Mod-ϕ convergence:
Normality zones and precise deviations

Valentin Féray

Pierre-Loïc Méliot

Ashkan Nikeghbali

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH — WINTERTHURERSTRASSE 190, 8057 ZÜRICH, SWITZERLAND
E-mail address, V. Féray: `valentin.feray@math.uzh.ch`

LABORATOIRE DE MATHÉMATIQUES, BÂTIMENT 425 — FACULTÉ DES SCIENCES D'ORSAY — UNIVERSITÉ PARIS-SUD, 91400 ORSAY, FRANCE
E-mail address, P.-L. Méliot: `pierre-loic.meliot@math.u-psud.fr`

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH — WINTERTHURERSTRASSE 190, 8057 ZÜRICH, SWITZERLAND
E-mail address, A. Nikeghbali: `ashkan.nikeghbali@math.uzh.ch`

ABSTRACT. In this book, we use the framework of mod-ϕ convergence to prove precise large or moderate deviations for quite general sequences of real valued random variables $(X_n)_{n \in \mathbb{N}}$, which can be lattice or non-lattice distributed. We establish precise estimates of the fluctuations $\mathbb{P}[X_n \in t_n B]$, instead of the usual estimates for the rate of exponential decay $\log(\mathbb{P}[X_n \in t_n B])$. Our approach provides us with a systematic way to characterise the normality zone, that is the zone in which the Gaussian approximation for the tails is still valid. Besides, the residue function measures the extent to which this approximation fails to hold at the edge of the normality zone.

The first chapters of the book are devoted to a proof of these abstract results and comparisons with existing results. We then propose new examples covered by this theory and coming from various areas of mathematics. In particular, we complete our theory of precise deviations by a concrete method of cumulants and dependency graphs, which applies to many examples of sums of "weakly dependent" random variables. The large number as well as the variety of examples hint at a universality class for second order fluctuations.
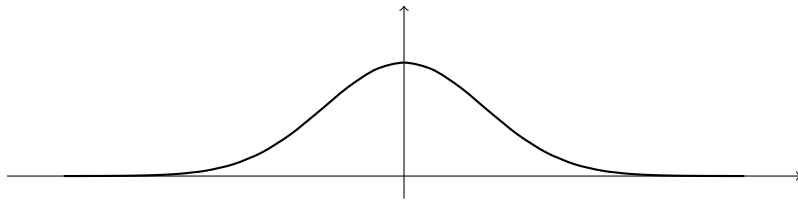
Contents

Preface

The central limit theorem ([dM56, Lap40, Lya00, Lya01, Lin22, Lév25]) for independent and identically distributed (i.i.d.) random variables is one of the most ubiquitous theorem in probability theory: suitably renormalised, the sum of such variables converges towards a standard normal variable.

We refer to [Fis11] for an historical account of this classical result. Roughly speaking, it shows the universality of fluctuations around the first-order limit given by the law of large numbers. Although independence is often used as a first approximation, many natural phenomena exhibit some complex dependency structure. Therefore, a large body of literature is devoted to relaxing the independence hypothesis. Thus, central limit theorems have been given for random variables with an underlying structure such as: martingales ([HH80, Chapter 3]), Markov chains (see *e.g.* [Cog72, IL71, GL78], and [Jon04] for a survey), mixing sequences (see [Ros56, Phi69, Dav73, LL96], and [Bra05] for a survey on mixing conditions), lattice models ([GJL75, EN78, New80]), m-dependence (*cf.* [RH48, Dia55, Ber73]), dependency graphs ([PL83, Jan88, BR89, Mik91]), exchangeable pairs (see [BC05, CGS11, Ros11]), determinantal point processes (*cf.* [HKPV09, Section 4.6]), *etc.*

Concretely, the central limit theorem for X_n expresses the limit of the tail probability when t is a fixed real number:

$$\lim_{n\to\infty} \mathbb{P}[X_n - \mathbb{E}[X_n] \geq t\sigma_n] = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-\frac{s^2}{2}} \, ds. \tag{1}$$

Such estimates are crucial in statistics to build confidence intervals. The convergence in law towards a Gaussian distribution is often complemented by other asymptotic results:

- the *speed of convergence* (uniformly in t) in Equation (1) (see [Ber41] and [Fel71, Chapter XVI]);
- the behaviour of the left-hand side of Eq. (1) when t tends to infinity together with n (*moderate and large deviation results*, see [DZ98]);
- a subsequent question is how fast can t grow, so that the limit given in Eq. (1) is still valid (*normality zone*);
- in another direction, *local limit theorems* describe the probability of $X_n - \mathbb{E}[X_n]$ to be in an interval of constant scale (see [Gne48]).

The canonical way to establish the central limit theorem for i.i.d. random variables is to use characteristic functions and Lévy's continuity theorem. This is also used in some of the above-mentioned extensions, in addition to other techniques, mainly based on the method of moments and Stein's method (see for instance the first chapters of [Str11]).

In this monograph, we focus on the characteristic function approach, for which we propose a *renormalisation theory* — called mod-ϕ convergence. If the characteristic function converges after a suitable renormalisation, we prove some precise moderate and large deviation results, which enables us to describe the normality zone. Results for the speed of convergence and local limit theorems will be discussed in a companion work [FMN16]. The idea of using estimates on characteristic functions to obtain central limit theorems and deviation probabilities is of course not new, but we provide here a general framework, together with many examples. These examples come from various mathematical fields: classical probability theory, number theory (statistics of additive arithmetic functions), combinatorics (statistics of random permutations), random matrix theory (characteristic polynomials of random matrices in compact Lie groups), graph theory (number of subgraphs in a random Erdős-Rényi graph), and non-commutative probability theory (asymptotics of random character values of symmetric groups).

Acknowledgements

CHAPTER 1

Introduction

1.1. Mod-ϕ convergence

The notion of mod-ϕ convergence has been studied in the articles [JKN11, DKN15, KN10, KN12, BKN14], in connection with problems from number theory, random matrix theory and probability theory. The main idea was to look for a natural renormalisation of the characteristic functions of random variables which do not converge in law (instead of a renormalisation of the random variables themselves). After this renormalisation, the sequence of characteristic functions converges to some non-trivial limiting function. Here is the definition of mod-ϕ convergence that we will use throughout this article (see Section 1.5 for a discussion on the different parts of this definition).

DEFINITION 1.1.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables, and let us denote by $\varphi_n(z) = \mathbb{E}[e^{zX_n}]$ their moment generating functions, which we assume to all exist in a strip*

$$\mathcal{S}_{(c,d)} = \{z,\ c < \operatorname{Re} z < d\},$$

with $c < d$ extended real numbers (we allow $c = -\infty$ and $d = +\infty$). We assume that there exists a non-constant infinitely divisible distribution ϕ with moment generating function $\int_{\mathbb{R}} e^{zx}\,\phi(dx) = \exp(\eta(z))$ that is well defined on $\mathcal{S}_{(c,d)}$, and an analytic function $\psi(z)$ that does not vanish on the real part of $\mathcal{S}_{(c,d)}$, such that locally uniformly in $z \in \mathcal{S}_{(c,d)}$,

$$\exp(-t_n\,\eta(z))\,\varphi_n(z) \to \psi(z), \tag{2}$$

where $(t_n)_{n\in\mathbb{N}}$ is some sequence going to $+\infty$. We then say that $(X_n)_{n\in\mathbb{N}}$ converges mod-ϕ on $\mathcal{S}_{(c,d)}$, with parameters $(t_n)_{n\in\mathbb{N}}$ and limiting function ψ. In the following we denote $\psi_n(z)$ the left-hand side of (2).

When ϕ is the standard Gaussian (resp. Poisson) distribution, we will speak of mod-Gaussian (resp. mod-Poisson) convergence. Besides, unless explicitely stated, we shall always assume that 0 belongs to the band of convergence $\mathcal{S}_{(c,d)}$, i.e., $c < 0 < d$. Under this assumption, Definition 1.1.1 implies mod-ϕ convergence in the sense of [JKN11, Definition 1.1] or [DKN15, Section 2].

It is immediate to see that mod-ϕ convergence implies a *central limit theorem* if the sequence of parameters t_n goes to infinity (see the remark after Theorem 3.3.1). But in fact there is much more information encoded in mod-ϕ convergence than merely the central limit theorem. Indeed, mod-ϕ convergence appears as a natural extension of the framework of sums of independent random variables (see Example 2.1.2): many interesting asymptotic results that hold for sums of independent random variables can also be established for sequences of random variables converging in the mod-ϕ sense ([JKN11, DKN15, KN10, KN12, BKN14]). For instance, under some general extra assumptions on the convergence in Equation (2), it is proved in [DKN15, KN12, FMN16] that one can establish local limit theorems for the random variables X_n. Then the local limit theorem of Stone appears as a special case of the local limit theorem for mod-ϕ

convergent sequences. But the latter also applies to a variety of situations where the random variables under consideration exhibit some dependence structure (*e.g.* the Riemann zeta function on the critical line, some probabilistic models of primes, the winding number for the planar Brownian motion, the characteristic polynomial of random matrices, finite fields *L*-functions, *etc.*). It is also shown in [BKN14] that mod-Poisson convergence (in fact mod-ϕ convergence for ϕ a lattice distribution) implies very sharp distributional approximation in the total variation distance (among other distances) for a large class of random variables. In particular, the total number of distinct prime divisors $\omega(n)$ of an integer n chosen at random can be approximated in the total variation distance with an arbitrary precision by explicitly computable measures.

Besides these quantitative aspects, mod-ϕ convergence also sheds some new light on the nature of some conjectures in analytic number theory. Indeed it is shown in [KN10] that the structure of the limiting function appearing in the moments conjecture for the Riemann zeta function by Keating and Snaith [KS00b] is shared by other arithmetic functions and that the limiting function ψ accounts for the fact that prime numbers do not behave independently of each other. More precisely, the limiting function ψ can be used to measure the deviation of the true result from what the probabilistic models based on a naive independence assumption would predict. One should note that these naive probabilistic models are usually enough to predict central limit theorems for arithmetic functions (*e.g.* the naive probabilistic model made with a sum of independent Bernoulli random variables to predict the Erdös-Kac central limit theorem for $\omega(n)$ or the stochastic zeta function to predict Selberg's central limit theorem for the Riemann zeta function) but fail to predict accurately mod-ϕ convergence by a factor which is contained in ψ. There is another example, where dependence appears in the limiting function ψ, while we have independence at the scale of central limit theorem: the log of the characteristic polynomial of a random unitary matrix, as a vector in $\mathbb{R}^2$, converges in the mod-Gaussian sense to a limiting function which is not the product of the limiting functions of each component considered individually although when properly normalised it converges to a Gaussian vector with independent components [KN12].

1.2. Theoretical results

The goal of this paper is to prove that the framework of mod-ϕ convergence as described in Definition 1.1.1 is suitable to obtain *precise* large and moderate deviation results for the sequence $(X_n)_{n\in\mathbb{N}}$ (throughout the paper, we call *precise deviation result* an equivalent of the deviation probability itself and not on its logarithm). Namely, our results are the following.

- We give equivalents for the quantity $\mathbb{P}[X_n \geq t_n x]$, where x is a fixed positive real number (see Theorem 3.2.2 for a lattice distribution ϕ and Theorem 4.2.1 for a non-lattice distribution). This can be viewed as an analogue (or an extension, see Section 4.5.2) of Bahadur-Rao theorem [BR60].

- We also consider probabilities of the kind $\mathbb{P}[X_n \in t_n B]$ where B is a Borelian set, and we give upper and lower bounds on this probability which coincide at first order for a *nice* Borelian set B, see Theorem 6.3.1. This result is an analogue of Ellis-Gärtner theorem [DZ98, Theorem 2.3.6] (see also the original papers [Gär77, Ell84]): we have stronger hypotheses than in Ellis-Gärtner theorem,

but also a more precise conclusion (the bounds involve the probability itself, not its logarithm).

- Besides, we give an equivalent for the probability $\mathbb{P}[X_n - \mathbb{E}[X_n] \geq s_n t_n]$, where $s_n = o(1)$, covering all intermediate scales between central limit theorem and deviations of order t_n (Theorem 3.3.1 in the lattice case, and Theorem 4.3.1 in the non-lattice case).

We also address the question of *normality zone*, *i.e.* the scale up to which the Gaussian approximation (coming from the central limit theorem) for the tail of the distribution of X_n is valid. In particular, our methods provide us with a systematic way to detect it and also explains how this approximation breaks at the edge of this zone; see Chapter 5. The problem of detecting the normality zone for sums of i.i.d. random variables has received some attention in the literature on limit theorems (originally in [Cra38], see also [IL71]). Our framework enables an extension of such results, going beyond the setting of independent random variables:

- we cover more situations, *e.g.* sums of dependent random variables with a sparse dependency graph, or integer valued random variables, such as random additive functions, satisfying mod-Poisson convergence;
- we describe the correction to the normal approximation needed at the edge of the normality zone.

An interesting fact in our deviation results is the appearance of the limiting function ψ in deviations at scale t_n. This means that, at smaller scales, a sequence X_n converging mod-ϕ behaves exactly as a sum of t_n i.i.d. variables with distribution ϕ. However, at scale t_n, this is not true any more and the limiting function ψ gives us exactly the correcting factor.

In particular, in the case of mod-Gaussian convergence, the scale t_n is the first scale where the equivalent given by the central limit theorem is not valid anymore. In this case, one often observes a symmetry breaking phenomenon which is explained by the appearance of function ψ; see Section 4.4.

A special case of mod-Gaussian convergence is the case where (2) is proved using bounds on the cumulants of X_n — see Section 5.1. This case is particularly interesting as:

- it contains a large class of examples, see below in Section 1.3;
- in this setting, one can obtain deviation results at a scale larger that t_n (typically, $o((t_n)^{5/4})$, see Proposition 5.2.1).

The arguments involved in the proofs of our deviation results are standard, but they nonetheless need to be carefully adapted to the framework of Definition 1.1.1: elementary complex analysis, the method of change of probability measure or tilting due to Cramér, or adaptations of Berry-Esseen type inequalities with smoothing techniques.

REMARK 1.2.1. We should here mention the work of Hwang [Hwa96], with some similarities with ours. Hwang works with hypotheses similar to Definition 1.1.1, except that the convergence takes place uniformly on all compact sets contained in a given disk centered at the origin (while we assume convergence in a strip; thus this is weaker than our hypothesis, see Remark 4.5.3 for a discussion on this point). Under this hypothesis (and an hypothesis of the convergence speed), Hwang obtains an

equivalent of the probability $\mathbb{P}[X_n - \mathbb{E}[X_n] \geq s_n t_n]$ with $s_n = o(1)$, and even gives some asymptotic expansion of this probability. However, Hwang does not give any deviation result at the scale t_n and hence, none of his results show the role of ψ in deviation probabilities. Besides, he has no results in the multi-dimensional setting.

1.3. Applications

After proving our abstract results, we provide a large set of (new) examples where these results can be applied. We have thus devoted the second half of the paper to examples, from a variety of different areas.

Chapter 7 contains examples where the moment generating function is explicit, or given as a path integral of an explicit function. First, in Section 7.2, we recover results of Radziwill [Rad09] on precise large deviations for additive arithmetic functions, by carefully recalling the principle of the Selberg-Delange method. The next examples — Sections 7.3 and 7.4 — involve the total number of cycles (resp. rises) for random permutations. The precise large deviation result in the case of cycles was announced in a recent paper of Nikeghbali and Zeindler [NZ13], where the mod-convergence was proved by the singularity analysis method. Finally, in Section 7.5, we compute deviation probabilities of the characteristic polynomial of random matrices in compact Lie groups. This completes previous results by Hughes, Keating and O'Connell [HKO01] on large deviations for the characteristic polynomial.

Surprisingly, mod-Gaussian convergence can also be established in some cases, even if neither the moment generating function nor an appropriate bivariate generating series is known explicitly. A first example of this situation is given in Chapter 8. We give a criterion based on the location of the zeroes of the probability generating function, which ensures mod-Gaussian convergence. We then apply this result to the number of blocks in a uniform random set-partition of size n. As a consequence, we obtain the normality zone for this statistics, refining the central limit theorem of Harper [Har67].

Our next examples lie in the framework in which mod-Gaussian convergence is obtained *via* bounds on cumulants (Section 5.1). In Chapter 9, we show that such bounds on cumulants typically arise for $X_n = \sum_{i=1}^{N_n} Y_i$, where the $Y_i's$ have a *sparse dependency graph* (references and details are provided in Chapter 9). With weak hypothesis on the second and third cumulants, this implies mod-Gaussian convergence of a renormalised version of X_n (Theorem 9.5.1). This allows us to provide new examples of variables converging in the mod-Gaussian sense.

