus r\mathbf{u}) \rangle d\eta(r, \mathbf{u})\right] \\ &= \mathbb{E}\left[\int_{E_{\text{pol}}} \langle r\mathbf{u}\mathbf{1}_{r\mathbf{u}, \mathbf{m}(\eta - \delta_{(r, \mathbf{u})})}, \nabla f(\mathbf{m}(\eta)) \rangle d\eta(r, \mathbf{u})\right], \\ &= \mathbb{E}[\langle \mathbf{m}(\eta), \nabla f(\mathbf{m}(\eta)) \rangle], \end{aligned}$$

giving us the announced identity. The penultimate equality comes from the fact that there is no point (r, \mathbf{u}) in η such that any coordinate of $r\mathbf{u}$ is greater than the corresponding coordinate of $\mathbf{m}(\eta)$. The last identity follows by observing that for every $j \in \llbracket 1, d \rrbracket$, the only pairs $(r, \mathbf{u}) \in \eta$ such that the j -th coordinate of $\mathbf{1}_{r\mathbf{u}, \mathbf{m}(\eta - \delta_{(r, \mathbf{u})})}$ is not null corresponds to the ones giving the j -th coordinate of $\mathbf{m}(\eta)$.

Next we give some properties of $\mathbf{P}_t^{\alpha, \nu}$ and the pseudo-inverse of $\mathcal{L}_{\alpha, \nu}$. In spite of the additional complexity caused by the multivariate setting, describing Stein's solutions g_h is not much harder than in the univariate setting.

Proposition 3.2.1 *Let $\alpha > 0$ and $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu)$. Set $h^* = h - \mathbb{E}[h(\mathbf{Z})]$ if h is $\mathbb{P}_{\mathbf{Z}}$ -integrable. We have the following:*

- $\mathbf{P}_t^{\alpha, \nu} h$ and $\mathcal{L}_{\alpha, \nu} h$ exist for all $h \in \text{Lip}_1(E_{\mathbf{0}}, \mathbb{R})$ if and only if $\alpha > 1$.
- For all $\mathbf{z} \in E_{\mathbf{0}}$, set $h_{\mathbf{z}} = \mathbb{1}_{(-\infty, \mathbf{z}]}$. Then $\mathbf{P}_t^\alpha h_{\mathbf{z}}$ is well defined.

In both cases, $t \mapsto \mathbf{P}_t^{\alpha, \nu} h^*(\mathbf{x})$ is integrable for all $\mathbf{x} \in E_0$, so that $\mathcal{L}_{\alpha, \nu}^{-1} h^*$ exists. In the second case, the pseudo-inverse of h_z^* equals:

$$\mathcal{L}_{\alpha, \nu}^{-1} h_z^*(\mathbf{x}) = \alpha (\max \log \mathbf{x} \mathbf{z}^{-1})_+ F_{\mathbf{Z}}(\mathbf{z}) - F_{\mathbf{Z}}(\mathbf{z}) \int_{\alpha (\max \log \mathbf{x} \mathbf{z}^{-1})_+}^{\infty} (F_{\mathbf{Z}}(\mathbf{z})^{-e^{-t}} - 1) dt. \quad (3.18)$$

Here the logarithm is taken coordinatewise and $\max \mathbf{x} = \max_{1 \leq j \leq d} x^j$.

Proof. The proof is quite similar to the one given in the univariate case.

- If h is 1-Lipschitz, then it is sub-linear and we have

$$\mathbb{E} \left[|h| \left(e^{-\frac{t}{\alpha}} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z} \right) \right] \leq C + \mathbb{E} [\|\mathbf{Z}\|_1] < +\infty$$

since all the coordinates of \mathbf{Z} have Fréchet distribution $\mathcal{F}(\alpha)$ and thus are integrable if $\alpha > 1$. Besides

$$\begin{aligned} \int_{E_{\text{pol}}} |f(\mathbf{x} \oplus r\mathbf{u}) - f(\mathbf{x})| \frac{\alpha}{r^{\alpha+1}} dr d\nu(\mathbf{u}) &\leq \int_{E_{\text{pol}}} \|\mathbf{x} \oplus r\mathbf{u} - \mathbf{x}\|_1 \frac{\alpha}{r^{\alpha+1}} dr d\nu(\mathbf{u}) \\ &= \sum_{j=1}^d \int_{E_{\text{pol}}} (ru^j - x^j)_+ \frac{\alpha}{r^{\alpha+1}} dr d\nu(\mathbf{u}) \\ &\leq \sum_{j=1}^d \int_{E_{\text{pol}}} u^j \frac{\alpha}{r^{\alpha}} dr d\nu(\mathbf{u}), \end{aligned}$$

and this quantity is finite since $\alpha > 1$ and ν is a finite measure on the compact set \mathbb{S}_+^{d-1} .

- Clearly $\mathbf{P}_t^{\alpha, \nu} h_z^*$ is well-defined because $\mathbb{1}_{(-\infty, z]}$ is bounded and measurable. The proof of identity (3.1) follows the same lines as (3.1):

$$\begin{aligned} \mathbf{P}_t^{\alpha, \nu} h_z^*(\mathbf{x}) &= - \int_0^{\infty} \mathbb{1}_{\{e^{-t/\alpha} \mathbf{x} \oplus (1 - e^{-t})^{1/\alpha} \mathbf{z} \leq \mathbf{z}\}} - \mathbb{P}(\mathbf{Z} \leq \mathbf{z}) dt \\ &= - \int_0^{\infty} \mathbb{P}((1 - e^{-t})^{1/\alpha} \mathbf{Z} \leq \mathbf{z}) \mathbb{1}_{\{e^{-t/\alpha} \mathbf{x} \leq \mathbf{z}\}} - \mathbb{P}(\mathbf{Z} \leq \mathbf{z}) dt \\ &= \alpha (\max \log \mathbf{x} \mathbf{z}^{-1})_+ F_{\mathbf{Z}}(\mathbf{z}) + \int_{\alpha (\max \log \mathbf{x} \mathbf{z}^{-1})_+}^{\infty} \mathbb{P}((1 - e^{-t})^{1/\alpha} \mathbf{Z} \leq \mathbf{z}) - \mathbb{P}(\mathbf{Z} \leq \mathbf{z}) dt. \end{aligned}$$

The last identity comes from noticing that

$$\begin{aligned} e^{-t/\alpha} \mathbf{x} \leq \mathbf{z} &\iff \log \mathbf{x} \leq \frac{t}{\alpha} \mathbf{1} + \log \mathbf{z} \\ &\iff t \geq \alpha \log x^j / y^j, \text{ for all } j \in \llbracket 1, d \rrbracket, \end{aligned}$$

which means that $t \geq \alpha (\max \log \mathbf{x} \mathbf{z}^{-1})_+$, the presence of the positive part coming from the fact that t is always non-negative. The relation

$$\mathbb{P}((1 - e^{-t})^{1/\alpha} \mathbf{Z} \leq \mathbf{z}) = F_{\mathbf{Z}}(\mathbf{z})^{1 - e^{-t}}$$

is a consequence of identity (1.20).

□

We now prove a counterpart to 2.3.

Proposition 3.2.2 *Let $\alpha > 1$ and h be a differentiable 1-Lipschitz function with respect to the norm $\|\cdot\|_1$.*

1. $\mathbf{P}_t^{\alpha,\nu}h$ is differentiable and for all $j \in \llbracket 1, d \rrbracket$

$$|\partial_j \mathbf{P}_t^{\alpha,\nu}h(\mathbf{x})| \leq e^{-\frac{t}{\alpha}} \mathbb{P}(Z \leq \gamma_t^{-1/\alpha} x^j), \quad t \geq 0, \quad \mathbf{x} \in E_0 \quad (3.19)$$

with $Z \sim \mathcal{F}(\alpha)$. In particular the gradient of $\mathbf{P}_t^{\alpha,\nu}h$ satisfies

$$\|\nabla \mathbf{P}_t^{\alpha,\nu}h(\mathbf{x})\|_1 \leq e^{-\frac{t}{\alpha}} \|\mathbf{p}_t(\mathbf{x})\|_1, \quad (3.20)$$

where $\mathbf{p}_t(\mathbf{x}) := (p_t(x^1), \dots, p_t(x^d)) = (\mathbb{P}(Z \leq \gamma_t^{-1/\alpha} x^1), \dots, \mathbb{P}(Z \leq \gamma_t^{-1/\alpha} x^d))$. As a result, $\mathbf{P}_t^{\alpha,\nu}h$ is $de^{-\frac{t}{\alpha}}$ -Lipschitz.

2. Furthermore assume that $\partial_j h$ is 1-Lipschitz for all $j \in \llbracket 1, d \rrbracket$, i.e. $h \in \text{Lip}_{[2]}(E_0, \mathbb{R})$. Then $\partial_j \mathbf{P}_t^{\alpha,\nu}h$ is $C_{1,\alpha}(t)$ -Lipschitz for all $j \in \llbracket 1, d \rrbracket$ and $t > 0$, with

$$C_{1,\alpha}(t) := (d-1)e^{-t/\alpha} + \alpha\gamma_t^{-1/\alpha}.$$

Proof. 1. Observe that the function $\mathbf{x} \mapsto e^{-\frac{t}{\alpha}} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z}$ is continuous and almost surely differentiable. Set:

$$\mathbf{Z}(\mathbf{x}) := e^{-\frac{t}{\alpha}} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z}.$$

For every $j \in \llbracket 1, d \rrbracket$, the j -th derivative of its k -th coordinate is given by:

$$\partial_j \mathbf{Z}^k(\mathbf{x}) = e^{-\frac{t}{\alpha}} \mathbb{1}_{\{x^j \geq \gamma_t^{1/\alpha} Z^j\}} \mathbb{1}_{\{k=j\}} \text{ a.s.}$$

Thus the chain rule yields:

$$\nabla h(e^{-\frac{t}{\alpha}} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z}) = e^{-\frac{t}{\alpha}} \begin{pmatrix} \partial_1 h(e^{-\frac{t}{\alpha}} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z}) \mathbb{1}_{\{x^1 \geq \gamma_t^{1/\alpha} Z^1\}} \\ \vdots \\ \partial_d h(e^{-\frac{t}{\alpha}} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z}) \mathbb{1}_{\{x^d \geq \gamma_t^{1/\alpha} Z^d\}} \end{pmatrix} \text{ a.s.}$$

Since $\mathbf{x} \mapsto \|\nabla h(\mathbf{x})\|_1$ is bounded, all the partial derivatives of h are also bounded. Using a dominated convergence argument and taking expectations, we get the announced result.

2. Now we deal with the second point. We must prove that $\partial_i \mathbf{P}_t^{\alpha,\nu}h$ is Lipschitz w.r.t. each coordinate. The partial derivative $\partial_j \mathbf{P}_t^{\alpha,\nu}h$ is $e^{-t/\alpha}$ -Lipschitz with respect to x^k for $k \neq j$ as a composition of a 1-Lipschitz functions, since $\partial_j h$ is assumed to be 1-Lipschitz, and a $e^{-t/\alpha}$ -Lipschitz function, the function $x^k \mapsto e^{-t/\alpha} \mathbf{x} \oplus (1 - e^{-t})^{\frac{1}{\alpha}} \mathbf{Z}$. If $k = j$, we must be careful because of the indicator function which is not Lipschitz. For ease of notations, assume $k = j = 1$. We want to prove that $\partial_1 \mathbf{P}_t^{\alpha,\nu}h$ is Lipschitz w.r.t. x^1 . For all $\mathbf{x} \in E_0$ and $\mathbf{y} = (y^1, x^2, \dots, x^d)$

we have

$$\begin{aligned}
 |\partial_1 \mathbf{P}_t^{\alpha, \nu} h(\mathbf{x}) - \partial_1 \mathbf{P}_t^{\alpha, \nu} h(\mathbf{y})| &= \left| \mathbb{E} \left[\partial_1 h(\mathbf{Z}(\mathbf{x})) \mathbf{1}_{\{x^1 \geq \gamma_t^{1/\alpha} Z^1\}} \right] - \mathbb{E} \left[\partial_1 h(\mathbf{Z}(\mathbf{y})) \mathbf{1}_{\{y^1 \geq \gamma_t^{1/\alpha} Z^1\}} \right] \right| \\
 &\leq \mathbb{E} \left[\left| \partial_1 h(\mathbf{Z}(\mathbf{x})) - \partial_1 h(\mathbf{Z}(\mathbf{y})) \right| \mathbf{1}_{\{x^1 \geq \gamma_t^{1/\alpha} Z^1\}} \right] \\
 &\quad + \mathbb{E} \left[\left| \partial_1 h(\mathbf{Z}(\mathbf{y})) \right| \left| \mathbf{1}_{\{x^1 \geq \gamma_t^{1/\alpha} Z^1\}} - \mathbf{1}_{\{y^1 \geq \gamma_t^{1/\alpha} Z^1\}} \right| \right] \\
 &\leq e^{-\frac{t}{\alpha}} |x^1 - y^1| + \mathbb{E} \left[\left| \mathbf{1}_{\{x^1 \geq \gamma_t^{1/\alpha} Z^1\}} - \mathbf{1}_{\{y^1 \geq \gamma_t^{1/\alpha} Z^1\}} \right| \right] \\
 &= e^{-\frac{t}{\alpha}} |x^1 - y^1| + |F_Z(\gamma_t^{-1/\alpha} x^1) - F_Z(\gamma_t^{-1/\alpha} y^1)|.
 \end{aligned}$$

The last equality comes from the monotony of $x \mapsto \mathbf{1}_{\{x \geq \gamma_t^{1/\alpha} Z^1\}}$. Besides F_Z is K_α -Lipschitz, with

$$K_\alpha = \alpha \left(\frac{\alpha e^{-1}}{\alpha + 1} \right)^{1 + \frac{1}{\alpha}} \leq \alpha.$$

Therefore $\partial_1 \mathbf{P}_t^{\alpha, \nu} h$ is $\gamma_t^{-1/\alpha} K_\alpha$ -Lipschitz.

□

As before, one can give general properties of the Stein's solutions. It appears their behaviour is not much different from what we have observed in dimension 1, even though the presence of dependent coordinates prevents certain simplifications to occur in higher dimension. For instance one does not have the representation formula (3.3) for the derivative of g_h in higher dimensions. However the ensuing bound on the derivative of g_h is still true.

Proposition 3.2.3 *Let h be a 1-Lipschitz function on \mathbb{R}^d .*

1. *The associated Stein solution g_h is α -Lipschitz and its derivative satisfies:*

$$|\partial_j g_h(\mathbf{x})| \leq \min(\alpha, (x^j)^\alpha), \quad \mathbf{x} \in E_0. \tag{3.21}$$

2. *Consequently, $\mathbf{D}_{\alpha, \nu} g_h(\mathbf{0})$ is well-defined and finite.*

3. *Furthermore, assume that $h \in \text{Lip}_{[2]}(E_0, \mathbb{R})$. Then $\partial_j g_h$ is $C_{2, \alpha}$, with*

$$C_{2, \alpha} := \int_0^\infty C_{1, \alpha}(t) dt.$$

Proof.

This is an immediate consequence of (3.19) and of the formula

$$g_h(\mathbf{x}) = - \int_0^\infty \mathbf{P}_t^{\alpha, \nu} h^*(\mathbf{x}) dt,$$

the inversion between the partial derivative ∂_j and the integral being justified as in dimension 1 (see the proof of proposition 3.4).

The argument works as in dimension 1, thanks to the previous item:

$$\begin{aligned}
\int_{E_{\text{pol}}} |\langle r\mathbf{v}\mathbf{1}_{r\mathbf{v}, \mathbf{Z}_n}, \nabla g_h(\mathbf{Z}_n) \rangle| d\rho_\alpha(r) d\nu(\mathbf{u}) &= \sum_{j=1}^d \int_{\{rv^j \geq x^j\}} rv^j |\partial_j g_h(\mathbf{Z}_n)| d\rho_\alpha(r) d\nu(\mathbf{u}) \\
&\leq \sum_{j=1}^d \int_{E_{\text{pol}}} rv^j \min(\alpha, r^\alpha) \frac{\alpha}{r^{\alpha+1}} dr d\nu(\mathbf{u}) \\
&= \sum_{j=1}^d \int_{\mathbb{S}_+^{d-1}} v^j d\nu(\mathbf{u}) \int_0^\infty \min(\alpha, r^\alpha) \frac{\alpha}{r^\alpha} dr \\
&< +\infty
\end{aligned}$$

The integral over \mathbb{S}_+^{d-1} equals 1, thanks to the moment constraint relation (1.18) satisfied by ν , and the integral over \mathbb{R}_+ is finite because $\alpha > 1$.

Using the notations of proposition 3.2.2, we already know that $\partial_j \mathbf{P}_t^{\alpha, \nu}$ is $C_{1, \alpha}(t)$ -differentiable. To justify the permutation between the partial derivative ∂_1 (say) and the integral sign, we write that

$$\begin{aligned}
\left| \partial_1 \mathbf{P}_t^{\alpha, \nu} h(\mathbf{x}) \mathbb{E} \left[\partial_1 h(\mathbf{Z}(\mathbf{x})) \mathbf{1}_{\{x^1 \geq \gamma_t^{1/\alpha} Z^1\}} \right] \right| &\leq \mathbb{P}(Z \leq \gamma_t^{-1/\alpha} x^1) \\
&= e^{-\frac{\gamma_t}{(x^1)^\alpha}} \\
&\leq e^{-\frac{\gamma_t}{b^\alpha}},
\end{aligned}$$

if we assume that $x^1 \in [a, b]$, for $a \leq b$ two non-negative numbers. This last function is integrable with respect to t . Hence we can write:

$$\begin{aligned}
|\partial_1 g_h(\mathbf{x})| &\leq \int_0^\infty |\partial_1 \mathbf{P}_t^{\alpha, \nu} h^*(\mathbf{x})| dt \\
&\leq \int_0^\infty C_{1, \alpha}(t) dt \\
&= \int_0^\infty (d-1)e^{-\frac{t}{\alpha}} + \alpha \gamma_t^{-1/\alpha} dt \\
&< +\infty,
\end{aligned}$$

because $\gamma_t^{-1/\alpha} \underset{t \rightarrow 0}{\sim} t^{-1/\alpha}$, and $1/\alpha \in (0, 1)$, as $\alpha > 1$. □

We can say more about $g_{\mathbf{z}}$, even in higher dimensions.

Proposition 3.2.4 *Let $\mathbf{z} \in E_0$. We have the following:*

1. $g_{\mathbf{z}}$ is non-decreasing and constant over $[0, \mathbf{z}]$:

$$g_{\mathbf{z}}(\mathbf{x}) = -F_{\mathbf{Z}}(\mathbf{z}) \int_0^\infty (F_{\mathbf{Z}}(\mathbf{z})^{-e^{-t}} - 1) dt, \quad \mathbf{x} \in [0, \mathbf{z}].$$

Furthermore we have the equivalent

$$g_{\mathbf{z}}(\mathbf{x}) \underset{x^j \rightarrow +\infty}{\sim} \alpha F_{\mathbf{Z}}(\mathbf{z}) \log x^j, \quad j \in \llbracket 1, d \rrbracket.$$

2. $g_{\mathbf{z}}$ is continuously differentiable w.r.t. x^j over $\mathbb{R}_+^* \setminus \{z^j, \max_{k \neq j} x^k / z^k\}$. Set

$$I_{\max} := \left\{ j \in \llbracket 1, d \rrbracket, \frac{x^j}{z^j} \geq \frac{x^k}{z^k}, \forall k \in \llbracket 1, d \rrbracket \right\}$$

Then one has:

$$\partial_j g_{\mathbf{z}}(\mathbf{x}) = \begin{cases} \frac{\alpha}{x^j} F_{\mathbf{Z}}(\mathbf{z})^{1 - \left(\frac{z^j}{x^j}\right)^\alpha} & \text{if } x^j > z^j \text{ and } j \in I_{\max} \\ 0 & \text{otherwise.} \end{cases} \quad (3.22)$$

Consequently, $\sup_{\mathbf{z} \in E_0} \mathbf{D}_0 g_{\mathbf{z}}(\mathbf{0})$ is finite.

3. $\partial_j g_{\mathbf{z}}$ satisfies the inequality

$$|\partial_j g_{\mathbf{z}}(\mathbf{x})| \leq \frac{\alpha}{x^j} \mathbb{1}_{\{x^j > z^j > 0\}}, \quad j \in \llbracket 1, d \rrbracket. \quad (3.23)$$

Thus it also satisfies for all $\mathbf{x}, \mathbf{y} \in E_0$:

$$|g_{\mathbf{z}}(\mathbf{x}) - g_{\mathbf{z}}(\mathbf{y})| \leq \alpha \left(\frac{1}{z^*} \|\mathbf{x} \oplus \mathbf{z} - \mathbf{y} \oplus \mathbf{z}\|_1 \right) \odot \|\log(\mathbf{x} \oplus \mathbf{z}) - \log(\mathbf{y} \oplus \mathbf{z})\|_1, \quad (3.24)$$

where z^* is the smallest non-null coordinate of \mathbf{z} . In particular $g_{\mathbf{z}}$ is α/z^* -Lipschitz.

Proof. 1. The proof of the first point is similar to the one given in the demonstration of proposition 3.1.5.

2. If j does not belong to I_{\max} then $g_{\mathbf{z}}(\mathbf{x})$ does not depend on x^j due to equality (3.18) and then $\partial_j g_{\mathbf{z}}(\mathbf{x}) = 0$. The same result occurs if x^j is not greater than z^j , just as in dimension 1. If both conditions are satisfied, then one can drop the max and the positive part to find:

$$\partial_j g_{\mathbf{z}}(\mathbf{x}) = \frac{\alpha}{z^j} F_{\mathbf{Z}}(\mathbf{z})^{1 - \left(\frac{z^j}{x^j}\right)^\alpha}.$$

Actually, our assumptions on x^j imply that $I_{\max} = \{j\}$.

3. The previous point implies that $|\partial_j g_{\mathbf{z}}(\mathbf{x})| \leq \alpha/x^j \mathbb{1}_{\{x^j > 0\}}$, for $F_{\mathbf{Z}}(\mathbf{z})$ goes to 0 exponentially fast when z^j tends to 0^+ . Besides, this partial derivative is null if x^j is less than z^j , hence the indicator function $\mathbb{1}_{\{x^j > z^j > 0\}}$. Just as in the proof given for the scalar case, we integrate each partial derivative to control every increment of $g_{\mathbf{z}}$. For the sake of clarity, we give the proof in dimension $d = 2$, the reasoning being the same in higher dimensions but with more cumbersome notations.

$$\begin{aligned} |g_{\mathbf{z}}(\mathbf{x}) - g_{\mathbf{z}}(\mathbf{y})| &\leq |g_{\mathbf{z}}(x^1, x^2) - g_{\mathbf{z}}(y^1, x^2)| + |g_{\mathbf{z}}(y^1, x^2) - g_{\mathbf{z}}(y^1, y^2)| \\ &\leq \alpha \int_{x^1}^{y^1} |\partial_1 g_{\mathbf{z}}(r, x^2)| dr + \alpha \int_{x^2}^{y^2} |\partial_1 g_{\mathbf{z}}(y^1, r)| dr. \end{aligned}$$

If y^1 is less than x^1 or y^2 is less than x^2 , simply exchange the bounds of integration. Notice inequality 3.23 depends only on the j -th coordinate. The rest of the proof follows the same lines as the one given in the univariate case (proposition 3.1.5).

□

3.2.2 Applications

In this subsection, unless specified otherwise, the reference norm $\|\cdot\|$ will be any ℓ^p -norm:

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^d |x^j|^p \right)^{\frac{1}{p}},$$

if $p \in [1, +\infty)$ and $\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq d} |x^j|$. This choice is rather arbitrary and serves only to make certain arguments easier to state later.

Distance between max-stable random vectors

Let \mathbf{Z}_1 and \mathbf{Z}_2 be two max-stable random vectors with distribution $\mathcal{MS}(\alpha_i, \nu_i)$ respectively, with $\alpha_i > 0$ and $i \in \{1, 2\}$. We wish to get an idea of how close are the distributions of \mathbf{Z}_1 and \mathbf{Z}_2 , and express this closeness in terms of α_i and ν_i .

We start by assuming that $\nu_1 = \nu_2$, i.e. \mathbf{Z}_1 and \mathbf{Z}_2 have the same angular measure ν but different stability indices α_1, α_2 .

Proposition 3.2.5 *Let α_1, α_2 be two positive numbers with $\alpha_1 < \alpha_2$, and ν a finite measure on \mathbb{S}_+^{d-1} satisfying the moment constraints (1.18). Let $Z_i \sim \mathcal{MS}(\alpha_i, \nu)$ for $i = 1, 2$.*

- There exists a constant $C_{\alpha_1, \alpha_2, \nu}^K$ such that:

$$d_K(\mathbf{Z}_1, \mathbf{Z}_2) \leq C_{\alpha_1, \alpha_2, \nu}^K \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right) |\alpha_1 - \alpha_2|.$$

Let $Z_1 \sim \mathcal{F}(\alpha_1)$ and μ_{α_1} the exponent measure of \mathbf{Z}_1 . A possible choice for the constant $C_{\alpha_1, \alpha_2, \nu_1}^K$ is:

$$C_{\alpha_1, \alpha_2, \nu}^K = d + \sum_{j=1}^d \int_{E_0} |\log y^j| \mathbb{P}(Z_1^{\alpha_2/\alpha_1} \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}). \quad (3.25)$$

- If furthermore both α_1 and α_2 are greater than 1, we have a constant $C_{\alpha_1, \alpha_2, \nu_1}^W > 0$ such that:

$$d_W(\mathbf{Z}_1, \mathbf{Z}_2) \leq C_{\alpha_1, \alpha_2, \nu}^W \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right) |\alpha_1 - \alpha_2|.$$

A possible choice for the constant $C_{\alpha_1, \alpha_2, \nu}$ is:

$$C_{\alpha_1, \alpha_2, \nu}^W = d\Gamma\left(1 - \frac{1}{\alpha_2}\right) + \sum_{j=1}^d \int_{E_0} (1 \oplus y^j) |\log y^j| \mathbb{P}(Z_1^{\alpha_2/\alpha_1} \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}). \quad (3.26)$$

Proof. - Set $h^* = h - \mathbb{E}[h(\mathbf{Z}_2)]$ for any $\mathbb{P}_{\mathbf{Z}_2}$ -integrable h , and take $g_{\mathbf{z}} = \mathcal{L}_{\alpha_2, \nu}^{-1} h_{\mathbf{z}}^*$. We have:

$$\begin{aligned} |F_{\mathbf{Z}_1}(\mathbf{z}) - F_{\mathbf{Z}_2}(\mathbf{z})| &= |\mathbb{E}[h_{\mathbf{z}}(\mathbf{Z}_1)] - \mathbb{E}[h_{\mathbf{z}}(\mathbf{Z}_2)]| \\ &= |\mathbb{E}[h_{\mathbf{z}}^*(\mathbf{Z}_1)]| \\ &= |\mathbb{E}[\mathcal{L}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)]|. \end{aligned}$$

Thanks to inequality (3.23), we see that $\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle$ and $\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)$ have finite expectations. Furthermore, the fact that $\mathbb{E}[\mathcal{L}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] = 0$ implies that:

$$\mathbb{E}[\mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] = \frac{1}{\alpha_1} \mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle],$$

so that

$$\begin{aligned} \alpha_2 \mathbb{E}[\mathcal{L}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] &= \alpha_2 \mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] - \mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle] \\ &= \alpha_2 \mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] + \alpha_2 \mathbb{E}[\mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] - \mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle] \\ &= \alpha_2 \mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] + \left(\frac{\alpha_2}{\alpha_1} - 1 \right) \mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle]. \end{aligned} \quad (3.27)$$

The second term is bounded by $d\alpha_2|1 - \alpha_2\alpha_1^{-1}|$, for

$$\mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle] = \sum_{j=1}^d \mathbb{E}[Z_1^j \partial_j g_{\mathbf{z}}(\mathbf{Z}_1)] \leq d\alpha_2,$$

thanks to inequality (3.23). As for the first one, comparing directly the operators $\mathbf{D}_{\alpha_2, \nu}$ and $\mathbf{D}_{\alpha_1, \nu}$ seems difficult. Instead, we start by giving a more convenient expression of $\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)$. Recall the notations introduced in subsection 2.3.2. With them, we can write:

$$\begin{aligned} \mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] &= \mathbb{E}\left[T_{\alpha_1}^{\alpha_2} \mathbf{D}_{\alpha_1, \nu} T_{\alpha_2}^{\alpha_1} g_{\mathbf{z}}(\mathbf{Z}_1)\right] \\ &= \mathbb{E}\left[\mathbf{D}_{\alpha_1, \nu} T_{\alpha_2}^{\alpha_1} g_{\mathbf{z}}(\mathbf{Z}_1^{\frac{\alpha_2}{\alpha_1}})\right] \\ &= \mathbb{E}[\mathbf{D}_{\alpha_1, \nu} T_{\beta-1} g_{\mathbf{z}}(\mathbf{Z}_1^{\beta})], \end{aligned}$$

where we set $\beta := \alpha_2/\alpha_1$. Let μ_{α_1} be the exponent measure of \mathbf{Z}_1 . From what precedes we get

$$\begin{aligned} \mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] &= \mathbb{E}[\mathbf{D}_{\alpha_1, \nu} T_{\beta-1} g_{\mathbf{z}}(\mathbf{Z}_1^{\beta}) - \mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] \\ &= \int_{E_0} \mathbb{E}[g_{\mathbf{z}}(\mathbf{Z}_1 \oplus \mathbf{y}^{1/\beta}) - g_{\mathbf{z}}(\mathbf{Z}_1) - g_{\mathbf{z}}(\mathbf{Z}_1 \oplus \mathbf{y}) + g_{\mathbf{z}}(\mathbf{Z}_1)] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &= \int_{E_0} \mathbb{E}[g_{\mathbf{z}}(\mathbf{Z}_1 \oplus \mathbf{y}^{1/\beta}) - g_{\mathbf{z}}(\mathbf{Z}_1 \oplus \mathbf{y})] \, d\mu_{\alpha_1}(\mathbf{y}). \end{aligned}$$

The triangle inequality and bound 3.24 yield:

$$\begin{aligned} &|\mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)]| \\ &\leq \int_{E_0} \mathbb{E}[|g_{\mathbf{z}}(\mathbf{Z}_1 \oplus \mathbf{y}^{1/\beta}) - g_{\mathbf{z}}(\mathbf{Z}_1 \oplus \mathbf{y})|] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &\leq \alpha_2 \int_{E_0} \mathbb{E}[\|\log(\mathbf{Z}_1 \oplus \mathbf{y}^{1/\beta}) - \log(\mathbf{Z}_1 \oplus \mathbf{y})\|_1] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &= \alpha_2 \sum_{j=1}^d \int_{E_0} \mathbb{E}[|\log(\mathbf{Z}_1 \oplus (y^j)^{1/\beta}) - \log(\mathbf{Z}_1 \oplus y^j)|] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &= \alpha_2 \sum_{j=1}^d \int_{E_0} \mathbb{E}[|\log(\mathbf{Z}_1 \oplus (y^j)^{1/\beta}) - \log(\mathbf{Z}_1 \oplus y^j)| \mathbf{1}_{\{y^j \geq Z_1^{\beta} \circledast Z_1\}}] \, d\mu_{\alpha_1}(\mathbf{y}), \end{aligned}$$

where $Z_1 \sim \mathcal{F}(\alpha_1)$. The indicator function in the last equality comes from noticing that if $y^j \leq Z_1$ and $y^j \leq Z_1^\beta$ at the same time, then the integrand vanishes. Finally, because the logarithm is a non-decreasing function, we know that $\log(x \oplus y) = \log x \oplus \log y$. Furthermore, $x \mapsto c \oplus x$ is 1-Lipschitz. Using this result with $c = \log Z_1$, we find:

$$\begin{aligned} |\mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_z(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_z(\mathbf{Z}_1)]| &\leq \alpha_2 \sum_{j=1}^d \int_{E_0} |\beta^{-1} \log y^j - \log y^j| \mathbb{P}(Z_1^\beta \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}) \\ &= |\alpha_2 - \alpha_1| \sum_{j=1}^d \int_{E_0} |\log y^j| \mathbb{P}(Z_1^\beta \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}). \end{aligned}$$

This last integral is bounded since the probability inside vanishes at exponential speed when y^j goes to 0. By dividing this bound by α_2 , we obtain the desired result.