- First, we consider subgraph count statistics X_γ in Erdös-Rényi random graph $G(n, p)$ (Theorem 10.0.1) for a fixed p between 0 and 1. Moderate deviation probabilities in this case are given and compared with the literature on the subject in Chapter 10. We are also able to determine the size of the normality zone of X_γ.

- In our last application in Chapter 11, we use the machinery of dependency graphs in non-commutative probability spaces, namely, the algebras $\mathbb{C}\mathfrak{S}(n)$ of the symmetric groups, all endowed with the restriction of a trace of the infinite symmetric group $\mathfrak{S}(\infty)$. The technique of cumulants still works and it gives the fluctuations of random integer partitions under so-called *central measures*

in the terminology of Kerov and Vershik. Thus, one obtains a central limit theorem and moderate deviations for the values of the random irreducible characters of symmetric groups under these measures.

The variety of the many examples that fall in the seemingly more restrictive setting of mod-Gaussian convergence makes it tempting to assert that it can be considered as a universality class for second order fluctuations.

REMARK 1.3.1. The idea of using bounds on cumulants to show moderate deviations for a family of random variable with some given dependency graph is not new — see in particular [DE13b]. Nevertheless, the bounds we obtain in Theorem 9.1.6 (and also in Theorem 9.1.7) are stronger than those which were previously known and, as a consequence, we obtain deviation results at a larger scale. Another advantage of our method is that it gives estimates of the deviation probability itself, and not only of its logarithm.

1.4. Forthcoming works

As an intermediate step for our deviation estimates, we give Berry-Esseen estimates for random variables that converge mod-ϕ (Proposition 4.1.1). These estimates are optimal for this setting, though they can be improved in some special cases, such as sums of independent or weakly dependent variables. In a companion paper [FMN16], we establish optimal Berry-Esseen bounds in these cases, providing a mod-ϕ alternative to Stein's method.

In another direction, in [FMN15], we extend some results of this paper to a multi-dimensional framework. This situation requires more care since the geometry of the Borel set B, when considering $\mathbb{P}[X_n \in t_n B]$, plays a crucial role.

1.5. Discussion on our hypotheses

The following remarks explain the role of each hypothesis of Definition 1.1.1. As we shall see later, some assumptions can be removed in some of our results (*e.g.*, the infinite-divisibility of the reference law), but Definition 1.1.1 provides a coherent setting where all the techniques presented in the paper do apply without further verification.

REMARK 1.5.1 (Analyticity). The existence of the relevant moment generating function on a strip is crucial in our proof, as we consider in Chapter 4 the Fourier transform of $\widetilde{X_n}$, obtained from X_n by an exponential change of measure. We also use respectively the existence of continuous derivatives up to order 3 for η and ψ on the strip $\mathcal{S}_{(c,d)}$, and the local uniform convergence of ψ_n and its first derivatives (say, up to order 3) toward those of ψ. By Cauchy's formula, the local uniform convergence of analytic functions imply those of their derivatives, so it provides a natural framework where convergence of derivatives are automatically verified.

Let us mention however that these assumptions of analyticity are a bit restrictive, as they imply that the X_n's and ϕ have moments of all order; in particular, ϕ cannot be any infinitely divisible distribution (for instance the Cauchy distribution is excluded). That

explains that the theory of mod-ϕ convergence was initially developed with characteristic functions on the real line rather than moment generating functions in a complex domain. With this somehow weaker hypothesis, one can find many examples for instance of mod-Cauchy convergence (see *e.g.* [DKN15, KNN15]), while the concept of mod-Cauchy convergence does not even make sense in the sense of Definition 1.1.1. We are unfortunately not able to give precise deviation results in this framework.

REMARK 1.5.2 (Infinite divisibility and non-vanishing of the terms of mod-ϕ convergence). The non-vanishing of ψ is a natural hypothesis since evaluations of ψ appear in many estimates of non-zero probabilities, and also in denominators in fractions, see for instance Lemma 4.2.5. The assumption that ϕ is an infinitely divisible distribution will be discussed in Section 4.5.2.

CHAPTER 2

Preliminaries

2.1. Basic examples of mod-convergence

Let us give a few examples of mod-ϕ convergence, which will guide our intuition throughout the paper. In these examples, it will be useful sometimes to precise the speed of convergence in Definition 1.1.1.

DEFINITION 2.1.1. *We say that the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges mod-ϕ at speed $O((t_n)^{-v})$ if the difference of the two sides of Equation (2) can be bounded by $C_K (t_n)^{-v}$ for any z in a given compact subset K of $\mathcal{S}_{(c,d)}$. We use the analoguous definition with the $o(\cdot)$ notation.*

EXAMPLE 2.1.2. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of centered, independent and identically distributed real-valued random variables, with $\mathbb{E}[e^{zY}] = \mathbb{E}[e^{zY_1}]$ analytic and non-vanishing on a strip $\mathcal{S}_{(c,d)}$, possibly with $c = -\infty$ and/or $d = +\infty$. Set $S_n = Y_1 + \cdots + Y_n$. If the distribution of Y is infinitely divisible, then S_n converges mod-Y towards the limiting function $\psi \equiv 1$ with parameter $t_n = n$.

But there is another mod-convergence hidden in this framework (we now drop the assumption of infinite divisibility of the law of Y). The cumulant generating series of S_n is

$$\log \mathbb{E}[e^{zS_n}] = n \log \mathbb{E}[e^{zY}] = n \sum_{r=2}^{\infty} \frac{\kappa^{(r)}(Y)}{r!} z^r,$$

which is also analytic on $\mathcal{S}_{(c,d)}$ — the coefficients $\kappa^{(r)}(Y)$ are the *cumulants* of the variable Y (see [FW32]). Let $v \geq 3$ be an integer such that $\kappa^{(r)}(Y) = 0$ for each integer r strictly between 3 and $v - 1$, and set $X_n = \frac{S_n}{n^{1/v}}$. It is always possible to take $v = 3$, but sometimes we can also consider higher value of v, for instance $v = 4$ as soon as Y is a symmetric random random variable, and has therefore its odd moments and cumulants that vanish. One has

$$\log \varphi_n(z) = n^{\frac{v-2}{v}} \frac{\kappa^{(2)}(Y)}{2} z^2 + \frac{\kappa^{(v)}(Y)}{v!} z^v + \sum_{r=v+1}^{\infty} \frac{\kappa^{(r)}(Y)}{r!\, n^{\frac{r}{v}-1}} z^r,$$

and locally uniformly on $\mathbb{C}$ the right-most term is bounded by $\frac{C}{n^{1/v}}$. Consequently,

$$\psi_n(z) = \exp\left(-n^{\frac{v-2}{v}} \frac{\sigma^2 z^2}{2}\right) \varphi_n(z) \to \exp\left(\frac{\kappa^{(v)}(Y)}{v!} z^v\right) + O(n^{-1/v}),$$

that is, $(X_n)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense with parameters $t_n = \sigma^2 n^{\frac{v-2}{v}}$, speed $O(n^{-1/v})$ and limiting function $\psi(z) = \exp(\kappa^{(v)}(Y) z^v / v!)$. Note that this first example was used in [KNN15] to characterise the set of limiting functions in the setting of mod-ϕ convergence.

Through this article, we shall commonly rescale random variables in order to get estimates of fluctuations at different regimes. In order to avoid any confusion, we provide the reader with the following scheme, which details each possible scaling, and for each scaling, the regimes of fluctuations that can be deduced from the mod-ϕ convergence, as well as their scope. We also underline or frame the scalings and regimes that will be studied in this paper, and give references for the other kinds of fluctuations.

scaling	mod-convergence		regime of fluctuations
	t_n	ϕ	
S_n	n	γ	$\frac{S_n}{n}$ — large deviations (*cf.* [BR60, DZ98])
$\frac{S_n}{n^{1/v}}$	$\sigma^2 n^{1-\frac{2}{v}}$	$\mathcal{N}_{\mathbb{R}}(0,1)$	$\frac{S_n}{n^{1-1/(v+1)}}$ moderate deviations $\frac{S_n}{n^{1-1/v}}$ normality zone $\frac{S_n}{n^{1/2}}$ local limit theorem (*cf.* [DKN15, KN12, FMN16])
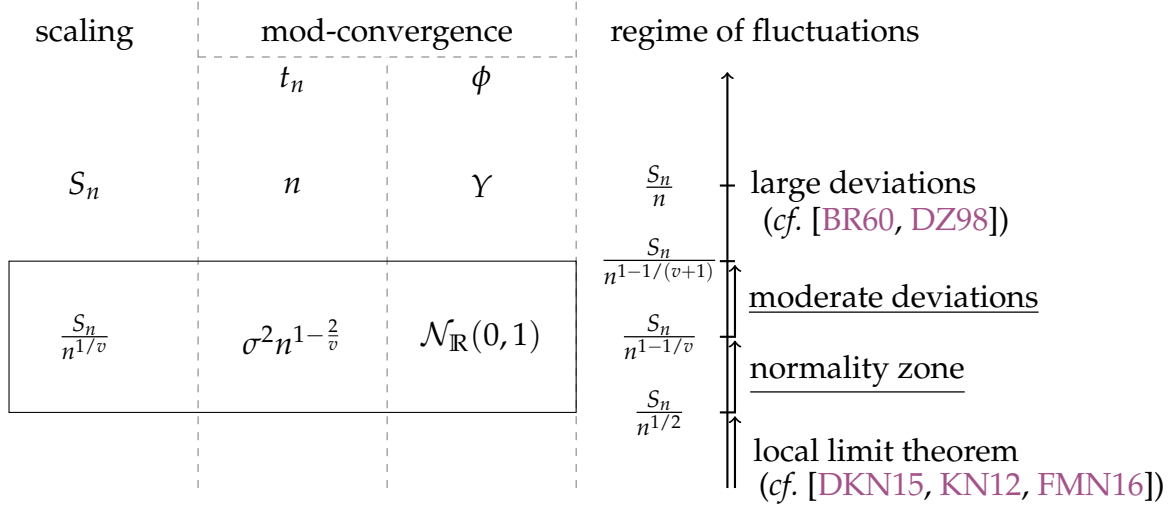
FIGURE 1. Panorama of the fluctuations of a sum of n i.i.d. random variables.

The content of this scheme will be fully explained in Chapter 4 (see in particular Section 4.4).

EXAMPLE 2.1.3. Denote X_n the number of disjoint cycles (including fixed points) of a random permutation chosen uniformly in the symmetric group $\mathfrak{S}(n)$. Feller's coupling (*cf.* [ABT03, Chapter 1]) shows that $X_n =_{\text{(law)}} \sum_{i=1}^{n} \mathcal{B}_{(1/i)}$, where $\mathcal{B}_p$ denotes a Bernoulli variable equal to 1 with probability p and to 0 with probability $1 - p$, and the Bernoulli variables are independent in the previous expansion. So,

$$\mathbb{E}[e^{zX_n}] = \prod_{i=1}^{n} \left(1 + \frac{e^z - 1}{i}\right) = e^{H_n(e^z - 1)} \prod_{i=1}^{n} \frac{1 + \frac{e^z - 1}{i}}{e^{\frac{e^z - 1}{i}}}$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i} = \log n + \gamma + O(n^{-1})$. The Weierstrass infinite product in the right-hand side converges locally uniformly to an entire function, therefore (see [WW27]),

$$\mathbb{E}[e^{zX_n}] e^{-(e^z - 1)\log n} \to e^{\gamma(e^z - 1)} \prod_{i=1}^{\infty} \frac{1 + \frac{e^z - 1}{i}}{e^{\frac{e^z - 1}{i}}} = \frac{1}{\Gamma(e^z)}$$

locally uniformly, *i.e.*, one has mod-Poisson convergence with parameters $t_n = \log n$ and limiting function $1/\Gamma(e^z)$. Moreover, the speed of convergence is a $O(n^{-1})$, hence, a $o((t_n)^{-v})$ for any integer v. We shall study generalisations of this example in Section 7.3.

scaling	mod-convergence		regime of fluctuations
	t_n	ϕ	
X_n	H_n	$\mathcal{P}(1)$	$\frac{X_n}{H_n} - 1$ — large deviations $(cf. \text{[Rad09]})$
$\frac{X_n - H_n}{(H_n)^{1/3}}$	$(H_n)^{1/3}$	$\mathcal{N}_{\mathbb{R}}(0,1)$	$\frac{X_n - H_n}{(H_n)^{3/4}}$ — moderate deviations; $\frac{X_n - H_n}{(H_n)^{2/3}}$ — normality zone; $\frac{X_n - H_n}{(H_n)^{1/2}}$ — local limit theorem
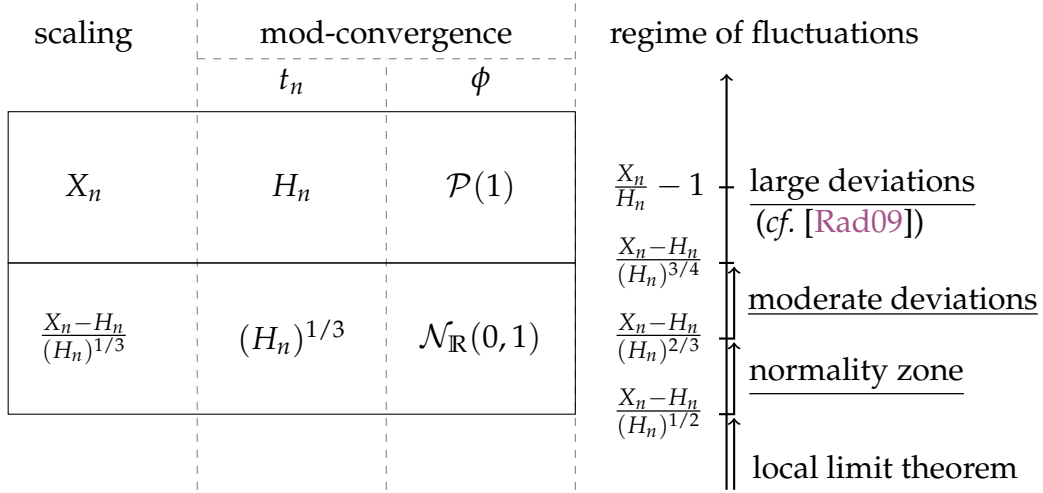
FIGURE 2. Panorama of the fluctuations of the number of cycles X_n of a random permutation of size n.

Once again, there is another mod-convergence hidden in this example. Indeed, consider $Y_n = \frac{X_n - H_n}{(H_n)^{1/3}}$. Its generating function has asymptotics

$$\mathbb{E}[e^{zY_n}] = e^{H_n\left(e^{\frac{z}{(H_n)^{1/3}}} - 1\right) - z(H_n)^{2/3}}(1 + o(1)) = e^{(H_n)^{1/3}\frac{z^2}{2}} \exp\left(\frac{z^3}{6}\right)(1 + o(1)).$$

Therefore, one has mod-Gaussian convergence of Y_n with parameters $t_n = (H_n)^{1/3}$ and limiting function $\exp(z^3/6)$.

This is in fact a particular case of a more general phenomenon: every sequence that converges mod-ϕ converges with a different rescaling in the mod-Gaussian sense.

PROPOSITION 2.1.4. *Assume X_n converges mod-ϕ with parameters t_n and limiting function ψ, where ϕ is not the Gaussian distribution. Let*

$$m = \min_{i \geq 3}\{i \mid \eta^{(i)}(0) \neq 0\}.$$

Then, the sequence of random variables $Y_n = (X_n - t_n\eta'(0))/(t_n)^{1/m}$ converges in the mod-Gaussian sense with parameters $(t_n)^{1-2/m}\eta''(0)$ towards the limiting function $\Psi(z) = \exp\left(\eta^{(m)}(0)z^m/m!\right)$.

PROOF. This follows from a simple computation

$$\mathbb{E}\left[\exp\left(\frac{z(X_n - t_n\eta'(0))}{(t_n)^{1/m}}\right)\right]$$
$$= \exp\left(\frac{-t_n\eta'(0))}{(t_n)^{1/m}}\right)\exp\left(t_n\,\eta\left(\frac{z}{(t_n)^{1/m}}\right)\right)\psi\left(\frac{z}{(t_n)^{1/m}}\right)(1 + o(1)).$$

The factor $\psi(\frac{z}{(t_n)^{1/m}})$ tends to 1 and we do a Taylor expansion of $\eta(\frac{z}{(t_n)^{1/m}})$. We get

$$\mathbb{E}\left[\exp\left(\frac{z(X_n - t_n\eta'(0))}{(t_n)^{1/m}}\right)\right] = \exp\left((t_n)^{1-2/m}\eta''(0)\frac{z^2}{2} + \eta^{(m)}(0)\frac{z^m}{m!} + o(1)\right)(1 + o(1)).$$

$\square$

Naturally, the mod-ϕ convergence gives more information than the implied mod-Gaussian convergence: our deviation results — Theorems 3.2.2 and 4.2.1 — for the former involve deviation probabilities of X_n at scale $O(t_n)$, while with the mod-Gaussian convergence, we get deviation probabilities of Y_n at scale $O((t_n)^{1-2/m})$, that is deviations of X_n at scale $O((t_n)^{1-1/m})$.