- We start from equality (3.27), but with $g_h = \mathcal{L}_{\alpha_2, \nu}^{-1} h^*$ instead of g_z , where h a 1-Lipschitz function on E_0 . As in the previous point, we have two terms to bound:

$$\begin{aligned} |\mathbb{E}[\langle \mathbf{Z}_1, \nabla g_h(\mathbf{Z}_1) \rangle]| &\leq \sum_{j=1}^d \mathbb{E}[Z^j |\partial_j g_h(\mathbf{Z})|] \\ &\leq \alpha_2 \sum_{j=1}^d \mathbb{E}[Z^j] \\ &= d\alpha_2 \Gamma\left(1 - \frac{1}{\alpha_2}\right). \end{aligned}$$

this inequality being a consequence of (3.21). As for the second term, we use the Lipschitz property of g_h , relying on the fact that both α_1 and α_2 are greater than 1.

$$\begin{aligned} |\mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_h(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_h(\mathbf{Z}_1)]| &\leq \int_{E_0} \mathbb{E}[|g_h(\mathbf{Z}_1 \oplus \mathbf{y}^{1/\beta}) - g_h(\mathbf{Z}_1 \oplus \mathbf{y})|] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &\leq \alpha_2 \int_{E_0} \mathbb{E}[\|\mathbf{Z}_1 \oplus \mathbf{y}^{1/\beta} - \mathbf{Z}_1 \oplus \mathbf{y}\|_1] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &= \alpha_2 \sum_{j=1}^d \int_{E_0} \mathbb{E}[|Z_1^j \oplus (y^j)^{1/\beta} - Z_1^j \oplus y^j| \mathbf{1}_{\{y^j \geq Z_1^\beta \odot Z_1\}}] \, d\mu_{\alpha_1}(\mathbf{y}) \\ &\leq \alpha_2 \sum_{j=1}^d \int_{E_0} |(y^j)^{\beta^{-1}} - y^j| \mathbb{P}(Z_1^\beta \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}) \\ &= \alpha_2 \sum_{j=1}^d \int_{E_0} y^j |(y^j)^{\beta^{-1}-1} - 1| \mathbb{P}(Z_1^\beta \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}). \end{aligned}$$

This quantity is finite because $\alpha_1 > 1$ and μ_{α_1} is the image measure of $\rho_1 \otimes \nu$ by $(r, \mathbf{u}) \mapsto (r\mathbf{u})^{1/\alpha_1}$, with ν a finite measure over the compact \mathbb{S}_+^{d-1} . Recall that $\alpha_1 < \alpha_2$, so that $\beta^{-1} < 1$, and thus $1 - \beta^{-1} \in (0, 1)$. We need to study the function $f : \delta \mapsto y^{-\delta} = \exp(-\delta \log y)$ over $(0, 1)$ for all $y \in \mathbb{R}_+^*$. If $y \in [1, +\infty)$, we have

$$|y^{-\delta} - 1| = 1 - \exp(-\delta \log y) \leq \delta \log y.$$

And if $y \in (0, 1]$, then $|f'(\delta)| = y^{-\delta} |\log y| \leq y^{-1} |\log y|$. Thus f is $(1 \oplus y^{-1}) |\log y|$ -Lipschitz over \mathbb{R}_+^* . Finally:

$$\begin{aligned} |\mathbb{E}[\mathbf{D}_{\alpha_2, \nu} g_h(\mathbf{Z}_1) - \mathbf{D}_{\alpha_1, \nu} g_h(\mathbf{Z}_1)]| &= \alpha_2 \sum_{j=1}^d \int_{E_0} y^j |(y^j)^{\beta-1-1} - 1| \mathbb{P}(Z_1^\beta \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}) \\ &\leq |\alpha_1 - \alpha_2| \sum_{j=1}^d \int_{E_0} (1 \oplus y^j) |\log y^j| \mathbb{P}(Z_1^\beta \odot Z_1 \leq y^j) \, d\mu_{\alpha_1}(\mathbf{y}), \end{aligned}$$

and this last integral is finite. □

Next we suppose the stability index is now the same α for both \mathbf{Z}_1 and \mathbf{Z}_2 but that they differ by their angular measures, ν_1 and ν_2 respectively.

Proposition 3.2.6 *Let α be a positive number and ν_1, ν_2 two finite measures on \mathbb{S}_+^{d-1} satisfying the moment constraints (1.18). Let also $\mathbf{Z}_i \sim \mathcal{MS}(\alpha, \nu_i)$ for $i = 1, 2$.*

- The following inequality holds:

$$d_K(\mathbf{Z}_1, \mathbf{Z}_2) \leq d d_{\text{TV}}(\nu_1, \nu_2).$$

- If furthermore α is greater than 1, then

$$d_W(\mathbf{Z}_1, \mathbf{Z}_2) \leq d \Gamma\left(1 - \frac{1}{\alpha}\right) d_{\text{TV}}(\nu_1, \nu_2).$$

Proof. - As before, set $h^* = h - \mathbb{E}[h(\mathbf{Z}_2)]$ for any $\mathbb{P}_{\mathbf{Z}_2}$ -integrable h , and $g_{\mathbf{z}} = \mathcal{L}_{\alpha_2, \nu}^{-1} h_{\mathbf{z}}^*$. We have

$$|F_{\mathbf{Z}_1}(\mathbf{z}) - F_{\mathbf{Z}_2}(\mathbf{z})| = |\mathbb{E}[\mathcal{L}_{\alpha, \nu_2} g_{\mathbf{z}}(\mathbf{Z}_1)]|.$$

We use once again that the term $\mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle]$ is common to both $\mathcal{L}_{\alpha, \nu_1}$ and $\mathcal{L}_{\alpha, \nu_2}$:

$$\begin{aligned} \mathbb{E}[\mathcal{L}_{\alpha_2, \nu} g_{\mathbf{z}}(\mathbf{Z}_1)] &= \mathbb{E}[\mathbf{D}_{\alpha, \nu_2} g_{\mathbf{z}}(\mathbf{Z}_1)] - \frac{1}{\alpha} \mathbb{E}[\langle \mathbf{Z}_1, \nabla g_{\mathbf{z}}(\mathbf{Z}_1) \rangle] \\ &= \mathbb{E}[\mathbf{D}_{\alpha, \nu_2} g_{\mathbf{z}}(\mathbf{Z}_1) - \mathbf{D}_{\alpha, \nu_1} g_{\mathbf{z}}(\mathbf{Z}_1)]. \end{aligned} \quad (3.28)$$

This time, comparing $\mathbf{D}_{\alpha, \nu_1}$ and $\mathbf{D}_{\alpha, \nu_2}$ is easier:

$$\begin{aligned} &\mathbf{D}_{\alpha, \nu_2} g_{\mathbf{z}}(\mathbf{x}) - \mathbf{D}_{\alpha, \nu_1} g_{\mathbf{z}}(\mathbf{x}) \\ &= \frac{1}{\alpha} \int_{E_{\text{pol}}} \langle r \mathbf{u}^{1/\alpha} \mathbf{1}_{r \mathbf{u}^{1/\alpha}, \mathbf{x}}, \nabla g_{\mathbf{z}}(\mathbf{x} \oplus r \mathbf{u}^{1/\alpha}) \rangle \, d\rho_{\alpha}(r) (d\nu_2(\mathbf{u}) - d\nu_1(\mathbf{u})) \\ &= \frac{1}{\alpha} \int_{\mathbb{R}_+^*} \int_{\mathbb{S}_+^{d-1}} \langle r \mathbf{u}^{1/\alpha} \mathbf{1}_{r \mathbf{u}^{1/\alpha}, \mathbf{x}}, \nabla g_{\mathbf{z}}(\mathbf{x} \oplus r \mathbf{u}^{1/\alpha}) \rangle (f_{\nu_2}(\mathbf{u}) - f_{\nu_1}(\mathbf{u})) \, d\nu_{1,2}(\mathbf{u}) \, d\rho_{\alpha}(r). \end{aligned}$$

for every $\mathbf{x} \in E_0$. Here $\nu_{1,2} := \nu_1 + \nu_2$, and $f_{\nu_i} = d\nu_i / d\nu_{1,2}$ for $i = 1, 2$, the density function of ν_i w.r.t. to $\nu_{1,2}$. Now, replacing \mathbf{x} by \mathbf{Z}_1 and taking expectations, we bound the scalar product in the integral by

$$|\langle r \mathbf{u}^{1/\alpha} \mathbf{1}_{r \mathbf{u}^{1/\alpha}, \mathbf{Z}_1}, \nabla g_{\mathbf{z}}(\mathbf{Z}_1 \oplus r \mathbf{u}^{1/\alpha}) \rangle| \leq \alpha \sum_{j=1}^d \mathbb{P}(Z \leq r(u^j)^{1/\alpha}) \leq d\alpha \mathbb{P}(Z \leq r),$$

because

$$r(u^j)^{1/\alpha} |\partial_j g_{\mathbf{z}}(\mathbf{Z}_1 \oplus r\mathbf{u}^{1/\alpha})| \leq \mathbb{1}_{\{r(u^j)^{1/\alpha} \geq Z_1^j\}},$$

thanks to inequality (3.23). We have also used that u^j is always in $[0, 1]$, as $\mathbf{u} \in \mathbb{S}_+^{d-1}$ (since the reference norm is a ℓ^p -norm). Besides, Z is a random variable with Fréchet distribution $\mathcal{F}(\alpha)$, which corresponds to the distribution of the marginals of \mathbf{Z}_1 . Finally we find:

$$\begin{aligned} |F_{\mathbf{Z}_1}(\mathbf{z}) - F_{\mathbf{Z}_2}(\mathbf{z})| &= |\mathbb{E}[\mathbf{D}_{\alpha, \nu_2} g_{\mathbf{z}}(\mathbf{Z}_1) - \mathbf{D}_{\alpha, \nu_1} g_{\mathbf{z}}(\mathbf{Z}_1)]| \\ &\leq d \int_{\mathbb{R}_+^*} \int_{\mathbb{S}_+^{d-1}} \mathbb{P}(Z \leq r) |f_{\nu_2}(\mathbf{u}) - f_{\nu_1}(\mathbf{u})| d\nu_{1,2}(\mathbf{u}) d\rho_{\alpha}(r) \\ &= d \int_{\mathbb{R}_+^*} e^{-\frac{1}{r\alpha}} \frac{\alpha}{r^{\alpha+1}} dr \int_{\mathbb{S}_+^{d-1}} |f_{\nu_2}(\mathbf{u}) - f_{\nu_1}(\mathbf{u})| d\nu_{1,2}(\mathbf{u}) \\ &= dd_{\text{TV}}(\nu_1, \nu_2). \end{aligned}$$

- We start from (3.28), with $g_h = \mathcal{L}_{\alpha, \nu_2} h^*$ instead of $g_{\mathbf{z}}$, where $h : E_{\mathbf{0}} \rightarrow \mathbb{R}$ is a 1-Lipschitz function. Then, we know from inequality (3.21) that:

$$\begin{aligned} |\langle r\mathbf{u}^{1/\alpha} \mathbf{1}_{r\mathbf{u}^{1/\alpha}, \mathbf{Z}_1}, \nabla g_h(\mathbf{Z}_1 \oplus r\mathbf{u}^{1/\alpha}) \rangle| &\leq \sum_{j=1}^d \min(\alpha, r^{\alpha} u^j) r(u^j)^{1/\alpha} \mathbb{P}(Z \leq r(u^j)^{1/\alpha}) \\ &\leq d\alpha r \mathbb{P}(Z \leq r). \end{aligned}$$

Consequently we find:

$$\begin{aligned} |h(\mathbf{Z}_1) - h(\mathbf{Z}_2)| &\leq d \int_{\mathbb{R}_+^*} e^{-\frac{1}{r\alpha}} \frac{\alpha}{r^{\alpha}} dr \int_{\mathbb{S}_+^{d-1}} |f_{\nu_2}(\mathbf{u}) - f_{\nu_1}(\mathbf{u})| d\nu_{1,2}(\mathbf{u}) \\ &= d\Gamma\left(1 - \frac{1}{\alpha}\right) d_{\text{TV}}(\nu_1, \nu_2). \end{aligned}$$

□

We combine those two results in the following corollary.

Corollary 3.2.7 *Let α_1, α_2 be two positive numbers with $\alpha_1 < \alpha_2$, and ν_1, ν_2 two finite measures on \mathbb{S}_+^{d-1} satisfying the moment constraints (1.18). Let $\mathbf{Z}_i \sim \mathcal{MS}(\alpha_i, \nu_i)$, for $i = 1, 2$.*

- By taking $C_{\alpha_1, \alpha_2, \nu_1}$ as in (3.25), the following inequality holds:

$$d_{\text{K}}(\mathbf{Z}_1, \mathbf{Z}_2) \leq C_{\alpha_1, \alpha_2, \nu_1} \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right) |\alpha_1 - \alpha_2| + dd_{\text{TV}}(\nu_1, \nu_2).$$

- If furthermore α_1, α_2 are greater than 1, then for $C_{\alpha_1, \alpha_2, \nu_1}$ given by (3.26), we have:

$$d_{\text{W}}(\mathbf{Z}_1, \mathbf{Z}_2) \leq C_{\alpha_1, \alpha_2, \nu_1} \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right) |\alpha_1 - \alpha_2| + d\Gamma\left(1 - \frac{1}{\alpha_2}\right) d_{\text{TV}}(\nu_1, \nu_2).$$

Proof. Simply bound $d_{\text{K}}(\mathbf{Z}_1, \mathbf{Z}_2)$ by $d_{\text{K}}(\mathbf{Z}_1, \mathbf{Z}_3) + d_{\text{K}}(\mathbf{Z}_3, \mathbf{Z}_2)$, with $\mathbf{Z}_3 \sim \mathcal{MS}(\alpha_2, \nu_1)$ and use propositions 3.2.5 and 3.2.6. □

Rate of convergence of the de Haan-LePage series

Let $n \in \mathbb{N}^*$ and $\phi = ((r_i, \mathbf{v}_i))_{i \geq 1}$ a configuration in E_0 . We arrange ϕ in decreasing order with respect to the first coordinate, *i.e.* $r_1 \geq r_2 \geq \dots$, and define

$$\mathbf{m}_n(\phi) := \bigoplus_{i=1}^n r_i \mathbf{v}_i.$$

Take \mathbf{Z} a max-stable random vector with distribution $\mathcal{MS}(\alpha, \nu)$. Define a measure ν_α on \mathbb{S}_+^{d-1} by

$$\nu_\alpha(B) := \nu(B^\alpha), \quad B \in \mathcal{B}(\mathbb{S}_+^{d-1}),$$

where $B^\alpha = \{x^\alpha, x \in B\}$. In other words, ν_α is the image-measure of ν by $\mathbf{u} \mapsto \mathbf{u}^{1/\alpha}$. Let $\eta = ((r_i, \mathbf{v}_i))_{i \geq 1}$ be a marked Poisson process on E_{pol} with intensity measure $\alpha r^{-(\alpha+1)} dr \otimes d\nu_\alpha(\mathbf{v})$. We know from [Davydov et al. \[2008\]](#) that when n goes to infinity, $\mathbf{m}_n(\eta)$ converges to $\mathbf{m}(\eta)$ in distribution. Our goal is to determine an estimation of the speed at which this convergence occurs.

Proposition 3.2.8 *Let $\eta = ((r_i, \mathbf{v}_i))_{i \geq 1}$ be a Poisson process on E_{pol} with intensity measure $\alpha r^{-(\alpha+1)} dr \otimes d\nu_\alpha(\mathbf{v})$, with $\alpha > 1$. Set*

$$\mathbf{Z}_n := \mathbf{m}_n(\eta) = \bigoplus_{i=1}^n r_i \mathbf{v}_i.$$

Then there exists a constant $C_\alpha > 0$ depending only on d and α such that

$$d_W(\mathbf{Z}_n, \mathbf{Z}) \leq C_\alpha \frac{1}{n}, \quad n \geq 2.$$

Proof. We first check that $\langle \mathbf{Z}_n, \nabla g_h(\mathbf{Z}_n) \rangle$ is integrable. Thanks to property [3.21](#), we know that g_h is α -Lipschitz for h is 1-Lipschitz; therefore

$$|\langle \mathbf{Z}_n, \nabla g_h(\mathbf{Z}_n) \rangle| \leq \alpha \sum_{j=1}^d Z_n^j \leq \alpha \sum_{j=1}^d Z^j$$

because by definition Z_n^j is dominated by $Z^j = \mathbf{m}(\eta)$, whose coordinates are integrable since $\alpha > 1$. The integrability of $\mathbf{D}_{\alpha, \nu} f(\mathbf{Z}_n)$ has already been checked in proposition [3.21](#).

To compare $\mathbb{E}[\mathbf{D}_{\alpha, \nu} g_h(\mathbf{Z}_n)]$ and $\alpha \mathbb{E}[\langle \mathbf{Z}_n, g_h(\mathbf{Z}_n) \rangle]$, we apply the Campbell-Mecke formula to:

$$\eta \mapsto \langle r \mathbf{v} \mathbf{1}_{r \mathbf{v}, \mathbf{m}_n(\eta)}, \nabla g_h(\mathbf{m}_n(\eta) \oplus r \mathbf{v}) \rangle,$$

so that we get

$$\begin{aligned} \mathbb{E}[\mathbf{D}_{\alpha, \nu} g_h(\mathbf{m}_n(\eta))] &= \int_{E_{\text{pol}}} \mathbb{E} \left[\langle r \mathbf{v} \mathbf{1}_{r \mathbf{v}, \mathbf{m}_n(\eta)}, \nabla g_h(\mathbf{m}_n(\eta) \oplus r \mathbf{v}) \rangle \right] \frac{\alpha}{r^{\alpha+1}} dr d\nu_\alpha(\mathbf{v}) \\ &= \mathbb{E} \left[\sum_{(r, \mathbf{v}) \in \eta} \langle r \mathbf{v} \mathbf{1}_{r \mathbf{v}, \mathbf{m}_n(\eta - \delta_{(r, \mathbf{v})})}, \nabla g_h(\mathbf{m}_n(\eta - \delta_{(r, \mathbf{v})}) \oplus r \mathbf{v}) \rangle \right] \\ &= \mathbb{E} \left[\sum_{(r, \mathbf{v}) \in \eta} \langle r \mathbf{v} \mathbf{1}_{r \mathbf{v}, \mathbf{m}_n(\eta_{r, \mathbf{v}})}, \nabla g_h(\mathbf{m}_n(\eta_{r, \mathbf{v}}) \oplus r \mathbf{v}) \rangle \right], \end{aligned} \tag{3.29}$$

with $\eta_{r, \mathbf{v}} := \eta - \delta_{(r, \mathbf{v})}$. Notice we do *not* have $\mathbf{m}_n(\eta) \oplus r \mathbf{v} = \mathbf{m}_n(\eta + \delta_{(r, \mathbf{v})})$ in general, because (r, \mathbf{v}) may not be one of the n first points of η ; in that case $\mathbf{m}_n(\eta) \oplus r \mathbf{v} = \mathbf{m}_n(\eta)$. To deal with this difficulty, we will need the following lemma.

Lemma 3.2.9 *With the above notations, η can be partitioned into three disjoint pieces:*

$$\eta = \eta_{<} \uplus \eta_{=} \uplus \eta_{>}, \quad (3.30)$$

where

$$\begin{aligned} \eta_{<} &:= \left\{ (r, \mathbf{v}) \in \eta, r\mathbf{v} \in [\mathbf{0}, \mathbf{m}_n(\eta)) \right\} \\ \eta_{=} &:= \left\{ (r, \mathbf{v}) \in \eta, \exists j \in \llbracket 1, d \rrbracket, rv^j = \mathbf{m}_n^j(\eta) \neq 0 \right\} \\ \eta_{>} &:= \left\{ (r, \mathbf{v}) \in \eta, r\mathbf{v} \notin [\mathbf{0}, \mathbf{m}_n(\eta)] \right\} \end{aligned}$$

Proof. Equality (3.30) indeed constitutes a partition of η : if a vector $r\mathbf{v}$ does not have all of its coordinates strictly less than the corresponding coordinates of $\mathbf{m}_n(\eta)$, then it means that one of its coordinates is equal or greater than its counterpart of $\mathbf{m}_n(\eta)$. Those two cases are incompatible since η is a Poisson process with diffuse intensity measure on E_{pol} . Indeed, assume there exist j_1, j_2 such that

$$rv^{j_1} = \mathbf{m}_n^{j_1}(\eta) \neq 0 \text{ and } rv^{j_2} > \mathbf{m}_n^{j_2}(\eta).$$

The second assumption implies that (r, \mathbf{v}) does not belong to the n first points of η . But then it would imply that one could find another (r_i, \mathbf{v}_i) in η , with $i \in \llbracket 1, n \rrbracket$, such that

$$rv^{j_1} = r_i v_i^{j_1}.$$

Because the measure $\alpha r^{-(\alpha+1)} dr$ is diffuse on \mathbb{R}_+^* , this situation can occur with positive probability only if both $v_i^{j_1}$ and v^{j_1} are equal to 0. This contradicts the fact that rv^{j_1} is not null. \square

A way to sum up the contents of the previous lemma is to say that each coordinate of $\mathbf{m}_n(\eta)$ comes from exactly one $r\mathbf{v}$, which does not prevent it from containing other coordinates of $\mathbf{m}_n(\eta)$. With this partition, we can decompose the sum in (3.29) into three pieces, denoted by $S_{<}$, $S_{=}$ and $S_{>}$, that we analyse now.

1. The first part corresponds to $\eta_{<}$ and is the easiest to deal with:

$$\sum_{(r, \mathbf{v}) \in \eta_{<}} \langle r\mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta_{r, \mathbf{v}})}, \nabla g_h(\mathbf{m}_n(\eta_{r, \mathbf{v}}) \oplus r\mathbf{v}) \rangle = 0,$$

because each term in the inner product vanishes. Indeed, due to the indicator functions, the sum is null as soon as all coordinates of $r\mathbf{v}$ are less than those of $\mathbf{m}_n(\eta_{r, \mathbf{v}})$. But since $(r, \mathbf{v}) \in \eta_{<}$, we already know that no coordinate of $r\mathbf{v}$ contributes to $\mathbf{m}_n(\eta)$, and thus to $\mathbf{m}_n(\eta_{r, \mathbf{v}})$. Thus

$$\mathbb{E}[S_{<}] = 0.$$

2. Next we take care of the sum over $\eta_{>}$. First observe that if $(r, \mathbf{v}) \in \eta_{>}$, then it cannot belong to the n first points of η and is not taken into account by \mathbf{m}_n . So we have

$$\mathbf{m}_n(\eta_{r, \mathbf{v}}) = \mathbf{m}_n(\eta).$$

Besides, assume $v^j \neq 0$. In particular, $rv^j \neq r_k v_k^j$ a.s. for all $k \in \llbracket 1, n \rrbracket$, because the Poisson process $(r_i)_{i \geq 1}$ has a diffuse measure. This means that:

$$\{rv^j \geq \mathbf{m}_n^j(\eta)\} = \{rv^j > \mathbf{m}_n^j(\eta)\} = \bigcap_{k=1}^n \{rv^j > r_k v_k^j\} \subseteq \bigcap_{k=1}^n \{v^j > v_k^j\}.$$

The inequality comes from noticing that since $r \in \eta_{>}$, so that it belongs to $(r_i)_{i \geq n+1}$, we have $r < r_k$ if $k \in \llbracket 1, n \rrbracket$, because the sequence $(r_i)_{i \geq 1}$ has been sorted in decreasing order. Thus, $rv^j \geq \mathbf{m}_n^j(\eta)$ is possible only if $v^j \geq v_k^j$ for all $k \in \llbracket 1, n \rrbracket$.

$$\begin{aligned} |\mathbb{E}[S_{=}]| &= \left| \mathbb{E} \left[\sum_{(r,\mathbf{v}) \in \eta_{>}} \langle r\mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta_{r,\mathbf{v}})}, \nabla g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \rangle \right] \right| \\ &= \left| \mathbb{E} \left[\sum_{(r,\mathbf{v}) \in \eta_{>}} \langle r\mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta)}, \nabla g_h(\mathbf{m}_n(\eta) \oplus r\mathbf{v}) \rangle \right] \right| \\ &\leq \sum_{j=1}^d \mathbb{E} \left[\sum_{(r,\mathbf{v}) \in \eta_{>}} rv^j |\partial_j g_h(\mathbf{m}_n(\eta) \oplus r\mathbf{v})| \mathbf{1}_{\{rv^j \geq \mathbf{m}_n^j(\eta)\}} \right] \\ &\leq \alpha \sum_{j=1}^d \mathbb{E} \left[\sum_{i=n+1}^{\infty} r_i \min(\alpha, (r_i v_i^j)^\alpha) \mathbf{1}_{\{v^j \neq 0\}} \mathbf{1}_{\bigcap_{k=1}^n \{v^j \geq v_k^j\}} \right] \\ &\leq \alpha \sum_{j=1}^d \sum_{i=n+1}^{\infty} \mathbb{E}[r_i \min(\alpha, r_i^\alpha)] \mathbb{P} \left(\bigcap_{k=1}^n \{v^j > v_k^j\} \right) \\ &\leq d\alpha \frac{1}{n} \sum_{i=n+1}^{\infty} \mathbb{E}[r_i^{\alpha+1}], \end{aligned}$$

thanks to inequality (3.21). The term $1/n$ comes from the basic bound:

$$\mathbb{P}(X_1 > X_2, \dots, X_1 > X_n) \leq \frac{1}{n},$$

where X_1, \dots, X_n are n i.i.d. random variables. Indeed, the marks $(\mathbf{v}_i)_{i \geq 1}$ are i.i.d. by definition. Notice it works even if the X_k are constant thanks to the strict inequalities. We have also used that the v^j are always less than $\mathbf{1}_{\{v^j \neq 0\}}$, because the reference norm is a ℓ^p -norm and $\mathbf{v} \in \mathbb{S}_+^{d-1}$. Set $\Gamma_i := r_i^{-\alpha}$. By definition of the r_i , which are the points of a Poisson process with intensity $\alpha r^{-(\alpha+1)}$, we know that $\Gamma_i \sim \Gamma(i, 1)$, the Gamma distribution with shape parameter i and scale parameter 1. This implies that:

$$\begin{aligned} \mathbb{E}[r_i^{\alpha+1}] &= \mathbb{E}[\Gamma_i^{-(1+1/\alpha)}] \\ &= \frac{1}{\Gamma(i)} \int_0^\infty x^{-(1+\frac{1}{\alpha})} x^{i-1} e^{-x} dx \\ &= \frac{\Gamma(i-1-1/\alpha)}{\Gamma(i)} \\ &\leq \left(i-1-\frac{1}{\alpha}\right)^{-(1+\frac{1}{\alpha})}, \end{aligned}$$

thanks to the inequality

$$\frac{\Gamma(x)}{\Gamma(y)} \leq \frac{x^{x-1}}{y^{y-1}} e^{y-x}, \quad y > x > 1 \tag{3.31}$$

which is a special case of a bound proved in Kečkić and Vasić [1971]. We have taken $x = i - 1 - 1/\alpha$ and $y = i$, for $i \geq n + 1$ and $n \geq 2$. This last bound is the general term of a convergent series. The integral test for convergence (aka the Maclaurin–Cauchy test) tells us that:

$$\sum_{i=n+1}^{\infty} \frac{1}{i^{1+\frac{1}{\alpha}}} \underset{n \rightarrow \infty}{\sim} \frac{1}{n^{\frac{1}{\alpha}}}.$$

In particular, by bounding $n/(n-1-1/\alpha)$ by 2, we find:

$$|\mathbb{E}[S_{>}]| = \left| \mathbb{E} \left[\sum_{(r,\mathbf{v}) \in \eta_{>}} \langle r\mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta_{r,\mathbf{v}})}, \nabla g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \rangle \right] \right| \leq 2d\alpha \frac{1}{n^{1+\frac{1}{\alpha}}}.$$

3. Finally, we treat the case of $\eta_{=}$. We will show that $\mathbb{E}[S_{=}] \simeq \mathbb{E}[\langle \mathbf{Z}_n, \nabla g_h(\mathbf{Z}_n) \rangle]$. For this, we forcefully make appear the second part of the generator times α . A consequence of lemma 3.30 is that it equals:

$$\langle \mathbf{m}_n(\eta), \nabla g_h(\mathbf{m}_n(\eta)) \rangle = \sum_{(r,\mathbf{v}) \in \eta_{=}} \langle r\mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta)}, \nabla g_h(\mathbf{m}_n(\eta)) \rangle.$$

Therefore the error we commit by making this substitution is:

$$\begin{aligned} S_{=} - \langle \mathbf{m}_n(\eta), \nabla g_h(\mathbf{m}_n(\eta)) \rangle &= \sum_{(r,\mathbf{v}) \in \eta_{=}} r \left[\langle \mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta_{r,\mathbf{v}})}, \nabla g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \rangle - \langle \mathbf{v} \mathbf{1}_{r\mathbf{v}, \mathbf{m}_n(\eta)}, \nabla g_h(\mathbf{m}_n(\eta)) \rangle \right] \\ &= \sum_{j=1}^d \sum_{(r,\mathbf{v}) \in \eta_{=}} rv^j \left[\partial_j g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \mathbf{1}_{\{rv^j \geq m_n^j(\eta_{r,\mathbf{v}})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbf{1}_{\{rv^j \geq m_n^j(\eta)\}} \right] \\ &= \begin{cases} \sum_{j=1}^d \sum_{(r,\mathbf{v}) \in \eta_{=}} rv^j \left[\partial_j g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \mathbf{1}_{\{rv^j \geq m_n^j(\eta_{r,\mathbf{v}})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbf{1}_{\{rv^j \geq m_n^j(\eta)\}} \right] \\ 0, \end{cases} \end{aligned}$$

the first case occurring as soon as one coordinate of $r_{n+1}\mathbf{v}_{n+1}$ is greater than its counterpart of $\mathbf{m}_n(\eta)$, while both terms in the subtraction are equal if $r_{n+1}\mathbf{v}_{n+1} \in [\mathbf{0}, \mathbf{m}_n(\eta)]$. Besides, we can bound rv^j by r_1 , and the partial derivatives by α . Gathering those arguments, we see that the error is bounded by:

$$\begin{aligned} &\mathbb{E} \left[\sum_{(r,\mathbf{v}) \in \eta_{=}} rv^j \left| \partial_j g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \mathbf{1}_{\{rv^j \geq m_n^j(\eta_{r,\mathbf{v}})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbf{1}_{\{rv^j \geq m_n^j(\eta)\}} \right| \right] \\ &= \mathbb{E} \left[\sum_{(r,\mathbf{v}) \in \eta_{=}} rv^j \left| \partial_j g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r\mathbf{v}) \mathbf{1}_{\{rv^j \geq m_n^j(\eta_{r,\mathbf{v}})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbf{1}_{\{rv^j \geq m_n^j(\eta)\}} \right| \mathbf{1}_{A_n} \right], \end{aligned}$$

where

$$A_n := \bigcup_{l=1}^d \bigcap_{k=1}^n \{v_{n+1}^l > v_k^l\}.$$

As a result, we get

$$\begin{aligned} |\mathbb{E}[S_{=}] - \mathbb{E}[\langle \mathbf{Z}_n, \nabla g_h(\mathbf{Z}_n) \rangle]| &\leq 2d\alpha \mathbb{E}[r_1 \mathbf{1}_{A_n}] \\ &\leq 2d^2\alpha \mathbb{E}[r_1 \mathbf{1}_{\bigcap_{k=1}^n \{v^j > v_k^j\}}] \\ &= 2d^2\alpha \mathbb{E}[r_1] \mathbb{P} \left(\bigcap_{k=1}^n \{v^j > v_k^j\} \right) \\ &\leq 2d^2\alpha \Gamma \left(1 - \frac{1}{\alpha} \right) \frac{1}{n} \end{aligned}$$

where r_1 has the Fréchet distribution $\mathcal{F}(\alpha)$. The presence of the term d^2 comes from the double sum: the sum over j has d terms, and the sum over $\eta_{=}$ has a random number of terms which is bounded by d .

□

Through a slightly finer analysis, it is possible to obtain a better rate of convergence in the previous result. This comes at the price of working with smoother test functions, namely twice Lipschitz functions (*i.e.* 1-Lipschitz functions with 1-Lipschitz partial derivatives).

Corollary 3.2.10 *We use the notations of proposition 3.2.8. There exists a constant $C_\alpha > 0$ depending only on d and α such that*

$$d_{[2]}(\mathbf{Z}_n, \mathbf{Z}) \leq C_\alpha \frac{1}{n^{1+\frac{1}{\alpha}}}, \quad n \geq 2.$$

Furthermore we have that $C_\alpha = O\left(\left(1 - \frac{1}{\alpha}\right)^{-2}\right)$ when α goes to 1^+ .

Proof. Because a doubly Lipschitz function h on E_0 is 1-Lipschitz by definition, all the arguments given in the proof of proposition 3.2.8 apply again. A careful examination shows that we lose the rate of convergence of $n^{-(1+1/\alpha)}$ when dealing with S_- . More precisely we must bound more accurately

$$\sum_{(r,v) \in \eta_n} rv^j |\partial_j g_h(\mathbf{m}_n(\eta_{r,v}) \oplus r\mathbf{v}) \mathbf{1}_{\{rv^j \geq \mathbf{m}_n^j(\eta_{r,v})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbf{1}_{\{rv^j \geq \mathbf{m}_n^j(\eta)\}}| \mathbf{1}_{A_n} \quad (3.32)$$

We define two subsets of $\llbracket 1, d \rrbracket$:

$$\begin{aligned} J_1 &:= \{j \in \llbracket 1, d \rrbracket, rv^j = \mathbf{m}_n^j(\eta)\} \\ J_2 &:= \{j \in \llbracket 1, d \rrbracket, r_{n+1}v_{n+1}^j < rv^j\} \end{aligned}$$

Recall that unless $v^j = 0$, in which case everything is null, we have $r_{n+1}v_{n+1}^j \neq rv^j$ almost surely. To make the rest of the proof clear, we distinguish all four possible cases, depending on whether j belongs to J_1 and/or J_2 , or not.

1. $j \in J_1 \cap J_2$ - In that case, both indicator functions are equal to 1. Now, because $\partial_j g_h$ is $C_{2,\alpha}$ -Lipschitz, we have:

$$\begin{aligned} rv^j |\partial_j g_h(\mathbf{m}_n(\eta_{r,v}) \oplus r\mathbf{v}) \mathbf{1}_{\{rv^j \geq \mathbf{m}_n^j(\eta_{r,v})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbf{1}_{\{rv^j \geq \mathbf{m}_n^j(\eta)\}}| & \\ & \leq C_{2,\alpha} rv^j \|\mathbf{m}_n(\eta_{r,v}) \oplus r\mathbf{v} - \mathbf{m}_n(\eta)\|_1 \\ & \leq C_{2,\alpha} r_1 \|\mathbf{m}_n(\eta_{r,v}) \oplus r\mathbf{v} - \mathbf{m}_n(\eta)\|_1 \\ & \leq 2d C_{2,\alpha} r_1 r_{n+1}. \end{aligned}$$

The coordinates of the vector $\mathbf{m}_n(\eta_{r,v}) \oplus r\mathbf{v} - \mathbf{m}_n(\eta)$ are either null or a factor of r_{n+1} by some $v^j \in [0, 1]$, hence the last inequality. To compute the expectation of $r_1 r_{n+1}$, we let E_1, \dots, E_{n+1} be $n+1$ i.i.d. random variables with the exponential distribution $\mathcal{E}(1)$. Then

$$(r_1, r_{n+1}) \stackrel{d}{=} (E_1^{-\frac{1}{\alpha}}, (E_1 + \dots + E_{n+1})^{-\frac{1}{\alpha}}) \leq (E_1, (E_2 + \dots + E_{n+1})^{-\frac{1}{\alpha}}).$$

Thus, by setting $\Gamma_n := \sum_{i=2}^{n+1} E_i$, we find

$$\begin{aligned} \mathbb{E}[r_1 r_{n+1}] &\leq \mathbb{E}[r_1] \mathbb{E}[\Gamma_n^{-\frac{1}{\alpha}}] \\ &= \Gamma\left(1 - \frac{1}{\alpha}\right) \frac{\Gamma(n - \frac{1}{\alpha})}{\Gamma(n)} \\ &\leq \Gamma\left(1 - \frac{1}{\alpha}\right) \left(n - \frac{1}{\alpha}\right)^{-\frac{1}{\alpha}} \\ &\leq 2\Gamma\left(1 - \frac{1}{\alpha}\right) \frac{1}{n^{\frac{1}{\alpha}}}. \end{aligned}$$

by using once again inequality (3.31). This gives us the term $n^{-1/\alpha}$ while the indicator function $\mathbb{1}_{A_n}$, which is independent of the $(r_i)_{i \geq 1}$, yields the term n^{-1} as before. Notice summing this bound makes w.r.t. j and (r, \mathbf{v}) appear a factor d^2 , while $C_{2,\alpha}$ also depends on d . Thus the final constant depends on d^4 .

2. $j \in J_1^c \cap J_2$ - If $j \notin J_1$, then the contribution of $r v^j$ to $\mathbf{m}_n^j(\eta)$ is ignored and thus the second indicator function in (3.32) vanishes. So does the first indicator function; otherwise $r v^j$ would have to be greater than both $r_{n+1} v_{n+1}^j$ and $\mathbf{m}_n^j(\eta_{r,\mathbf{v}}) = \mathbf{m}_n^j(\eta)$. This contradicts $j \notin J_1$. Consequently, both indicator functions are null.
3. $j \in J_1 \cap J_2^c$ - Under the assumption that $j \notin J_2$, the first indicator function is equal to 0. It also implies that $r v^j < r_{n+1} v_{n+1}^j \leq r_{n+1}$. The second term is not null and bounded by a constant, so that

$$r v^j |\partial_j g_h(\mathbf{m}_n(\eta_{r,\mathbf{v}}) \oplus r \mathbf{v}) \mathbb{1}_{\{r v^j \geq \mathbf{m}_n^j(\eta_{r,\mathbf{v}})\}} - \partial_j g_h(\mathbf{m}_n(\eta)) \mathbb{1}_{\{r v^j \geq \mathbf{m}_n^j(\eta)\}}| \leq 2\alpha r_{n+1}$$

4. $j \in J_1^c \cap J_2^c$ - As seen previously, $j \notin J_1$ is enough to make both indicator functions vanish.

To prove the estimate on the constant, recall from proposition that $C_{2,\alpha}$ depends on the integral of $\gamma_t^{-1/\alpha}$, and observe that:

$$\int_0^\infty \gamma_t^{-\frac{1}{\alpha}} dt = \int_0^1 \gamma_t^{-\frac{1}{\alpha}} dt + \int_1^\infty \gamma_t^{-\frac{1}{\alpha}} dt \leq \int_0^1 t^{-\frac{1}{\alpha}} dt + \int_1^\infty \gamma_t^{-\frac{1}{\alpha}} dt = \left(1 - \frac{1}{\alpha}\right)^{-1} + \int_1^\infty \gamma_t^{-\frac{1}{\alpha}} dt.$$

Because $\Gamma(x) \underset{x \rightarrow 0^+}{\sim} \frac{1}{x}$, we see that the constant in the bound for the case $j \in J_1 \cap J_2$ is of order $(1 - 1/\alpha)^{-2}$, hence concluding the proof. case \square

We make two observations: first the bound of theorem 3.2.10 gets better as α is closer to 1 but in exchange the constant C_α explodes. Second we had to resort to the distance $d_{[2]}$ to obtain this rate. Unlike the Kolmogorov distance, it is not invariant by monotonous bijective transformations and so we cannot deduce rates of convergence when $\alpha \in (0, 1]$. Given the properties of the Stein's solution for the Kolmogorov distance, it is not obvious to find directly a bound for this distance. We bring a partial solution to both problems by using proposition 1.4.2 proved in Gaunt and Li [2023]:

Corollary 3.2.11 *Let $\alpha \in \mathbb{R}_+^*$ and assume that ν is such that $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu)$ has a bounded density with respect to the Lebesgue measure on \mathbb{R}^d . Then there exist a constant $C > 0$*

depending only d such that the de Haan-LePage series $(\mathbf{Z}_n)_{n \geq 1}$ satisfies:

$$d_{\text{Kol}}(\mathbf{Z}_n, \mathbf{Z}) \leq C \left(\frac{\log n}{n} \right)^{\frac{2}{3}}, \quad n \geq 2.$$

Proof. Let $\mathbf{Z}' \sim \mathcal{MS}(\beta, \nu)$ and notice that \mathbf{Z}' has the same distribution as $\mathbf{Z}^{\alpha/\beta}$ for every positive β . The same goes for the partial "sum" \mathbf{Z}'_n , which has the same law as $\mathbf{Z}_n^{\alpha/\beta}$.

Now, thanks to proposition 1.4.2 and corollary 3.2.10, we know that there exists a constant C independent of n and β such that

$$d_{\text{Kol}}(\mathbf{Z}'_n, \mathbf{Z}') \leq C \left(1 - \frac{1}{\beta} \right)^{-\frac{2}{3}} \frac{1}{n^{\frac{1}{3}(1+\frac{1}{\beta})}}$$

for n greater than 2 and any $\beta \in (1, +\infty)$. Because $\mathbf{x} \mapsto \mathbf{x}^{\alpha/\beta}$ is a non-decreasing function on E_0 , and the Kolmogorov distance is invariant by non-decreasing transformations, the left-hand side is also equal to $d_{\text{Kol}}(\mathbf{Z}_n, \mathbf{Z})$. In particular, it does not depend on β . Thus we are free to optimize the right-hand side w.r.t. β . Taking $\beta^{-1} = 1 - (\log n)^{-1}$, we find the announced result. \square

Remark 3.2.12. To the best of our knowledge, those are the first results quantifying the speed of convergence of de Haan-LePage series. In the case of α -stable distributions, the literature is richer: Ledoux and Paulauskas [1996], Davydov and Nagaev [1999], Bentkus et al. [1996, 2000] managed to obtain a rate of convergence in total variation distance for all $\alpha \in (0, 2)$. None relies on Stein's method to prove their result. \blacksquare

We conclude this section by giving an heuristic way to prove Berry-Esseen bounds for the multivariate extreme value theorem, *i.e.* how to bound the distance between a renormalized coordinatewise maximum of random vectors and a max-stable distribution. Using the same arguments as in dimension 1 should be possible, but at the cost of a great loss of generality because random vectors with a density with respect to Lebesgue measure are less common. For instance, it is uncommon to work with continuous random variables with no density. A famous example is the Cantor distribution. In higher dimensions, even max-stable random vectors may not admit a density, depending on their angular measures. To circumvent this difficult, we propose the following argument. Assume $\mathbf{X}_1, \dots, \mathbf{X}_n$ are n i.i.d. random vectors with *continuous* common distribution function $F_{\mathbf{X}}$. Set

$$\mathbf{Z}_{n,k} := a_k^{-1} \bigoplus_{i=1}^n \mathbf{X}_i$$

for some renormalizing sequence $(a_n)_{n \geq 1}$ which is such that:

$$n\mathbb{P}(a_n^{-1} \mathbf{X}_1 \not\leq \mathbf{x}) \xrightarrow{n \rightarrow \infty} \mu[\mathbf{0}, \mathbf{x}]^c.$$

where μ satisfies the equality $\mu(tA) = t^{-\alpha} \mu(A)$ for every positive t and $A \in \mathcal{B}(E_0)$. That assumption is equivalent to the convergence of $(\mathbf{Z}_{n,n})_{n \geq 1}$ to a max-stable random vector with exponent measure μ (see Resnick [2006] page 211). For commodity, we will assume that \mathbf{X}_1 has its support in E_0 . Let:

$$\eta_{n,k} := \{a_k^{-1} \mathbf{X}_1, \dots, a_k^{-1} \mathbf{X}_n\}.$$

The point process $\eta_{n,k}$ is a binomial process with n points and intensity measure $a_k F_{\mathbf{X}}(a_k d\mathbf{x})$. In this context, it will be easier to work directly with the exponent measure μ rather than with its

polar decomposition $\alpha r^{-(\alpha+1)} dr \otimes d\nu(\mathbf{u})$. Recall that

$$\mathbf{D}_\mu f(\mathbf{x}) = \int_{E_0} (f(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x})) d\mu(\mathbf{y}) = \frac{1}{\alpha} \int_{E_0} \langle \mathbf{y} \mathbf{1}_{\mathbf{y}, \mathbf{x}}, \nabla g(\mathbf{y} \oplus \mathbf{x}) \rangle d\mu(\mathbf{y}).$$

As before, we want to show that for nice functions g , we have:

$$\mathbb{E}[\langle \mathbf{Z}_{n,n}, \nabla g(\mathbf{Z}_{n,n}) \rangle] \simeq \alpha \mathbb{E}[\mathbf{D}_\mu g(\mathbf{Z}_{n,n})].$$

One way to do so without relying on the assumption that $\mathbf{Z}_{n,n}$ has a density consists in using the Campbell-Mecke's formula for binomial processes. It states that for any nice function φ on $E_0 \times \mathfrak{N}_{E_0}$, we have:

$$\mathbb{E} \left[\int_{E_0} \varphi(\mathbf{y}, \eta_{n,n} - \delta_{\mathbf{y}}) d\eta_{n,n}(\mathbf{y}) \right] = \int_{E_0} \mathbb{E}[\varphi(\mathbf{y}, \eta_{n-1,n})] d\mu_n(\mathbf{y}). \quad (3.33)$$

where $\delta_{\mathbf{y}}$ is the Dirac measure in \mathbf{y} and μ_n is the intensity measure of the binomial process $a_n^{-1} \eta_n$. In our context, $\eta_{n,n} = \sum_{i=1}^n \delta_{a_n^{-1} \mathbf{X}_i}$ and so, $\mu_n[\mathbf{0}, \mathbf{x}]^c = n \mathbb{P}(\mathbf{X} \not\leq a_n \mathbf{x})$. We apply the formula (3.33) to the map:

$$\varphi(\mathbf{y}, \phi) = \langle \mathbf{y} \mathbf{1}_{\mathbf{y}, \mathbf{m}(\phi)}, \nabla g(\mathbf{y} \oplus \mathbf{m}(\phi)) \rangle.$$

This yields an integration-by-parts formula for $\mathbf{Z}_n = \mathbf{m}(\eta_n)$:

$$\begin{aligned} \mathbb{E} \left[\int_{E_0} \varphi(\mathbf{y}, \eta_n - \delta_{\mathbf{y}}) d\eta_n(\mathbf{y}) \right] &= \mathbb{E} \left[\int_{E_0} \varphi(\mathbf{y}, \eta_n - \delta_{\mathbf{y}}) d\eta_n(\mathbf{y}) \right] \\ &= \int_{E_0} \mathbb{E}[\langle \mathbf{y} \mathbf{1}_{\mathbf{y}, \mathbf{Z}_{n-1,n}}, \nabla g(\mathbf{y} \oplus \mathbf{Z}_{n-1,n}) \rangle] d\mu_n(\mathbf{y}). \end{aligned}$$

The left-hand side simplifies greatly, since only one point $\mathbf{y} \in \eta_n$ can give one (or more) coordinate j of \mathbf{Z}_n . Indeed, no other point in η_n can reach the maximum since we assumed the \mathbf{Z}_i to be diffuse. In that case we have $Z_n^j = y^j$ and thus the previous identity takes the form

$$\begin{aligned} \mathbb{E}[\mathbf{Z}_n g'(\mathbf{Z}_n)] &= \int_{E_0} \mathbb{E}[\langle \mathbf{y} \mathbf{1}_{\mathbf{y}, \mathbf{Z}_{n-1,n}}, \nabla g(\mathbf{y} \oplus \mathbf{Z}_{n-1,n}) \rangle] d\mu_n(\mathbf{y}) \\ &= \sum_{j=1}^d \int_{E_0} y^j \mathbb{E}[\partial_j g(\mathbf{y} \oplus \mathbf{Z}_{n-1,n}) \mathbf{1}_{\{Z_{n-1,n}^j \leq y^j\}}] d\mu_n(\mathbf{y}). \end{aligned}$$

On the other hand, we have

$$\alpha \mathbb{E}[\mathbf{D}_\mu g(\mathbf{Z}_n)] = \sum_{j=1}^d \int_{E_0} y^j \mathbb{E}[\partial_j g(\mathbf{y} \oplus \mathbf{Z}_{n,n}) \mathbf{1}_{\{Z_{n,n}^j \leq y^j\}}] d\mu(\mathbf{y}).$$

As a result, the expectation of $\mathcal{L}_\mu g(\mathbf{Z}_n)$ becomes:

$$\begin{aligned} \alpha \mathbb{E}[\mathcal{L}_\mu g(\mathbf{Z}_n)] &= \alpha \mathbb{E}[\mathbf{D}_\mu g(\mathbf{Z}_n)] - \mathbb{E}[\mathbf{Z}_n g'(\mathbf{Z}_n)] \\ &= \sum_{j=1}^d \int_{E_0} y^j \left(\mathbb{E}[\partial_j g(\mathbf{y} \oplus \mathbf{Z}_{n,n}) \mathbf{1}_{\{Z_{n,n}^j \leq y^j\}}] d\mu(\mathbf{y}) - \mathbb{E}[\partial_j g(\mathbf{y} \oplus \mathbf{Z}_{n-1,n}) \mathbf{1}_{\{Z_{n-1,n}^j \leq y^j\}}] d\mu_n(\mathbf{y}) \right). \end{aligned}$$

In dimension 1, the term inside g is simply y and is not random anymore. Then, assuming that the X_i have a density, one can rewrite μ_n in terms of $nF_X^{n-1}(a_n r)f_X(a_n r)$, effectively recovering the well-known density of a maximum of n i.i.d. random variables. In the general case, there are two sources of error: replacing the exponent measure μ by μ_n , and replacing $Z_{n,n}$ by $Z_{n-1,n}$. Notice also that $Z_{n-1,n}$ depends only on $n - 1$ random variables, making the above formula a "max" counterpart to the well-known leave-one-out approach used to apply Stein's method to sums of random variables.

4 Stein's method for the Gumbel and Weibull distributions

Contents

4.1	Properties of the Stein solution	143
4.1.1	The case of the Gumbel distribution	143
4.1.2	The case of the positive Weibull distribution	147
4.2	Application to the coupon collector problem	153
4.2.1	Bounding the error on $[\varepsilon, \infty)$	158
4.2.2	Bounding the error on $[0, \varepsilon]$	166

This chapter is dedicated to the application of Stein's method to the Gumbel distribution as well as the Weibull distribution. Unlike the previous chapter, we focus only on the univariate case, as this where the most important differences with the Fréchet distribution lie.

4.1 Properties of the Stein solution

4.1.1 The case of the Gumbel distribution

We start by proving that \mathcal{L}_0 is a Stein operator for the Gumbel distribution: it characterizes this law in the sense of the next proposition.

Proposition 4.1.1 *Let X be a random variable on \mathbb{R} such that $\mathbb{E}[e^{-X}]$ is finite. Then X has the Gumbel distribution $\mathcal{G}(0, 1)$ if and only if*

$$\mathbb{E}[f'(X)] = \mathbb{E}[\mathbf{D}_0 f(X)],$$

for all $f \in \text{Lip}(\mathbb{R}, \mathbb{R})$.

Proof. The reciprocal means that if X has the Gumbel distribution, then $\mathbb{E}[\mathcal{L}_0 f(X)] = 0$. This can be proved by a generic argument regarding semi-groups:

$$\begin{aligned} \mathbb{E}[\mathcal{L}_0 f(X)] &= \mathbb{E}\left[\lim_{t \rightarrow 0} \frac{d}{dt} \mathbf{P}_t^0 f(X)\right] \\ &= \lim_{t \rightarrow 0} \frac{d}{dt} \mathbb{E}[\mathbf{P}_t^0 f(X)] \\ &= \lim_{t \rightarrow 0} \frac{d}{dt} \mathbb{E}[\mathbf{P}_0^0 f(X)] \\ &= 0, \end{aligned}$$

because the Gumbel distribution is a stationary measure for $(\mathbf{P}_t^0)_{t \geq 0}$.

Now we turn to the direct implication: we first prove that the Laplace transform $L_X : \lambda \mapsto \mathbb{E}[e^{-\lambda X}]$ of X is defined on a neighbourhood of 0. The assumption on X and relation (2.29) yield:

$$\mathbb{E}[f'(X)] = \mathbb{E}[e^{-X} f'(X + Y)],$$

where $Y \sim \mathcal{E}(1)$ and is independent of X . By assumption, we already know that $L_X(1)$ is finite, and thus L_X is well-defined on $[0, 1]$. Now take $f = x \mapsto e^{tx}$, for some $t \in [0, 1)$. Then the previous identity gives:

$$L_X(-t) = \mathbb{E}[e^{-(1-t)X} e^{tY}] = L_X(1-t) \frac{1}{1-t}.$$

Because t is less than 1, the right-hand side is well defined, so the left-hand side is also well-defined. Thus so is L_X is well defined (at least) on $(-1, 1]$. Actually, by taking $f = x \mapsto e^{-x}$, it is easily seen that L_X is in fact defined on $(-1, 2]$. By recurrence, one shows that L_X is well-defined on $] -1, \infty)$.

To determine L_X , we apply the Bohr-Mollerup theorem (see Artin [1964]) to $f : \lambda \mapsto L_X(\lambda - 1)$, which is well-defined on \mathbb{R}_+^* thanks to the previous discussion. This theorem states that if f is defined on \mathbb{R}_+^* , is log-convex, satisfies $f(1) = 1$ and $f(x+1) = xf(x)$ for all positive x , then f is the Gamma function. Clearly, $f(1) = L_X(0) = 1$. It is also well-known that the logarithm of the Laplace transform of a probability measure is convex (this is an easy consequence of Hölder's inequality). Finally, the assumption gives for all positive λ :

$$\begin{aligned} f(\lambda) &= L_X(\lambda - 1) \\ &= \mathbb{E}[e^{-\lambda X} e^{-(\lambda-1)Y}] \\ &= L_X(\lambda) \mathbb{E}[e^{-(\lambda-1)Y}] \\ &= f(\lambda + 1) \frac{1}{\lambda}, \end{aligned}$$

which is the required functional equation. Therefore

$$L_X(\lambda) = \Gamma(\lambda + 1), \quad \lambda \in (-1, +\infty).$$

An easy computation shows that this is indeed the Laplace transform of the Gumbel distribution. Since they coincide on a neighbourhood of 0, this proves that X has the Gumbel distribution. □

We will mainly work with this expression of the generator \mathcal{L}_0 of $(\mathbf{P}_t^0)_{t \geq 0}$:

$$\mathcal{L}_0 f(x) = -f'(x) + \int_x^\infty f'(r) e^{-r} dr, \quad x \in \mathbb{R}, \quad f \in \mathcal{S}_\Lambda.$$

As in the previous chapter, we will mainly deal with Lipschitz test functions or with the indicator functions of the sets $(-\infty, z]$ for all non-negative z . The next proposition ensures that \mathbf{P}_t^0 and \mathcal{L}_0^{-1} are defined for those functions. Most of the arguments are the same as in the Fréchet case so we skip the proofs, unless necessary.

Proposition 4.1.2 *Let $Z \sim \mathcal{G}(0, 1)$. Set $h^* = h - \mathbb{E}[h(Z)]$ if h is \mathbb{P}_Z -integrable. We have the following:*

- $\mathbf{P}_t^0 h$ and $\mathcal{L}_0 h$ exist for all $h \in \text{Lip}_1(\mathbb{R}_+, \mathbb{R})$.

- For all $z \in \mathbb{R}$, set $h_z = \mathbb{1}_{(-\infty, z]}$. Then $\mathbf{P}_t^0 h_z$ is well defined.

In both cases, $t \mapsto \mathbf{P}_t^0 h^*(x)$ is integrable for all $x \in \mathbb{R}$, so that $\mathcal{L}_0^{-1} h^*$ exists. In the second case, the pseudo-inverse of h_z^* equals:

$$\mathcal{L}_0^{-1} h_z^*(x) = (x - z)_+ e^{-e^{-z}} - \int_{(x-z)_+}^{\infty} \left(e^{-(1-e^{-t})e^{-z}} - e^{-e^{-z}} \right) dt. \quad (4.1)$$

We next list some properties of those Stein solutions, starting with the Lipschitz case.