2.2. Legendre-Fenchel transforms

We now present the definition and some properties of the Legendre-Fenchel transform, a classical tool in large deviation theory (see *e.g.* [DZ98, Section 2.2]) that we shall use a lot in this paper. The Legendre-Fenchel transform is the following operation on (convex) functions:

DEFINITION 2.2.1. *The Legendre-Fenchel transform of a function η is defined by:*

$$F(x) = \sup_{h \in \mathbb{R}} (hx - \eta(h)).$$

This is an involution on convex lower semi-continuous functions.

Assume that η is the logarithm of the moment generating series of a random variable. In this case, η is a convex function (by Hölder's inequality). Then F is always non-negative, and the unique h maximizing $hx - \eta(h)$, if it exists, is then defined by the implicit equation $\eta'(h) = x$ (note that h depends on x, but we have chosen not to write $h(x)$ to make notation lighter). This implies the following useful identities:

$$F(x) = xh - \eta(h) \quad ; \quad F'(x) = h \quad ; \quad F''(x) = h'(x) = \frac{1}{\eta''(h)}.$$

EXAMPLE 2.2.2. If $\eta(z) = mz + \frac{\sigma^2 z^2}{2}$ (Gaussian variable with mean m and variance σ^2), then

$$h = \frac{x - m}{\sigma^2} \quad ; \quad F_{\mathcal{N}_\mathbb{R}(m,\sigma^2)}(x) = \frac{(x - m)^2}{2\sigma^2}$$

whereas if $\eta(z) = \lambda(e^z - 1)$ (Poisson law with parameter λ), then

$$h = \log \frac{x}{\lambda} \quad ; \quad F_{\mathcal{P}(\lambda)}(x) = \begin{cases} x \log \frac{x}{\lambda} - (x - \lambda) & \text{if } x > 0, \\ +\infty & \text{otherwise} \end{cases}.$$

FIGURE 3. The Legendre-Fenchel transforms of a Gaussian law and of a Poisson law.

2.3. Gaussian integrals

Some computations involving the Gaussian density are used several times throughout the paper, so we decided to present them together here.

LEMMA 2.3.1 (Gaussian integrals).

(1) moments:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{x^2}{2}} x^{2m} \, dx = (2m-1)!! = (2m-1)(2m-3)\cdots 3\,1,$$

and the odd moments vanish.

(2) Fourier transform: with $g(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$, one has

$$g^*(\zeta) = \int_{\mathbb{R}} g(x) \, e^{ix\zeta} \, dx = e^{-\frac{\zeta^2}{2}}.$$

More generally, with the Hermite polynomials $H_r(x) = (-1)^r e^{\frac{x^2}{2}} \frac{\partial^r}{\partial x^r}(e^{-\frac{x^2}{2}})$, one has

$$(g\,H_r)^*(\zeta) = (i\zeta)^r \, e^{-\frac{\zeta^2}{2}}.$$

(3) tails: if $a \to +\infty$, then

$$\int_0^\infty e^{-\frac{(y+a)^2}{2}} \, dy = \frac{e^{-\frac{a^2}{2}}}{a}\left(1 - \frac{1}{a^2} + O\left(\frac{1}{a^4}\right)\right)$$

$$\int_0^\infty y\, e^{-\frac{(y+a)^2}{2}} \, dy = \frac{e^{-\frac{a^2}{2}}}{a^2}\left(1 + O\left(\frac{1}{a^2}\right)\right)$$

$$\int_0^\infty y^2\, e^{-\frac{(y+a)^2}{2}} \, dy = O\left(\frac{e^{-\frac{a^2}{2}}}{a^3}\right)$$

$$\int_0^\infty y^3\, e^{-\frac{(y+a)^2}{2}} \, dy = O\left(\frac{e^{-\frac{a^2}{2}}}{a^2}\right)$$

In particular, the tail of the Gaussian distribution is $\frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-\frac{x^2}{2}} \, dx \simeq \frac{1}{a\sqrt{2\pi}} e^{-\frac{a^2}{2}}$.

(4) complex transform: for $\beta > 0$,

$$\int_{\mathbb{R}} \frac{e^{-\frac{\beta^2}{2}}}{2\pi} \frac{e^{-\frac{w^2}{2}}}{\beta + iw} \, dw = \int_\beta^\infty \frac{e^{-\frac{\alpha^2}{2}}}{\sqrt{2\pi}} \, d\alpha = \mathbb{P}[\mathcal{N}_{\mathbb{R}}(0,1) \geq \beta].$$

PROOF. Recall that the generating series of Hermite polynomials ([Sze75, Chapter 5]) is

$$\sum_{r=0}^\infty H_r(x) \frac{t^r}{r!} = e^{\frac{x^2}{2}} \sum_{r=0}^\infty \frac{(-t)^r}{r!} \frac{\partial^r}{\partial x^r}\left(e^{-\frac{x^2}{2}}\right) = e^{\frac{x^2}{2}} e^{-\frac{(x-t)^2}{2}} = e^{-\frac{t^2}{2}+tx}.$$

Integrating against $g(x)\,e^{ix\zeta}\,dx$ yields

$$\sum_{r=0}^{\infty}(g(x)\,H_r(x))^*(\zeta)\,\frac{t^r}{r!}=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-\frac{(x-t)^2}{2}+i\zeta x}\,dx$$

$$=\frac{e^{i\zeta t}}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-\frac{y^2}{2}+i\zeta y}\,dy=e^{i\zeta t-\frac{\zeta^2}{2}}=\sum_{r=0}^{\infty}(i\zeta)^r e^{-\frac{\zeta^2}{2}}\frac{t^r}{r!}$$

whence the identity (2) for Fourier transforms.

With $r=0$, one gets the Fourier transform of the Gaussian $g^*(\zeta)=e^{-\frac{\zeta^2}{2}}$, hence the moments (1) by derivation at $\zeta=0$. The estimate of tails (3) is obtained by an integration by parts; notice that similar techniques yield the tails of distributions $x^m\,e^{-x^2/2}\,dx$ with $m\geq 1$. Finally, as for the complex transform (4), remark that

$$F(\beta)=\int_{\mathbb{R}}\frac{e^{-\frac{\beta^2}{2}}}{2\pi}\frac{e^{-\frac{w^2}{2}}}{\beta+iw}\,dw=\frac{1}{2i\pi}\int_{\Gamma=\beta+i\mathbb{R}}\frac{e^{\frac{(z-\beta)^2-\beta^2}{2}}}{z}\,dz,$$

the second integral being along the complex curve $\Gamma=\beta+i\mathbb{R}$. By standard complex analysis arguments, this integral is the same along any line $\Gamma'=\beta'+i\mathbb{R}$ (for $\beta'>0$). Namely

$$F(\beta)=\frac{1}{2i\pi}\int_{\Gamma'=\beta'+i\mathbb{R}}\frac{e^{\frac{(z-\beta)^2-\beta^2}{2}}}{z}\,dz.$$

Since $\lim_{\beta\to+\infty}F(\beta)=0$,

$$F(\beta)=-\int_{\beta}^{\infty}F'(\alpha)\,d\alpha=\int_{\beta}^{\infty}\left(\frac{1}{2i\pi}\int_{\Gamma'=\beta'+i\mathbb{R}}e^{\frac{(z-\alpha)^2-\alpha^2}{2}}\,dz\right)d\alpha.$$

Again, the integration line Γ' in the second integral can be replaced by $\Gamma=\alpha+i\mathbb{R}$ and we get

$$F(\beta)=\int_{\beta}^{\infty}\left(\frac{1}{2i\pi}\int_{\Gamma=\alpha+i\mathbb{R}}e^{\frac{(z-\alpha)^2-\alpha^2}{2}}\,dz\right)d\alpha.=\int_{\beta}^{\infty}\frac{e^{-\frac{\alpha^2}{2}}}{\sqrt{2\pi}}\,d\alpha,$$

which is the tail $\mathbb{P}[\mathcal{N}_{\mathbb{R}}(0,1)\geq\beta]$ of a standard Gaussian law. $\qquad\square$

Also, there will be several instances of the Laplace method for asymptotics of integrals, but each time in a different setting; so we found it more convenient to reprove it each time.

CHAPTER 3

Fluctuations in the case of lattice distributions

3.1. Lattice and non-lattice distributions

If ϕ is an infinitely divisible distribution, recall that its characteristic function writes uniquely as

$$\int_{\mathbb{R}} e^{iux}\,\phi(dx) = \exp\left(imu - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}\setminus\{0\}} \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right)\Pi(dx)\right), \quad (3)$$

where Π is the Lévy measure of ϕ, and is required to integrate $1 \wedge x^2$ (see [Kal97, Chapter 13]). If $\sigma^2 > 0$, then ϕ has a normal component and its support

$$\mathrm{supp}(\phi) = \big(\text{smallest closed subset } S \text{ of } \mathbb{R} \text{ with } \phi(S) = 1\big)$$

is the whole real line, since ϕ can be seen as the convolution of some probability measure with a non-degenerate Gaussian law. Suppose now $\sigma^2 = 0$, and, set

$$\gamma = m - \int_{\mathbb{R}\setminus\{0\}} \frac{x}{1+x^2}\,\Pi(dx),$$

which is the shift parameter of ϕ; note the integral above is not always convergent, so that γ is not always defined.

LEMMA 3.1.1. *[SH04, Chapter 4, Theorem 8.4]*

(1) *If γ is well-defined and finite, and if $\Pi([-\varepsilon,\varepsilon] \setminus \{0\}) = 0$ for some $\varepsilon > 0$, then*

$$\mathrm{supp}(\phi) = \gamma + \overline{\mathbb{N}[\mathrm{supp}(\Pi)]},$$

where $\mathbb{N}[S]$ is the semigroup generated by a part S of $\mathbb{R}$ (the set of all sums of elements of S, including the empty sum 0), and $\overline{\mathbb{N}[S]}$ is its closure.

(2) *Otherwise, the support of ϕ is either $\mathbb{R}$, or the half-line $[\gamma, +\infty)$, or the half-line $(-\infty, \gamma]$.*

Recall that an additive subgroup of $\mathbb{R}$ is either discrete of type $\lambda\mathbb{Z}$ with $\lambda \geq 0$, or dense in $\mathbb{R}$. We call an infinitely divisible distribution *discrete*, or of type *lattice*, if $\sigma^2 = 0$, if γ is well-defined and finite, and if the subgroup $\mathbb{Z}[\mathrm{supp}(\Pi)]$ is discrete. Otherwise, we say that ϕ is a non-lattice infinitely divisible distribution.

PROPOSITION 3.1.2. *An infinitely divisible distribution ϕ is of type lattice if and only if one of the following equivalent assertions is satisfied:*

(1) *Its support is included in a set $\gamma + \lambda\mathbb{Z}$ for some parameters γ and $\lambda > 0$.*

(2) *For some parameter $\lambda > 0$, the characteristic function has modulus $|\phi(e^{iux})| = 1$ if and only if $u \in \frac{2\pi}{\lambda}\mathbb{Z}$.*

If both hold and if λ is chosen maximal in (1), then the parameters λ in (1) and (2) coincide.

Moreover, an infinitely divisible distribution ϕ is of type non-lattice if and only if $|\phi(e^{iux})| < 1$ for all $u \neq 0$.

PROOF. In the following we exclude the case of a degenerate Dirac distribution $\phi = \delta_\gamma$, which is trivial. We can also assume that $\sigma^2 = 0$: otherwise, ϕ is of type non-lattice and with support $\mathbb{R}$, and the inequality $|\phi(e^{iux})| < 1$ for $u \neq 0$ is true for any non-degenerate Gaussian law, and therefore by convolution for every infinitely divisible law with parameter $\sigma^2 \neq 0$.

Suppose ϕ of type lattice. Then, since $\mathbb{Z}[\operatorname{supp}(\Pi)] = \lambda\mathbb{Z}$ for some $\lambda > 0$, the semi-group $\mathbb{N}[\operatorname{supp}(\Pi)] \subset \lambda\mathbb{Z}$ is discrete, and hence closed. It thus follows from Lemma 3.1.1 that
$$\operatorname{supp}(\phi) = \gamma + \mathbb{N}[\operatorname{supp}(\Pi)] \subset \gamma + \lambda\mathbb{Z}.$$
Conversely, if $\operatorname{supp}(\phi)$ is included in a shifted lattice $\gamma + \lambda\mathbb{Z}$, then the second case of Lemma 3.1.1 is excluded, so γ is well-defined and finite, and then
$$\operatorname{supp}(\phi) = \gamma + \overline{\mathbb{N}[\operatorname{supp}(\Pi)]}.$$
But $\operatorname{supp}(\phi) \subset \gamma + \lambda\mathbb{Z}$, so this forces $\mathbb{N}[\operatorname{supp}(\Pi)] \subset \lambda\mathbb{Z}$, and therefore $\mathbb{Z}[\operatorname{supp}(\Pi)] \subset \lambda\mathbb{Z}$. Hence, ϕ is of type lattice. We have proved that the first assertion is indeed equivalent to the definition of a lattice infinitely divisible distribution.

The equivalence of the two assertions (1) and (2) is then a general fact on probability measures ϕ on the real line. If ϕ is such a measure, let $G_\phi = \{u \in \mathbb{R} \mid |\phi(e^{iux})| = 1\}$. We claim that G_ϕ is an additive subgroup of $\mathbb{R}$. Indeed, if $u \neq 0$, then

$$u \in G_\phi \iff \left| \int_{\mathbb{R}} e^{iux} \phi(dx) \right| = 1$$

$$\iff \text{ the phase of } e^{iux} \text{ is constant } \phi\text{-almost surely}$$

$$\iff \phi \text{ is supported on a set } \gamma_u + \frac{2\pi}{u}\mathbb{Z}.$$

Suppose that $u \neq 0$ and $v \neq 0$ belong to G_ϕ. Then,

$$\operatorname{supp}(\phi) \subset \left(\gamma_u + \frac{2\pi\mathbb{Z}}{u} \right) \cap \left(\gamma_v + \frac{2\pi\mathbb{Z}}{v} \right),$$

and the right-hand side of this formula is again a (shifted) discrete subgroup $\gamma_w + \frac{2\pi\mathbb{Z}}{w}$, with $\frac{k}{u} = \frac{l}{v} = \frac{1}{w}$ for some non-zero integers k and l. In particular,

$$(k+l)w = kw + lw = u + v \quad ; \quad \frac{1}{w} = \frac{k+l}{u+v} \quad ; \quad \operatorname{supp}(\phi) \subset \gamma_w + \frac{2\pi\mathbb{Z}}{w} \subset \gamma_w + \frac{2\pi\mathbb{Z}}{u+v},$$

so $u + v \in G_\phi$ and G_ϕ is indeed a subgroup of $\mathbb{R}$.

If G_ϕ is discrete and writes as $p\mathbb{Z}$ with $p > 0$, then ϕ is supported on a lattice $\gamma + \lambda\mathbb{Z}$ with $\lambda = \frac{2\pi}{p}$, and $|\phi(e^{iu})| = 1$ if and only if $u \in \frac{2\pi\mathbb{Z}}{\lambda}$. Otherwise, G_ϕ cannot be a dense subgroup of $\mathbb{R}$, because then by continuity of $u \mapsto \phi(e^{iu})$, we would have $G_\phi = \mathbb{R}$, which implies that ϕ is a Dirac, and this case has been excluded. So, the only other possibility is $G_\phi = 0$, which is the last statement of the proposition. $\qquad\square$

In the remaining of this chapter, we place ourselves in the setting of Definition 1.1.1, and we suppose that the X_n's and the (non-constant) infinitely divisible distribution ϕ both take values in the lattice $\mathbb{Z}$, and furthermore, that ϕ has period 2π (in other words, the lattice $\mathbb{Z}$ is *minimal* for ϕ). In particular, for every $u \in (0, 2\pi)$, $|\exp(\eta(\mathrm{i}u))| < 1$, since by the previous discussion the period of the characteristic function of a $\mathbb{Z}$-valued infinitely divisible distribution is also the smallest $u > 0$ such that $|\phi(\mathrm{e}^{\mathrm{i}ux})| = 1$. For more details on (discrete) infinitely-divisible distributions, we refer to the aforementioned textbook [SH04], and also to [Kat67] and [Fel71, Chapter XVII].

3.2. Deviations at the scale $O(t_n)$

LEMMA 3.2.1. *Let X be a $\mathbb{Z}$-valued random variable whose generating function $\varphi_X(z) = \mathbb{E}[\mathrm{e}^{zX}]$ converges absolutely in the strip $\mathcal{S}_{(c,d)}$, with $c < 0 < d$. For $k \in \mathbb{Z}$,*

$$\forall h \in (c, d), \ \ \mathbb{P}[X = k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{-k(h+\mathrm{i}u)} \, \varphi_X(h + \mathrm{i}u) \, du;$$

$$\forall h \in (0, d), \ \ \mathbb{P}[X \geq k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathrm{e}^{-k(h+\mathrm{i}u)}}{1 - \mathrm{e}^{-(h+\mathrm{i}u)}} \, \varphi_X(h + \mathrm{i}u) \, du.$$

PROOF. Since

$$\varphi_X(h + \mathrm{i}u) = \sum_{k \in \mathbb{Z}} \mathbb{P}[X = k] \, \mathrm{e}^{k(h+\mathrm{i}u)},$$

$\mathbb{P}[X = k] \, \mathrm{e}^{kh}$ is the k-th Fourier coefficient of the 2π-periodic and smooth function $u \mapsto \varphi_X(h + \mathrm{i}u)$; this leads to the first formula. Then, assuming also $h > 0$,

$$\mathbb{P}[X \geq k] = \sum_{l=k}^{\infty} \mathbb{P}[X = l] = \sum_{l=k}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{-l(h+\mathrm{i}u)} \, \varphi_X(h + \mathrm{i}u) \, du,$$

and the sum of the moduli of the functions on the right-hand side is dominated by the integrable function $\frac{\mathrm{e}^{-kh}}{1-\mathrm{e}^{-h}} \, \varphi_X(h)$; so by Lebesgue's dominated convergence theorem, one can exchange the integral and the summation symbol, which yields the second equation. □

We now work under the assumptions of Definition 1.1.1, with a lattice infinitely divisible distribution ϕ. Furthermore, we assume that the convergence is at speed $O((t_n)^{-v})$, on a strip $\mathcal{S}_{(c,d)}$ containing 0. Note that necessarily $\eta(0) = 0$ and $\psi(0) = 1$. A simple computation gives also the following approximation formulas:

$$\mathbb{E}(X_n) = \varphi'_n(0) = t_n \eta'(0) + \psi'(0) \sim t_n \eta'(0) = t_n \eta'(0) + O(1);$$

$$\mathrm{Var}(X_n) = \varphi''_n(0) - (\varphi'_n(0))^2 = t_n \eta''(0) + O(1).$$

THEOREM 3.2.2. *Let x be a real number in the interval $(\eta'(c), \eta'(d))$, and h defined by the implicit equation $\eta'(h) = x$. We assume $t_n x \in \mathbb{N}$.*

(1) The following expansion holds:

$$\mathbb{P}[X_n = t_n x] = \frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta''(h)}} \left(\psi(h) + \frac{a_1}{t_n} + \frac{a_2}{(t_n)^2} + \cdots + \frac{a_{v-1}}{(t_n)^{v-1}} + O\left(\frac{1}{(t_n)^v}\right) \right)$$

$$= \exp(-t_n F(x)) \sqrt{\frac{F''(x)}{2\pi t_n}} \left(\psi(F'(x)) + \frac{a_1}{t_n} + \cdots + \frac{a_{v-1}}{(t_n)^{v-1}} + O\left(\frac{1}{(t_n)^v}\right) \right),$$

for some numbers a_k.

(2) Similarly, if x is a real number in the range of $\eta'_{|(0,d)}$, then

$$\mathbb{P}[X_n \geq t_n x] = \frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta''(h)}} \frac{1}{1 - e^{-h}} \left(\psi(h) + \frac{b_1}{t_n} + \cdots + \frac{b_{v-1}}{(t_n)^{v-1}} + O\left(\frac{1}{(t_n)^v}\right) \right),$$

for some numbers b_k.

Both a_k and b_k are rational fractions in the derivatives of η and ψ at h, that can be computed explicitly — see Remark 3.2.5.

PROOF. With the notations of Definition 1.1.1, the first equation of Lemma 3.2.1 becomes

$$\mathbb{P}[X_n = t_n x] = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-t_n x\,(h+iu)}\, \varphi_n(h + iu)\, du$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-t_n x h}\, e^{t_n(\eta(h+iu) - iux)}\, \psi_n(h + iu)\, du$$

$$= \frac{e^{-t_n F(x)}}{2\pi} \int_{-\pi}^{\pi} e^{t_n(\eta(h+iu) - \eta(h) - iu\eta'(h))}\, \psi_n(h + iu)\, du. \qquad (4)$$

The last equality uses the facts that $xh = F(x) + \eta(h)$ and $x = \eta'(h)$. We perform the Laplace method on (4), and to this purpose we split the integral in two parts. Fix $\delta > 0$, and denote $q_\delta = \max_{u \in (-\pi,\pi) \setminus (-\delta,\delta)} |\exp(\eta(h + iu) - \eta(h))|$. This is strictly smaller than 1, since

$$\exp(\eta(h + iu) - \eta(h)) = \frac{\mathbb{E}[e^{(h+iu)X}]}{\mathbb{E}[e^{hX}]} = \mathbb{E}_{\mathbb{Q}}[e^{iuX}]$$

is the characteristic function of X under the new probability $d\mathbb{Q}(\omega) = \frac{e^{hX(\omega)}}{\mathbb{E}[e^{hX}]} d\mathbb{P}(\omega)$ (and X has minimum lattice $\mathbb{Z}$). Note that Lemma 3.3.3 hereafter is a more precise version of this inequality.

As a consequence, if $I_{(-\delta,\delta)}$ and $I_{(-\delta,\delta)^c}$ denote the two parts of (4) corresponding to $\int_{-\delta}^{\delta}$ and $\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}$, then

$$|I_{(-\delta,\delta)^c}| \leq \frac{e^{-t_n F(x)}}{2\pi} \int_{(-\delta,\delta)^c} (q_\delta)^{t_n} |\psi_n(h + iu)|\, du \leq 2\,(e^{-F(x)} q_\delta)^{t_n} \max_{u \in (-\pi,\pi)} |\psi(h + iu)|$$

for n big enough, since ψ_n converges uniformly towards ψ on the compact set $K = h + i[-\pi, \pi]$. Since $q_\delta < 1$, for any $\delta > 0$ fixed, $I_{(-\delta,\delta)^c}\, e^{t_n F(x)}$ goes to 0 faster than any negative power of t_n, so $I_{(-\delta,\delta)^c}$ is negligible in the asymptotics (recall that $F(x)$ is non-negative by definition, as $\eta(0) = 0$).

As for the other part, we can first replace ψ_n by ψ up to a $(1 + O((t_n)^{-v}))$, since the integral is taken on a compact subset of $\mathcal{S}_{(c,d)}$. We then set $u = \frac{w}{\sqrt{t_n \eta''(h)}}$:

$$I_{(-\delta,\delta)} = \frac{e^{-t_n F(x)}\left(1 + O\left(\frac{1}{(t_n)^v}\right)\right)}{2\pi \sqrt{t_n \eta''(h)}} \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} \psi\left(h + \frac{iw}{\sqrt{t_n\eta''(h)}}\right) e^{t_n \Delta_n(w) - \frac{w^2}{2}} \, dw, \quad (5)$$

where $\Delta_n(w)$ is the Taylor expansion

$$\eta(h + iu) - \eta(h) - \eta'(h)(iu) - \frac{\eta''(h)}{2}(iu)^2$$

$$= \sum_{k=3}^{2v+1} \frac{\eta^{(k)}(h)}{k!}\left(\frac{iw}{\sqrt{t_n\eta''(h)}}\right)^k + O\left(\frac{1}{(t_n)^{v+1}}\right)$$

$$= \frac{1}{t_n}\left(-\frac{w^2}{\eta''(h)}\sum_{k=1}^{2v-1}\frac{\eta^{(k+2)}(h)}{(k+2)!}\left(\frac{iw}{\sqrt{t_n\eta''(h)}}\right)^k + O\left(\frac{1}{(t_n)^v}\right)\right).$$

We also replace ψ by its Taylor expansion

$$\psi\left(h + \frac{iw}{\sqrt{t_n\eta''(h)}}\right) = \sum_{k=0}^{2v-1}\frac{\psi^{(k)}(h)}{k!}\left(\frac{iw}{\sqrt{t_n\eta''(h)}}\right)^k + O\left(\frac{1}{(t_n)^v}\right).$$

Thus, if one defines α_k by the equation

$$f_n(w) := \left(\sum_{k=0}^{2v-1}\frac{\psi^{(k)}(h)}{k!}\left(\frac{iw}{\sqrt{t_n\eta''(h)}}\right)^k\right)\exp\left(-\frac{w^2}{\eta''(h)}\sum_{k=1}^{2v-1}\frac{\eta^{(k+2)}(h)}{(k+2)!}\left(\frac{iw}{\sqrt{t_n\eta''(h)}}\right)^k\right)$$

$$= \sum_{k=0}^{2v-1}\frac{\alpha_k(w)}{(t_n)^{k/2}} + O\left(\frac{1}{(t_n)^v}\right),$$

then one can replace $\psi(h + iu)\, e^{t_n \Delta_n(w)}$ by $f_n(w)$ in Equation (5). Moreover, observe that each coefficient $\alpha_k(w)$ writes as

$$\alpha_k(w) = \alpha_{k,0}(h)\left(\frac{w}{\sqrt{\eta''(h)}}\right)^k + \alpha_{k,1}(h)\left(\frac{w}{\sqrt{\eta''(h)}}\right)^{k+2} + \cdots + \alpha_{k,r}(h)\left(\frac{w}{\sqrt{\eta''(h)}}\right)^{k+2r}$$

with the $\alpha_{k,r}(h)$'s polynomials in the derivatives of ψ and η at point h. So,

$$I_{(-\delta,\delta)} = \left(1 + O\left(\frac{1}{(t_n)^v}\right)\right)\frac{e^{-t_n F(x)}}{\sqrt{2\pi t_n \eta''(h)}}\left(\sum_{k=0}^{2v-1}\int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}}\frac{\alpha_k(w)}{(t_n)^{k/2}}\frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw\right).$$

For any power w^m,

$$\left|\int_{-\infty}^{\infty} w^m \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw - \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} w^m \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\, dw\right|$$

is smaller than any negative power of t_n as n goes to infinity (see Lemma 2.3.1, (3) for the case $m = 0$): indeed, by integration by parts, one can expand the difference as $e^{-\delta^2 t_n \eta''(h)/2} R_m(\sqrt{t_n})$, where R_m is a rational fraction that depends on m, h, δ and on the order of the expansion needed. Therefore, one can take the full integrals in the

previous formula. On the other hand, the odd moments of the Gaussian distribution vanish. One concludes that

$$\mathbb{P}[X_n = t_n x] = \frac{\mathrm{e}^{-t_n F(x)}}{\sqrt{2\pi t_n \eta''(h)}} \left(\sum_{k=0}^{v-1} \frac{1}{(t_n)^k} \left(\int_{\mathbb{R}} \alpha_{2k}(w) \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}} \, dw \right) + O\left(\frac{1}{(t_n)^v} \right) \right),$$

and each integral $\int_{\mathbb{R}} \alpha_{2k}(w) \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{2\pi}} \, dw$ is equal to

$$\frac{\alpha_{2k,0}(h) \, (2k-1)!!}{(\eta''(h))^k} + \cdots + \frac{\alpha_{2k,r}(h) \, (2k+2r-1)!!}{(\eta''(h))^{k+r}}$$

where $(2m-1)!!$ is the $2m$-th moment of the Gaussian distribution (*cf.* Lemma 2.3.1, (1)). This ends the proof of the first part of our Theorem, the second formula coming from the identities $h = F'(x)$ and $\eta''(h) = \frac{1}{F''(x)}$. The second part is exactly the same, up to the factor

$$\frac{1}{1 - \mathrm{e}^{-h-iu}} = \frac{1}{1 - \mathrm{e}^{-h}} \left(\frac{1 - \mathrm{e}^{-h}}{1 - \mathrm{e}^{-h-\frac{iw}{\sqrt{t_n \eta''(h)}}}} \right)$$

in the integrals. $\square$

REMARK 3.2.3. For $x > \eta'(0)$, the first term of the expansion

$$\frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta''(h)}}$$

is the leading term in the asymptotics of $\mathbb{P}[Y_{t_n} = t_n x]$, where $(Y_t)_{t \in \mathbb{R}_+}$ is the Lévy process associated to the analytic function $\eta(z)$. Thus, the residue ψ measures the difference between the distribution of X_n and the distribution of Y_{t_n} in the interval $(t_n \eta'(0), t_n \eta'(d))$.

REMARK 3.2.4. If the convergence is faster than any negative power of t_n, then one can simplify the statement of the theorem as follows: as formal power series in t_n,

$$\sqrt{2\pi t_n \eta''(h)} \, \exp(t_n F(x)) \, \mathbb{P}[X_n = t_n x] = \int_{\mathbb{R}} f_n(w) \, \mathrm{e}^{-\frac{w^2}{2}} \, dw,$$

i.e., the expansions of both sides up to any given power $O\left(\frac{1}{(t_n)^v} \right)$ agree.

REMARK 3.2.5. As mentioned in the statement of the theorem, the proof also gives an algorithm to obtain formulas for a_k and b_k. More precisely, denote

$$\Delta_n(w) = t_n \left(\eta\left(h + \frac{iw}{\sqrt{t_n \eta''(h)}} \right) - \eta(h) - \eta'(h) \frac{iw}{\sqrt{t_n \eta''(h)}} + \frac{w^2}{2t_n} \right)$$

$$f_n(w) = \psi\left(h + \frac{iw}{\sqrt{t_n \eta''(h)}} \right) \exp(t_n \Delta_n(w)) = \sum_{k=0}^{\infty} \frac{\alpha_k(w)}{(t_n)^{k/2}},$$

the last expansion holding in a neighbourhood of zero. The coefficient $\alpha_{2k}(w)$ is an even polynomial in w with valuation $2k$ and coefficients which are polynomials in the

derivatives of ψ and η at h, and in $\frac{1}{\eta''(h)}$. Then,

$$a_k = \int_{\mathbb{R}} \alpha_{2k}(w) \, \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}} \, dw,$$

and in particular,

$$a_0 = \psi(h);$$

$$a_1 = -\frac{1}{2}\frac{\psi''(h)}{\eta''(h)} + \frac{1}{24}\frac{\psi(h)\,\eta^{(4)}(h) + 4\,\psi'(h)\,\eta^{(3)}(h)}{(\eta''(h))^2} - \frac{15}{72}\frac{\psi(h)\,(\eta^{(3)}(h))^2}{(\eta''(h))^3}.$$

the b_k's are obtained by the same recipe as the a_k's, but starting from the power series

$$g_n(w) = \frac{1 - \exp(-h)}{1 - \exp\left(-h - \frac{iw}{\sqrt{t_n\eta''(h)}}\right)} \, f_n(w).$$

EXAMPLE 3.2.6. Suppose that $(X_n)_{n\in\mathbb{N}}$ is mod-Poisson convergent, that is to say that $\eta(z) = e^z - 1$. The expansion of Theorem 3.2.2 reads then as follows:

$$\mathbb{P}[X_n = t_n x] = \frac{e^{t_n(x-1-x\log x)}}{\sqrt{2\pi x t_n}}\left(\psi(h) + \frac{\psi'(h) - 3\psi''(h) - \psi(h)}{6x t_n} + O\left(\frac{1}{(t_n)^2}\right)\right)$$

with $h = \log x$. For instance, if X_n is the number of cycles of a random permutation in $\mathfrak{S}(n)$, then the discussion of Example 2.1.3 shows that for $x > 0$ such that $x \log n \in \mathbb{N}$,

$$\mathbb{P}[X_n = x(\log n)] = \frac{n^{-(x\log x - x + 1)}}{\sqrt{2\pi x \log n}} \frac{1}{\Gamma(x)}\left(1 + O(1/\log n)\right).$$

Similarly, for $x > 1$ such that $x \log n \in \mathbb{N}$, one has

$$\mathbb{P}[X_n \geq x(\log n)] = \frac{n^{-(x\log x - x + 1)}}{\sqrt{2\pi x \log n}} \frac{x}{x-1} \frac{1}{\Gamma(x)}\left(1 + O(1/\log n)\right).$$

As the speed of convergence is very good in this case, precise expansions in $1/\log n$ to any order could also be given.

3.3. Central limit theorem at the scales $o(t_n)$ and $o((t_n)^{2/3})$

The previous paragraph has described in the lattice case the fluctuations of $(X_n)_{n\in\mathbb{N}}$ in the regime $O(t_n)$, with a result akin to large deviations. In this section, we establish in the same setting an extended central limit theorem, for fluctuations of order up to $o(t_n)$. In particular, for fluctuations of order $o((t_n)^{2/3})$, we obtain the usual central limit theorem. Hence, we describe the panorama of fluctuations drawn on Figure 1.