Lemma 4.1.3 *Let $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$. Then $\mathbf{P}_t^0 h$ is 1-Lipschitz and its derivative is 2-Lipschitz for all $t \geq 0$.*

Proof. We know that $(\mathbf{P}_t^0 h)'(x)$ is equal to $h'(x - t)e^{-\gamma t e^{-x}}$. Because h' is bounded by 1, we see that $\mathbf{P}_t^0 h$ is 1-Lipschitz. Besides, $(\mathbf{P}_t^0 h)'$ writes as the product of two bounded 1-Lipschitz functions, so it is 2-Lipschitz. \square

Proposition 4.1.4 *1. Let h be a 1-Lipschitz function on \mathbb{R} . The associated Stein solution satisfies:*

$$g'_h(x) = - \int_0^{\infty} e^{-\gamma t e^{-x}} h'(x - t) dt, \quad x \in \mathbb{R}. \quad (4.2)$$

As a result, we have

$$|g'_h(x)| \leq \min(2 + |x|, e^x), \quad x \in \mathbb{R}. \quad (4.3)$$

2. Furthermore, assume that $h \in \text{Lip}_{[2]}(\mathbb{R}_+, \mathbb{R})$. Then g'_h is C -Lipschitz, with C a positive constant.

Proof. 1. Thanks to (2.25), we have for $h \in \mathcal{S}_\Lambda$,

$$(\mathbf{P}_t^0 h)'(x) = e^{-\gamma t e^{-x}} h'(x - t), \quad x \in \mathbb{R}.$$

The absolute value of $(\mathbf{P}_t^0 h)'(x)$ is less than $e^{-\gamma t e^{-b}}$ if $x \in [a, b]$, for any $a, b \in \mathbb{R}$ such that $a \leq b$. Therefore we can exchange the derivation w.r.t. x and the integral to get:

$$|g'_h(x)| \leq \int_0^{\infty} e^{-\gamma t e^{-x}} e^{-t} dt.$$

The changes of variable $u = e^t - 1$ and $t = e^{-x} u$ yield:

$$\int_0^{\infty} e^{-\gamma t e^{-x}} dt = \int_0^{\infty} \frac{e^x}{e^{xt} + 1} e^{-t} dt = \int_0^1 \frac{e^x}{e^{xt} + 1} e^{-t} dt + \int_1^{\infty} \frac{e^x}{e^{xt} + 1} e^{-t} dt$$

When x is non-negative, we can bound the first integral by $\log(e^x + 1) \leq x + e^{-x}$ and the second by 1, while if x is negative, we bound the whole integral by e^x . This yields the announced bound. \square

In the case of indicator functions, the Stein's solution given by the Gumbel semi-group behaves in the same way as in the Fréchet case, except that $x^{-\alpha}$ is replaced by e^{-x} and $x \in \mathbb{R}$. Other than that, the proofs are exactly the same.

Proposition 4.1.5 *Let $z \in \mathbb{R}_+^*$. Denote by g_z the pseudo-inverse of $h_z^* = h_z - \mathbb{P}(Z \leq z)$. We have the following:*

1. g_z is non-decreasing and constant over $(-\infty, z]$:

$$g_z(x) = - \int_0^\infty \left(e^{-(1-e^{-t})e^{-z}} - e^{-e^{-z}} \right) dt, \quad x \in (-\infty, z].$$

Furthermore we have the equivalent

$$g_z(x) \underset{x \rightarrow +\infty}{\sim} x e^{-e^{-z}}.$$

2. g_z is continuously differentiable over $\mathbb{R} \setminus \{z\}$ and its derivative equals:

$$g'_z(x) = \begin{cases} e^{-(e^{-z}-e^{-x})} & \text{if } x > z \\ 0 & \text{otherwise.} \end{cases} \quad (4.4)$$

3. g'_z satisfies the inequality

$$|g'_z(x)| \leq \mathbf{1}_{\{x > z\}}, \quad x \in \mathbb{R}. \quad (4.5)$$

Thus it also satisfies for all x, y :

$$|g_z(x) - g_z(y)| \leq |x \oplus z - y \oplus z|. \quad (4.6)$$

In particular, g_z is 1-Lipschitz.

4. We have the inequality:

$$0 \leq \mathbf{D}_0 g_z(x) \leq e^{-x}, \quad x \in \mathbb{R}.$$

More precisely

$$\mathbf{D}_0 g_z(x) = e^{-e^{-z}} (e^{e^x \odot e^z} - 1), \quad x \in \mathbb{R}. \quad (4.7)$$

Consequently, $\sup_{z \in \mathbb{R}} \lim_{x \rightarrow -\infty} \mathbf{D}_0 g_z(x)$ is finite.

With those tools at hand, we can now give general bounds on the distance to the Gumbel distribution.

Proposition 4.1.6 *Let W be a random variable such that there exists some positive K for which F_W is absolutely continuous on $[K, +\infty)$. Assume Z a random variable with the Gumbel distribution $\mathcal{G}(0, 1)$. Let also \mathcal{H} be a space of test-functions.*

- If $\mathcal{H} = \text{Lip}_1(\mathbb{R}, \mathbb{R})$, assume that there exists some $p > 1$ such that $|W|^p$ is integrable and

let q be the conjugate of p . Then there exists a positive constant C such that we have

$$d_W(W, Z) \leq \left(C + \mathbb{E}[|W|^p]^{\frac{1}{p}} \right) F_W(K)^{\frac{1}{q}} + \int_K^\infty r \min(2+r, e^r) F_W(r) \left| \frac{f_W(r)}{F_W(r)} - e^{-r} \right| dr. \quad (4.8)$$

- If $\mathcal{H} = \{h_z = x \mapsto \mathbb{1}_{(-\infty, z]}(x), z \in \mathbb{R}\}$, then

$$d_K(W, Z) \leq 2F_W(K) + \sup_{z > K} \int_z^\infty e^{-(e^{-z} - e^{-r})} F_W(r) \left| \frac{f_W(r)}{F_W(r)} - e^{-r} \right| dr. \quad (4.9)$$

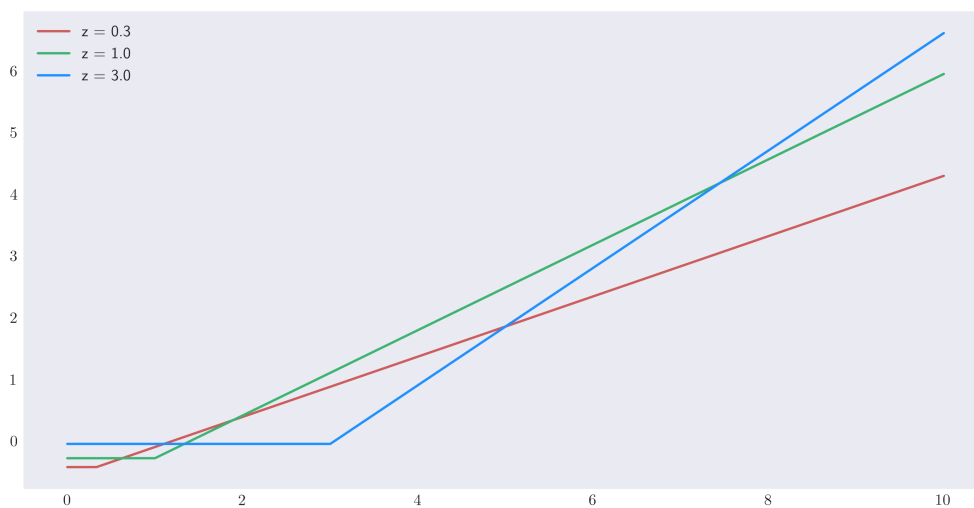


Figure 4.1: Graphs of g_z over $[1, 10]$ for $z \in \{\frac{1}{3}, 1, 3\}$ - Gumbel case

Let X be a random variable in the domain of attraction of the Gumbel distribution. Unlike the Fréchet case, where it is enough to assume that F_X behaves like $x^{-\alpha}$ to standardize the normalizing constants, such a thing is much less easy to do in the Gumbel case. Therefore, a counterpart to proposition 3.1.9 does not seem accessible in this context.

4.1.2 The case of the positive Weibull distribution

Next we adapt the results proven for the Fréchet distribution to the positive, min-stable, Weibull distribution.

Proposition 4.1.7 *Let X be a random variable on \mathbb{R}_+ such that $\mathbb{E}[X]$ is finite. Then X has the Weibull distribution $\mathcal{W}_+(\alpha)$ if and only if*

$$\mathbb{E}[X f'(X)] = -\alpha \mathbb{E}[\mathbf{D}_{-\alpha, +} f(X)],$$

for all $f \in \mathcal{C}^1(\mathbb{R}_+^*)$ such that there exists a constant C for which $|f'(z)| \leq Cz$ for z small enough.

The generator $\mathcal{L}_{-\alpha,+}$ of $(\mathbf{P}_t^{-\alpha,+})_{t \geq 0}$ is also equal to

$$\mathcal{L}_{-\alpha,+}f(x) = \frac{1}{\alpha}xf'(x) + \int_0^x f'(r)r^\alpha dr, \quad x > 0, \quad f \in \mathcal{S}_\Phi.$$

We now express the Stein's solutions provided by the Weibull process and give their main properties.

Proposition 4.1.8 *Let $\alpha > 0$ and $Z \sim \mathcal{W}_+(\alpha)$. Set $h^* = h - \mathbb{E}[h(Z)]$ if h is \mathbb{P}_Z -integrable. We have the following:*

- $\mathbf{P}_t^{-\alpha,+}h$ and $\mathcal{L}_{-\alpha,+}h$ exist for all $h \in \text{Lip}_1(\mathbb{R}_+, \mathbb{R})$.
- For all $z \in \mathbb{R}_+$, set $h_z = \mathbb{1}_{(-\infty, z]}$. Then $\mathbf{P}_t^{-\alpha,+}h_z$ is well defined.

In both cases, $t \mapsto \mathbf{P}_t^{-\alpha,+}h^*(x)$ is integrable for all $x \in \mathbb{R}_+^*$, so that $\mathcal{L}_{-\alpha,+}^{-1}h^*$ exists. In the second case, the pseudo-inverse of h_z^* equals:

$$\mathcal{L}_{-\alpha,+}^{-1}h_z^*(x) = \alpha(\log xz^{-1})_+e^{-z^\alpha} - \int_{\alpha(\log xz^{-1})_+}^{\infty} \left(e^{-(1-e^{-t})z^\alpha} - e^{-z^\alpha} \right) dt. \quad (4.10)$$

Proof. We only prove equality (4.10); as before we have

$$\begin{aligned} \mathcal{L}_{-\alpha,+}^{-1}h_z^*(x) &= - \int_0^{\infty} \mathbf{P}_t^{-\alpha,+}h_z^*(x) dt \\ &= - \int_0^{\infty} [\mathbb{P}(e^{t/\alpha}x \odot (1 - e^{-t})^{-1/\alpha}Z \leq z) - \mathbb{P}(Z \leq z)] dt \\ &= - \int_0^{\infty} \mathbb{1}_{\{e^{t/\alpha}x \leq z\}} [\mathbb{P}((1 - e^{-t})^{-1/\alpha}Z \leq z) - \mathbb{P}(Z \leq z)] - \mathbb{1}_{\{e^{t/\alpha}x > z\}} \mathbb{P}(Z \leq z) dt \\ &= \alpha(\log xz^{-1})_+e^{-z^\alpha} - \int_{\alpha(\log xz^{-1})_+}^{\infty} \left(e^{-(1-e^{-t})z^\alpha} - e^{-z^\alpha} \right) dt, \end{aligned}$$

because $e^{t/\alpha}x \leq z$ for t nonnegative is equivalent to $t \geq \alpha(\log xz^{-1})_+$. \square

Here we gather some properties of those Stein solutions, starting with the Lipschitz case.

Proposition 4.1.9 *Let $\alpha > 0$.*

1. *Let h be a 1-Lipschitz function on \mathbb{R}_+^* . The associated Stein solution g_h is α -Lipschitz and its derivative satisfies:*

$$g'_h(x) = - \int_0^{\infty} e^{-\frac{t}{\alpha}} e^{-\gamma t x^\alpha} h'(e^{-\frac{t}{\alpha}}x) dt, \quad x > 0. \quad (4.11)$$

As a result, we have

$$|g'_h(x)| \leq \min(\alpha, x^{-\alpha}), \quad x > 0. \quad (4.12)$$

2. *Consequently, $\lim_{x \rightarrow +\infty} \mathbf{D}_{-\alpha,+}g_h(x)$ is finite.*

3. Furthermore, assume that $h \in \text{Lip}_{[2]}(\mathbb{R}_+^*, \mathbb{R})$, i.e. h and h' are 1-Lipschitz on \mathbb{R}_+^* . Then g'_h is C_α -Lipschitz, with C_α a positive constant depending only on α .

The Stein's solution for test functions equal to h_z takes a shape similar to the one given in the Fréchet case.

Proposition 4.1.10 Let $z \in \mathbb{R}_+^*$. Denote by g_z the inverse of $h_z^* = h_z - \mathbb{P}(Z \leq z)$. We have the following:

1. g_z is non-decreasing and constant over $[0, z]$:

$$g_z(x) = - \int_{\alpha(\log x z^{-1})_+}^{\infty} \left(e^{-(1-e^{-t})z^\alpha} - e^{-z^\alpha} \right) dt, \quad x \in [0, z].$$

Moreover, we have the equivalent

$$g_z(x) \underset{x \rightarrow +\infty}{\sim} \alpha e^{-z^\alpha} \log x.$$

2. g_z is continuously differentiable over $\mathbb{R}_+^* \setminus \{z\}$ and its derivative equals:

$$g'_z(x) = \begin{cases} \frac{\alpha}{x} e^{-(z^\alpha - x^\alpha)} & \text{if } x > z \\ 0 & \text{otherwise.} \end{cases} \quad (4.13)$$

3. g'_z satisfies the inequality

$$|g'_z(x)| \leq \frac{\alpha}{x} \mathbf{1}_{\{x > z\}}, \quad x > 0. \quad (4.14)$$

Thus it also satisfies for all positive x, y :

$$|g_z(x) - g_z(y)| \leq \alpha z |x \oplus z - y \oplus z|. \quad (4.15)$$

In particular, g_z is αz^{-1} -Lipschitz.

4. We have the inequality:

$$0 \leq \mathbf{D}_{-\alpha, +} g_z(x) \leq x^\alpha, \quad x > 0.$$

More precisely

$$\mathbf{D}_{-\alpha, +} g_z(x) = e^{-z^\alpha} (e^{x^\alpha \odot z^\alpha} - 1), \quad x > 0. \quad (4.16)$$

Consequently, $\sup_{z \in \mathbb{R}_+} \lim_{x \rightarrow +\infty} \mathbf{D}_{-\alpha, +} g_z(x)$ is finite.

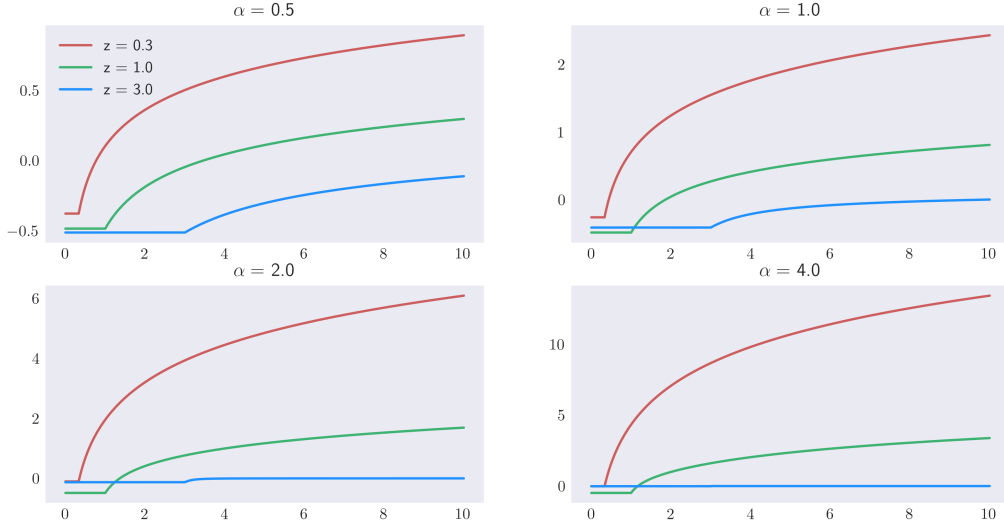


Figure 4.2: Graphs of g_z over $[1, 10]$ for $z \in \{\frac{1}{3}, 1, 3\}$ and $\alpha \in \{\frac{1}{2}, 1, 2, 4\}$ - positive Weibull case

Compared to figure 3.1, the order of the curves is inverted: the curve corresponding to the biggest α is the lowest, while it was the highest for large enough z in the Fréchet case.

As in the two previous cases, it is possible to deduce from the properties of the Stein's solutions general bounds on the Wasserstein and Kolmogorov distance.

Proposition 4.1.11 *Let W be a random variable such that there exists some positive K for which \bar{F}_W is absolutely continuous on $[0, K]$. Assume that Z is a random variable with the Weibull distribution $\mathcal{W}_+(\alpha)$ for some $\alpha > 0$. Let also \mathcal{H} be a space of test-function.*

- If $\mathcal{H} = \text{Lip}_1(\mathbb{R}, \mathbb{R})$, take $\alpha > 1$ and assume there exists $p \in (1, \alpha)$ such that $|W|^p$ is integrable. Let q be the conjugate of p i.e. $p^{-1} + q^{-1} = 1$. Then there exists a positive constant C_α such that we have

$$d_W(W, Z) \leq \left(C_\alpha + \mathbb{E}[|W|^p]^{\frac{1}{p}} \right) \bar{F}_W(K)^{\frac{1}{q}} + \frac{1}{\alpha} \int_0^K r \min(\alpha, r^{-\alpha}) \bar{F}_W(r) \left| \frac{f_W(r)}{\bar{F}_W(r)} - \alpha r^{\alpha-1} \right| dr. \quad (4.17)$$

- If $\mathcal{H} = \{h_z = x \mapsto \mathbb{1}_{(-\infty, z]}(x), z \in \mathbb{R}\}$, then

$$d_K(W, Z) \leq 2\bar{F}_W(K) + \sup_{z < K} \int_0^z e^{-(z^\alpha - r^\alpha)} \bar{F}_W(r) \left| \frac{f_W(r)}{\bar{F}_W(r)} - \alpha r^{\alpha-1} \right| dr. \quad (4.18)$$

We can also use the previous result to give Berry-Esseen-type bounds on the distance between a renormalized maximum of n i.i.d. random variables and a positive Weibull distribution, as we did in the Fréchet case.

Let $(X_i)_{i \geq 1}$ be a sequence of i.i.d. random variables and set

$$M_n := \min_{1 \leq i \leq n} X_i \text{ and } Z_n := \frac{M_n - b_n}{a_n}, \quad n \geq 1$$

where $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$ are sequences of real numbers, with $a_n > 0$ for all n . For $(Z_n)_{n \geq 1}$ to converge to the Weibull distribution, F_X must have finite left end-point x_0 and be regularly varying at x_0 with order α , *i.e.* to write as

$$F_X(x) = (x - x_0)^\alpha L(x - x_0), \quad x \geq x_0$$

with L a slowly-varying function at 0, for x large enough. Just as for the Fréchet distribution, we can give a general bound for the distance between an (almost) arbitrary distribution and the Weibull law.

Proposition 4.1.12 *Let $(X_i)_{i \geq 1}$ be a sequence of i.i.d. random variables with common distribution F_X and left end-point $x_0 > -\infty$. Suppose also there exists $\alpha > 0$ for which we have*

$$F_X(x) \underset{x \rightarrow x_0}{\sim} (x - x_0)^\alpha.$$

We further assume that there is a positive K such that \bar{F}_X is differentiable over $[0, K]$ and which satisfies:

$$\delta(r - x_0)^\alpha \leq F_X(r) \leq \delta'(r - x_0)^\alpha, \quad r \geq K,$$

for some $\delta \in (0, 1]$. Take $a_n = n^{-1/\alpha}$ and $b_n = x_0$, so that $Z_n = n^{1/\alpha}(M_n - x_0)$. Then:

- *Assume that that X_+ , the positive part of X , is integrable. Then there exists a constant C_{α, F_X}^W depending only on α and F_X such that:*

$$d_W(Z_n, Z) \leq C_{\alpha, F_X}^W \left[\frac{1}{n} + n^{1-\frac{1}{\alpha}} \int_0^K e^{-\frac{\delta}{2}nr^\alpha} |r^{-(\alpha-1)} f_X(r) - \alpha|r^\alpha \, dr \right], \quad n \geq 2. \quad (4.19)$$

- *There exists another constant C_{α, F_X}^K depending only on α and F_X such that:*

$$d_K(Z_n, Z) \leq C_{\alpha, F_X}^K \left[\frac{1}{n} + n \int_0^K e^{-\frac{\delta}{2}nr^\alpha} |r^{-(\alpha-1)} f_X(r) - \alpha|r^{\alpha-1} \, dr \right], \quad n \geq 2. \quad (4.20)$$

The previous result implies that the distance between Z_n and the Weibull distribution $\mathcal{W}_+(\alpha)$ is controlled by the gap between $x^{-(\alpha+1)} f_X(x)$ and α , weighted by an exponential term which offsets the contributions the error committed when x is large. This naturally leads us to a counterpart of corollary 3.1.12 for the Weibull distribution:

Corollary 4.1.13 *Assume that F_X is a distribution function satisfying the assumptions of proposition 4.1.12. If it also satisfies*

$$|r^{-(\alpha-1)} f_X(r) - \alpha| \lesssim r^\gamma, \quad r \leq K \quad (4.21)$$

for some $\gamma > 0$, then

- in the Wasserstein case, there exists C_{α, F_X}^W such that:

$$d_W(Z_n, Z) \leq C_{\alpha, F_X}^W \left[\frac{1}{n} + \frac{1}{n^\alpha} \right], \quad n \geq 1, \quad (4.22)$$

- in the Kolmogorov case, there exists another C_{α, F_X}^K such that:

$$d_K(Z_n, Z) \leq C_{\alpha, F_X}^K \left[\frac{1}{n} + \frac{1}{n^\alpha} \right], \quad n \geq 1. \quad (4.23)$$

Notice that the equivalent of lemma 3.1.8 in this setting is simply the equality:

$$\int_0^\infty r^{\alpha+\beta} e^{-\gamma r^\alpha} dr = \frac{1}{\alpha} \frac{1}{\gamma^{1+\frac{\beta+1}{\alpha}}} \Gamma\left(1 + \frac{\beta+1}{\alpha}\right). \quad (4.24)$$

For instance, we find the n^{-1} term in the Kolmogorov bounds by computing:

$$n \int_0^K e^{-\frac{\delta}{2} nr^\alpha} r^{2\alpha-1} dr \leq n \int_0^\infty e^{-\frac{\delta}{2} nr^\alpha} r^{2\alpha-1} dr = \frac{4}{\delta^2 \alpha} \frac{1}{n},$$

which corresponds to taking $\gamma = \delta n/2$ and $\beta = \alpha - 1$ in (4.24).

Example 4.1.14. Now we give a few examples of application of the last results.

1. **Beta distribution** - The Beta distribution with shape parameters $\alpha, \beta > 0$ has density

$$f_X(r) = \frac{r^{\alpha-1}(1-r)^{\beta-1}}{B(\alpha, \beta)} \mathbb{1}_{[0,1]}(r),$$

with $B(\alpha, \beta)$ the Beta function. It generalizes the uniform distribution (and its positive powers), as well as the arcsine distribution. Using the same notations as in the examples given in the Fréchet case (see 3.1.13), we find

$$\begin{aligned} |r^{-(\alpha-1)} f_{\sigma^{-1}X}(r) - \alpha| &= |\sigma r^{-(\alpha-1)} f_X(\sigma r) - \alpha| \\ &= \left| \frac{\sigma^{-\alpha}}{B(\alpha, \beta)} (1 - \sigma r)^{\beta-1} - \alpha \right|. \end{aligned}$$

When r goes to 0, this quantity vanishes if and only if

$$\sigma = (\alpha B(\alpha, \beta))^{\frac{1}{\alpha}}.$$

For that choice of σ and $r \in [0, 1/2]$, we find

$$\begin{aligned} |r^{-(\alpha-1)} f_{\sigma^{-1}X}(r) - \alpha| &= |\sigma r^{-(\alpha-1)} f_X(\sigma r) - \alpha| \\ &= \alpha |(1 - \sigma r)^{\beta-1} - 1| \\ &\lesssim \alpha \sigma r, \end{aligned}$$

because $r \mapsto (1 - \sigma r)^{\beta-1}$ is Lipschitz on $[0, 1/2]$. The last inequality implies we can take $\gamma = 1$ in corollary 4.1.13, thus yielding a rate of convergence of $n^{-1/\alpha}$ in both Kolmogorov and Weierstrass distances if $\alpha \geq 1$, and n^{-1} otherwise.

2. **Gamma distribution** - The Gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$ has density

$$f_X(r) = \frac{\lambda^\alpha}{\Gamma(\alpha)} r^{\alpha-1} e^{-\lambda r} \mathbb{1}_{\mathbb{R}_+}(r).$$

It generalizes the exponential distribution, the χ^2 distribution, as well as the Erlang distribution. With σ equal to

$$\sigma = \lambda^{-1} (\Gamma(\alpha + 1))^{\frac{1}{\alpha}}$$

we find that

$$|r^{-(\alpha-1)} f_{\sigma^{-1}X}(r) - \alpha| = \alpha |e^{-r} - 1| \leq \alpha r.$$

This implies that $\sigma^{-1}Z_n$ converges to the positive Weibull distribution at rate $\max(n^{-1/\alpha}, n^{-1})$ in Kolmogorov and Wasserstein distances.

4.2 Application to the coupon collector problem

In this section, we use the Gumbel semi-group and Stein's method to find a rate of convergence in the classic coupon collector problem. Let us describe its simplest version. A collection is constituted of n distinct items which can be found in each box of a brand of cereals with the same probability $1/n$. We want to complete this collection by buying one box per week until we get all items at least once. It has been proved in Erdős and Rényi [1961] that

$$Z_n := \frac{T_n - n \log n}{n} \xrightarrow[n \rightarrow \infty]{d} \mathcal{G}(0, 1),$$

where T_n is the first (random) time at which all the n coupons have been found at least once. The previous result is well known in the literature and has applications in many fields, such as computer science or engineering (see for instance Boneh and Hofri [1997], Mitzenmacher and Upfal [2017]), and we want to estimate its rate of convergence to evaluate the error made when replacing Z_n by a Gumbel distribution.

A convenient way to study the distribution of T_n consists in introducing the random variables τ_i for $i \in \llbracket 1, n \rrbracket$, where τ_i^n is the number of weeks needed to get the i -th new item while $i - 1$ distinct ones have already been collected. With this definition:

$$T_n = \sum_{i=1}^n \tau_i^n.$$

Furthermore, it can be easily proved that the τ_i^n are independent and have a geometric distribution with parameter $(n - i + 1)/n$ respectively. Notice T_n can be expressed as a maximum of n random variables: we number each coupon from 1 to n and introduce T_i^n the number of weeks spent before finding the coupon number i . Then we have:

Proposition 4.2.1 Set $S_i^n := \sum_{j=1}^i \tau_j^n$ and $(T_{(1)}^n, \dots, T_{(n)}^n)$ the order statistics of the T_i^n . Then:

1. We have the equality in distribution

$$S_i^n \stackrel{d}{=} T_{(i)}^n, \quad i \in \llbracket 1, n \rrbracket.$$

In particular:

$$T_n = \max(T_1^n, \dots, T_n^n).$$

2. The random variables T_1^n, \dots, T_n^n are exchangeable.

Proof. 1. By definition, S_i^n is the number of weeks needed to get i distinct items for the first time. This is also the moment when item j is found for the first time, for a certain (random) j . Consequently, it equals one of the T_j^n , and it is the i -th smallest (by definition it is greater than S_1^n, \dots, S_{i-1}^n). This indeed means that $S_i^n = T_{(i)}^n$.

2. The enumeration of the coupons is arbitrary and each coupon has the same probability of being found. So, for any permutation $\sigma \in \mathfrak{S}_n$, we have:

$$(T_{\sigma(1)}^n, \dots, T_{\sigma(n)}^n) \stackrel{d}{=} (T_1^n, \dots, T_n^n),$$

which is the definition of exchangeability. □

Of course the T_1^n, \dots, T_n^n cannot be independent, because they *always* take distinct values per definition.

As noticed in remark 3.1.4, we cannot work with the Wasserstein distance or the Kolmogorov distance: in both cases g'_h is not defined everywhere, while Z_n is a discrete random variable whose atoms lie in the set

$$\left\{ \frac{k - n \log n}{n}, k \in \mathbb{N}^* \right\}.$$

Therefore we will work with test functions $h \in \text{Lip}_{[2]}(\mathbb{R}_+^*, \mathbb{R})$, *i.e.* with the distance $d_{[2]}$. Let us add that we found the proof given in this section before studying the properties of g_h . Due to its length, we have not reformulated it to take into account those results. Instead we give the original proof which relies on direct manipulations of $(\mathbf{P}_t^0)_{t \geq 0}$.

Our approach relies on identity (2.29): Z has the Gumbel distribution $\mathcal{G}(0, 1)$ if and only if

$$\mathbb{E}[h'(Z)] = \mathbb{E}[e^{-Z} h'(Z + Y)], \quad (4.25)$$

where $Y \sim \mathcal{E}(1)$ and is independent of Z .

Stein's method is usually applied (but not restricted!) to the study of sums of random variables $S_n = \sum_{i=1}^n X_i^n$. The generator approach is one of the many ways to put this method into application and in that context, one tries to make appear the generator associated with the limit distribution, for example the classical Ornstein-Uhlenbeck generator of the normal distribution in the case of Berry-Esseen's theorem. At some point it becomes necessary to compare S_n with S_{n-1} , because we will work with Lipschitz test-functions. We will do something similar in our case, but we need to be careful: even though we manipulate sums of independent random variables, they have to be defined on the same probability space for different n .