THEOREM 3.3.1. *Consider a sequence* $(X_n)_{n\in\mathbb{N}}$ *that converges mod-ϕ, with a reference infinitely divisible law ϕ that is a lattice distribution. Assume* $y = o((t_n)^{1/6})$. *Then,*

$$\mathbb{P}\left[X_n \geq t_n\eta'(0) + \sqrt{t_n\eta''(0)}\, y\right] = \mathbb{P}[\mathcal{N}_{\mathbb{R}}(0,1) \geq y]\,(1 + o(1)).$$

order of fluctuations

$O(t_n)$ — large deviations ($\eta'(0) < x$): $\mathbb{P}[\frac{X_n}{t_n} \geq x] \simeq \frac{\exp(-t_n F(x))}{\sqrt{2\pi t_n \eta'(x)}} \frac{1}{1-e^{-F'(x)}} \psi(F'(x))$;

extended central limit theorem ($(t_n)^{1/6} \lesssim y \ll (t_n)^{1/2}$): $\mathbb{P}[\frac{X_n - t_n\eta'(0)}{\sqrt{t_n\,\eta''(0)}} \geq y] \simeq \frac{\exp(-t_n F(x))}{F'(x)\sqrt{2\pi t_n\eta'(x)}}$;

$O((t_n)^{2/3})$

central limit theorem ($y \ll (t_n)^{1/6}$):

$$\mathbb{P}[\frac{X_n - t_n\eta'(0)}{\sqrt{t_n\,\eta''(0)}} \geq y] \simeq \mathbb{P}[\mathcal{N}_{\mathbb{R}}(0,1) \geq y].$$
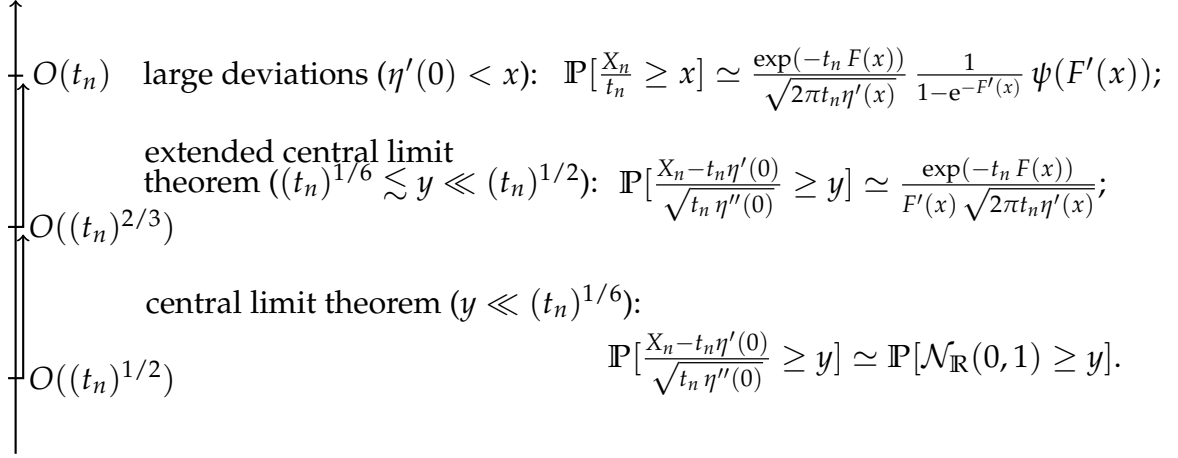
$O((t_n)^{1/2})$

FIGURE 1. Panorama of the fluctuations of a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ that converges modulo a lattice distribution (with $x = \eta'(0) + \sqrt{\eta''(0)/t_n}\, y$).

On the other hand, assuming $y \gg 1$ and $y = o((t_n)^{1/2})$, if $x = \eta'(0) + \sqrt{\eta''(0)/t_n}\, y$ and h is the solution of $\eta'(h) = x$, then

$$\mathbb{P}\left[X_n \geq t_n\eta'(0) + \sqrt{t_n\eta''(0)}\, y\right] = \frac{e^{-t_n F(x)}}{h\sqrt{2\pi t_n\,\eta''(h)}}(1 + o(1));$$

$$= \frac{e^{-t_n F(x)}}{y\sqrt{2\pi}}(1 + o(1)). \tag{6}$$

REMARK 3.3.2. The case $y = O(1)$, which is the classical central limit theorem, follows immediately from the assumptions of Definition 1.1.1, since by a Taylor expansion around 0 of η the characteristic functions of the rescaled r.v.

$$Y_n = \frac{X_n - t_n\eta'(0)}{\sqrt{t_n\eta''(0)}}$$

converge pointwise to $e^{-\frac{\zeta^2}{2}}$, the characteristic function of the standard Gaussian distribution. In the first statement, the improvement here is the weaker assumption $y = o((t_n)^{1/6})$.

As we shall see, the ingredients of the proof are very similar to the ones in the previous paragraph. We start with a technical lemma of control of the module of the Fourier transform of the reference law ϕ.

LEMMA 3.3.3. *Consider a non-constant infinitely divisible law ϕ, of type lattice, and with convergent moment generating function $\int e^{zx}\,\phi(dx) = e^{\eta(z)}$ on a strip $\mathcal{S}_{(c,d)}$ with $c < 0 < d$. We assume without loss of generality that ϕ has minimal lattice $\mathbb{Z}$. Then, there exists a constant $D > 0$ only depending on ϕ, and an interval $(-\varepsilon, \varepsilon) \subset (c, d)$, such that for all $h \in (-\varepsilon, \varepsilon)$ and all δ small enough,*

$$q_\delta = \max_{u\in[-\pi,\pi]\setminus(-\delta,\delta)} |\exp(\eta(h + iu) - \eta(h))|$$

is smaller than $1 - D\,\delta^2$.

PROOF. Denote X a random variable under the infinitely divisible distribution ϕ. We claim that there exist two consecutive integers n and $m = n - 1$ with $\mathbb{P}[X = n] \neq 0$ and $\mathbb{P}[X = m] \neq 0$. Indeed, under our hypotheses, if Π is the Lévy measure of ϕ, then

$$\mathbb{Z} = \mathbb{Z}[\mathrm{supp}(\Pi)] = \mathbb{N}[\mathrm{supp}(\Pi)] - \mathbb{N}[\mathrm{supp}(\Pi)],$$

so there exist a and b in $\mathbb{N}[\mathrm{supp}(\Pi)]$ such that $b - a = 1$. However, $\mathrm{supp}(\phi) = \gamma + \mathbb{N}[\mathrm{supp}(\Pi)]$ for some $\gamma \in \mathbb{Z}$, so $n = \gamma + b$ and $m = \gamma + a$ satisfy the claim.

Now, we have seen that $\exp(\eta(h + iu) - \eta(h))$ can be interpreted as the characteristic function of X under the new probability measure $dQ = \frac{e^{hX}}{\mathbb{E}[e^{hX}]}\,d\mathbb{P}$. So, for any u,

$$|\exp(\eta(h + iu) - \eta(h))|^2 = \left|\mathbb{E}_Q[e^{iuX}]\right|^2 = \sum_{n,m \in \mathbb{Z}} Q[X = n]\,Q[X = m]\,e^{iu(n-m)}$$

$$= \sum_{k \in \mathbb{Z}} \left(\sum_{n-m=k} Q[X = n]\,Q[X = m] \right) \cos ku.$$

Fix two integers n and $m = n - 1$ such that $\mathbb{P}[X = n] \neq 0$ and $\mathbb{P}[X = m] \neq 0$. Then one also has $Q[X = n] \neq 0$, $Q[X = m] \neq 0$, and there exists $D > 0$ such that

$$Q[X = n]\,Q[X = m] \geq 15\,D > 0$$

for h small enough (Q tends to $\mathbb{P}$ for $h \to 0$). As $\cos u \leq 1 - \frac{u^2}{5}$ for all $u \in (-\pi, \pi)$,

$$|\exp(\eta(h + iu) - \eta(h))|^2 \leq 1 + 15\,D\,(\cos u - 1) \leq 1 - 3\,D\,u^2;$$

$$q_\delta \leq \sqrt{1 - 3\,D\,\delta^2} \leq 1 - D\,\delta^2 \text{ for } \delta \text{ small enough.}$$

This concludes the proof of the Lemma.   □

PROOF OF THEOREM 3.3.1. Notice that $\eta''(0) \neq 0$ since this is the variance of the law ϕ, assumed to be non-trivial. Set $x = \eta'(0) + s$, and assume $s = o(1)$. The analogue of Equation (4) reads in our setting

$$\mathbb{P}[X_n \geq t_n x] = \frac{e^{-t_n F(x)}}{2\pi} \int_{-\pi}^{\pi} \frac{e^{t_n(\eta(h+iu) - \eta(h) - iu\eta'(h))}}{1 - e^{-h-iu}}\,\psi_n(h + iu)\,du. \tag{7}$$

Since $h'(x) = \frac{1}{\eta''(x)}$, one has $h = \frac{s}{\eta''(0)} + O(s^2)$. The same argument as in the proof of Theorem 3.2.2 shows that the integral over $(-\delta, \delta)^c$ is bounded by $C\,\delta\,(q_\delta)^{t_n}$, where $q_\delta < 1$, and $C\,\delta$ (with C a constant independent from s and δ) comes from the computation of

$$\max_{u \in (-\delta, \delta)^c} \left| \frac{\psi(h + iu)}{1 - e^{-h-iu}} \right|.$$

In the following we shall need to make δ go to zero sufficiently fast, but with $\delta\sqrt{t_n\eta''(0)}$ still going to infinity. Thus, set $\delta = (t_n)^{-2/5}$, so that in particular $(t_n)^{-1/2} \ll \delta \ll (t_n)^{-1/3}$. Notice that $I_{(-\delta, \delta)^c}\,e^{t_n F(x)}$ still goes to zero faster than any power of t_n; indeed,

$$(q_\delta)^{t_n} \leq \left(1 - \frac{D}{(t_n)^{4/5}}\right)^{t_n} \leq e^{-D\,(t_n)^{1/5}}$$

by Lemma 3.3.3. The other part of (7) is

$$
\frac{\mathrm{e}^{-t_n F(x)}}{2\pi\sqrt{t_n\eta''(h)}} \int_{-\delta\sqrt{t_n\eta''(h)}}^{\delta\sqrt{t_n\eta''(h)}} \psi\left(h + \frac{iw}{\sqrt{t_n\eta''(h)}}\right) \mathrm{e}^{t_n\Delta_n(w)} \frac{\mathrm{e}^{-\frac{w^2}{2}}}{1 - \mathrm{e}^{-h-\frac{iw}{\sqrt{t_n\eta''(h)}}}}\, dw,
$$

up to a factor $(1 + o(1))$. Let us analyze each part of the integral:

- The difference between $\psi\left(h + \frac{iw}{\sqrt{t_n\eta''(h)}}\right)$ and $\psi(0)$ is bounded by

$$
\max_{z\in[-s,s]+\mathrm{i}[-\delta,\delta]} |\psi(z) - \psi(0)| = o(1)
$$

  by continuity of ψ, so one can replace the term with ψ by the constant $\psi(0) = 1$, up to factor $(1 + o(1))$.

- The term $\Delta_n(w)$ has for Taylor expansion

$$
\frac{\eta^{(3)}(h)}{6}\left(\frac{iw}{\sqrt{t_n\eta''(h)}}\right)^3 + O\left(\frac{1}{(t_n)^2}\right),
$$

  so $t_n\Delta_n(w)$ is bounded by a $O(t_n\delta^3)$, which is a $o(1)$ since $\delta \ll (t_n)^{-1/3}$. So again one can replace $\mathrm{e}^{t_n\Delta_n(w)}$ by the constant 1.

- The Taylor expansion of $\left(1 - \mathrm{e}^{-h-\frac{iw}{\sqrt{t_n\eta''(h)}}}\right)^{-1}$ is $\frac{1}{h+\frac{iw}{\sqrt{t_n\eta''(h)}}}(1 + o(1))$. Hence,

$$
\mathbb{P}\left[X_n \geq t_n(\eta'(0) + s)\right] = \frac{\mathrm{e}^{-t_n F(x)}}{2\pi}\left(\int_{\mathbb{R}} \frac{\mathrm{e}^{-\frac{w^2}{2}}}{\sqrt{t_n\eta''(h)}\, h + iw}\, dw\right)(1 + o(1))
$$

$$
= \mathrm{e}^{-t_n F(x) + \frac{h^2\, t_n\,\eta''(h)}{2}}\, \mathbb{P}\left[\mathcal{N}_{\mathbb{R}}(0,1) \geq h\sqrt{t_n\eta''(h)}\right](1 + o(1)).
$$

  Indeed, setting $\beta = h\sqrt{t_n\,\eta''(h)}$, this leads directly to the computation done in Lemma 2.3.1, (4).

Hence, we have shown so far that

$$
\mathbb{P}\left[X_n \geq t_n(\eta'(0) + s)\right] = \mathrm{e}^{-t_n F(\eta'(0)+s)}\, \mathrm{e}^{\frac{\beta^2}{2}}\, \mathbb{P}[\mathcal{N}_{\mathbb{R}}(0,1) \geq \beta](1 + o(1)), \tag{8}
$$

with $\beta = h\sqrt{t_n\,\eta''(h)}$.

We now set $y = s\sqrt{t_n/\eta''(0)} = o(t_n^{1/2})$, and we consider the following regimes. If $y \gg 1$ (and *a fortiori* if y is of order bigger than $O(t_n)^{1/6}$), then $s \gg (t_n)^{-1/2}$, so $h \gg (t_n)^{-1/2}$ and $\beta \gg 1$. We can then use Lemma 2.3.1, (3) to replace in Equation (8) the function of β by the tail-estimate of the Gaussian:

$$
\mathbb{P}\left[X_n \geq t_n\eta'(0) + \sqrt{t_n\,\eta''(0)}y\right] = \frac{\mathrm{e}^{-t_n\, F(x)}}{h\sqrt{2\pi t_n\,\eta''(h)}}(1 + o(1)). \tag{9}
$$

Recall that $h = \frac{s}{\eta''(0)}(1 + O(s))$, so that the denominator above can be approximated as follows:

$$
h\sqrt{t_n\,\eta''(h)} = \frac{s}{\eta''(0)}(1 + O(s))\sqrt{t_n\,(\eta''(0) + O(s))} = y(1 + O(s)) = y(1 + o(1)).
$$

This completes the proof of the second part of the theorem.

Suppose on the opposite that $y = o((t_n)^{1/6})$, or, equivalently, $s = o((t_n)^{-1/3})$. Let us then see how everything is transformed.

- By making a Taylor expansion around $\eta'(0)$ of the Legendre-Fenchel transform, we get (recall that $x = \eta'(0)$ implies $h = 0$)

$$F(x) = F(\eta'(0)) + F'(\eta'(0))\, s + \frac{F''(\eta'(0))}{2}\, s^2 + O(s^3) = \frac{y^2}{2t_n} + o((t_n)^{-1}), \qquad (10)$$

so $e^{-t_n F(\eta'(0)+s)} \simeq e^{-\frac{y^2}{2}}$.

- As above,

$$\beta = h\sqrt{t_n\, \eta''(h)} = y\,(1 + O(s)) = y\,(1 + o((t_n)^{-1/3}))$$

Consequently, $\beta^2 = y^2(1 + o((t_n)^{-1/3})) = y^2 + o(1)$, so $e^{\frac{\beta^2}{2}}$ can be replaced safely by $e^{\frac{y^2}{2}}$, which compensates the previous term.

- Finally, fix y, and denote $F_y(\lambda) = \mathbb{P}[\mathcal{N}_\mathbb{R}(0,1) \geq \lambda y]$. Then, for $|\lambda|$ say between $\frac{1}{2}$ and 2,

$$|F'_y(\lambda)| = \left| \frac{y}{\sqrt{2\pi}}\, e^{-\frac{\lambda^2 y^2}{2}} \right| \leq \max_{y \in \mathbb{R}} \left| \frac{y}{\sqrt{2\pi}}\, e^{-\frac{y^2}{8}} \right| = C < +\infty;$$

as a consequence,

$$|\mathbb{P}[\mathcal{N}_\mathbb{R}(0,1) \geq \beta] - \mathbb{P}[\mathcal{N}_\mathbb{R}(0,1) \geq y]| = \left| F_y(1 + o((t_n)^{-1/3})) - F_y(1) \right|$$

$$\leq \frac{C}{(t_n)^{1/3}} = o(1).$$

This ends the proof of Theorem 3.3.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

REMARK 3.3.4. Equation (8) is the probabilistic counterpart of the number-theoretic results of [Kub72, Rad09], see in particular Theorems 2.1 and 2.2 in [Rad09]. In Section 7.2, we shall explain how to recover the precise large deviation results of [Rad09] for arithmetic functions whose Dirichlet series can be studied with the Selberg-Delange method.