Therefore we need to define the random variables $(\tau_i^n)_{1 \leq i \leq n}$ and $(\tau_i^{n-1})_{1 \leq i \leq n-1}$ in such a way that they take close values while having the desired distributions. So first we introduce a coupling between those two sets of random variables. Put simply, the τ_{i-1}^{n-1} are built using the same exponential random variables Y_i as τ_i^n , for $i \geq 2$ (the case $i = 1$ is trivial: $\tau_1^n = 1$ a.s.). The coupling itself relies on the well-known inversion method:

Lemma 4.2.2 Let Y_2, \dots, Y_n be i.i.d. random variables with exponential distribution $\mathcal{E}(1)$. Set $\tau_1^n = \tau_1^{n-1} = 1$ and:

$$\tau_i^n := \left\lceil -\frac{Y_i}{\log(1-p_i^n)} \right\rceil \text{ and } \tau_{i-1}^{n-1} := \left\lceil -\frac{Y_i}{\log(1-p_{i-1}^{n-1})} \right\rceil, \quad i \in \llbracket 2, n \rrbracket$$

where we set:

$$p_i^n := \frac{n-i+1}{n}, \quad n \in \mathbb{N}^*, \quad i \in \llbracket 1, n \rrbracket.$$

Then $\tau_i^n \sim \mathcal{G}(p_i^n)$ and $\tau_i^{n-1} \sim \mathcal{G}(p_i^{n-1})$. Furthermore:

$$\tau_{i-1}^{n-1} \leq \tau_i^n, \quad i \in \llbracket 2, n \rrbracket.$$

Proof. This is simply a consequence of the fact that if $Y \sim \mathcal{E}(1)$, then:

$$\left\lceil \frac{Y}{\lambda} \right\rceil \sim \mathcal{G}(e^{-\lambda}), \quad \lambda > 0.$$

The second part of the lemma is proved by noticing that $p \mapsto \lceil -y/\log(1-p) \rceil$ is non-increasing on $(0, 1)$ for all $y \in \mathbb{R}_+$ and that $p_i^n \leq p_{i-1}^{n-1}$:

$$p_i^n = \frac{n-i+1}{n} \leq \frac{n-i+1}{n-1} = p_{i-1}^{n-1}.$$

□

Now we can compare Z_n and Z_{n-1} . We will need in particular to know how much Z_n is close to Z_{n-1} in norm \mathbf{L}^1 :

Lemma 4.2.3 There exists a positive constant C such that:

$$\|Z_n - Z_{n-1}\|_{\mathbf{L}^1} \leq C \frac{\log n}{n}, \quad n \geq 2.$$

Proof. Notice that since $\tau_1^n = 1$ a.s., we have for all $n \geq 2$:

$$\begin{aligned} Z_n - Z_{n-1} &= \frac{1}{n} \sum_{i=1}^n \tau_i^n - \log n - \frac{1}{n-1} \sum_{i=1}^{n-1} \tau_i^{n-1} + \log(n-1) \\ &= \frac{1}{n} \sum_{i=1}^{n-1} \tau_{i+1}^n - \frac{1}{n-1} \sum_{i=1}^{n-1} \tau_i^{n-1} + \frac{1}{n} + \log\left(1 - \frac{1}{n}\right) \\ &= \frac{1}{n} \sum_{i=1}^{n-1} (\tau_{i+1}^n - \tau_i^{n-1}) - \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \tau_i^{n-1} + \frac{1}{n} + \log\left(1 - \frac{1}{n}\right) \end{aligned}$$

By the triangle inequality, and because $\tau_{i+1}^n \geq \tau_i^{n-1}$ a.s., we find:

$$\begin{aligned}
\|Z_n - Z_{n-1}\|_{\mathbf{L}^1} &\leq \frac{2}{n-1} + \frac{1}{n} \sum_{i=1}^{n-1} \|\tau_{i+1}^n - \tau_i^{n-1}\|_{\mathbf{L}^1} + \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \|\tau_i^{n-1}\|_{\mathbf{L}^1} \\
&= \frac{2}{n-1} + \frac{1}{n} \sum_{i=1}^{n-1} \mathbb{E}[\tau_{i+1}^n - \tau_i^{n-1}] + \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \mathbb{E}[\tau_i^{n-1}] \\
&= \frac{2}{n-1} + \frac{1}{n} \sum_{i=1}^{n-1} \left(\frac{n}{n-i} - \frac{n-1}{n-i} \right) + \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \frac{n-1}{n-i} \\
&= \frac{2}{n-1} + \frac{1}{n} \sum_{i=1}^{n-1} \frac{1}{i} + \frac{1}{n} \sum_{i=1}^{n-1} \frac{1}{i} \\
&= \frac{2}{n-1} + \frac{2}{n} H_{n-1},
\end{aligned}$$

where H_n denotes the n -th harmonic number. □

Notation 1: Set $\alpha_n := 1 - 1/n$ and $\beta_n := -(n-1) \log \alpha_n$, as well as:

$$\delta_n := \log \left(1 - \frac{1}{n} \right) - \frac{1}{n} \log(n-1).$$

We will make constant use of the following inequalities in the sequel for $n \geq 2$:

$$-1 \leq \delta_n \leq \frac{1}{n} \leq 0 \leq \alpha_n < \beta_n < 1. \quad (4.26)$$

The next proposition is basically a Stein's identity for Z_n , which is expressed using both of Z_n and Z_{n-1} . Observe that when n goes to ∞ , it approaches the characteristic relation (4.25) of the Gumbel distribution.

Proposition 4.2.4 *Let f be a function in $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ and $n \geq 2$. We have the identity :*

$$\mathbb{E}[f'(Z_n)] = C_n \mathbb{E} \left[e^{-\beta_n Z_{n-1}} f' \left(\alpha_n x + \frac{G_n}{n} + \delta_n \right) \right], \quad (4.27)$$

where the τ_i^{n-1} are i.i.d., respectively with distribution $\text{Geom}(\frac{n-i}{n-1})$ for each $i \in \llbracket 1, n-1 \rrbracket$, the random variable G_n has distribution $\text{Geom}(\frac{1}{n})$ and is independent of the τ_i^{n-1} as well as of the τ_i^n for all i , and

$$C_n := n \left(1 - \frac{1}{n} \right)^{(n-1) \log(n-1)} \geq \frac{n}{n-1} \geq 1.$$

Proof. The idea is to do a change of probability: we replace the distribution of $(\tau_1^n, \dots, \tau_{n-1}^n, \tau_n^n)$ by the distribution of $(\tau_1^{n-1}, \dots, \tau_{n-1}^{n-1}, \tau_n^n)$. Doing so makes appear a density function which will behave as $e^{-Z_n} \stackrel{d}{=} e^{-Z}$ when n goes to infinity, where $Z \sim \mathcal{G}(0, 1)$. As for τ_n^n , this term is singled out because of the following well-known result:

$$\frac{\tau_n^n}{n} \stackrel{d}{=} \frac{1}{n} \text{Geom} \left(\frac{1}{n} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{E}(1).$$

This will give us the exponential random variable which appears in the generator of the Gumbel distribution. Now denote by P_j^n the joint distribution of $(\tau_1^n, \dots, \tau_j^n)$ for all $j \in \llbracket 1, n \rrbracket$ and $(t_i)_{1 \leq i \leq j} \in (\mathbb{N}^*)^j$:

$$\begin{aligned} P_j^n(t_1, \dots, t_j) &:= \mathbb{P}\left(\bigcap_{i=1}^j \{\tau_i^n = t_i\}\right) \\ &= \prod_{i=1}^j \left(\frac{n-i+1}{n}\right) \left(1 - \frac{n-i+1}{n}\right)^{t_i-1} \\ &= \prod_{i=1}^j \left(\frac{n-i+1}{n}\right) \left(\frac{i-1}{n}\right)^{t_i-1}. \end{aligned}$$

Furthermore, the density of $(\tau_1^n, \dots, \tau_{n-1}^n)$ with respect to the distribution of $(\tau_1^{n-1}, \dots, \tau_{n-1}^{n-1})$ is equal to:

$$\begin{aligned} \frac{P_{n-1}^n(t_1, \dots, t_{n-1})}{P_{n-1}^{n-1}(t_1, \dots, t_{n-1})} &= \frac{\prod_{i=1}^{n-1} \left(\frac{n-i+1}{n}\right) \left(\frac{i-1}{n}\right)^{t_i-1}}{\prod_{i=1}^{n-1} \left(\frac{n-i}{n-1}\right) \left(\frac{i-1}{n-1}\right)^{t_i-1}} \\ &= \left(1 - \frac{1}{n}\right)^{n-1} \frac{n!}{(n-1)!} \frac{\prod_{i=1}^{n-1} \left(\frac{1}{n}\right)^{t_i-1}}{\prod_{i=1}^{n-1} \left(\frac{1}{n-1}\right)^{t_i-1}} \\ &= n \left(1 - \frac{1}{n}\right)^{\sum_{i=1}^{n-1} t_i} \\ &= n \left(1 - \frac{1}{n}\right)^{(n-1) \left(\frac{1}{n-1} \sum_{i=1}^{n-1} t_i\right)} \\ &= n \left(1 - \frac{1}{n}\right)^{(n-1) \log(n-1)} \left(1 - \frac{1}{n}\right)^{(n-1) \left(\frac{1}{n-1} \sum_{i=1}^{n-1} (t_i - \log(n-1))\right)} \end{aligned}$$

Using that equality, we perform a change of probability to get rid of the $(n-1)$ first τ_i^n and replace them by the τ_i^{n-1} . Since τ_n^n is independent of τ_i^n for all $i \leq n-1$, we can replace τ_n^n by a random variable G_n with distribution $\text{Geom}\left(\frac{1}{n}\right)$. This substitution is essential, as τ_n^n is **not** independent of τ_{n-1}^{n-1} , since both have been constructed by using the same exponential random variable Y_n . In other words, we have used *two* couplings of τ_n^n and τ_{n-1}^{n-1} : the first one, described in lemma 4.2.2, and a trivial one, with independent geometric random variables. The purpose of the first was to compare

Z_n and Z_{n-1} , while the second will make appear the random variable Y of lemma 4.25 at the limit.

$$\begin{aligned}
\mathbb{E}[f'(Z_n)] &= \mathbb{E}\left[f'\left(\frac{1}{n} \sum_{i=1}^{n-1} \tau_i^n - \log n + \frac{\tau_n^n}{n}\right)\right] \\
&= \mathbb{E}\left[f'\left(\frac{1}{n} \sum_{i=1}^{n-1} \tau_i^n - \log n + \frac{G_n}{n}\right)\right] \\
&= \mathbb{E}\left[\frac{P_{n-1}^n(\tau_1^{n-1}, \dots, \tau_{n-1}^{n-1})}{P_{n-1}^{n-1}(\tau_1^{n-1}, \dots, \tau_{n-1}^{n-1})} f'\left(\frac{1}{n} \sum_{i=1}^{n-1} \tau_i^{n-1} - \log n + \frac{G_n}{n}\right)\right] \\
&= n \left(1 - \frac{1}{n}\right)^{(n-1) \log(n-1)} \mathbb{E}\left[\left(1 - \frac{1}{n}\right)^{(n-1)Z_{n-1}} f'\left(\frac{1}{n} \sum_{i=1}^{n-1} \tau_i^{n-1} - \log n + \frac{G_n}{n}\right)\right] \\
&= C_n \mathbb{E}\left[\left(1 - \frac{1}{n}\right)^{(n-1)Z_{n-1}} f'\left(\frac{1}{n} \sum_{i=1}^{n-1} \tau_i^{n-1} - \log n + \frac{G_n}{n}\right)\right] \\
&= C_n \mathbb{E}\left[\left(1 - \frac{1}{n}\right)^{(n-1)Z_{n-1}} f'\left(\left(1 - \frac{1}{n}\right)Z_{n-1} + \frac{G_n}{n} + \log\left(1 - \frac{1}{n}\right) - \frac{1}{n} \log(n-1)\right)\right].
\end{aligned}$$

Notice that when n goes to infinity, the term C_n before the expectation converges to 1 (as fast as $\log n/n$ goes to 0).

□

Notation 2: The following notation will be convenient in the sequel:

$$g_n(x) := e^{-\beta_n x} \mathbb{E}\left[f'\left(\alpha_n x + \frac{G_n}{n} + \delta_n\right)\right]. \quad (4.28)$$

We will apply identity (4.27) to $f = \mathbf{P}_t^0 h$, which is indeed of class \mathcal{C}^1 . By using the previous results, we will find a bound to control the rate of convergence of $\mathbb{E}[\mathcal{L}_0 f(Z_n)] = \mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]$ to 0. Because that bound will be integrated over $[0, +\infty)$, we must ensure it is an integrable function of t . To be more precise, the term $\gamma_t^{-1} = (e^t - 1)^{-1}$ will appear throughout the next subsection. It is integrable on $[\varepsilon, \infty)$ for all $\varepsilon > 0$, but not when $\varepsilon = 0$. This problem also appears with the standard Ornstein-Uhlenbeck semi-group, at least in dimension 2 or higher (see Decreusefond [2015]). A simple way to fix it consists in finding two different bounds, one for t in $[0, \varepsilon]$ and the other for t in $[\varepsilon, \infty)$. Optimizing in ε then concludes the proof. The rest of this section is subdivided into two subsections, one for each bound.

4.2.1 Bounding the error on $[\varepsilon, \infty)$

The goal of this section is to prove that the following proposition holds:

Proposition 4.2.5 *Let $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$. There exists a constant $C > 0$ such that:*

$$|\mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]| \leq C \frac{1 + |\log \gamma_t| \log n}{\gamma_t n}, \quad t \in [\varepsilon, \infty), \quad n \in \mathbb{N}^*.$$

To do so, we introduce the decomposition on which relies this subsection (and the next):

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_0 f(Z_n)] &= -\mathbb{E}[f'(Z_n)] + \mathbb{E}[e^{-Z_n} f'(Z_n + Y)] \\
&= -C_n \mathbb{E}[g_n(Z_{n-1})] + \mathbb{E}[g(Z_n)] \\
&= (1 - C_n) \mathbb{E}[g_n(Z_{n-1})] + \mathbb{E}[g(Z_n)] - \mathbb{E}[g_n(Z_{n-1})] \\
&= (1 - C_n) \mathbb{E}[g_n(Z_{n-1})] + (\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z_{n-1})]) + (\mathbb{E}[g(Z_{n-1})] - \mathbb{E}[g_n(Z_{n-1})]) \\
&= (1) + (2) + (3), \tag{4.29}
\end{aligned}$$

where the second equality is a consequence of proposition 4.25. Next we deal with (1) and (2). To do so, we give some properties of g and $\mathbb{E}[g_n(Z_{n-1})]$:

Lemma 4.2.6 *Let $h_t = \mathbf{P}_t^0 h$, where $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$, so that*

$$g(x) = e^{-x} \mathbb{E}[e^{-\gamma_t e^{-(x+Y)}} h'(x + Y - t)],$$

where $Y \sim \mathcal{E}(1)$. We couple G_n with Y in the following way

$$G_n := \frac{1}{n} \left\lceil \frac{-Y}{\log \alpha_n} \right\rceil \sim \text{Geom}\left(\frac{1}{n}\right).$$

Then the followings hold true.

- The function g is $(e\gamma_t)^{-1}$ -Lipschitz.
- The sequence $(\mathbb{E}[g_n(Z_{n-1})])_{n \geq 1}$ is bounded by $C\gamma_t^{-1}$ for some $C > 0$. As a result,

$$(C_n - 1) |\mathbb{E}[g_n(Z_{n-1})]| \leq \frac{C}{\gamma_t} (C_n - 1) = \frac{1}{\gamma_t} O\left(\frac{\log n}{n}\right), \quad t \geq 0, \quad n \in \mathbb{N}^*.$$

Proof. - It is enough to notice that g' admits a simple expression thanks to the change of variable $y' = y + x$:

$$g(x) = e^{-x} \int_0^\infty e^{-\gamma_t e^{-(x+z)}} h'(x + z - t) e^{-z} dz = \int_x^\infty e^{-z} e^{-\gamma_t e^{-z}} h'(z - t) dz.$$

Differentiating with respect to x , we find that $g'(x) = -e^{-x} e^{-\gamma_t e^{-x}} h'(x - t)$, so that

$$|g'(x)| \leq e^{-x} e^{-\gamma_t e^{-x}} \leq \frac{e^{-1}}{\gamma_t}.$$

- Recall that g_n is given by (4.28), and that the constants α_n, β_n and δ_n have been defined just before (4.26). We know that the double exponential term is a non-decreasing function, and that $\lceil y \rceil \leq y + 1$, for every non-negative y . Besides, as $-n \log \alpha_n \geq 1$ and $\delta_n \leq -1/n$, we have:

$$-\frac{Y}{n \log \alpha_n} + \delta_n \leq \frac{G_n}{n} + \delta_n \leq -\frac{Y}{n \log \alpha_n} + \delta_n + \frac{1}{n} \leq Y. \tag{4.30}$$

Moreover, since $Z_{n-1} \geq 1 - \log(n-1) \geq -\log n$ a.s., and as x will be replaced by Z_{n-1} in the sequel, we only need to work with $x \geq -\log n$:

$$\begin{aligned} |g_n(x)| &\leq e^{-\beta_n x} \mathbb{E} \left[e^{-\gamma t e^{-(\alpha_n x + \frac{G_n}{n} + \delta_n)}} \right] \\ &\leq e^{-\beta_n x} \mathbb{E} \left[e^{-\gamma t e^{-(\alpha_n x + Y)}} \right] \\ &= \frac{1}{\gamma t} e^{-(\beta_n - \alpha_n)x} (1 - e^{-\gamma t e^{-\alpha_n x}}) \\ &\leq \frac{1}{\gamma t} e^{-(\beta_n - \alpha_n)x} \\ &\leq \frac{1}{\gamma t} n^{\beta_n - \alpha_n}. \end{aligned}$$

as for all positive γ we have $\mathbb{E}[e^{-\gamma e^{-Y}}] = (1 - e^{-\gamma})/\gamma$ if Y has exponential distribution $\mathcal{E}(1)$. This last term is bounded w.r.t. to n , because $n^{\beta_n - \alpha_n} \underset{n \rightarrow \infty}{\sim} 1 + \log n / (2n)$. Finally, notice that:

$$0 \leq C_n - 1 = n \left(1 - \frac{1}{n}\right)^{(n-1) \log(n-1)} - 1 \leq n e^{-(1 - \frac{1}{n}) \log(n-1)} - 1 = e^{\log(1 + \frac{1}{n-1}) + \frac{1}{n} \log n} - 1 \underset{n \rightarrow \infty}{\sim} \frac{\log n}{n}.$$

The intermediary result, available for x greater than $-\log n$

$$e^{-\beta_n x} \mathbb{E} \left[e^{-\gamma t e^{-(\alpha_n x + \frac{G_n}{n} + \delta_n)}} \right] \leq \frac{C}{\gamma t}, \quad (4.31)$$

will be useful on its own in the sequel. □

This lemma has two immediate conclusions: the first item furnishes a bound for (1), while the second one directly settles the case of (2) in (4.29):

$$|\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z_{n-1})]| \leq \frac{e^{-1}}{\gamma t} \|Z_n - Z_{n-1}\|_{\mathbf{L}^1} \leq \frac{C \log n}{\gamma t n}.$$

As for (3), we will need to decompose it further:

$$\begin{aligned} &\mathbb{E}[g(Z_{n-1})] - \mathbb{E}[g_n(Z_{n-1})] \\ &= \left(\mathbb{E}[g(Z_{n-1})] - \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} h'(Z_{n-1} + Y - t) \right] \right) \\ &\quad + \left(\mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} h'(Z_{n-1} + Y - t) \right] - \mathbb{E}[g_n(Z_{n-1})] \right) \\ &= (3.1) + (3.2). \end{aligned}$$

Let us start with (3.2). Recall that δ_n is non-positive:

$$\begin{aligned} &\left| \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} h'(Z_{n-1} + Y - t) \right] - \mathbb{E}[g_n(Z_{n-1})] \right| \\ &\leq \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} \left| h'(Z_{n-1} + Y - t) - h' \left(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n - t \right) \right| \right] \\ &\leq \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} \left| (1 - \alpha_n) Z_{n-1} + Y - \frac{G_n}{n} - \delta_n \right| \right] \\ &\leq \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} \left(\frac{1}{n} |Z_{n-1}| + \left| Y - \frac{G_n}{n} \right| - \delta_n \right) \right]. \end{aligned}$$

That last expression is also equal to:

$$\begin{aligned} & \frac{1}{n} \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma_t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} (|Z_{n-1}| - \delta_n) \right] + \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma_t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} - \delta_n)}} \left| Y - \frac{G_n}{n} \right| \right] \\ & \leq \frac{C}{n\gamma_t} (\mathbb{E}[|Z_{n-1}|] - \delta_n) + \left(\mathbb{E} \left[e^{-2\beta_n Z_{n-1}} e^{-2\gamma_t e^{-(\alpha_n x + \frac{G_n}{n} + \delta_n)}} \right] \right)^{\frac{1}{2}} \left\| Y - \frac{G_n}{n} \right\|_{\mathbf{L}^2} \\ & \leq \frac{C}{n\gamma_t} (\mathbb{E}[|Z_{n-1}|] - \delta_n) + \frac{C}{\gamma_t} (\mathbb{E}[e^{-\beta_n Z_{n-1}}])^{\frac{1}{2}} \left\| Y - \frac{G_n}{n} \right\|_{\mathbf{L}^2}. \end{aligned}$$

We have used successively inequality (4.31), Cauchy-Schwarz inequality and once more inequality (4.31). The sequences $(\mathbb{E}[|Z_{n-1}|])_{n \geq 1}$ and $(\mathbb{E}[e^{-Z_{n-1}}])_{n \geq 1}$ are bounded. This is easily seen for the first one, for instance by recalling that:

$$\mathbb{V}(Z_n) = \frac{\pi^2}{6} + o(1).$$

As for the second sequence, notice that since $\beta_n \leq 1$, we have:

$$\mathbb{E}[e^{-\beta_n Z_{n-1}}] \leq (\mathbb{E}[e^{-Z_{n-1}}])^{\beta_n} \leq 1,$$

thanks to the next statement:

Lemma 4.2.7 *For all $\lambda \geq 0$, the sequence $(\mathbb{E}[e^{-\lambda Z_n}])_{n \geq 1}$ is bounded. Moreover, we have the asymptotic expansion:*

$$\mathbb{E}[e^{-\lambda Z_n}] - 1 \underset{n \rightarrow \infty}{\sim} -\frac{\log n}{2n}. \quad (4.32)$$

Proof. For all $n \geq 1$, we have by definition of Z_n :

$$\mathbb{E}[e^{-\lambda Z_n}] = n^\lambda \prod_{i=1}^n \frac{1 - \frac{i-1}{n}}{e^{\frac{\lambda}{n}} - \frac{i-1}{n}} = n^\lambda n! \prod_{i=1}^n \frac{1}{ne^{\frac{\lambda}{n}} - i + 1} \leq n^\lambda n! \prod_{i=1}^n \frac{1}{n - i + 1 + \lambda} = n^\lambda n! \prod_{i=1}^n \frac{1}{i + \lambda},$$

thanks to the inequality $ne^{\lambda/n} \geq n + \lambda$. Therefore:

$$\mathbb{E}[e^{-\lambda Z_n}] = n^\lambda \prod_{i=1}^n \frac{i}{i + \lambda} \leq e^{-\lambda \left(\sum_{i=1}^n \frac{1}{i+\lambda} - \log n \right)} = e^{-\lambda \left(\sum_{i=1}^n \frac{1}{i} - \log n - \lambda \sum_{i=1}^n \frac{1}{i(i+\lambda)} \right)},$$

and using the well-known asymptotic expansion $\sum_{i=1}^n i^{-1} = \log n + \gamma + o(1)$, where $\gamma = 0.577 \dots$ is the so-called gamma constant, we conclude that the right-hand side converges for every non-negative

λ. As for the second part of the statement

$$\begin{aligned}
\mathbb{E}[e^{-Z_n}] &= nn! \prod_{i=1}^n \frac{1}{ne^{1/n} - i + 1} \\
&= \frac{nn!}{(n+1)!} \prod_{i=1}^n \frac{n-i+2}{ne^{1/n} - i + 1} \\
&= \frac{n}{n+1} \prod_{i=1}^n \frac{n-i+2}{ne^{1/n} - i + 1} \\
&= \frac{n}{n+1} \prod_{i=1}^n \frac{n-i+2}{(n(e^{1/n} - 1) - 1) + n - i + 2} \\
&= \frac{n}{n+1} \prod_{i=1}^n \left(1 + \frac{n(e^{1/n} - 1) - 1}{i+1}\right)^{-1} \\
&= \frac{n}{n+1} \exp\left(-\sum_{i=1}^n \log\left(1 + \frac{n(e^{1/n} - 1) - 1}{n-i+2}\right)\right)
\end{aligned}$$

Recall that:

$$n(e^{1/n} - 1) - 1 \underset{n \rightarrow \infty}{\sim} \frac{1}{2n}.$$

Therefore, a standard result about positive divergent series yields:

$$\mathbb{E}[e^{-Z_n}] - 1 \underset{n \rightarrow \infty}{\sim} \frac{n}{n+1} e^{-\frac{1}{2n} H_n} - 1 \underset{n \rightarrow \infty}{\sim} e^{-\frac{1}{2n} \log n} - 1 \underset{n \rightarrow \infty}{\sim} -\frac{1}{2n} \log n,$$

which concludes the proof. □

As a result, we need to compare Y and G_n/n . We already know that the latter converges weakly to Y . To get a rate of convergence, we use the fact that G_n depends of Y in a simple way:

Lemma 4.2.8 *Let $Y \sim \mathcal{E}(1)$ and G_n defined as before. Then $(G_n/n)_{n \geq 1}$ converges to Y in \mathbf{L}^2 norm, with rate of convergence at least $\sqrt{2}/n$:*

$$\left\| Y - \frac{G_n}{n} \right\|_{\mathbf{L}^2} \leq \frac{\sqrt{2}}{n}, \quad n \geq 2.$$

Proof. To determine the rate of convergence in \mathbf{L}^2 norm, we will need to compute $\mathbb{E}[Y[Y/\lambda]]$ for

some $\lambda > 0$. Set $u_n := \int_0^{n\lambda} ye^{-y} dy$

$$\begin{aligned}
\mathbb{E}\left[Y\left[\frac{Y}{\lambda}\right]\right] &= \int_0^\infty y\left[\frac{y}{\lambda}\right]e^{-y} dy \\
&= \sum_{n=0}^\infty (n+1) \int_{n\lambda}^{(n+1)\lambda} ye^{-y} dy \\
&= \sum_{n=0}^\infty (n+1)(u_n - u_{n+1}) \\
&= \sum_{n=0}^\infty u_n + (nu_n - (n+1)u_{n+1}) \\
&= \sum_{n=0}^\infty u_n \\
&= \frac{\lambda e^{-\lambda}}{(1-e^{-\lambda})^2} + \frac{1}{1-e^{-\lambda}}.
\end{aligned}$$

Evaluating this identity at $\lambda = -\log(1 - 1/n) > 0$, we get that

$$\begin{aligned}
\mathbb{E}\left[\left|Y - \frac{G_n}{n}\right|^2\right] &= 2 + 2 - \frac{1}{n} - \frac{2}{n}\mathbb{E}\left[Y\left[\frac{Y}{\lambda}\right]\right] \\
&= 4 - \frac{1}{n} + 2(n-1)\log\left(1 - \frac{1}{n}\right) - 2 \\
&= 2 - \frac{1}{n} + 2(n-1)\log\left(1 - \frac{1}{n}\right) \\
&\leq \frac{2}{n^2}.
\end{aligned}$$

The last inequality comes from the well-known inequality:

$$\log\left(1 - \frac{1}{x}\right) \leq -\frac{1}{x} - \frac{1}{2x^2}, \quad x \geq 1.$$

□

And so (3.2) is bounded by $C/(\gamma_t n)$, for some constant $C > 0$ independent of t . Now we move onto (3.1). Before bounding this term, we show briefly how we can replace $e^{-\beta_n Z_{n-1}}$ by $e^{-Z_{n-1}}$.

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}\right)e^{-\gamma_t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} h'(Z_{n-1} + Y - t)\right]\right| \\
&\leq \mathbb{E}\left[\left|e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}\right|e^{-\gamma_t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}\right] \\
&\leq \mathbb{E}\left[\left|e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}\right|e^{-\gamma_t e^{-(\alpha_n Z_{n-1} + Y)}}\right] \\
&= \mathbb{E}\left[\left|e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}\right|\varphi_n(Z_{n-1})\right],
\end{aligned}$$

where we set $\varphi_n(x) := \mathbb{E}\left[e^{-\gamma_t e^{-(\alpha_n x + Y)}}\right]$ for all $x \geq -\log n$. The change of variable $y' = y - \log \gamma_t$ shows that:

$$\varphi_n(x) = \frac{1}{\gamma_t} \int_{-\log \gamma_t}^\infty e^{-y} e^{-e^{-(\alpha_n x + y)}} dy \leq \frac{1}{\gamma_t} \int_{\mathbb{R}} e^{-y} e^{-e^{-(\alpha_n x + y)}} dy = \frac{1}{\gamma_t} e^{\alpha_n x}.$$

Thus, noticing that $z \mapsto e^{-Z_{n-1}z}$ is Lipschitz on $[\beta_n - \alpha_n, 1 - \alpha_n]$, with Lipschitz constant less than $|Z_{n-1}|(e^{-(\beta_n - \alpha_n)Z_{n-1}} + e^{-(1 - \alpha_n)Z_{n-1}})$ almost surely, we can replace β_n by 1:

$$\begin{aligned} \left| \mathbb{E} \left[(e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}) \varphi_n(Z_{n-1}) \right] \right| &\leq \frac{C}{\gamma t} \mathbb{E} \left[|e^{-(\beta_n - \alpha_n)Z_{n-1}} - e^{-(1 - \alpha_n)Z_{n-1}}| \right] \\ &\leq \frac{C}{\gamma t} (1 - \beta_n) \mathbb{E} \left[|Z_{n-1}| (e^{-(\beta_n - \alpha_n)Z_{n-1}} + e^{-(1 - \alpha_n)Z_{n-1}}) \right]. \end{aligned}$$

We have used Cauchy-Schwarz inequality and lemma 4.2.7 to bound the remaining expectation. As a result, we can exchange (3.1) for

$$\begin{aligned} \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}) h'(Z_{n-1} + Y - t) \right] \\ \leq \mathbb{E} \left[e^{-Z_{n-1}} |e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}| \right]. \end{aligned} \quad (4.33)$$

As tempting as it may be, we *cannot* exploit (yet) the fact that $x \mapsto e^{-\gamma t e^{-x}}$ is Lipschitz, because doing so would yield an expression without the term γt^{-1} , which is crucial to find an integrable bound w.r.t. t on $[\varepsilon, +\infty)$. Set $\varepsilon_n := -n \log \alpha_n \geq 1$. Then the above expression is bounded by:

$$\begin{aligned} &\mathbb{E} \left[e^{-Z_{n-1}} |e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}}| \right] \\ &\quad + \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}) \right] \\ &= \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}} - e^{-\gamma t e^{-(Z_{n-1}+Y)}}) \mathbf{1}_{\{Z_{n-1} \leq 0\}} \right] \\ &\quad + \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}}) \mathbf{1}_{\{Z_{n-1} > 0\}} \right] \\ &\quad + \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}) \right] \\ &\leq \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}} - e^{-\gamma t e^{-(Z_{n-1}+Y)}}) \mathbf{1}_{\{Z_{n-1} \leq 0\}} \right] \\ &\quad + \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}}) \mathbf{1}_{\{Z_{n-1} > 0\}} \right] \\ &\quad + \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{Y}{\varepsilon_n} + \delta_n)}}) \right]. \end{aligned}$$

We have used the inequality (4.30) at the last line, which amounts to saying that $G_n/n \geq Y/\varepsilon_n$. Unfortunately, in spite of the monotony of the double exponential and the inequality $G_n/n + \delta_n \leq Y$, we could not remove the absolute values in the first expression, because it is not true that $\alpha_n x \leq x$ if $x < 0$. This explains the need for the above inequality. With the absolute value out of the picture, the two first terms become easy to estimate as we know that for $\alpha \in \{\alpha_n, 1\}$, :

$$e^{-x} \mathbb{E} [e^{-\gamma t e^{-(\alpha x + Y)}}] = \frac{1}{\gamma t} e^{-(1-\alpha)x} (1 - e^{-\gamma t e^{-\alpha x}}), \quad x \in \mathbb{R}$$

so that for all $x \in \mathbb{R}$:

$$\begin{aligned} \mathbb{E} [e^{-x} (e^{-\gamma t e^{-(\alpha_n x + Y)}} - e^{-\gamma t e^{-(x + Y)}})] &= \frac{1}{\gamma t} ((e^{-(1-\alpha_n)x} - 1) + (e^{-\gamma t e^{-x}} - e^{-(1-\alpha_n)x} e^{-\gamma t e^{-\alpha_n x}})) \\ &= \frac{1}{\gamma t} ((e^{-(1-\alpha_n)x} - 1)(1 - e^{-\gamma t e^{-\alpha_n x}}) + (e^{-\gamma t e^{-x}} - e^{-\gamma t e^{-\alpha_n x}})) \end{aligned}$$

Now we can bound the first two terms by putting the indicator functions together again:

$$\begin{aligned} \mathbb{E}\left[e^{-Z_{n-1}}|e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}}\right] \\ \leq \frac{1}{\gamma t} \mathbb{E}\left[|e^{-(1-\alpha_n)Z_{n-1}} - 1| + |e^{-\gamma t e^{-Z_{n-1}}} - e^{-\gamma t e^{-\alpha_n Z_{n-1}}}\right]. \end{aligned}$$

Controlling these two expectations is simple: first the function $z \mapsto e^{-Z_{n-1}z}$ is Lipschitz on $[0, 1 - \alpha_n]$, with Lipschitz constant less than $|Z_{n-1}|(1 + e^{-(1-\alpha_n)Z_{n-1}})$ almost surely. Second, remark that the function $z \mapsto e^{-\gamma t e^{-z}}$ is e^{-1} -Lipschitz. Thanks to Cauchy-Schwartz inequality, we know that

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}[|Z_{n-1}| e^{-\lambda Z_{n-1}}] < +\infty, \quad \lambda \geq 0.$$

Therefore, we can write:

$$\mathbb{E}\left[e^{-Z_{n-1}}|e^{-\gamma t e^{-(Z_{n-1}+Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1}+Y)}}\right] \leq \frac{C}{\gamma t n}.$$

The third term is a source of troubles because Y is divided by $\varepsilon_n = -n \log \alpha_n \geq 1$. A change of variable gives:

$$e^{-x} \mathbb{E}\left[e^{-\gamma t e^{-(\alpha_n x + \frac{Y}{\varepsilon_n} + \delta_n)}\right] = \varepsilon_n e^{-x} \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma t e^{-(\alpha_n x + y + \delta_n)}} dy, \quad x \in \mathbb{R}.$$

We need to compare this integral to $e^{-x} \int_0^\infty e^{-y} e^{-\gamma t e^{-(\alpha_n x + y + \delta_n)}} dy$:

Lemma 4.2.9 *There exists a constant $C > 0$ such that for all $n \geq 2$, we have for all positive t and for all real numbers x :*

$$\begin{aligned} e^{-x} \left| \int_0^\infty e^{-y} e^{-\gamma t e^{-(\alpha_n x + y)}} dy - \varepsilon_n \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma t e^{-(\alpha_n x + y + \delta_n)}} dy \right| \\ \leq C \frac{|x| + |\log \gamma t| \log n}{\gamma t n} e^{-\frac{1}{n}x}. \end{aligned}$$

In particular

$$\begin{aligned} \mathbb{E}\left[e^{-Z_{n-1}} \left| \int_0^\infty e^{-y} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + y)}} dy - \varepsilon_n \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma t e^{-(\alpha_n Z_{n-1} + y + \delta_n)}} dy \right| \right] \\ \leq C \frac{1 + |\log \gamma t| \log n}{\gamma t n}. \end{aligned}$$

Proof. The arguments are similar to those given above, but certain new difficulties arise nonetheless.

- First we replace $\varepsilon_n = -n \log(1 - 1/n)$ by 1 by using that $\varepsilon_n \geq 1$. Recall that δ_n is negative, so that:

$$0 \leq (\varepsilon_n - 1) \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma t e^{-(\alpha_n x + y + \delta_n)}} dy \leq (\varepsilon_n - 1) \int_0^\infty e^{-y} e^{-\gamma t e^{-(\alpha_n x + y)}} dy \leq \frac{e^{\alpha_n x}}{\gamma t} (\varepsilon_n - 1).$$

Next we take care of the $e^{-\varepsilon_n y}$ part in the integral. As before, observe that $z \mapsto e^{-yz}$ is Lipschitz on $[1, \varepsilon_n]$, with Lipschitz constant ye^{-y} :

$$\begin{aligned} 0 \leq \int_0^\infty (e^{-y} - e^{-\varepsilon_n y}) e^{-\gamma t e^{-(\alpha_n x + y + \delta_n)}} dy \\ \leq (\varepsilon_n - 1) \int_0^\infty ye^{-y} e^{-\gamma t e^{-(\alpha_n x + y + \delta_n)}} dy = (\varepsilon_n - 1) e^{-\tau} \mathbb{E}[(Z + \tau) \mathbf{1}_{\{Z + \tau \geq 0\}}]. \end{aligned}$$

where $\tau := \log \gamma_t - \alpha_n x - \delta_n \in \mathbb{R}$ and Z is a random variable with standard Gumbel distribution. The right-hand side makes appear $|\log \gamma_t|$ in the bound:

$$\begin{aligned} 0 &\leq \int_0^\infty (e^{-y} - e^{-\varepsilon_n y}) e^{-\gamma_t e^{-(\alpha_n x + y + \delta_n)}} dy \\ &\leq e^{-\tau} (C + |\log \gamma_t| - \alpha_n |x|) (\varepsilon_n - 1) = \frac{e^{\alpha_n x + \delta_n}}{\gamma_t} (C + |\log \gamma_t| + \alpha_n |x|) (\varepsilon_n - 1) \end{aligned}$$

- As for the final term, recall once more that $\delta_n = \log(1 - 1/n) - \log(n - 1)/n \leq 0$:

$$\begin{aligned} 0 &\leq \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y)}} dy - \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y + \delta_n)}} dy \\ &= \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y)}} dy - e^{\varepsilon_n \delta_n} \int_{\delta_n}^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y)}} dy \\ &\leq \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y)}} dy - e^{\varepsilon_n \delta_n} \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y)}} dy \\ &= (1 - e^{\varepsilon_n \delta_n}) \int_0^\infty e^{-\varepsilon_n y} e^{-\gamma_t e^{-(\alpha_n x + y)}} dy \\ &\leq \frac{e^{\alpha_n x}}{\gamma_t} (1 - e^{\varepsilon_n \delta_n}). \end{aligned}$$

The second inequality of the statement is obtained by replacing x by Z_{n-1} and integrating w.r.t. this random variable. Recall that the sequence $(\mathbb{E}[|Z_{n-1}| e^{-\lambda Z_{n-1}}])_{n \geq 1}$ is bounded for all non-negative λ .

□

4.2.2 Bounding the error on $[0, \varepsilon]$

The next proposition bounds $\mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]$ by an integrable term w.r.t. t in the neighbourhood of 0:

Proposition 4.2.10 *Let $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$. Then there exists a constant $C > 0$ such that:*

$$|\mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]| \leq C \frac{\log n}{n}, \quad t \geq 0, \quad n \in \mathbb{N}^*.$$

Proof. We have already found a bound depending on γ_t^{-1} , so we have to obtain another bound, constant w.r.t. t . This boils down to bounding the same four terms as before, but independently of t , which makes the task much faster and easier. First, we have

$$\begin{aligned} (1 - C_n) |\mathbb{E}[g_n(Z_{n-1})]| &\leq \mathbb{E} \left[e^{-\beta_n Z_{n-1}} e^{-\gamma_t e^{-(\alpha_n Z_{n-1} + \frac{C_n}{n} + \delta_n)}} \right] \\ &\leq (1 - C_n) \mathbb{E}[e^{-\beta_n Z_{n-1}}] \\ &\lesssim (1 - C_n), \end{aligned}$$

and we already know that this last term goes to 0 as fast as $\log n/n$. The second term is dealt with in the next lemma:

Lemma 4.2.11 Let $g : x \mapsto e^{-x} \mathbb{E}[e^{-\gamma t} h'(x + Y - t)]$, where $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$ and Y is a random variable with exponential distribution $\mathcal{E}(1)$. Then there exists $C > 0$ such that:

$$|\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z_{n-1})]| \leq C \frac{\log n}{n}, \quad t \geq 0, n \geq 2.$$

Proof. We have that:

$$\begin{aligned} & |\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z_{n-1})]| \\ & \leq \mathbb{E} \left[e^{-Z_n} |e^{-\gamma t e^{-(Z_n+Y)}} h'(Z_n + Y - t) - e^{-\gamma t e^{-(Z_{n-1}+Y)}} h'(Z_{n-1} + Y - t)| \right] \\ & \quad + \mathbb{E} [|e^{-Z_n} - e^{-Z_{n-1}}| e^{-\gamma t e^{-(Z_{n-1}+Y)}} |h'(Z_{n-1} + Y - t)|] \\ & \leq \mathbb{E}[e^{-Z_n} |Z_n - Z_{n-1}|] + \mathbb{E}[|e^{-Z_n} - e^{-Z_{n-1}}|]. \end{aligned}$$

To exploit the fact that $\tau_{i+1}^n \geq \tau_i^{n-1}$, recall that $\tau_1^n = 1$ a.s. and introduce:

$$Z'_n := \frac{1}{n-1} \sum_{i=1}^n \tau_i^n - \log(n-1) = \frac{1}{n-1} \sum_{i=1}^{n-1} \tau_{i+1}^n + \frac{1}{n-1} - \log(n-1), \quad n \geq 2.$$

Clearly $Z'_n \geq Z_{n-1}$ and $Z'_n \geq Z_n$ since

$$Z_n - Z'_n = -\frac{1}{n(n-1)} \sum_{i=1}^n \tau_i^n + \log\left(1 - \frac{1}{n}\right) \leq 0.$$

Therefore:

$$\begin{aligned} |\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z_{n-1})]| & \leq \mathbb{E}[e^{-Z_n} |Z_n - Z_{n-1}|] + \mathbb{E}[|e^{-Z_n} - e^{-Z_{n-1}}|] \\ & \leq \mathbb{E}[e^{-Z_n} (Z'_n - Z_n)] + \mathbb{E}[e^{-Z_n} (Z'_n - Z_{n-1})] \\ & \quad + \mathbb{E}[e^{-Z_n} - e^{-Z'_n}] + \mathbb{E}[e^{-Z_{n-1}} - e^{-Z'_n}] \\ & = (1) + (2) + (3) + (4). \end{aligned}$$

We start with (1). Introduce the notation:

$$Z_{n \setminus i} := \frac{1}{n} \sum_{\substack{j=1 \\ j \neq i}}^n \tau_j^n - \log n, \quad i \in \llbracket 1, n \rrbracket.$$

Now conditioning on Y_i , the i -th exponential random variable in the definitions of τ_{i+1}^n and τ_i^{n-1} , we can write:

$$\begin{aligned} \mathbb{E}[e^{-Z_n} (Z'_n - Z_n)] & = \frac{1}{n(n-1)} \sum_{i=1}^n \mathbb{E}[e^{-Z_n} \tau_i^n] - \log\left(1 - \frac{1}{n}\right) \mathbb{E}[e^{-Z_n}] \\ & \leq \frac{1}{n(n-1)} \sum_{i=1}^n \mathbb{E}[e^{-Z_n} \tau_i^n] + \frac{1}{n-1} \\ & = \frac{1}{n(n-1)} \sum_{i=1}^n \mathbb{E}[e^{-Z_{n \setminus i}}] \mathbb{E}[e^{-\frac{1}{n} \tau_i^n} \tau_i^n] + \frac{1}{n-1} \\ & \leq \frac{1}{n(n-1)} \sum_{i=1}^n \mathbb{E}[e^{-Z_{n \setminus i}}] \mathbb{E}[\tau_i^n] + \frac{1}{n-1} \\ & = \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[e^{-Z_{n \setminus i}}] \frac{1}{n-i+1} + \frac{1}{n-1}. \end{aligned}$$

That last expectation is bounded with respect to n :

$$\begin{aligned}
\mathbb{E}[e^{-Z_{n \setminus i}}] &= n \prod_{\substack{j=1 \\ j \neq i}}^n \frac{n-j+1}{ne^{1/n} - j + 1} \\
&= \frac{nn!}{n-i+1} \prod_{\substack{j=1 \\ j \neq i}}^n \frac{1}{ne^{1/n} - j + 1} \\
&\leq \frac{nn!}{n-i+1} \prod_{\substack{j=1 \\ j \neq i}}^n \frac{1}{n-j+2} \\
&= \frac{nn!}{n-i+1} \times \frac{n-i+2}{(n+1)!} \\
&= \frac{n-i+2}{n-i+1} \times \frac{n}{n+1} \\
&\leq 2.
\end{aligned}$$

Consequently we have

$$\begin{aligned}
\mathbb{E}[e^{-Z_n}(Z'_n - Z_n)] &\leq \frac{1}{n-1} + \frac{1}{n(n-1)} \sum_{i=1}^n \mathbb{E}[e^{-Z_{n \setminus i}}] \mathbb{E}[e^{-\frac{1}{n} \tau_i^n} \tau_i^n] \\
&\leq \frac{1}{n-1} + \frac{2}{n-1} H_n.
\end{aligned}$$

Those arguments apply again to (2) and give (almost) the same bound:

$$\begin{aligned}
\mathbb{E}[e^{-Z_n}(Z'_n - Z_{n-1})] &= \frac{1}{n-1} + \frac{1}{n-1} \sum_{i=1}^{n-1} \mathbb{E}[e^{-Z_n}(\tau_{i+1}^n - \tau_i^{n-1})] \\
&= \frac{1}{n-1} + \frac{1}{n-1} \sum_{i=1}^{n-1} \mathbb{E}[e^{-Z_{n \setminus i}}] \mathbb{E}[e^{-\tau_{i+1}^n}(\tau_{i+1}^n - \tau_i^{n-1})] \\
&\leq \frac{1}{n-1} + \frac{2}{n-1} \sum_{i=1}^{n-1} \frac{1}{n-i} \\
&= \frac{1}{n-1} + \frac{2}{n-1} H_{n-1}.
\end{aligned}$$

Finally we deal with (3) and (4). This last part essentially boils down to determining how fast $\mathbb{E}[e^{-Z_n}]$ goes to 1, a question whose answer has already been given in lemma 4.32, where we found that

$$\mathbb{E}[e^{-Z_n}] - 1 \underset{n \rightarrow \infty}{\sim} -\frac{\log n}{2n},$$

which concludes the proof. □

As for the third term, we rely once more on the fact that $z \mapsto e^{-Z_{n-1}z}$ is $|Z_{n-1}|(e^{-\beta_n Z_{n-1}} + e^{-Z_{n-1}})$ -

Lipschitz almost surely, when $z \in [\beta_n, 1]$:

$$\begin{aligned} & \left| \mathbb{E} \left[(e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}) e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}} h'(Z_{n-1} + Y - t) \right] \right| \\ & \leq \mathbb{E} [|e^{-\beta_n Z_{n-1}} - e^{-Z_{n-1}}|] \\ & \leq (1 - \beta_n) \mathbb{E} [|Z_{n-1}| (e^{-\beta_n Z_{n-1}} + e^{-Z_{n-1}})] \\ & \leq C(1 - \beta_n). \end{aligned}$$

Now, as before, we use that $x \mapsto e^{-\gamma t e^{-x}}$ is e^{-1} -Lipschitz on \mathbb{R} , so that the fourth and final term satisfies:

$$\begin{aligned} & \left| \mathbb{E} \left[e^{-Z_{n-1}} (e^{-\gamma t e^{-(Z_{n-1} + Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}) h'(Z_{n-1} + Y - t) \right] \right| \\ & \leq \mathbb{E} \left[e^{-Z_{n-1}} |e^{-\gamma t e^{-(Z_{n-1} + Y)}} - e^{-\gamma t e^{-(\alpha_n Z_{n-1} + \frac{G_n}{n} + \delta_n)}}| \right] \\ & \leq e^{-1} \mathbb{E} \left[e^{-Z_{n-1}} ((1 - \alpha_n) |Z_{n-1}| + (Y - \frac{G_n}{n} - \delta_n)) \right]. \end{aligned}$$

□

Putting proposition 4.2.5 and 4.2.10 together, we finally get:

Proposition 4.2.12 *Let $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$. Then there exists a constant C such that:*

$$|\mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]| \leq C \min \left(1, \frac{1 + |\log \gamma t|}{\gamma t} \right) \frac{\log n}{n}, \quad t \geq 0, \quad n \geq 2.$$

By combining all the previous results, we find finally that:

Proposition 4.2.13 *Let Z be a random variable with Gumbel distribution. Then there exists a constant $C > 0$ such that*

$$d_{[2]}(Z_n, Z) \leq C \frac{\log n}{n}, \quad n \geq 1.$$

Proof. We already know that for every $h \in \text{Lip}_{[2]}(\mathbb{R}, \mathbb{R})$

$$\mathbb{E}[h(Z_n)] - \mathbb{E}[h(Z)] = - \int_0^\infty \mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)] dt.$$

Using proposition 4.2.12 and taking an arbitrary $\varepsilon > 0$, we obtain:

$$\begin{aligned} |\mathbb{E}[h(Z_n)] - \mathbb{E}[h(X)]| & \leq \left| \int_0^\infty \mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)] dt \right| \\ & \leq \int_0^\varepsilon |\mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]| dt + \int_\varepsilon^\infty |\mathbb{E}[\mathcal{L}_0 \mathbf{P}_t^0 h(Z_n)]| dt \\ & \leq C \frac{\log n}{n} \varepsilon + C \frac{\log n}{n} \int_\varepsilon^\infty \frac{1 + |\log \gamma t|}{\gamma t} dt \\ & = C \frac{\log n}{n} \left(\varepsilon + \int_\varepsilon^\infty \frac{1 + |\log \gamma t|}{\gamma t} dt \right). \end{aligned}$$

Since the rate of convergence is the same for $t \in [0, \varepsilon]$ and for $t \in [\varepsilon, +\infty)$, we do not need to optimize in ε , as it would only affect the constant, not the speed of convergence. Let us just mention that the function $\varepsilon \mapsto 1 - |1 + \log \gamma_\varepsilon|/\gamma_\varepsilon$ admits one unique zero around $\varepsilon = 0.7$, and it is a global minimum. That concludes the proof. \square

Remark 4.2.14. To the best of our knowledge, only one paper so far has given similar uniform bounds for the coupon collector problem: [Sándor \[1993\]](#). The author finds the same rate as ours, but in the Kolmogorov distance and through a different method. Sándor's PhD student has developed this topic in her doctoral thesis [Pósfai \[2010\]](#). She has made use of Stein-Chen method, among other techniques, to find rates of convergence in several limit theorems related to the coupon collector problem. \blacksquare

5 Connections with Poisson stochastic analysis

Contents

5.1	Stochastic analysis for max-id distributions	171
5.1.1	Covariance identities	171
5.1.2	Poincaré inequalities	175
5.2	Another quantization of the Poisson process	178

The last chapter aims at analysing more thoroughly the connections between max-id random variables and Poisson processes. The fact that a max-id random variable can be seen as the maximum of a certain Poisson process allows us to exploit well-known elements of stochastic analysis for those point process to prove new results for max-id distributions. This is the subject of the first section. We will also see that several limit theorems involving max-stable distributions can actually be considered as special cases of limit theorems for point processes. Although this idea is not new, we introduce a specific tool in the second section to quantify it through Stein's method.

5.1 Stochastic analysis for max-id distributions

In this section, we use adapt several results known to be true for the Poisson process to max-id distributions. The proofs are usually very short since the main argument is always the same: any max-id distribution is equal in distribution to the componentwise maximum of a certain Poisson process. For ease of notations, when dealing with max-stable distributions, we will restrict ourselves to the case $\alpha = \alpha \mathbf{1}$, with $\alpha > 0$. All statements can be readily adapted to general max-stable distributions, although with more involved notations.

5.1.1 Covariance identities

There exist several alternative expressions for the covariance of two functionals of a Poisson process. To convert them into covariance identities for max-id distributions, we will use the following equality, which has already played an important role in the definition of $(\mathbf{P}_t^{\alpha, \nu})_{t \geq 0}$:

$$\mathbf{m}(\phi + \delta_{\mathbf{x}}) = \mathbf{m}(\phi) \oplus \mathbf{x}. \quad (5.1)$$

available for all $\mathbf{x} \in E_{\ell}$ and configuration $\phi \in \mathfrak{N}_{E_{\ell}}$. Denote by $\Delta_{\mathbf{y}} f(\mathbf{x}) := f(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x})$. If ϕ is a configuration on E_{ℓ} and $\mathbf{x} = \mathbf{m}(\phi)$, we get from (5.1) that:

$$\Delta_{\mathbf{y}} f(\mathbf{x}) = D_{\mathbf{y}}^+ \bar{f}(\phi),$$

where $\bar{f} = f \circ \mathbf{m}$. More generally, denote by $\Delta_{\mathbf{y}_1, \dots, \mathbf{y}_n}$ the composition $\Delta_{\mathbf{y}_1} \circ \dots \circ \Delta_{\mathbf{y}_n}$, so that

$$\Delta_{\mathbf{y}_1, \dots, \mathbf{y}_n} f(\mathbf{x}) = D_{\mathbf{y}_1, \dots, \mathbf{y}_n}^+ \bar{f}(\phi), \quad \mathbf{y}_1, \dots, \mathbf{y}_n \in E_{\ell}. \quad (5.2)$$

We sum up this line of reasoning in the following diagram:

$$\begin{array}{ccc} \mathfrak{N}_{E_\ell} & \xrightarrow{\mathfrak{m}} & E_\ell \\ & \searrow \bar{f} & \downarrow f \\ & & \mathbb{R} \end{array}$$

We can now use the Fock space representation (1.5) to derive covariance identities for max-id random vectors.

Proposition 5.1.1 *Let \mathbf{Z} be a max-id random vector with exponent measure μ on E_ℓ . Let $f, g \in \mathbf{L}^2(E_\ell, \mathbb{P}_{\mathbf{Z}})$. Set:*

$$T_n f(\mathbf{x}_1, \dots, \mathbf{x}_n) := \mathbb{E}[\Delta_{\mathbf{x}_1, \dots, \mathbf{x}_n} f(\mathbf{Z})]$$

and for all $u, v \in \mathbf{L}^2(E_\ell^n, \mu^{\otimes n})$:

$$\langle u, v \rangle_n := \int_{E_\ell^n} u(\mathbf{x}_1, \dots, \mathbf{x}_n) v(\mathbf{x}_1, \dots, \mathbf{x}_n) d^n \mu(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

We have the following identity:

$$\text{Cov}(f(\mathbf{Z}), g(\mathbf{Z})) = \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n f, T_n g \rangle_n. \quad (5.3)$$

Proof. We apply identity (1.5) to the Poisson process η with σ -finite intensity measure μ and to the functionals $\bar{f} = f \circ \mathfrak{m}$ and $\bar{g} = g \circ \mathfrak{m}$. Clearly \bar{f} and \bar{g} belong to $\mathbf{L}^2(\mathfrak{N}_{E_\ell}, \mathbb{P}_\eta)$. Finally we identify the terms inside the series in (1.5) by using identity (5.2) and the fact that

$$\mathbf{Z} \stackrel{d}{=} \mathfrak{m}(\eta) = \bigoplus_{\mathbf{x} \in \eta} \mathbf{x}.$$

□

In dimension 1, it is possible to greatly simplify this identity thanks to the next lemma:

Lemma 5.1.2 *Let $f : \bar{\mathbb{R}} \rightarrow \mathbb{R}$ and $x, r_1, \dots, r_n \in \mathbb{R}$ for some $n \geq 1$. Set $r_{(n)} := \min(r_1, \dots, r_n)$. We have*

$$\Delta_{r_1, \dots, r_n} f(x) = (-1)^{n-1} \Delta_{x \odot r_{(n)}} f(x) = \begin{cases} (-1)^{n-1} \Delta_{r_{(n)}} f(x) & \text{if } x \leq r_{(n)} \\ 0 & \text{otherwise.} \end{cases} \quad (5.4)$$

Proof. We prove this result by induction on n . The case $n = 1$ is trivial. Assume the proposition

holds true for some n . Then we have:

$$\begin{aligned} \Delta_{r_1, \dots, r_n, r_{n+1}} f(x) &= \Delta_{r_{n+1}} \Delta_{r_1, \dots, r_n} f(x) \\ &= (-1)^{n-1} \Delta_{r_{n+1}} \Delta_{x \odot r(1)} f(x) \\ &= (-1)^{n-1} \left[f(x \oplus (x \odot r_{(n)}) \oplus r_{n+1}) - f(x \oplus (x \odot r_{(n)})) - f(x \oplus r_{n+1}) + f(x) \right] \\ &= (-1)^{n-1} \left[f(x \oplus r_{(n)} \oplus r_{n+1}) - f(x \oplus r_{(n)}) - f(x \oplus r_{n+1}) + f(x) \right] \\ &= (-1)^n \Delta_{x \odot r_{(n+1)}} f(x). \end{aligned}$$

The last identity is checked by distinguishing cases depending on the rank of r_{n+1} w.r.t. to x and $r_{(n)}$. □

Remark 5.1.3. Unfortunately this identity does not seem to generalize well in higher dimensions. Troubles appear already in dimension 2 when considering examples like $\Delta_{\mathbf{y}_1, \mathbf{y}_2} f(\mathbf{x})$ with $\mathbf{y}_1 = (r, 0)$, $\mathbf{y}_2 = (0, r)$ and $\mathbf{x} = (1, 1)$ for all $r > 1$. Iterated applications of $\Delta_{(r,0)}$ and $\Delta_{(0,r)}$ for different values of r will not simplify as in dimension 1. This is because the vectors $(r, 0)$ and $(0, r)$ are "intertwined". Of course, if one has

$$\mathbf{y}_1 \geq \mathbf{y}_2 \cdots \geq \mathbf{y}_n$$

for instance, in the sense of the order defined in (1.17) *i.e.* each coordinate of \mathbf{y}_i is more than the corresponding coordinate of \mathbf{y}_{i+1} , then equality (5.4) still holds although it is too specific to be of much use:

$$\Delta_{\mathbf{y}_1, \dots, \mathbf{y}_n} f(\mathbf{x}) = (-1)^{n-1} \Delta_{\mathbf{x} \odot \mathbf{y}_{(n)}} f(\mathbf{x}).$$

■

Relation (5.4) allows us to give a more convenient expression of $T_n f$.