The following corollary gives a more explicit form of Theorem 3.3.1, depending on the order of magnitude of y.

COROLLARY 3.3.5. *If $y = o((t_n)^{1/4})$, then one has*

$$\mathbb{P}[X_n \geq t_n \eta'(0) + \sqrt{t_n \eta''(0)}\, y] = \frac{(1 + o(1))}{y\sqrt{2\pi}}\, e^{-\frac{y^2}{2}} \exp\left(\frac{\eta'''(0)}{6\,(\eta''(0))^{3/2}} \frac{y^3}{\sqrt{t_n}} \right). \qquad (11)$$

More generally, if $y = o((t_n)^{1/2 - 1/m})$, then one has

$$\mathbb{P}[X_n \geq t_n \eta'(0) + \sqrt{t_n \eta''(0)}\, y] = \frac{(1 + o(1))}{y\sqrt{2\pi}} \exp\left(-\sum_{i=2}^{m-1} \frac{F^{(i)}(\eta'(0))}{i!} \frac{(\eta''(0))^{i/2}\, y^i}{(t_n)^{(i-2)/2}} \right).$$
$$(12)$$

PROOF. As above in Equation (10), we write $s = y\sqrt{\eta''(0)/t_n}$ and $x = \eta'(0) + s$ and do a Taylor expansion of F around $\eta'(0)$:

$$F(x) = \sum_{i=0}^{m-1} \frac{F^{(i)}(\eta'(0))}{i!} \left(y\sqrt{\frac{\eta''(0)}{t_n}} \right)^i + O(s^m).$$

Note that $F(\eta(0)) = F'(\eta'(0)) = 0$. Because of the hypothesis $y = o((t_n)^{1/2-1/m})$, we have $t_n O(s^m) = o(1)$. Therefore, plugging the equation above in Equation (6), we get (12).

Observing that $F''(\eta'(0)) = 1/\eta''(0)$ and $F'''(\eta'(0)) = \frac{-\eta'''(0)}{\eta''(0)^3}$, we get the first equation. $\qquad\square$

To summarise, in the lattice case, mod-ϕ convergence implies a large deviation principle (Theorem 3.2.2) and a precised central limit theorem (Theorem 3.3.1), and these two results cover a whole interval of possible scalings for the fluctuations of the sequence $(X_n)_{n\in\mathbb{N}}$. As we shall see in the next Chapter 4, the same holds for non-lattice reference distributions.

Fluctuations in the non-lattice case

In this chapter, we prove the analogues of Theorems 3.2.2 and 3.3.1 when ϕ is not lattice-distributed; hence, by Proposition 3.1.2, $|e^{\eta(iu)}| < 1$ for any $u \neq 0$. In this setting, assuming ϕ absolutely continuous w.r.t. the Lebesgue measure, there is a formula equivalent to the one given in Lemma 3.2.1, namely,

$$\mathbb{P}[X \geq x] = \lim_{R \to \infty} \left(\frac{1}{2\pi} \int_{-R}^{R} \frac{e^{-x(h+iu)}}{h+iu} \varphi_X(h+iu) \, du \right) \tag{13}$$

if $\varphi_X(h) = \mathbb{E}[e^{hX}] < +\infty$ for $h > 0$ (see [Fel71, Chapter XV, Section 3]). However, in order to manipulate this formula as in Chapter 3, one would need strong additional assumptions of integrability on the characteristic functions of the random variables X_n. Thus, instead of Equation (13), our main tool will be a Berry-Esseen estimate (see Proposition 4.1.1 hereafter), which we shall then combine with techniques of tilting of measures (Lemma 4.2.5) similar to those used in the classical theory of large deviations (see [DZ98, p. 32]).

4.1. Berry-Esseen estimates

As explained above, we start by establishing some Berry-Esseen estimates in the setting of mod-ϕ convergence.

PROPOSITION 4.1.1 (Berry-Esseen expansion). *We place ourselves under the assumptions of Definition 1.1.1, with ϕ non-lattice infinitely divisible law, and the strip $\mathcal{S}_{(c,d)}$ that contains 0. Denote*

$$g(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

the density of a standard Gaussian variable, and $F_n(x) = \mathbb{P}[X_n \leq t_n \eta'(0) + \sqrt{t_n \eta''(0)}\, x]$. One has

$$F_n(x) = \int_{-\infty}^{x} \left(1 + \frac{\psi'(0)}{\sqrt{t_n \eta''(0)}} y + \frac{\eta'''(0)}{6\sqrt{t_n(\eta''(0))^3}} (y^3 - 3y) \right) g(y) \, dy + o\left(\frac{1}{\sqrt{t_n}} \right)$$

with the $o(\cdot)$ uniform on $\mathbb{R}$.

PROOF. We use the same arguments as in the proof of [Fel71, Theorem XVI.4.1], but adapted to the assumptions of Definition 1.1.1. Given an integrable function f, or more generally a distribution, its Fourier transform is $f^*(\zeta) = \int_{\mathbb{R}} e^{i\zeta x} f(x) \, dx$. Consider a probability law $F(x) = \int_{-\infty}^{x} f(y) \, dy$ with vanishing expectation $(f^*)'(0) = 0$; and $G(x) = \int_{-\infty}^{x} g(y) \, dy$ a m-Lipschitz function with g^* continuously differentiable and

$$(g^*)'(0) = 0 \quad ; \quad \lim_{y \to -\infty} G(y) = 0 \quad ; \quad \lim_{y \to +\infty} G(y) = 1.$$

Then Feller's Lemma [Fel71, Lemma XVI.3.2] states that, for any $x \in \mathbb{R}$ and any $T > 0$,

$$|F(x) - G(x)| \le \frac{1}{\pi} \int_{-T}^{T} \left| \frac{f^*(\zeta) - g^*(\zeta)}{\zeta} \right| d\zeta + \frac{24m}{\pi T}.$$

Notice that this is true even when f is a distribution. Define the auxiliary variables

$$Y_n = \frac{X_n - t_n \eta'(0)}{\sqrt{t_n \eta''(0)}}.$$

We shall apply Feller's Lemma to the functions

$$F_n(x) = \text{cumulative distribution function of } Y_n;$$

$$G_n(x) = \int_{-\infty}^{x} \left(1 + \frac{\psi'(0)}{\sqrt{t_n \eta''(0)}} y + \frac{\eta'''(0)}{6\sqrt{t_n (\eta''(0))^3}} (y^3 - 3y) \right) g(y) \, dy.$$

Note that each G_n is clearly a Lipschitz function (with a uniform Lipschitz constant, *i.e.* that does not depend on n). Besides, by Lemma 2.3.1, (2), the Fourier transform corresponding to the distribution function G_n is, setting $z = i\zeta$,

$$g_n^*(\zeta) = e^{\frac{z^2}{2}} \left(1 + \frac{\psi'(0)\,z}{\sqrt{t_n \eta''(0)}} + \frac{\eta'''(0)\,z^3}{6\sqrt{t_n (\eta''(0))^3}} \right). \tag{14}$$

Consider now $f_n^*(\zeta)$: if $z = i\zeta$, then

$$f_n^*(\zeta) = \mathbb{E} \left[e^{z \left(\frac{X_n - t_n \eta'(0)}{\sqrt{t_n \eta''(0)}} \right)} \right] = \exp\left(-z \sqrt{\frac{t_n}{\eta''(0)}} \, \eta'(0) \right) \times \varphi_n \left(\frac{z}{\sqrt{t_n \eta''(0)}} \right)$$

$$= \exp\left(t_n \left(\eta \left(\frac{z}{\sqrt{t_n \eta''(0)}} \right) - \eta'(0) \frac{z}{\sqrt{t_n \eta''(0)}} \right) \right) \times \psi_n \left(\frac{z}{\sqrt{t_n \eta''(0)}} \right)$$

But

$$\psi_n \left(\frac{z}{\sqrt{t_n \eta''(0)}} \right) = \left(1 + \frac{\psi_n'(0)\,z}{\sqrt{t_n \eta''(0)}} + o\left(\frac{z}{\sqrt{t_n}} \right) \right) = \left(1 + \frac{\psi'(0)\,z}{\sqrt{t_n \eta''(0)}} + o\left(\frac{z}{\sqrt{t_n}} \right) \right)$$

where the o is uniform in n because of the local uniform convergence of the analytic functions ψ_n to ψ (and hence, of ψ_n' and ψ_n'' to ψ' and ψ''). Thus

$$f_n^*(\zeta) = \exp\left(\frac{z^2}{2} + \frac{\eta'''(0)\,z^3}{6\sqrt{t_n (\eta''(0))^3}} + |z|^2 o\left(\frac{z}{\sqrt{t_n}} \right) \right) \times \left(1 + \frac{\psi'(0)\,z}{\sqrt{t_n \eta''(0)}} + o\left(\frac{z}{\sqrt{t_n}} \right) \right)$$

$$= e^{\frac{z^2}{2}} \left(1 + \frac{\psi'(0)\,z}{\sqrt{t_n \eta''(0)}} + \frac{\eta'''(0)\,z^3}{6\sqrt{t_n (\eta''(0))^3}} + (1 + |z|^2)\, o\left(\frac{z}{\sqrt{t_n}} \right) \right). \tag{15}$$

Beware that in the previous expansions, the $o(\cdot)$ is

$$o\left(\frac{z}{\sqrt{t_n}} \right) = \frac{|z|}{\sqrt{t_n}} \varepsilon \left(\frac{z}{\sqrt{t_n}} \right) \quad \text{with } \lim_{t \to 0} \varepsilon(t) = 0.$$

In particular, z might still go to infinity in this situation. To make everything clear we will continue to use the notation $\varepsilon(t)$ in the following. Fix $0 < \delta < \Delta$ and take

$T = \Delta\sqrt{t_n}$. Comparing (14) and (15) and using Feller's lemma, we get:

$$|F_n(x) - G_n(x)| \leq \frac{1}{\pi} \int_{-\Delta\sqrt{t_n}}^{\Delta\sqrt{t_n}} \left| \frac{f_n^*(\zeta) - g_n^*(\zeta)}{\zeta} \right| d\zeta + \frac{24m}{\Delta\pi\sqrt{t_n}}$$

$$\leq \frac{1}{\pi\sqrt{t_n}} \int_{-\delta\sqrt{t_n}}^{\delta\sqrt{t_n}} e^{-\frac{\zeta^2}{2}} (1 + |\zeta|^2)\, \varepsilon\left(\frac{\zeta}{\sqrt{t_n}} \right) d\zeta + \frac{24m}{\Delta\pi\sqrt{t_n}}$$

$$+ \frac{1}{\pi\delta\sqrt{t_n}} \int_{[-\Delta\sqrt{t_n},\Delta\sqrt{t_n}]\setminus[-\delta\sqrt{t_n},\delta\sqrt{t_n}]} |f_n^*(\zeta) - g_n^*(\zeta)|\, d\zeta. \quad (16)$$

In the right-hand side, the first part is of the form $\frac{\varepsilon'(\delta)}{\sqrt{t_n}}$ when $\lim_{\delta\to 0} \varepsilon'(\delta) = 0$, while the second part is smaller than $\frac{M}{\Delta\sqrt{t_n}}$ for some constant M.

Let us show that the last integral goes to zero faster than any power of t_n. Indeed, for $|\zeta| \in [\delta\sqrt{t_n}, \Delta\sqrt{t_n}]$,

$$|f_n^*(\zeta)| = \left| \varphi_n\left(\frac{i\zeta}{\sqrt{t_n\eta''(0)}} \right) \right| \leq \left| \psi_n\left(\frac{i\zeta}{\sqrt{t_n\eta''(0)}} \right) \right| \times \left| \exp\left(t_n\, \eta\left(\frac{i\zeta}{\sqrt{t_n\eta''(0)}} \right) \right) \right|$$

The first part is bounded by a constant $K(\Delta)$ because of the uniform convergence of ψ_n towards ψ on the complex segment $[-i\Delta/\sqrt{\eta''(0)}, i\Delta/\sqrt{\eta''(0)}]$. The second part can be bounded by

$$\left(\max_{\frac{\delta}{\sqrt{\eta''(0)}} \leq |u| \leq \frac{\Delta}{\sqrt{\eta''(0)}}} |\exp(\eta(iu))| \right)^{t_n},$$

but the maximum is a constant $q_{\delta,\Delta}$ strictly smaller than 1, because η is not lattice distributed. This implies that in the domain $[-\Delta\sqrt{t_n}, \Delta\sqrt{t_n}] \setminus [-\delta\sqrt{t_n}, \delta\sqrt{t_n}]$, one has the bound

$$|f_n^*(\zeta)| \leq K(\Delta)(q_{\delta,\Delta})^{t_n}.$$

The explicit expression (14) shows that the same kind of bound holds for $|g_n^*(\zeta)|$. We shall use the notation $\widetilde{K}(\Delta)$ and $\widetilde{q}_{\delta,\Delta}$ for constants valid for both $|f_n^*(\zeta)|$ and $|g_n^*(\zeta)|$. Thus the third summand in the bound (16) is bounded by

$$\frac{4\Delta}{\pi\delta}\, \widetilde{K}(\Delta)\, (\widetilde{q}_{\delta,\frac{1}{\delta}})^{t_n}.$$

Fix $\varepsilon > 0$, then δ such that $\varepsilon(\delta) < \varepsilon$ and $M\delta < \varepsilon$. Take $\Delta = \frac{1}{\delta}$; we get

$$|F_n(x) - G_n(x)| \leq \frac{2\varepsilon}{\sqrt{t_n}} + \frac{4}{\pi\delta^2}\, \widetilde{K}(\delta^{-1})\, (\widetilde{q}_{\delta,\Delta})^{t_n} \leq \frac{3\varepsilon}{\sqrt{t_n}}$$

for t_n large enough. This completes the proof of the proposition. $\square$

REMARK 4.1.2. Proposition 4.1.1 gives an approximation for the Kolmogorov distance between the law μ_n and the normal law. Indeed, assume to simplify that the

reference law ϕ is the Gaussian law. Then, $\eta''(0) = 1$ and $\eta'''(0) = 0$, and one computes

$$d_{\mathrm{Kol}}(\mu_n, \mathcal{N}_{\mathbb{R}}(0,1)) = \frac{1}{\sqrt{t_n}} \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{x} \psi'(0)\, y\, g(y)\, dy \right| + o\left(\frac{1}{\sqrt{t_n}} \right)$$
$$= \frac{|\psi'(0)|}{\sqrt{2\pi t_n}} + o\left(\frac{1}{\sqrt{t_n}} \right).$$

This makes explicit the bound given by Theorem 1 in [Hwa98]. If $\psi'(0) \neq 0$ (e.g., as in Lemma 4.2.5), we get an equivalent of the Kolmogorov distance. However, if $\psi'(0) = 0$, then the estimate $d_{\mathrm{Kol}} = o(1/\sqrt{t_n})$ may not be optimal. Indeed, in the case of a scaled sum of i.i.d. random variables, $t_n = n^{1/3}$ and one obtains the bound

$$d_{\mathrm{Kol}}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i, \mathcal{N}_{\mathbb{R}}(0,1) \right) = o\left(\frac{1}{n^{1/6}} \right),$$

which is not as good as the classical Berry-Esseen estimate $O(\frac{1}{n^{1/2}})$. There is a way to modify the arguments in order to get such optimal estimates, by controlling the zone of mod-convergence. We refer to [FMN16], where such "optimal" computations of Kolmogorov distances is performed.

4.2. Deviations at scale $O(t_n)$

THEOREM 4.2.1. *Suppose ϕ non-lattice, and consider as before a sequence $(X_n)_{n \in \mathbb{N}}$ that converges mod-ϕ on a band $\mathcal{S}_{(c,d)}$ with $c < 0 < d$. If $x \in (\eta'(0), \eta'(d))$, then*

$$\mathbb{P}[X_n \geq t_n x] = \frac{\exp(-t_n F(x))}{h \sqrt{2\pi t_n \eta''(h)}}\, \psi(h)\, (1 + o(1))$$

where as usual h is defined by the implicit equation $\eta'(h) = x$.

REMARK 4.2.2. By applying the result to $(-X_n)_{n \in \mathbb{N}}$, one gets similarly

$$\mathbb{P}[X_n \leq t_n x] = \frac{\exp(-t_n F(x))}{|h| \sqrt{2\pi t_n \eta''(h)}}\, \psi(h)\, (1 + o(1))$$

for $x \in (\eta'(c), \eta'(0))$, with h defined by the implicit equation $\eta'(h) = x$.

REMARK 4.2.3. Our Theorem 4.2.1 should be compared with [Hwa96, Theorem 1], which studies another regime of deviations in the mod-ϕ setting, namely, when h goes to zero (or equivalently, $x \to \eta'(0)$). We shall also look at this regime in our Theorem 4.3.1.

REMARK 4.2.4. The main difference between Theorems 3.2.2 and 4.2.1 is the replacement of the factor $\psi(h)/(1 - \mathrm{e}^{-h})$ by $\psi(h)/h$; the same happens with Bahadur-Rao's estimates when going from lattice distributions to non-lattice distributions.

LEMMA 4.2.5. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables that converges mod-ϕ with parameters $(t_n)_{n\in\mathbb{N}}$ and limiting function ψ, on a strip $\mathcal{S}_{(c,d)}$ that does not necessarily contain 0. For $h \in (c, d)$, we make the exponential change of measure*

$$\mathbb{Q}[dy] = \frac{e^{hy}}{\varphi_{X_n}(h)} \, \mathbb{P}[X_n \in dy],$$

and denote $\widetilde{X}_n$ a random variable following this law. The sequence $(\widetilde{X}_n)_{n\in\mathbb{N}}$ converges mod-$\widetilde{\phi}$, where $\widetilde{\phi}$ is the infinitely divisible distribution with characteristic function $e^{\eta(z+h)-\eta(h)}$. The parameters of this new mod-convergence are again $(t_n)_{n\in\mathbb{N}}$, and the limiting function is

$$\widetilde{\psi}(z) = \frac{\psi(z+h)}{\psi(h)}.$$

The new mod-$\widetilde{\phi}$ convergence occurs in the strip $\mathcal{S}_{(c-h,d-h)}$.

PROOF. Obvious since $\varphi_{\widetilde{X}_n}(z) = \varphi_{X_n}(z+h) / \varphi_{X_n}(h)$. □

PROOF OF THEOREM 4.2.1. Fix $h \in (c, d)$, and consider the sequence $(\widetilde{X}_n)_{n\in\mathbb{N}}$ of Lemma 4.2.5. All the assumptions of Proposition 4.1.1 are satisfied, so, the distribution function $F_n(u)$ of

$$\frac{\widetilde{X}_n - t_n\eta'(h)}{\sqrt{t_n\eta''(h)}}$$

is

$$G_n(u) = \int_{-\infty}^{u} \left(1 + \frac{\psi'(h)}{\psi(h)\sqrt{t_n\eta''(h)}} y + \frac{\eta'''(h)}{\sqrt{t_n(\eta''(h))^3}} (y^3 - 3y)\right) g(y) \, dy$$

up to a uniform $o(1/\sqrt{t_n})$. Then,

$$\mathbb{P}[X_n \geq t_n\eta'(h)] = \int_{y=t_n\eta'(h)}^{\infty} \varphi_{X_n}(h) \, e^{-hy} \, \mathbb{Q}(dy)$$

$$= \varphi_{X_n}(h) \int_{u=0}^{\infty} e^{-h\left(t_n\eta'(h) + \sqrt{t_n\eta''(h)}\, u\right)} \, dF_n(u)$$

$$= \psi_n(h) \, e^{-t_n F(x)} \int_{u=0}^{\infty} e^{-h\sqrt{t_n\eta''(h)}\, u} \, dF_n(u), \quad (\text{as } F(x) = h\eta'(h) - \eta(h)).$$

To compute the integral I, we choose the primitive $F_n(u) - F_n(0)$ of $dF_n(u)$ that vanishes at $u = 0$, and we make an integration by parts. Notice that we now need $h > 0$ (hence, $x > \eta'(0)$) in order to manipulate some of the terms below:

$$I = h\sqrt{t_n\eta''(h)} \int_{u=0}^{\infty} e^{-h\sqrt{t_n\eta''(h)}\, u} \, (F_n(u) - F_n(0)) \, du$$

$$= h\sqrt{t_n\eta''(h)} \int_{u=0}^{\infty} e^{-h\sqrt{t_n\eta''(h)}\, u} \left(G_n(u) - G_n(0) + o\left(\frac{1}{\sqrt{t_n}}\right)\right) du$$

$$\simeq h\sqrt{t_n\eta''(h)} \iint_{0\leq y\leq u} e^{-h\sqrt{t_n\eta''(h)}\,u}\left(1+\frac{\psi'(h)\,y}{\psi(h)\,\sqrt{t_n\eta''(h)}}+\frac{\eta'''(h)\,(y^3-3y)}{\sqrt{t_n(\eta''(h))^3}}\right)g(y)\,dy\,du$$

$$\simeq \int_{y=0}^{\infty} e^{-h\sqrt{t_n\eta''(h)}\,y}\left(1+\frac{\psi'(h)}{\psi(h)\,\sqrt{t_n\eta''(h)}}\,y+\frac{\eta'''(h)}{\sqrt{t_n(\eta''(h))^3}}\,(y^3-3y)\right)g(y)\,dy$$

$$\simeq \frac{e^{\frac{h^2\,t_n\eta''(h)}{2}}}{\sqrt{2\pi}}\int_{y=0}^{\infty}e^{-\frac{(y+h\sqrt{t_n\eta''(h)})^2}{2}}\left(1+\frac{\psi'(h)}{\psi(h)\,\sqrt{t_n\eta''(h)}}\,y+\frac{\eta'''(h)}{\sqrt{t_n(\eta''(h))^3}}\,(y^3-3y)\right)dy,$$

where on the three last lines the symbol $\simeq$ means that the remainder is a $o((t_n)^{-1/2})$. By Lemma 2.3.1, (3), the only contribution in the integral that is not a $o((t_n)^{-1/2})$ is

$$\frac{e^{\frac{h^2\,t_n\eta''(h)}{2}}}{\sqrt{2\pi}}\int_{y=0}^{\infty}e^{-\frac{(y+h\sqrt{t_n\eta''(h)})^2}{2}}\,dy = \frac{1}{h\sqrt{2\pi t_n\eta''(h)}}+o\left(\frac{1}{\sqrt{t_n}}\right).$$

This ends the proof since $\psi_n(h)\to\psi(h)$ locally uniformly.   □

4.3. Central limit theorem at the scales $o(t_n)$ and $o((t_n)^{2/3})$

As in the lattice case, one can also prove from the hypotheses of mod-convergence an extended central limit theorem:

THEOREM 4.3.1. *Consider a sequence $(X_n)_{n\in\mathbb{N}}$ that converges mod-ϕ with limiting distribution ψ and parameters t_n, where ϕ is a non-lattice infinitely divisible law that is absolutely continuous w.r.t. Lebesgue measure. Let $y = o((t_n)^{1/6})$. Then,*

$$\mathbb{P}\left[X_n\geq t_n\eta'(0)+\sqrt{t_n\eta''(0)}\,y\right] = \mathbb{P}[\mathcal{N}_{\mathbb{R}}(0,1)\geq y]\,(1+o(1)).$$

On the other hand, assume $y\gg 1$ and $y = o((t_n)^{1/2})$. If $x = \eta'(0)+\sqrt{\eta''(0)/t_n}\,y$ and h is the solution of $\eta'(h) = x$, then

$$\mathbb{P}\left[X_n\geq t_n\eta'(0)+\sqrt{t_n\eta''(0)}\,y\right] = \frac{e^{-t_n\,F(x)}}{h\sqrt{2\pi t_n\,\eta''(h)}}\,(1+o(1)) = \frac{e^{-t_n\,F(x)}}{y\sqrt{2\pi}}\,(1+o(1)).$$

As in the proof of Theorem 3.3.1, we need to control the modulus of the Fourier transform of the reference law ϕ. Thus, let us state the non-lattice analogue of Lemma 3.3.3:

LEMMA 4.3.2. *Consider a non-constant infinitely divisible law ϕ, of type non-lattice, with a convergent moment generating function in a strip $\mathcal{S}_{(c,d)}$ with $c < 0 < d$. We also assume that ϕ is absolutely continuous w.r.t. the Lebesgue measure. Then, there exists a constant $D > 0$ only depending on ϕ, and an interval $(-\varepsilon,\varepsilon)\subset(c,d)$, such that for all $h\in(-\varepsilon,\varepsilon)$, and all δ small enough,*

$$q_\delta = \max_{u\in\mathbb{R}\setminus(-\delta,\delta)}|\exp(\eta(h+iu)-\eta(h))|\leq 1-D\,\delta^2.$$

REMARK 4.3.3. One can give a sufficient condition on the Lévy-Khintchine representation of ϕ to ensure the absolute continuity with respect to the Lebesgue measure; *cf.* [SH04, Chapter 4, Theorem 4.23]. Hence, it is the case if $\sigma^2 > 0$, or if $\sigma^2 = 0$ and if the absolutely continuous part of the Lévy measure Π has infinite mass.

REMARK 4.3.4. Let us explain why we need to add the assumption of absolute continuity with respect to Lebesgue measure, which is a strictly stronger hypothesis than being non-lattice. The hypotheses on the infinitely divisible law ϕ imply that it has finite variance, and therefore, that the Lévy-Khintchine representation of the Fourier transform given by Equation (3) can be replaced by a Kolmogorov representation. This representation actually holds for the complex moment generating function (see [SH04, Chapter 4, Theorem 7.7]):

$$\eta(z) = mz + \sigma^2 \int_{\mathbb{R}} \frac{e^{zx} - 1 - zx}{x^2} K(dx)$$

where K is a probability measure on $\mathbb{R}$, and where the fraction in the integral is extended by continuity at $x = 0$ by the value $-\frac{z^2}{2}$. As a consequence,

$$|\exp(\eta(h + iu) - \eta(h))| = \exp\left(\sigma^2 \int_{\mathbb{R}} \frac{e^{hx}(\cos ux - 1)}{x^2} K(dx)\right) \le 1.$$

This expression can be expanded in series of u as

$$1 - \frac{\sigma^2 u^2}{2} \int_{\mathbb{R}} e^{hx} K(dx) + O_h(u^3).$$

Therefore, Lemma 4.3.2 holds as soon as one can show that

$$\sup_{h \in (-\varepsilon, \varepsilon)} \limsup_{|u| \to \infty} |\exp(\eta(h + iu) - \eta(h))| < 1,$$

because one has a bound of type $1 - D u^2$ in a neighbourhood of zero. Unfortunately, for general probability measures, the Riemann-Lebesgue lemma does not apply, and even for $h = 0$, it is unclear whether for a general exponent η the Cramér condition (C)

$$\limsup_{|u| \to \infty} |\exp(\eta(iu))| < 1$$

is satisfied (see [Pet95] for more discussion and references on condition (C)). We refer to [Wol83, Theorem 2], where it is shown that *decomposable* probability measures enjoy this property. This difficulty explains why one has to restrict oneself to absolutely continuous measures in the non-lattice case, in order to use the Riemann-Lebesgue lemma. In the following we provide an *ad hoc* proof of Lemma 4.3.2 in the absolutely continuous cases, that does not rely on the Kolmogorov representation.

PROOF OF LEMMA 4.3.2. We shall adapt the arguments of Lemma 3.3.3 from the discrete to the continuous case. Though the density f cannot be supported on a compact segment (*cf.* Lemma 3.1.1 and the classification of the possible supports of an infinitely divisible law), one can work as if it were the case, thanks to the following calculation:

$$
\begin{aligned}
|\exp(\eta(h + iu) - \eta(h))| &= \left|\frac{\phi(e^{(h+iu)x})}{\phi(e^{hx})}\right| \\
&= \frac{|\phi_{<a}(e^{(h+iu)x}) + \int_a^b e^{(h+iu)x} f(x)\,dx + \phi_{>b}(e^{(h+iu)x})|}{\phi_{<a}(e^{hx}) + \int_a^b e^{hx} f(x)\,dx + \phi_{>b}(e^{hx})} \\
&\le \frac{\phi_{<a}(e^{hx}) + \phi_{>b}(e^{hx}) + |\int_a^b e^{(h+iu)x} f(x)\,dx|}{\phi_{<a}(e^{hx}) + \phi_{>b}(e^{hx}) + \int_a^b e^{hx} f(x)\,dx}
\end{aligned}
$$

where $\phi_{<a}$ (respectively, $\phi_{>b}$) is the measure $\mathbb{1}_{x<a}\,\phi(dx)$ (resp., $\mathbb{1}_{x>b}\,\phi(dx)$). Therefore, it suffices to show:

$$\max_{u\in(-\delta,\delta)^{\mathsf{c}}} \frac{|\int_a^b e^{(h+iu)x}\,f(x)\,dx|}{\int_a^b e^{hx}\,f(x)\,dx} \leq 1 - D\,\delta^2$$

for δ and h small enough. This reduction to a compact support will be convenient later in the computations.

Set $g_h(x) = \dfrac{e^{hx}\,f(x)}{\int_a^b e^{hx}\,f(x)\,dx}$ and

$$F(h,u) = \left|\int_a^b g_h(x)\,e^{iux}\,dx\right|^2 = \iint_{[a,b]^2} g_h(x)g_h(y)\,e^{iu(x-y)}\,dx\,dy$$

$$= \iint_{[a,b]^2} g_h(x)g_h(y)\,\cos(u(x-y))\,dx\,dy$$

$$= \int_{t=-(b-a)}^{b-a} \left(\int_{x=\max(a,t+a)}^{\min(b,t+b)} g_h(x)g_h(x-t)\,dx\right)\cos ut\,dt.$$

The problem is to show that

$$\sup_{h\in(-\varepsilon,\varepsilon)}\;\sup_{u\in(-\delta,\delta)^{\mathsf{c}}} F(h,u) \leq 1 - D\,\delta^2$$

for some constant D. With h fixed, by the Riemann-Lebesgue lemma applied to the integrable function

$$m(t) = \int_{x=\max(a,t+a)}^{\min(b,t+b)} g_h(x)g_h(x-t)\,dx,$$

the limit as $|u|$ goes to infinity of $F(h,u)$ is 0. On the other hand, if $u\neq 0$, then $F(h,u) < F(h,0) = 1$. Indeed, suppose the opposite: then $\cos ut = 1$ almost surely w.r.t. the measure $m(t)\,dt$. This means that this measure $m(t)\,dt$ is concentrated on the lattice $\frac{2\pi}{|u|}\,\mathbb{Z}$, which is impossible for a measure continuous with respect to the Lebesgue measure. Combining these two observations, one sees that for any $\delta > 0$,

$$\sup_{u\in(-\delta,\delta)^{\mathsf{c}}} F(h,u) \leq C_{(h,\delta)} < 1$$

for some constant $C_{(h,\delta)}$. Since all the terms considered depend smoothly on h, for h small enough, one can even take a uniform constant C_δ:

$$\forall\delta > 0,\; \exists C_\delta < 1 \text{ such that } \sup_{h\in(-\varepsilon,\varepsilon)}\;\sup_{u\in(-\delta,\delta)^{\mathsf{c}}} F(h,u) \leq C_\delta. \tag{17}$$

On the other hand, notice that

$$\frac{\partial F(h,u)}{\partial u} = -\iint_{[a,b]^2} g_h(x)g_h(y)\,(x-y)\,\sin(u(x-y))\,dx\,dy.$$

However, if $u(b-a) \leq \frac{\pi}{2}$, then $(x-y)\sin(u(x-y)) \geq \frac{2u}{\pi}(x-y)^2$ over the whole domain of integration, so,

$$\frac{\partial F(h,u)}{\partial u} \leq -\frac{2B_h}{\pi}\,u$$

where $B_h = \iint_{[a,b]^2} g_h(x) g_h(y)\,(x-y)^2\,dx\,dy$. By integration,

$$F(h,u) \leq 1 - \frac{B_h}{\pi}\,u^2 \quad \text{for all } u \leq \frac{\pi}{2(b-a)}.$$

Again, by continuity of the constant B_h w.r.t. h, one can take a uniform constant :

$$\exists B > 0 \text{ such that for all } u \leq \frac{\pi}{2(b-a)}, \quad \sup_{h \in (-\varepsilon, \varepsilon)} F(h,u) \leq 1 - B\,u^2. \tag{18}$$

The two assertions (17) (with $\delta = \frac{\pi}{2(b-a)}$) and (18) enable one to conclude, with

$$D = \inf\left(B, \frac{1-C_\delta}{\delta^2} \right), \quad \text{where } \delta = \frac{\pi}{2(b-a)}. \qquad \square$$

We also refer to [Ess45, Theorem 6] for a general result on the Lebesgue measure of the set of points such that the characteristic function of a distribution is larger in absolute value than $1 - \delta^2$.