Proposition 5.1.4 *Let Z be a max-id random variable on (ℓ, ∞) . Let $f, g \in \mathbf{L}^2((\ell, \infty], \mathbb{P}_Z)$. Then*

$$\text{Cov}(f(Z), g(Z)) = \int_{\ell}^{\infty} \mathbb{E}[\Delta_r f(Z)] \mathbb{E}[\Delta_r g(Z)] \frac{d\mu(r)}{F_Z(r)}. \tag{5.5}$$

Proof. Thanks to (5.4), we have

$$\begin{aligned} T_n f(r_1, \dots, r_n) T_n g(r_1, \dots, r_n) &= \mathbb{E}[\Delta_{r_1, \dots, r_n} f(Z)] \mathbb{E}[\Delta_{r_1, \dots, r_n} g(Z)] \\ &= \mathbb{E}[\Delta_{r_{(n)}} f(Z)] \mathbb{E}[\Delta_{r_{(n)}} g(Z)]. \end{aligned}$$

Summing over the n possibilities for the value of $r_{(n)}$ and integrating over E_{ℓ}^n w.r.t. the exponent

measure $d\mu(r) = \log F_Z(dr)$, we find:

$$\begin{aligned}
\langle T_n f, T_n g \rangle_n &= \sum_{i=1}^n \int_{E_\ell^n} \mathbb{E}[\Delta_{r_i} f(Z)] \mathbb{E}[\Delta_{r_i} g(Z)] \mathbb{1}_{\{r^{(n)}=r_i\}} d\mu(r_1) \dots d\mu(r_n) \\
&= n \int_{E_\ell^n} \mathbb{E}[\Delta_{r_n} f(Z)] \mathbb{E}[\Delta_{r_n} g(Z)] \mathbb{1}_{\{r^{(n)}=r_n\}} d\mu(r_1) \dots d\mu(r_n) \\
&= n \int_{E_\ell} \mathbb{E}[\Delta_{r_n} f(Z)] \mathbb{E}[\Delta_{r_n} g(Z)] \int_{r_n}^{\infty} \dots \int_{r_n}^{\infty} 1 d\mu(r_1) \dots d\mu(r_{n-1}) d\mu(r_n) \\
&= n \int_{E_\ell} \mathbb{E}[\Delta_{r_n} f(Z)] \mathbb{E}[\Delta_{r_n} g(Z)] (-\log F(r_n))^{n-1} d\mu(r_n) \\
&= n \int_{E_\ell} \mathbb{E}[\Delta_r f(Z)] \mathbb{E}[\Delta_r g(Z)] (-\log F(r))^{n-1} d\mu(r),
\end{aligned}$$

since $\mu[x, \infty] = -\log F(x)$ due to equality (1.27). Finally, dividing the previous result by $n!$ and summing from $n = 1$ to ∞ , we obtain:

$$\begin{aligned}
\text{Cov}(f(Z), g(Z)) &= \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n f, T_n g \rangle_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{E_\ell} \mathbb{E}[\Delta_r f(Z)] \mathbb{E}[\Delta_r g(Z)] (-\log F(r))^n d\mu(r) \\
&= \int_{E_\ell} \mathbb{E}[\Delta_r f(Z)] \mathbb{E}[\Delta_r g(Z)] e^{-\log F(r)} d\mu(r),
\end{aligned}$$

hence the desired result. \square

Example 6. We give an example of application of the previous covariance identity. Let $Z \sim \mathcal{F}(\alpha)$ with $\alpha > 2$. Then Z admits a variance and we have by a direct computation that

$$\mathbb{V}(Z) = \Gamma\left(1 - \frac{2}{\alpha}\right) - \Gamma\left(1 - \frac{1}{\alpha}\right)^2.$$

On the other hand, identity (5.5) yields another expression for $\mathbb{V}(Z)$ from which we deduce:

$$\int_0^\infty e^u \left[u^{-\frac{1}{\alpha}} e^{-u} - \Gamma\left(1 - \frac{1}{\alpha}, u\right) \right]^2 du = \Gamma\left(1 - \frac{2}{\alpha}\right) - \Gamma\left(1 - \frac{1}{\alpha}\right)^2,$$

where $\Gamma(x, u)$ denotes the lower-incomplete Gamma function: $\Gamma(x, u) = \int_u^\infty t^{x-1} e^{-t} dt$.

The relation (5.3) is not the only covariance identity one can prove for max-stable distributions. Several other such identities for the Poisson process can be found in Last and Penrose [2017] and Decreusefond [2022]. However none of them translate well to the max-id setting, in the sense that the underlying Poisson process does not vanish in the final identity. We will detail this point in the next section. Nonetheless, there is at least one corollary of those covariance identities which admits a max-id counterpart. We say a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is *non-decreasing* if

$$\mathbf{x} \leq \mathbf{y} \implies f(\mathbf{x}) \leq f(\mathbf{y}),$$

where $\mathbf{x} \leq \mathbf{y}$ means that $x^j \leq y^j$ for all $j \in \llbracket 1, d \rrbracket$.

Proposition 5.1.5 *Let \mathbf{Z} be a max-id random vector with exponent measure μ on E_ℓ and A some subset of E_ℓ . Assume that f, g belong to $\mathbf{L}^2(E_\ell, \mathbb{P}_{\mathbf{Z}})$ and are non-decreasing on A and non-increasing on $E_\ell \setminus A$. Then:*

$$\mathbb{E}[f(\mathbf{Z})g(\mathbf{Z})] \geq \mathbb{E}[f(\mathbf{Z})]\mathbb{E}[g(\mathbf{Z})]. \tag{5.6}$$

Proof. This is a direct consequence of the Harris-FKG inequality (Last and Penrose [2017] page 217) which states that if \bar{f}, \bar{g} belong to $\mathbf{L}^2(\mathfrak{N}_{E_\ell}, \mathbb{P}_\eta)$ and are non-decreasing on some set $B \subseteq \mathfrak{N}_{E_\ell}$ and non-increasing on its complementary, then

$$\mathbb{E}[\bar{f}(\eta)\bar{g}(\eta)] \geq \mathbb{E}[\bar{f}(\eta)]\mathbb{E}[\bar{g}(\eta)].$$

Here, being nondecreasing means that $\bar{f}(\phi + \delta_x) \geq \bar{f}(\phi)$ for all $x \in E_\ell$ and $\phi \in \mathfrak{N}_{E_\ell}$. The application \mathbf{m} is clearly non-decreasing, so we can apply this result to $\bar{f} = f \circ \mathbf{m}$ and $\bar{g} = g \circ \mathbf{m}$. \square

An immediate consequence of this inequality is that max-id random vectors are *associated*: if \mathbf{Z} is a max-id random vector, then for any non-increasing $f, g : E_\ell \rightarrow \mathbb{R}$, we have

$$\text{Cov}(f(\mathbf{Z}), g(\mathbf{Z})) \geq 0.$$

An alternative proof of this result can be found in Resnick [1987] (page 299). It consists in proving first that Poisson processes are associated and then applying \mathbf{m} to transfer this result to max-id random vectors, thus following a line of reasoning similar to the one we use in this section. A more direct approach which does not rely on Poisson processes is given in Marshall and Olkin [1983].

Remark 5.1.6. Observe that this approach allows us to prove that positive α -stable random vectors are also associated, since they are equal in distribution to a non-decreasing transformation of a Poisson process (the sum of its points). See Samorodnitsky and Taqqu [1994] page 204 for a more thorough statement with a completely different proof. More generally, every random variable equal in distribution to a non-decreasing transformation of a Poisson process is associated. For more on the notion of associativity and its applications, see the original paper introducing it: Esary et al. [1967]. \blacksquare

5.1.2 Poincaré inequalities

We have seen in the second chapter that the univariate Fréchet, Weibull and Gumbel semi-groups all satisfy a Poincaré inequality. We now generalize this result in higher dimensions by using different arguments.

Proposition 5.1.7 *Let \mathbf{Z} be a max-id random vector with exponent measure μ supported by $E_\ell = [\ell, +\infty] \setminus \{\ell\}$ for some $\ell \in [-\infty, +\infty)$, and $f \in \mathbf{L}^2(E_\ell, \mathbb{P}_{\mathbf{Z}})$. We have:*

$$\mathbb{V}(f(\mathbf{Z})) \leq \int_{E_\ell} \mathbb{E}[(f(\mathbf{Z} \oplus \mathbf{x}) - f(\mathbf{Z}))^2] d\mu(\mathbf{x}). \tag{5.7}$$

Proof. Recall equation (1.26):

$$\mathbf{Z} \stackrel{d}{=} \bigoplus_{k=1}^{\infty} \mathbf{Z}_k = \mathbf{m}(\eta),$$

where $\eta = \sum_k \delta_{\mathbf{Z}_k}$ is a Poisson process on E_ℓ with intensity measure $d\mu$ on E_ℓ , for some $\ell \in [-\infty, +\infty)^d$. As a result, the Poincaré's inequality for the Poisson process (1.6) applied to the functional $\bar{f} := f \circ \mathbf{m} \in \mathbf{L}^2(\mathfrak{N}_{E_\ell}, \mathbb{P}_\eta)$ yields:

$$\begin{aligned} \mathbb{V}(f(\mathbf{Z})) &= \mathbb{V}(\bar{f}(\eta)) \leq \int_{E_\ell} \mathbb{E}[(\bar{f}(\eta + \delta_{\mathbf{x}}) - \bar{f}(\eta))^2] d\mu(\mathbf{x}) \\ &= \int_{E_\ell} \mathbb{E}[(f(\mathbf{Z} \oplus \mathbf{x}) - f(\mathbf{Z}))^2] d\mu(\mathbf{x}), \end{aligned}$$

thanks to identity (5.1). Alternatively, one could have proved this result by using the covariance identity (5.3), just like the original Poincaré inequality is demonstrated in Last and Penrose [2017] (page 193). \square

Remark 5.1.8. The method we have used made the proof of this result is extremely short. It can be applied to different situation and only requires a counterpart of (1.26) to work. A general framework where such a decomposition exists is the context of stable distributions on convex cones (Davydov et al. [2008]) where it is known as a LePage series. For example, one can prove that α -stable distributions admit such a decomposition (see Samorodnitsky and Taqqu [1994]). When α belongs to $(0, 1)$, an instance of this result consists in taking a Poisson process η on \mathbb{R}_+^* with intensity measure $\alpha r^{-(\alpha+1)}$ and setting

$$\mathfrak{s}(\eta) = \sum_{x \in \eta} x = \int_{\mathbb{R}_+^*} x d\eta(x).$$

Such a random series does converge and its distribution is α -stable. The rest of the proof would look formally the same and would yield a Poincaré inequality for stable distributions. The latter is well-known and has been proved through other means in the more general context of infinitely divisible distributions: with stochastic calculus for Lévy processes in Chen [1985], a covariance representation for infinitely divisible random vectors in Houdré et al. [1998], and semi-group arguments for self-decomposable random vectors in Arras and Houdré [2019]. \blacksquare

At first sight, inequality (5.7) seems different from the Poincaré's inequality (2.22) we have proved in dimension 1 because of its right-hand side. Actually the latter is equal to $\mathcal{E}_{\Phi_{\alpha,\nu}}$, the energy form associated to $(\mathbf{P}_t^{\alpha,\nu})_{t \geq 0}$, as we shall see now.

Lemma 5.1.9 *The carré du champ operator associated to $(\mathbf{P}_t^{\alpha,\nu})_{t \geq 0}$ is denoted by*

$$\Gamma_{\Phi_{\alpha,\nu}}(f, g) = \frac{1}{2} (\mathcal{L}_{\alpha,\nu}(fg) - f\mathcal{L}_{\alpha,\nu}g - g\mathcal{L}_{\alpha,\nu}f), \quad f, g \in \mathcal{S}_{\alpha,d}.$$

We have for $\mathbf{x} \in E_0$:

$$\Gamma_{\Phi_{\alpha,\nu}}(f, g)(\mathbf{x}) = \frac{1}{2} \int_{E_0} (f(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x})) (g(\mathbf{x} \oplus \mathbf{y}) - g(\mathbf{x})) d\mu(\mathbf{y}).$$

Consequently, if $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu)$, we have:

$$\mathcal{E}_{\Phi_{\alpha,\nu}}(f) = \frac{1}{2} \int_{E_0} \mathbb{E}[(f(\mathbf{Z} \oplus \mathbf{x}) - f(\mathbf{Z}))^2] d\mu(\mathbf{x}).$$

Proof. The purpose of the *carré du champ* operator is to measure how far the generator $\mathcal{L}_{\alpha,\nu}$ is from being a derivation, *i.e.* from satisfying the Leibniz rule $(fg)' = fg' + gf'$. We know that

$$\mathcal{L}_{\alpha,\nu} = \mathfrak{d}_{\alpha,d} + \mathbf{D}_{\alpha,\nu},$$

where $\mathfrak{d}_{\alpha,d}f(\mathbf{x}) = -\alpha^{-1}\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle$ is the generator of the d -dimensional dilation semi-group $(\mathfrak{p}_t^{\alpha,d})_{t \geq 0}$ defined by

$$\mathfrak{p}_t^{\alpha,d}f(\mathbf{x}) = f(e^{-\frac{t}{\alpha}}\mathbf{x}).$$

One easily checks that $\mathfrak{d}_{\alpha,d}$ is a derivation and thus does not contribute to the *carré du champ* operator:

$$\begin{aligned} 2\Gamma_{\Phi_{\alpha,\nu}}(f, g)(\mathbf{x}) &= (\mathcal{L}_{\alpha,\nu}(fg) - f\mathcal{L}_{\alpha,\nu}g - g\mathcal{L}_{\alpha,\nu}f)(\mathbf{x}) \\ &= ((\mathfrak{d}_{\alpha,d} + \mathbf{D}_{\alpha,\nu})(fg) - f(\mathfrak{d}_{\alpha,d} + \mathbf{D}_{\alpha,\nu})g - g(\mathfrak{d}_{\alpha,d} + \mathbf{D}_{\alpha,\nu})f)(\mathbf{x}) \\ &= (\mathbf{D}_{\alpha,\nu}(fg) - f\mathbf{D}_{\alpha,\nu}g - g\mathbf{D}_{\alpha,\nu}f)(\mathbf{x}), \end{aligned}$$

which yields the result, as:

$$\begin{aligned} (\mathbf{D}_{\alpha,\nu}(fg) - f\mathbf{D}_{\alpha,\nu}g - g\mathbf{D}_{\alpha,\nu}f)(\mathbf{x}) &= \int_{E_0} (f(\mathbf{x} \oplus \mathbf{y})g(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x})g(\mathbf{x})) \, d\mu(\mathbf{y}) \\ &\quad - f(\mathbf{x}) \int_{E_0} (g(\mathbf{x} \oplus \mathbf{y}) - g(\mathbf{x})) \, d\mu(\mathbf{y}) \\ &\quad - g(\mathbf{x}) \int_{E_0} (f(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x})) \, d\mu(\mathbf{y}) \\ &= \int_{E_0} (f(\mathbf{x} \oplus \mathbf{y}) - f(\mathbf{x}))(g(\mathbf{x} \oplus \mathbf{y}) - g(\mathbf{x})) \, d\mu(\mathbf{y}). \end{aligned}$$

The second result stems from the fact $\mathcal{E}_{\Phi_{\alpha,\nu}}(f) = \mathbb{E}[\Gamma_{\Phi_{\alpha,\nu}}(f, f)(\mathbf{Z})]$. \square

We deduce from the previous propositions that $(\mathbf{P}_t^{\alpha,\nu})_{t \geq 0}$ satisfies a Poincaré inequality with constant 2, the same constant we found in dimension 1 in (2.22).

Proposition 5.1.10 *Let $\alpha > 0$ and $\mathbf{Z} \sim \mathcal{MS}(\alpha, \nu)$ be a max-stable random vector with angular measure ν , and $f \in \mathbf{L}^2(E_0, \Phi_{\alpha,\nu})$. Then we have:*

$$\mathbb{V}(f(\mathbf{Z})) \leq 2\mathcal{E}_{\alpha,\nu}(f) = \int_{E_0} \mathbb{E}\left[\left(f(\mathbf{Z} \oplus r\mathbf{u}^{\frac{1}{\alpha}}) - f(\mathbf{Z})\right)^2\right] \frac{\alpha}{r^{\alpha+1}} \, dr \, d\nu(\mathbf{u}). \quad (5.8)$$

Thus $(\mathbf{P}_t^{\alpha,\nu})_{t \geq 0}$ converges exponentially fast to the stationary measure $\Phi_{\alpha,\nu}$ in $\mathbf{L}^2(E_0, \Phi_{\alpha,\nu})$:

$$\|\mathbf{P}_t^{\alpha,\nu}f - \mathbb{E}[f(\mathbf{Z})]\|_{\mathbf{L}^2(E_0, \Phi_{\alpha,\nu})} \leq e^{-\frac{t}{2}} \|f - \mathbb{E}[f(\mathbf{Z})]\|_{\mathbf{L}^2(E_0, \Phi_{\alpha,\nu})},$$

for all $t \geq 0$ and $f \in \mathcal{S}_{\alpha,d}$.

Proof. The first part of the result comes readily from the previous lemma. The second part is once again a consequence of 1.1.12. \square

What we have proved is sometimes called a first-order Poincaré inequality. We now turn to the second-order Poincaré inequality; it quantifies the distance between a functional of a max-id random vector and the normal distribution. We prove it as a direct corollary of a famous result by Last,

Peccati and Schulte. Let η be a Poisson process over E_ℓ with intensity measure μ . Denote by F a Poisson functional. Assume that \bar{f} is centered and has unit variance, and set:

$$\begin{aligned}\gamma_1 &:= 2 \left(\int_{E_\ell^3} (\mathbb{E}[(D_{\mathbf{x}}^+ \bar{f})^2 (D_{\mathbf{y}}^+ \bar{f})^2])^{\frac{1}{2}} (\mathbb{E}[(D_{\mathbf{x},z}^+ \bar{f})^2 (D_{\mathbf{y},z}^+ \bar{f})^2])^{\frac{1}{2}} d^3 \mu(\mathbf{x}, \mathbf{y}, \mathbf{z}) \right)^{\frac{1}{2}} \\ \gamma_2 &:= \left(\int_{E_\ell^3} \mathbb{E}[(D_{\mathbf{x},z}^+ \bar{f})^2 (D_{\mathbf{y},z}^+ \bar{f})^2] d^3 \mu(\mathbf{x}, \mathbf{y}, \mathbf{z}) \right)^{\frac{1}{2}} \\ \gamma_3 &:= \int_{E_\ell} \mathbb{E}[|D_{\mathbf{x}} \bar{f}|^3] d\mu(\mathbf{x}).\end{aligned}$$

The result of Last, Peccati and Schulte states that if μ is σ -finite, then:

$$d_W(\bar{f}, N) \leq \gamma_1 + \gamma_2 + \gamma_3,$$

where N is a random variable with Gaussian distribution. A similar result controlling the Kolmogorov distance instead of the Wasserstein distance between \bar{f} and N is also given in their paper, but the bound is more involved. The important point for us here is that all the γ_i depend on the underlying Poisson process η only through \bar{f} . By choosing $\bar{f} = f \circ \mathbf{m}$, where $f : E_\ell \rightarrow \mathbb{R}$, one obtains a bound on the Wasserstein distance between a functional f of a max-id random vector with σ -finite exponent measure μ and a Gaussian random variable (an exponent measure is always σ -finite). The bound itself involves only the distribution of \mathbf{Z} , and makes no reference to the Poisson process η anymore, hence its interest. In other words, we have deduced a second-order Poincaré inequality for max-id distributions from its original Poisson counterpart.

5.2 Another quantization of the Poisson process

This section is mainly of an heuristic nature and merely introduces new objects without studying them in a fully rigorous manner.

It originates from the following observation: when a sequence of random variables $(Z_n)_{n \geq 1}$ converges in distribution to an extreme value distribution, it is often the case that Z_n is the maximum of some random set. The most basic occurrence of this is when Z_n is a renormalized maximum of n i.i.d. positive random variables X_1, \dots, X_n with regularly varying tails, say with index $-\alpha$. One can actually prove that the random set $\{X_1, \dots, X_n\}$ converges in distribution to a Poisson process with intensity measure $\alpha r^{-(\alpha+1)}$ on \mathbb{R}_+^* , see [Durrett \[2019\]](#). Remember that the Fréchet distribution $\mathcal{F}(\alpha)$ is equal in law to the maximum of such a process.

Another example is the coupon collector problem studied previously: Z_n is the maximum of n exchangeable random variables, each with geometric distribution $\text{Geom}(1/n)$, and converges in law to a Gumbel distribution, which can be seen as the maximum of a Poisson process with intensity measure e^{-r} on \mathbb{R} . Two recent papers ([Glavaš and Mladenović \[2018\]](#), [Iliencko \[2019\]](#)) show that the set of the arrival times of the coupons $\{T_1^n, \dots, T_n^n\}$ converges to a Poisson process with intensity measure e^{-r} .

We give one last example, taken from [Soshnikov \[2004\]](#). The largest eigenvalue of a Wigner random matrix whose entries have heavy-tails converges in law to a Fréchet distribution when the dimension n goes to ∞ . And indeed, the random set of the positive eigenvalues of such matrices, properly renormalized, converges in law to a Poisson process with intensity measure $\alpha r^{-(\alpha+1)}$ on \mathbb{R}_+^* .

We have already given tools to quantify the speed of convergence to the Fréchet and Gumbel distributions in each of those examples, and even treat the first two in this thesis. But it is certainly more instructive to know at which speed the whole process will converge to the limiting Poisson process, although it will most likely be at a slower rate than its maximum. To reach that goal, we need to quantize Poisson processes. Of course the Glauber semi-group already does that, and it even enjoys on a Mehler’s formula:

$$P_t f(\phi) = \mathbb{E}[f(e^{-t} \circ \phi + (1 - e^{-t}) \circ \eta)], \phi \in \mathfrak{N}_E.$$

However this formula relies only on the stability property shared by all Poisson processes, whatever their intensity measure is:

$$p \circ \eta_1 \cup (1 - p) \circ \eta_2 \stackrel{d}{=} \eta_1,$$

where η_1, η_2 are two independent Poisson processes with the same intensity measure. But the Poisson processes which interest us are those whose intensity measure μ satisfies a homogeneity condition given in the next definition:

Definition 5.2.1. Let d be a positive integer and μ a Radon measure on \mathbb{R}^d . For any α , we say that μ is α -homogeneous if for all $A \in \mathcal{B}(\mathbb{R}^d)$

$$\mu(tA) = t^{-\alpha} \mu(A), \quad t \in \mathbb{R}_+^*,$$

when $\alpha \in \mathbb{R}^*$, or

$$\mu(A + t) = e^{-t} \mu(A), \quad t \in \mathbb{R},$$

if $\alpha = 0$.

For instance, the intensity measure of any simple max-stable random vector is α -homogeneous with $\alpha > 0$. The case $\alpha = 0$ corresponds to Gumbel marginals, while $\alpha < 0$ is associated to Weibull marginals. Notice that thanks to the polar decomposition (1.23), any α -homogeneous measure is finite in neighborhoods of $+\infty$. Any Poisson process whose intensity measure fulfills that homogeneity condition for some $\alpha \neq 0$ then possesses the following stability property:

$$e^{-\frac{t}{\alpha}} \eta_1 \cup (1 - e^{-t})^{\frac{1}{\alpha}} \eta_2 \stackrel{d}{=} \eta_1$$

where the thinning operation \circ has been replaced by the multiplication of configurations by scalars, as described in definition 1.2.3. Indeed, the left-hand side is a Poisson process whose intensity measure is

$$e^{-t} \mu + (1 - e^{-t}) \mu = \mu,$$

thanks to the homogeneity property of μ . In the case $\alpha = 0$, we have instead

$$(\eta_1 - t) \cup (\eta_2 + \log(1 - e^{-t})) \stackrel{d}{=} \eta_1.$$

This suggests us the next two definitions. Before giving them, we must define what is a differentiable function on $\mathfrak{N}_{\mathbb{R}^d}$. Because α -homogeneous measures have an accumulation point at either $\mathbf{0}$ or $-\infty$, we cannot assume our configurations will be finite. Thus, we will work with functionals $f(\phi)$ where $f : (\mathbb{R}^d)^{\mathbb{N}} \rightarrow \mathbb{R}$ is a symmetric function on the space of sequences, *i.e.* for every bijective application σ from \mathbb{N}^* to \mathbb{N}^* , we have:

$$f(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \dots) = f(\mathbf{x}_1, \mathbf{x}_2, \dots)$$

if $\phi = (\mathbf{x}_n)_{n \geq 1}$. Such a functional is by definition a function of an infinite number of variables, but is symmetric w.r.t. each of them. We will say it is *differentiable* if f is differentiable w.r.t. to its first variable. A much more rigorous definition of what a differentiable functional of configurations on manifolds can be found in [Albeverio et al. \[1998\]](#). Our definition amounts to theirs in our context. Examples of differentiable functionals include:

- The coordinate-wise sum function \mathfrak{s} restricted to the subset of summable families of $\mathbb{R}^{\mathbb{N}}$:

$$\mathfrak{s}(\phi) = \sum_{n=1}^{\infty} x_n,$$

is differentiable, with derivative

$$\partial_n \mathfrak{s}(\phi) = 1.$$

- The cardinal function restricted to finite configurations is another example:

$$\partial_n \text{card}(\phi) = 0.$$

The fact that all its partial derivatives vanish can be explained by the fact that moving the points of ϕ does not change the size of ϕ .

- On the other hand, if A is any non-empty subset of \mathbb{R}^d , then

$$\phi \mapsto \text{card}(A \cap \phi)$$

is not differentiable everywhere, because moving a point of ϕ too far could make it leave A . However, if λ is a measure on $\mathbb{R}^{\mathbb{N}}$ such that

$$\lambda\left(\{(\mathbf{x}_n)_{n \geq 1} \in (\mathbb{R}^d)^{\mathbb{N}}, \mathbf{x}_n \in \partial A \text{ for some } n\}\right) = 0,$$

where ∂A is the border of A , then one has again

$$\partial_n \text{card}(A \cap \phi) = 0 \text{ } \lambda\text{-a.e.}$$

- The coordinate-wise maximum function \mathfrak{m} is also differentiable λ -a.e. if λ puts zero mass on configurations which possess the same point several times:

$$\lambda\left(\{(\mathbf{x}_n)_{n \geq 1} \in (\mathbb{R}^d)^{\mathbb{N}}, x_n^j = x_m^j \text{ for some } j \in \llbracket 1, d \rrbracket \text{ and } n \neq m\}\right) = 0,$$

in which case we find

$$\partial_n \mathfrak{m}(\phi) = (\mathbb{1}_{\{x_n^1 \geq x_m^1, m \in \mathbb{N}^*\}}, \dots, \mathbb{1}_{\{x_n^d \geq x_m^d, m \in \mathbb{N}^*\}}) \text{ } \lambda\text{-a.e.}$$

To compute the generator of $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$, we need to find a set of test functions regular enough to ensure the objects we work with are well-defined. However, it will be difficult to find a set large enough to include both the maximum function and the sum function for instance. Indeed, depending on the value of α , the sum of the points of a Poisson process with intensity $\alpha r^{-(\alpha+1)}$ will not converge a.s., while its maximum will always be finite. So, instead of trying to be as general as possible, we will inspire ourselves from the max-stable setting. To keep things simple, we will restrict ourselves to the case $\alpha > 0$.

Definition 5.2.2. Let $\mathfrak{t} : \mathfrak{N}_{E_0} \rightarrow \mathbb{R}^d$ be an application on the space of configurations \mathfrak{N}_{E_0} .

- We will say it is a *compact transformation* if there exists a point $\mathbf{x}_0 \in E_0$ such that

$$\mathfrak{t}(\phi) = \mathfrak{t}(\phi|_{[\mathbf{0}, \mathbf{x}_0]^c}), \quad \phi \in \mathfrak{N}_{E_0}$$

i.e. $\mathfrak{t}(\phi)$ only depends on the points of ϕ which belong to $[\mathbf{0}, \mathbf{x}_0]^c$.

- Let μ be an α -homogeneous measure on E_0 . If \mathfrak{t} is also μ -a.e. differentiable, then we say that \mathfrak{t} is a *μ -differentiable compact transform*.

The assumption that \mathfrak{t} takes its values in \mathbb{R}^d implies that $\mathfrak{t}(\eta)$ is always well-defined as soon as η is a configuration of E_0 . The adjective "compact" comes from the fact that sets of the form $[\mathbf{0}, \mathbf{x}_0]^c$ are compact in E_0 .

Definition 5.2.3. Let \mathfrak{t} be a μ -differentiable compact transform. We define the class of functions:

$$\bar{\mathcal{S}}_{\mu, \tau} := \{ \bar{f} = f \circ \mathfrak{t}, f \in \mathcal{S}_\Phi \}.$$

Definition 5.2.4 (α -stable Poisson semi-group). Let $\alpha \in \mathbb{R}$. Define the *α -stable Poisson semi-group* $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$ on the algebra $\mathcal{A} = \bar{\mathcal{S}}_{\mu, \tau}$ by

$$\bar{\mathbf{P}}_t^\mu f(\phi) := \begin{cases} \mathbb{E}[\bar{f}(e^{-\frac{t}{\alpha}} \phi \cup (1 - e^{-t})^{\frac{1}{\alpha}} \eta)], & \text{if } \alpha \in \mathbb{R}^* \\ \mathbb{E}[\bar{f}((\phi - t) \cup (\eta + \log(1 - e^{-t})))], & \text{if } \alpha = 0 \end{cases} \quad (5.9)$$

where η is a Poisson process whose intensity measure is α -homogeneous.

The definition of $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$ through a Mehler formula automatically yields the semi-group property, the ergodicity property and the fact that the distribution of η is an invariant measure for $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$.

Proposition 5.2.5 *Let μ be an α -homogeneous measure on \mathbb{R}^d , for some $\alpha \in \mathbb{R}$. Then the family of operators $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$ is a semi-group on $\bar{\mathcal{S}}_{\mu, \tau}$. The distribution of a Poisson process with intensity measure μ on \mathbb{R}^d is an invariant measure of $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$. Besides, this semi-group is ergodic.*

Proof. The proof is very similar to the ones we have given for the max-stable semi-groups. Everything relies on the fact that for every non-negative a and $\phi_1, \phi_2 \in \mathfrak{N}_{\mathbb{R}^d}$:

$$\begin{aligned} a(\phi_1 \cup \phi_2) &= a\phi_1 \cup a\phi_2 \\ \phi_1 \cup \phi_2 + a &= (\phi_1 + a) \cup (\phi_2 + a) \end{aligned}$$

and the aforementioned stability relations. □

What makes the previous proposition central in the study of stable laws is that it generates the other semi-groups associated to those distributions.

- Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a bounded function and $\bar{f} = f \circ \mathfrak{m}$. Then we have the relation:

$$\mathbf{P}_t^\alpha f(x) = \bar{\mathbf{P}}_t^\mu \bar{f}(\{x\}),$$

with $d\mu(r) = \alpha r^{-(\alpha+1)} dr$ and $\alpha > 0$.

- Let $\alpha \in (0, 1)$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded function, as well as $\bar{f} = f \circ \mathfrak{s}$. Then the operators P_t^α defined as:

$$P_t^\alpha f(x) = \bar{\mathbf{P}}_t^\mu \bar{f}(\{x\}),$$

constitute a Markov semi-group whose invariant measure is the distribution of $\mathfrak{s}(\eta)$, where η is a Poisson process on \mathbb{R}_+^* with intensity measure $d\mu(r) = \alpha r^{-(\alpha+1)} dr$ and $\alpha > 0$. The distribution of $\mathfrak{s}(\eta)$ is α -stable.

As in chapter 2, we will need a preliminary lemma to compute the generator of $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$. We will assume $\alpha = 1$ for ease of notations. We will also note $\phi \cup \mathbf{y}$ instead of $\phi \cup \{\mathbf{y}\}$:

Lemma 5.2.6 *Let η be a Poisson process on E_0 with intensity measure μ . Assume μ is 1-homogeneous. We have for $f \in \mathcal{S}_{\mu, \tau}$ and $\phi \in \mathfrak{N}_{E_0}$:*

$$\begin{aligned} \bar{\mathbf{P}}_t^\mu \bar{f}(\phi) &= e^{-\gamma_t \mu[\mathbf{0}, \mathbf{x}_0]^c} \bar{f}(e^{-t} \phi) \\ &+ e^{-\gamma_t \mu[\mathbf{0}, \mathbf{x}_0]^c} \gamma_t \int_{[\mathbf{0}, \mathbf{x}_0]^c} \bar{f}(e^{-t}(\phi \cup \mathbf{y})) d\mu(\mathbf{y}) \\ &+ e^{-\gamma_t \mu[\mathbf{0}, \mathbf{x}_0]^c} \sum_{n=2}^{\infty} \frac{\gamma_t^n}{n!} \int_{([\mathbf{0}, \mathbf{x}_0]^c)^n} \bar{f}(e^{-t}(\phi \cup \mathbf{y}_1 \cup \dots \cup \mathbf{y}_n)) d\mu(\mathbf{y}_1) \dots d\mu(\mathbf{y}_n). \end{aligned} \tag{5.10}$$

Proof. The proof of this lemma is similar to the one given in chapter 2: we want to apply identity (1.2) to $\phi' \mapsto \bar{f}(e^{-t}(\phi \cup \gamma_t \phi'))$. But \bar{f} depends on η only through its restriction to $[\mathbf{0}, \mathbf{x}_0]^c$. So the previous function depends only on $\gamma_t \eta$ restricted to $[\mathbf{0}, \mathbf{x}_0]^c$, which is still a Poisson process with intensity measure $\gamma_t \mu$, due to the 1-homogeneity property of μ . \square

Thanks to that lemma, we can compute the generator $\bar{\mathcal{L}}_\mu$ of $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$. The proof follows the same line of reasoning as before.

Proposition 5.2.7 *The semi-group $(\bar{\mathbf{P}}_t^\mu)_{t \geq 0}$ has generator $\bar{\mathcal{L}}_\mu$, given by:*

$$\bar{\mathcal{L}}_\mu \bar{f}(\phi) = -\langle \phi, \nabla \bar{f}(\phi) \rangle_{[\mathbf{0}, \mathbf{x}_0]^c} + \int_{[\mathbf{0}, \mathbf{x}_0]^c} (\bar{f}(\phi \cup \mathbf{y}) - \bar{f}(\phi)) d\mu(\mathbf{y}), \quad \bar{f} \in \bar{\mathcal{S}}_{\mu, \tau},$$

for any $\phi \in \mathfrak{N}_{E_0}$, where

$$\langle \phi, \nabla \bar{f}(\phi) \rangle_{[\mathbf{0}, \mathbf{x}_0]^c} := \sum_{x \in \phi \cap [\mathbf{0}, \mathbf{x}_0]^c} x \partial_x \bar{f}(\phi).$$

Bibliography

- [1] S Albeverio, Yu.G Kondratiev, and M Röckner. Analysis and geometry on configuration spaces: The Gibbsian case. *Journal of Functional Analysis*, 157(1):242–291, 1998. ISSN 0022-1236. doi: <https://doi.org/10.1006/jfan.1997.3215>. URL <https://www.sciencedirect.com/science/article/pii/S0022123697932158>. (Cited on p. 180).
- [2] A. Anastasiou, A. Barp, F. Briol, B. Ebner, R. Gaunt, F. Ghaderinezhad, J. Gorham, A. Gretton, C. Ley, Q. Liu, L. Mackey, C. Oates, G. Reinert, and Y. Swan. Stein’s method meets statistics: A review of some recent developments. 05 2021. (Cited on p. 14 and 24).
- [3] B. Arras and C. Houdré. On some operators associated with non-degenerate symmetric α -stable probability measures. *Potential Analysis*, 59(4):1643–1694, 2023. (Cited on p. 15, 25, and 64).
- [4] B. Arras and C. Houdré. On Stein’s method for multivariate self-decomposable laws. *Electronic Journal of Probability*, 2019. (Cited on p. 176).
- [5] B. Arras and Y. Swan. A stroll along the gamma. *Stochastic Processes and their Applications*, 127, 11 2015. doi: 10.1016/j.spa.2017.03.012. (Cited on p. 14 and 24).
- [6] E. Artin. *The Gamma function*. Dover Publications, 1964. (Cited on p. 144).
- [7] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer International Publishing, Cham, 2014. ISBN 9783319002262 9783319002279. doi: 10.1007/978- — 3- — 319- — 00227- — 9. URL <https://link.springer.com/10.1007/978- — 3- — 319- — 00227- — 9>. (Cited on p. 31 and 33).
- [8] A. D Barbour. Stein’s method for diffusion approximations. *Probability theory and related fields*, 84(3):297–322, 1990. (Cited on p. 14 and 24).
- [9] A. D Barbour and T. C Brown. Stein’s method and point process approximation. *Stochastic Processes and their Applications*, 43(1):9–31, 1992. (Cited on p. 14 and 24).
- [10] C. Bartholome and Y. Swan. Rates of convergence towards the Fréchet distribution. 2013. URL [arXiv:1311.3896](https://arxiv.org/abs/1311.3896). (Cited on p. 72 and 117).
- [11] V. Bentkus, F. Götze, and V. Paulauskas. Bounds for the accuracy of Poissonian approximations of stable laws. *Stochastic Processes and their Applications*, 65(1):55–68, 1996. ISSN 0304-4149. doi: [https://doi.org/10.1016/S0304- — 4149\(96\)00101- — 9](https://doi.org/10.1016/S0304- — 4149(96)00101- — 9). URL <https://www.sciencedirect.com/science/article/pii/S0304414996001019>. (Cited on p. 139).
- [12] V. Bentkus, A. Juozulynas, and Vygantas Paulauskas. Lévy-lepage series representation of

- stable vectors: Convergence in variation. *Journal of Theoretical Probability*, 14, 01 2000. doi: 10.1023/A:1017520702943. (Cited on p. 139).
- [13] H. Bergström. On some expansions of stable distribution functions. *Arkiv för Matematik*, 2, 01 1953. (Cited on p. 120).
- [14] A. Boneh and M. Hofri. The coupon-collector problem revisited — a survey of engineering problems and computational methods. *Communications in Statistics. Stochastic Models*, 13(1):39–66, 1997. doi: 10.1080/15326349708807412. URL <https://doi.org/10.1080/15326349708807412>. (Cited on p. 153).
- [15] D. Chafai. Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue. *ESAIM: Probability and Statistics*, 10:317–339, 2006. doi: 10.1051/ps:2006013. (Cited on p. 101).
- [16] L. H. Y. Chen. On the convergence of Poisson binomial to Poisson distributions. *The Annals of Probability*, pages 178–180, 1974. (Cited on p. 14 and 24).
- [17] L. H. Y. Chen. Poincaré-type inequalities via stochastic integrals. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 69:251–277, 1985. (Cited on p. 176).
- [18] L. H. Y. Chen and A. Xia. Stein’s method, Palm theory and Poisson process approximation. *The Annals of Probability*, 32, 07 2004. doi: 10.1214/009117904000000027. (Cited on p. 14 and 24).
- [19] L.H.Y. Chen, L. Goldstein, and Q.M. Shao. *Normal Approximation by Stein’s Method*. Probability and Its Applications. Springer Berlin Heidelberg, 2010. ISBN 9783642150074. (Cited on p. 13 and 24).
- [20] P. Chen, I. Nourdin, L. Xu, X. Yang, and R. Zhang. Non-integrable stable approximation by Stein’s method. *Journal of Theoretical Probability*, pages 1–50, 2022. (Cited on p. 14, 24, 113, and 117).
- [21] P. Chen, I. Nourdin, L. Xu, and X. Yang. Multivariate stable approximation by stein’s method. *Journal of Theoretical Probability*, 37(1):446–488, 2024. (Cited on p. 14 and 24).
- [22] J. P. Cohen. Convergence rates for the ultimate and penultimate approximations in extreme-value theory. *Advances in Applied Probability*, 14(4):833–854, 1982. ISSN 00018678. URL <http://www.jstor.org/stable/1427026>. (Cited on p. 11 and 22).
- [23] L. Coutin and L. Decreusefond. Donsker’s theorem in Wasserstein-1 distance. *Electron. Commun. Probab.*, 25:13 pp., 2020. doi: 10.1214/20-__ECP308. URL https://doi.org/10.1214/20-__ECP308. (Cited on p. 14 and 24).
- [24] L. Coutin, L. Decreusefond, and L. Huang. New approaches to CLT for stable random variables. *To be released*, 2024. (Cited on p. 14 and 24).
- [25] Y. Davydov and A. Nagaev. Theoretical aspects of simulation of random vectors having a symmetric stable distribution. *preprint*, 1999. (Cited on p. 139).
- [26] Y. Davydov, I. Molchanov, and S. Zuyev. Strictly stable distributions on convex cones. *Electronic Journal of Probability*, 13:no. 11, 259–321, 2008. doi: 10.1214/EJP.v13-__487.

- URL <https://doi.org/10.1214/EJP.v13-487>. (Cited on p. 16, 18, 26, 27, 28, 50, 133, and 176).
- [27] Y. Davydov, I. Molchanov, and S. Zuyev. Stability for random measures, point processes and discrete semigroups. *Bernoulli*, 17(3):1015 – 1043, 2011. doi: 10.3150/10- BEJ301. URL <https://doi.org/10.3150/10- BEJ301>. (Cited on p. 17 and 27).
- [28] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2007. ISBN 9780387344713. (Cited on p. 39 and 48).
- [29] L. de Haan and S. Resnick. Second-order regular variation and rates of convergence in extreme-value theory. *The Annals of Probability*, 24(1):97–124, 1996. ISSN 00911798. URL <http://www.jstor.org/stable/2244834>. (Cited on p. 11 and 22).
- [30] L. Decreusefond. The Stein-Dirichlet-Malliavin method. *ESAIM: Proceedings*, page 11, 2015. URL <https://partage.imt.fr/index.php/s/FGmZiZMLXz3FJfN>. (Cited on p. 158).
- [31] L. Decreusefond. *Selected topics in Malliavin calculus*, volume 10 of *Bocconi Springer Ser.* Springer, 2022. ISBN 978-3-031-01310-2; 978-3-031-01311-9. doi: 10.1007/978-3-031-01311-9. URL <https://partage.imt.fr/index.php/s/mrdNT6pXqjHF2yW>. (Cited on p. 18, 28, 34, 35, 82, and 174).
- [32] C. Döbler. *New developments in Stein’s method with applications*. PhD thesis, Dissertation, Bochum, Ruhr-Universität, 2012. (Cited on p. 111).
- [33] C. Dombry, S. Engelke, and M. Oesting. Asymptotic properties of the maximum likelihood estimator for multivariate extreme value distributions. 2016. URL <https://api.semanticscholar.org/CorpusID:88514381>. (Cited on p. 55).
- [34] R. M. Dudley. *Convergence of Laws on Separable Metric Spaces*, page 385–438. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2002. (Cited on p. 59).
- [35] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019. ISBN 9781108473682. URL <https://books.google.fr/books?id=b22MDwAAQBAJ>. (Cited on p. 42, 55, and 178).
- [36] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events: for Insurance and Finance*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013. ISBN 9783540609315. URL <https://books.google.fr/books?id=BXOI2pICfJUC>. (Cited on p. 11, 21, 39, 42, and 44).
- [37] P. Erdős and A. Rényi. On a classical problem of probability theory. *Publ. Math. Inst. Hung. Acad. Sci.*, pages 215–220, 1961. (Cited on p. 153).
- [38] J. D. Esary, F. Proschan, and D. W. Walkup. Association of Random Variables, with Applications. *The Annals of Mathematical Statistics*, 38(5):1466 – 1474, 1967. doi: 10.1214/aoms/1177698701. URL <https://doi.org/10.1214/aoms/1177698701>. (Cited on p. 175).
- [39] S.N. Ethier and T.G. Kurtz. *Markov Processes : Characterizations and Convergence*. Wiley, 1986. (Cited on p. 31).

- [40] J. Faraut. *Analysis on Lie Groups: An Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008. (Cited on p. 74).
- [41] A. Feidt. *Stein's Method for Multivariate Extremes*. Theses, Université de Zürich, 2013. URL <https://arxiv.org/abs/1310.2564>. (Cited on p. 11, 22, 44, and 117).
- [42] R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):180–190, 1928. doi: 10.1017/S0305004100015681. (Cited on p. 11 and 21).
- [43] M. Fréchet. Sur la loi de probabilité de l'écart maximum. In *Annales de la société Polonaise de Mathématique*, volume 6, pages 93–116, 1927. (Cited on p. 11 and 21).
- [44] J-F. Le Gall. *Brownian motion, martingales and stochastic calculus*. Springer, 1st ed. edition, 2016. (Cited on p. 91 and 93).
- [45] R. E. Gaunt and S. Li. Bounding Kolmogorov distances through Wasserstein and related integral probability metrics. *Journal of Mathematical Analysis and Applications*, 522(1): 126985, June 2023. ISSN 0022-247X. doi: 10.1016/j.jmaa.2022.126985. URL <https://www.sciencedirect.com/science/article/pii/S0022247X22009994>. (Cited on p. 61 and 138).
- [46] L. Glavaš and P. Mladenović. New limit results related to the coupon collector's problem. *Studia Scientiarum Mathematicarum Hungarica*, 55:115–140, 03 2018. doi: 10.1556/012.2018.55.1.1385. (Cited on p. 178).
- [47] C. M. Goldie and R. L. Smith. Slow variation with remainder: theory and applications. *The Quarterly Journal of Mathematics*, 38(1):45–71, 03 1987. ISSN 0033-5606. doi: 10.1093/qmath/38.1.45. URL <https://doi.org/10.1093/qmath/38.1.45>. (Cited on p. 116).
- [48] L. Goldstein and G. Reinert. Stein's method and the zero bias transformation with application to simple random sampling. *The Annals of Applied Probability*, 7, 12 2001. doi: 10.1214/aoap/1043862419. (Cited on p. 62).
- [49] J. Gorham and L. Mackey. Measuring sample quality with Stein's method. *Advances in neural information processing systems*, 28, 2015. (Cited on p. 15 and 25).
- [50] J. Gorham, A. Duncan, S. Vollmer, and L. Mackey. Measuring sample quality with diffusions. *The Annals of Applied Probability*, 29, 11 2016. doi: 10.1214/19- AAP1467. (Cited on p. 15 and 25).
- [51] F Gotze. On the rate of convergence in the multivariate CLT. *The Annals of Probability*, pages 724–739, 1991. (Cited on p. 14 and 24).
- [52] E.J. Gumbel. *Statistics of Extremes*. Columbia University Press, 1958. doi: 10.7312/gumb92958. (Cited on p. 11 and 21).
- [53] B.C. Hall. *Quantum Theory for Mathematicians*. Graduate Texts in Mathematics. Springer New York, 2013. ISBN 9781461471165. (Cited on p. 73 and 74).
- [54] W. J. Hall and Jon A. Wellner. The rate of convergence in law of the maximum of an exponential sample. *Statistica Neerlandica*, 33(3):151–154, 1979. doi: <https://doi.org/10.1111/j>

- 1467- — 9574.1979.tb00671.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467- — 9574.1979.tb00671.x>. (Cited on p. 11, 21, and 117).
- [55] C. Houdré, V. Pérez-Abreu, and D. Surgailis. Interpolation, correlation identities, and inequalities for infinitely divisible variables. *Journal of Fourier Analysis and Applications*, 4: 651–668, 1998. (Cited on p. 176).
- [56] A. Iliencko. Convergence of point processes associated with coupon collector’s and Dixie cup problems. 04 2019. (Cited on p. 178).
- [57] J. D. Kečkić and P. M. Vasić. Some inequalities for the gamma function. *Publications de l’Institut Mathématique*, 11(31):107–114, 1971. (Cited on p. 135).
- [58] H. Kusumoto and A. Takeuchi. Remark on rates of convergence to extreme value distributions via the Stein equations. *Springer - Extremes*, 2020. (Cited on p. 72, 78, 86, and 117).
- [59] G. Last and M. Penrose. *Lectures on the Poisson Process*. Cambridge University Press, 2017. (Cited on p. 34, 36, 174, 175, and 176).
- [60] G Last and M. D Penrose. Poisson process, Fock space representation, chaos expansion and covariance inequalities. *Probability theory and related fields*, 150(3):663–690, 2011. (Cited on p. 18, 28, and 34).
- [61] G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler’s formula, second order Poincaré inequalities and stabilization. *Probability theory and related fields*, 165:667–723, 2016. (Cited on p. 18 and 28).
- [62] M.R. Leadbetter, G. Lindgren, and H. Rootzen. *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer New York, 2012. ISBN 9781461254492. (Cited on p. 12 and 39).
- [63] M. Ledoux and Vygantas Paulauskas. A rate of convergence in the Poissonian representation of stable distributions. *Lithuanian Mathematical Journal*, 36:388–399, 01 1996. doi: 10.1007/BF02986862. (Cited on p. 139).
- [64] C. Ley, G. Reinert, and Y. Swan. Stein’s method for comparison of univariate distributions. *Probability Surveys*, 14(none), 2017-01. doi: 10.1214/16- — PS278. (Cited on p. 13, 14, 24, and 25).
- [65] F. Lust-Piquard. Ornstein–Uhlenbeck semi-groups on stratified groups. *Journal of Functional Analysis*, 258(6):1883–1908, 2010. ISSN 0022-1236. doi: <https://doi.org/10.1016/j.jfa.2009.11.012>. URL <https://www.sciencedirect.com/science/article/pii/S0022123609004911>. (Cited on p. 16 and 26).
- [66] Malcolm AH MacCallum. On the classification of the real four-dimensional lie algebras. In *On Einstein’s Path: essays in honor of Engelbert Schucking*, pages 299–317. Springer, 1999. (Cited on p. 74 and 75).
- [67] A. W. Marshall and I. Olkin. Domains of Attraction of Multivariate Extreme Value Distributions. *The Annals of Probability*, 11(1):168 – 177, 1983. doi: 10.1214/aop/1176993666. URL <https://doi.org/10.1214/aop/1176993666>. (Cited on p. 48 and 175).

- [68] G. Mijoule, M. Raič, G. Reinert, and Y. Swan. Stein’s density method for multivariate continuous distributions. *Electronic Journal of Probability*, 28:1–40, 2023. (Cited on p. 13 and 23).
- [69] M. Mitzenmacher and E. Upfal. *Probability and computing : randomization and probabilistic techniques in algorithms and data analysis*. CUP, 2nd ed. edition, 2017. (Cited on p. 153).
- [70] A. Müller. Integral probability metrics and their generating classes of functions. *Advances in Applied Probability*, 29(2):429–443, 1997. ISSN 00018678. URL <http://www.jstor.org/stable/1428011>. (Cited on p. 59).
- [71] M. Namiki. *Stochastic quantization*, volume 9. Springer Science & Business Media, 1992. (Cited on p. 14 and 24).
- [72] I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus*. CUP, 1st ed. edition, 2012. (Cited on p. 13, 23, 59, and 67).
- [73] E. Omev and S.T. Rachev. Rates of convergence in multivariate extreme value theory. *Journal of Multivariate Analysis*, 38(1):36–50, 1991. ISSN 0047-259X. doi: [https://doi.org/10.1016/0047-259X\(91\)90030-6](https://doi.org/10.1016/0047-259X(91)90030-6). URL <https://www.sciencedirect.com/science/article/pii/0047259X91900306>. (Cited on p. 11 and 22).
- [74] I. Papastathopoulos and J. A. Tawn. Extended generalised pareto models for tail estimation. *Journal of Statistical Planning and Inference*, 143(1):131–143, 2013. ISSN 0378-3758. doi: <https://doi.org/10.1016/j.jspi.2012.07.001>. URL <https://www.sciencedirect.com/science/article/pii/S0378375812002388>. (Cited on p. 45).
- [75] G. Parisi and Y. S. Wu. Perturbation theory without gauge fixing. *Scientia Sinica. Zhongguo Kexue*, 24(4):483–496, 1981. ISSN 0582-236x. (Cited on p. 14 and 24).
- [76] A. Pósfai. *Approximation theorems related to the coupon collector’s problem*. PhD thesis, Szegedi Tudományegyetem (Hungary), 2010. (Cited on p. 170).
- [77] N. Privault. *Stochastic Analysis in Discrete and Continuous Settings*. Springer, 1st ed. edition, 2009. (Cited on p. 34 and 36).
- [78] S. I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer New York, 1987. doi: <https://doi.org/10.1007/978-0-387-75953-1>. (Cited on p. 34, 39, 41, 42, 44, 45, 48, 55, 102, 114, 116, and 175).
- [79] S. I. Resnick. *Heavy-Tail Phenomena*. Springer New York, 2006. doi: <https://doi.org/10.1007/978-0-387-45024-7>. (Cited on p. 34, 39, 45, and 139).
- [80] S. I. Resnick and R. Roy. Random USC functions, max-stable processes and continuous choice. *The Annals of Applied Probability*, 1(2):267–292, 1991. ISSN 10505164. URL <http://www.jstor.org/stable/2959768>. (Cited on p. 56).
- [81] L.C.G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales*. Cambridge University Press, 2nd ed. edition, April 2000. (Cited on p. 31 and 91).
- [82] H. Rootzén. Attainable rates of convergence of maxima. *Statistics & Probability Letters*, 2(4):219–221, 1984. ISSN 0167-7152. doi: [https://doi.org/10.1016/0167-7152\(84\)90019-1](https://doi.org/10.1016/0167-7152(84)90019-1).

- URL <https://www.sciencedirect.com/science/article/pii/0167715284900191>. (Cited on p. 114).
- [83] N. Ross. Fundamentals of Stein's method. *Probability Surveys*, 8:210 – 293, 2011. doi: 10.1214/11-PS182. URL <https://doi.org/10.1214/11-PS182>. (Cited on p. 13 and 24).
- [84] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes*. Routledge, 1994. doi: <https://doi.org/10.1201/9780203738818>. (Cited on p. 56, 90, 175, and 176).
- [85] R. L. Smith. Uniform rates of convergence in extreme-value theory. *Advances in Applied Probability*, 14(3):600–622, 1982. doi: 10.2307/1426676. (Cited on p. 11, 22, and 116).
- [86] A. Soshnikov. Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails. *Electronic Communications in Probability*, 9, 06 2004. doi: 10.1214/ECP.v9-1112. (Cited on p. 178).
- [87] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, pages 583–602, 1972. (Cited on p. 12, 22, and 23).
- [88] F. W. Steutel and K. van Harn. Discrete analogues of self-decomposability and stability. *The Annals of Probability*, 7(5):893–899, 1979. ISSN 00911798, 2168894X. URL <http://www.jstor.org/stable/2243313>. (Cited on p. 17 and 27).
- [89] S.A. Stoev and M.S. Taqqu. Extremal stochastic integrals: a parallel between max-stable processes and alpha-stable processes. *Extremes*, 8:237–266, 2005. (Cited on p. 56, 57, and 90).
- [90] C. Sándor. A rate of convergence for coupon collectors. 1993. ISSN 0001-6969. (Cited on p. 170).
- [91] A. Vasseur. *Asymptotic analysis of some point processes*. Theses, Télécom ParisTech, December 2017. URL <https://pastel.hal.science/tel-03413164>. (Cited on p. 18 and 28).
- [92] W. Vervaat. *Random upper semicontinuous functions and extremal processes*. Report. Centrum voor Wiskunde en Informatica. Stichting Mathematisch Centrum. MS. CWI, 1988. URL <https://books.google.fr/books?id=khPjcQAACAAJ>. (Cited on p. 56).
- [93] C. Villani. *Optimal transport – Old and new*, volume 338, pages xxii+973. Springer, 01 2008. doi: 10.1007/978-3-540-71050-9. (Cited on p. 62 and 106).
- [94] F. Wong and J. Collins. Evidence that coronavirus superspreading is fat-tailed. *Proceedings of the National Academy of Sciences*, 117, 09 2020. doi: 10.1073/pnas.2018490117. (Cited on p. 11 and 21).
- [95] L. Xu. Approximation of stable law in Wasserstein-1 distance by Stein's method. *The Annals of Applied Probability*, 29(1):pp. 458–504, 2019. ISSN 10505164, 21688737. URL <https://www.jstor.org/stable/26581794>. (Cited on p. 117).
- [96] V. M. Zolotarev. Probability metrics. *Theory of Probability & Its Applications*, 28(2):278–302, 1984. (Cited on p. 59).
- [97] V. M. Zolotarev. *One-dimensional stable distributions*, volume 65. American Mathematical Soc., 1986. (Cited on p. 120).

Titre : Méthode de Stein pour les lois d'extremum

Mots clés : Théorie des valeurs extrêmes, Méthode de Stein, Approche par générateur, Processus de Markov, Processus de Poisson, Inégalités fonctionnelles, Cônes convexes

Résumé : La théorie des valeurs extrêmes étudie la probabilité de survenance d'événements extrêmes, tels les inondations, les sécheresses ou encore les crises financières. Une part importante de cette théorie repose sur les théorèmes limites, comme celui des valeurs extrêmes, ou de Pickands-Balkema-de Haan. Afin d'appliquer ces théorèmes avec précision et approcher raisonnablement la loi des données extrêmes, inconnue en général, par son modèle limite, il faut pouvoir quantifier la vitesse de convergence de ces théorèmes. Une façon de faire est d'utiliser l'approche par générateur de la méthode de Stein. Aussi, dans cette thèse nous introduisons et étudions une famille de semi-groupes de Markov spécialement construits pour admettre les lois d'extremum comme mesure invariante. Pour ce faire, la définition choisie repose sur une formule de Mehler, elle-même

conséquence des relations de stabilité satisfaites par les lois max-stables. L'avantage principal de cette construction est que les semi-groupes ainsi définis disposent automatiquement de propriétés similaires à celles du semi-groupe d'Ornstein-Uhlenbeck (propriété de commutation, inégalité de Poincaré, identités de covariance, etc.). Nous appliquons ensuite ces résultats à l'obtention de bornes générales sur les distances à une loi d'extremum, puis nous spécialisons ces bornes dans différents contextes pour obtenir des taux explicites. Enfin le dernier chapitre porte sur les processus de Poisson dont la mesure intensité satisfait une propriété d'homogénéité. Nous étudions comment les propriétés bien connues de ces processus se traduisent en nouveaux résultats pour les lois max-stables, éclairant ainsi d'une autre manière le contenu des chapitres précédents.

Title : Stein's method for extreme value distributions

Keywords : Extreme value theory, Stein's method, Generator approach, Markov processes, Poisson processes, Functional inequalities, Convex cones

Abstract : Extreme value theory deals with the probability of occurrence of extreme events, such as floods, droughts or financial crises. An important part of that theory relies on limit theorems, such as the extreme value theorem, or the Pickands-Balkema-de Hann theorem. In order to apply those theorems accurately and approximate efficiently the usually unknown distribution of the extreme data by its limit model, one needs to quantify the speed of convergence of those theorems. A manner of doing so is to use the generator approach of Stein's method. That is why in this thesis we construct a family of Markov semi-groups whose invariant measure is an extreme value distribution. We do so *via* a Mehler's formula, which re-

lies itself on the stability property satisfied by max-stable distributions. Thanks to this definition, the semi-groups satisfy similar properties to the usual Ornstein-Uhlenbeck semi-group (commutation rule, Poincaré's inequality, covariance identities, etc.). We then proceed to apply those results to the generator approach of Stein's method to deduce rates of convergence to extreme value distributions in various settings. The last chapter focuses on Poisson processes whose intensity measure satisfies an homogeneity assumption and how their standard properties translate into new results for max-stable distributions, thus shedding a new light on the contents of the previous chapters.