PROOF OF THEOREM 4.3.1. The proof is now exactly the same as in the lattice case (Theorem 3.3.1). Indeed, the conclusions of the technical Lemma 4.3.2 hold, and on the other hand, the equivalents for $\mathbb{P}[X_n \geq t_n x]$ in the lattice and non-lattice cases (Theorems 3.2.2 and 4.2.1) differ only by the fact that $1 - e^{-h}$ is replaced by h. But in the proof of Theorem 3.3.1, the quantity $1 - e^{-h}$ is approximated by h, so everything works the same way as in the non-lattice case. $\square$

As in the non-lattice case, we have the following corollary (with the exact same statement and proof):

COROLLARY 4.3.5. *If $y = o((t_n)^{1/4})$, then one has*

$$\mathbb{P}\left[X_n \geq t_n \eta'(0) + \sqrt{t_n \eta''(0)}\,y\right] = \frac{(1+o(1))}{y\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,\exp\left(\frac{\eta'''(0)}{6\,(\eta''(0))^{3/2}}\,\frac{y^3}{\sqrt{t_n}} \right). \tag{19}$$

More generally, if $y = o((t_n)^{1/2 - 1/m})$, then one has

$$\mathbb{P}\left[X_n \geq t_n \eta'(0) + \sqrt{t_n \eta''(0)}\,y\right] = \frac{(1+o(1))}{y\sqrt{2\pi}}\,\exp\left(-\sum_{i=2}^{m-1} \frac{F^{(i)}(\eta'(0))}{i!}\,\frac{(\eta''(0))^{i/2}\,y^i}{(t_n)^{(i-2)/2}} \right). \tag{20}$$

Hence, one can again describe all the fluctuations of X_n from order $O(\sqrt{t_n})$ up to order $O(t_n)$, see Figure 1.

To conclude this paragraph, let us mention an application that looks similar to the law of the iterated logarithm, and that also works in the lattice case. Consider a sequence $(X_n)_{n \in \mathbb{N}}$ converging mod-ϕ with parameters t_n such that $t_n \gg (\log n)^3$. We also assume that the random variables X_n are defined on the same probability space, and we look for sequences γ_n such that almost surely,

$$\limsup_{n \to \infty} \left(\frac{X_n - t_n \eta'(0)}{\gamma_n} \right) \leq 1.$$

order of fluctuations

$O(t_n)$ large deviations $(\eta'(0) < x)$: $\mathbb{P}[\frac{X_n}{t_n} \geq x] \simeq \frac{\exp(-t_n F(x))}{F'(x)\sqrt{2\pi t_n \eta'(x)}} \psi(F'(x))$;

extended central limit
theorem $((t_n)^{1/6} \lesssim y \ll (t_n)^{1/2})$: $\mathbb{P}[\frac{X_n - t_n\eta'(0)}{\sqrt{t_n\,\eta''(0)}} \geq y] \simeq \frac{\exp(-t_n F(x))}{F'(x)\,\sqrt{2\pi t_n \eta'(x)}}$;

$O((t_n)^{2/3})$

central limit theorem $(y \ll (t_n)^{1/6})$:

$$\mathbb{P}[\frac{X_n - t_n\eta'(0)}{\sqrt{t_n\,\eta''(0)}} \geq y] \simeq \mathbb{P}[\mathcal{N}_\mathbb{R}(0,1) \geq y].$$
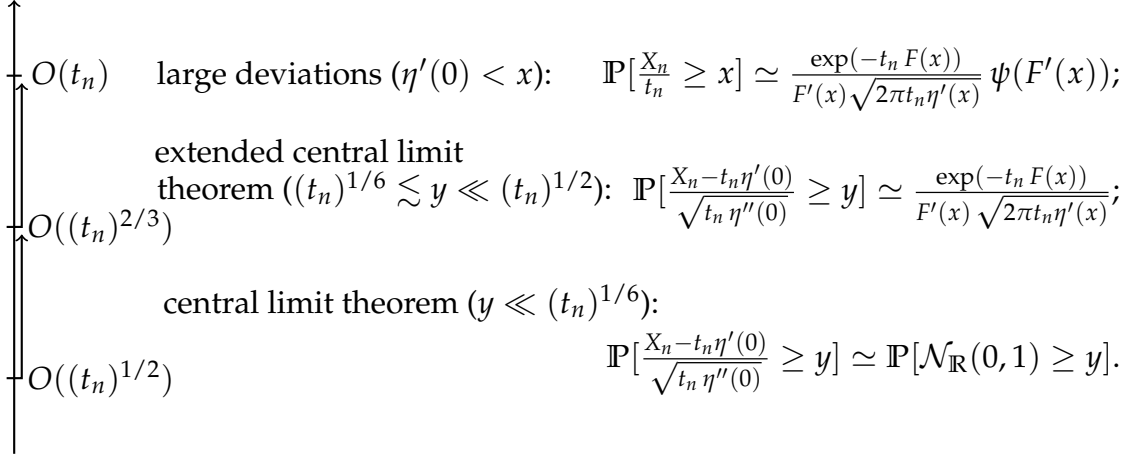
$O((t_n)^{1/2})$

FIGURE 1. Panorama of the fluctuations of a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ that converges modulo an absolutely continuous distribution (with $x = \eta'(0) + \sqrt{\eta''(0)/t_n}\,y$).

Unlike in the usual law of the iterated logarithm, we do not make any assumption of independence. Such assumptions are common in this setting, or at least some conditional independence (for instance, a law of the iterated logarithm can be stated for martingales); see the survey [Bin86] or [Pet75, Chapter X].

On the one hand we have a less precise result: we only obtain an upper bound, which is not tight in the case of sums of i.i.d. variable since we have a $\sqrt{\log(n)}$ factor instead the usual $\sqrt{\log(\log(n))}$ factor. On the other hand, our result does not depend at all on the way one realises the random variables X_n. In other words, for every possible coupling of the variables X_n, the following holds:

PROPOSITION 4.3.6. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence that converges mod-ϕ with parameters t_n, where ϕ is either a non-constant lattice distribution, or a non-lattice distribution that is absolutely continuous with respect to the Lebesgue measure. We assume*

$$\lim_{n\to\infty} \frac{t_n}{(\log n)^3} = +\infty.$$

Then,

$$\limsup_{n\to\infty} \frac{X_n - t_n\eta'(0)}{\sqrt{2\eta''(0)\,t_n \log n}} \leq 1 \quad \text{almost surely.}$$

PROOF. Notice the term $\log n$ instead of $\log\log n$ for the usual law of iterated logarithm. One computes

$$\mathbb{P}\left[X_n - t_n\eta'(0) \geq \sqrt{2(1+\varepsilon)\,\eta''(0)\,t_n \log n}\right].$$

Set $y = \sqrt{2(1+\varepsilon)\log n}$. Due to the hypotheses on t_n, one has $y = o((t_n)^{1/6})$ and one can apply Theorem 4.3.1: using the classical equivalent

$$\mathbb{P}[\mathcal{N}_\mathbb{R}(0,1) \geq y] \sim \frac{e^{-y^2/2}}{y\sqrt{2\pi}},$$

we get

$$\mathbb{P}\left[X_n - t_n\eta'(0) \geq \sqrt{2(1+\varepsilon)\eta''(0)t_n\log n}\right] \simeq \frac{e^{-(1+\varepsilon)\log n}}{\sqrt{4\pi(1+\varepsilon)\log n}} \leq \frac{1}{n^{1+\varepsilon}}.$$

for n large enough. For any $\varepsilon > 0$, this is summable, so by the Borel-Cantelli lemma, one has almost surely

$$X_n - t_n\eta'(0) < \sqrt{2(1+\varepsilon)\,\eta''(0)\,t_n\log n} \quad \text{for } n \text{ large enough.}$$

Since this is true for every ε, one has (almost surely):

$$\limsup_{n\to\infty} \frac{X_n - t_n\eta'(0)}{\sqrt{2\eta''(0)\,t_n\log n}} \leq 1.$$

$\square$

4.4. Normality zones for mod-ϕ and mod-Gaussian sequences

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables that converges mod-ϕ (we do not assume ϕ non-lattice for the moment). Then we have seen that

$$Y_n = \frac{X_n - t_n\eta'(0)}{\sqrt{t_n\eta''(0)}}$$

satisfies a central limit theorem: for all fixed y,

$$\lim_{n\to\infty} \frac{\mathbb{P}[Y_n \geq y]}{(2\pi)^{-1/2}\int_y^\infty e^{-\frac{s^2}{2}}\,ds} = 1. \tag{21}$$

The question that we address here is the question of the *normality zone*: we want to identify the maximal scale u_n such that Equation (21) holds for $y = o(u_n)$. The results of the previous sections allows to identify this scale and to describe what happens for $y = O(u_n)$.

Suppose that ϕ is a lattice distribution, or a non-lattice distribution that is absolutely continuous with respect to the Lebesgue measure. From Theorem 3.3.1 or Theorem 4.3.1, we get that if $y = o((t_n)^{1/6})$, then $\mathbb{P}[Y_n \geq y]$ is given by the Gaussian distribution. Assume $\eta'''(0) \neq 0$. Then, the previous result is optimal, because of Equations (11) and (19): for $y = c(t_n)^{1/6}$ the second factor is different from 1 and the approximation by the Gaussian tail is no longer valid. Thus, if $\eta'''(0) \neq 0$, then the normality zone for the sequence $(Y_n)_{n\in\mathbb{N}}$ is $o((t_n)^{1/6})$. In particular, one has the same asymptotics and normality zone than in the case of the sum of t_n i.i.d. variables of law ϕ; see [Cra38] and [Pet54].

The only case where this comparison does not give us the normality zone is the case of mod-Gaussian convergence, that we shall discuss now.

PROPOSITION 4.4.1. *Assume that $(X_n)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense, with a non-trivial limiting function (i.e., $\psi \not\equiv 1$). Then the normality zone for $Y_n = X_n/\sqrt{t_n}$ is $o(\sqrt{t_n})$.*

PROOF.   • Let $y = o((t_n)^{1/2})$. Set $x = h = y/\sqrt{t_n}$ as in Theorem 4.3.1 in the mod-Gaussian case. Then, the second part of Theorem 4.3.1 states that

$$\mathbb{P}[Y_n \geq y] = \mathbb{P}[X_n \geq t_n x] = \frac{e^{-\frac{t_n x^2}{2}}}{h\sqrt{2\pi t_n}}\,(1 + o(1)) = \frac{e^{-\frac{y^2}{2}}}{y\sqrt{2\pi}}\,(1 + o(1)).$$

Thus, the normality zone is at least $o(\sqrt{t_n})$ in this case.

• Set now $y = x\sqrt{t_n}$ for a fixed $x > 0$. Then, Theorem 4.2.1 states that

$$\mathbb{P}[Y_n \geq y] = \mathbb{P}[X_n \geq t_n\,x] = \frac{e^{-\frac{y^2}{2}}}{y\sqrt{2\pi}}\,\psi(x)\,(1 + o(1)). \tag{22}$$

Similarly, if $y = -x\sqrt{t_n}$ for a fixed $x > 0$, we get (see Remark 4.2.2)

$$\mathbb{P}[Y_n \leq y] = \mathbb{P}[X_n \leq -t_n\,x] = \frac{e^{-\frac{y^2}{2}}}{|y|\sqrt{2\pi}}\,\psi(-x)\,(1 + o(1)). \tag{23}$$

In particular, if $\psi(x)$ is not identically equal to 1, then the approximation (21) of $\mathbb{P}[Y_n \geq y]$ by the Gaussian tail breaks and the normality zone is exactly $o(\sqrt{t_n})$.   □

As seen in the proof, from a simple application of Theorem 4.2.1, the residue ψ describes how to correct the Gaussian tail to find an equivalent for $\mathbb{P}[Y_n \geq y]$. A standard and interesting case is the case where the limiting function is $\psi = \exp(Lz^v)$, where $L \neq 0$ is a real number and v a positive integer ($v \geq 3$). This might seem restrictive, but we will see in the examples that ψ is very often of this type — Examples 7.1.1, 7.1.2, 2.1.2 and Theorems 11.2.4, 10.0.1, 8.2.1 and 9.5.1. If v is odd ($v = 3$ is a common case), comparing Equations (22) and (23) shows the following phenomenon: the negative and positive deviations of Y_n at order $O(\sqrt{t_n})$ have different asymptotic behaviour, one being larger than the other one depending on the sign of L. In other words, our results reveal a *breaking of symmetry* at the edge of the normality zone.

REMARK 4.4.2. This breaking of symmetry also occurs in multi-dimensional setting. In particular, in two dimensions, the residue allows to compute the distribution of the angle of a sum of i.i.d. random variables at the edge of the normality zone, see [FMN15].

4.5. Discussion and refinements

4.5.1. Bahadur-Rao theorem and Cramér-Petrov expansion. We consider here the case of a sum $S_n = Y_1 + \cdots + Y_n$ of i.i.d. random variables such that $Y = Y_1$ has an infinitely divisible distribution of Lévy exponent η. Then S_n *converges mod-Y* with parameters n and limiting function $\psi = 1$; see Example 2.1.2. In this case, Theorems 4.2.1 and 3.2.2 correspond to Bahadur-Rao estimates

$$\mathbb{P}[S_n \geq n\,x] \simeq \begin{cases} \frac{\exp(-n\,F(x))}{(1-e^{-h})\sqrt{2\pi n\,\eta''(h)}} & \text{in the lattice case (assume } \mathbb{Z} \text{ is the minimal lattice);} \\ \frac{\exp(-n\,F(x))}{h\sqrt{2\pi n\,\eta''(h)}} & \text{in the non-lattice case,} \end{cases}$$

where $\eta(h) = \log \mathbb{E}[e^{hY}]$ and F is the Legendre-Fenchel transform of η; see Theorem 3.7.4 in [DZ98], and also the papers [BR60, Ney83, Ilt95].

In the same setting, Theorems 4.3.1 and 3.3.1 correspond to Cramér-Petrov expansion [Pet95] (in the non-lattice case, we assume in addition that the law of Y is absolutely continuous with respect to Lebesgue measure). To the best of our knowledge, the link with the Legendre-Fenchel transform is new.

4.5.2. On the infinite divisibility of ϕ. As above, consider the case of a sum $S_n = Y_1 + \cdots + Y_n$ of i.i.d. random variables, but with the law of Y_1 not necessarily infinitely divisible. In this case, $\mathbb{E}[e^{zS_n}] = (\mathbb{E}[e^{zY_1}])^n$, but, if $\mathbb{E}[e^{zY_1}]$ vanishes for some complex value of z, one cannot write this as $\exp(n\eta(z))$ as in Definition 1.1.1.

The proofs of our large deviation results — Theorems 4.2.1 and 3.2.2 — can nevertheless be adapted to this setting. For the extended central limit theorem — Theorems 3.3.1 and 4.3.1 — we would need to assume the estimate given by Lemma 3.3.3 or Lemma 4.3.2. This is satisfied in particular if:

- either Y takes its values in $\mathbb{Z}$, and there are two consecutive integers $n, m = n - 1$ such that $\mathbb{P}[Y = n] \neq 0$ and $\mathbb{P}[Y = m] \neq 0$;

- or, Y has a component absolutely continuous w.r.t. Lebesgue measure.

Since these are classical results and since our method are close to the usual ones, we do not give details on how to adapt our proof to the non infinitely divisible setting.

REMARK 4.5.1. For Bahadur-Rao theorem, it should be noticed that the assumption that $\mathbb{Z}$ is the minimal lattice is necessary. For instance, if one considers a sum S_n of n independent Bernoulli random variables with $\mathbb{P}[B = 1] = \mathbb{P}[B = -1] = 1/2$, then the estimate above is not true, because S_n has always the same parity as n. This is related to the fact that $\mathbb{E}[e^{zB}]$ has modulus 1 at $z = i\pi$.

4.5.3. Quasi powers. Mod-ϕ convergent is reminiscent of the quasi-power theory developed by Hwang [Hwa96, Hwa98] — see also [FS09, Chapter IX].

DEFINITION 4.5.2. *[Hwa96] A sequence (X_n) of random variables satisfy the quasi-power hypothesis if*

$$\mathbb{E}[e^{zX_n}] = e^{\phi(n)u(z)+v(z)}(1 + O(\kappa_n^{-1})), \tag{24}$$

where $\phi(n), \kappa_n \to \infty$, $u(z)$ and $v(z)$ are analytic functions for $|s| < \rho$ (with $u''(0) \neq 0$) and the O symbol is uniform in the disk $D(0, \rho)$.

Clearly, any sequence converging mod-ϕ satisfies this hypothesis, taking $\phi(n) = t_n$, $v(z) = \ln(\psi(z))$, $u(z) = \eta(z)$ (since $\psi(0) = 1$, a determination of the ln always exists on a sufficiently small neighbourhood of the origin). A major difference between mod-convergence and the quasi-power framework is that we assume that $\eta(z)$ is the Lévy exponent of an infinitely divisible distribution, while Hwang does not have any hypothesis on $u(z)$ (except $u''(0) \neq 0$). The fact that $\exp(\eta(z)) = \int_{\mathbb{R}} e^{zx} \phi(dx)$ is important to study deviations at scale $O(t_n)$, since we used the inequality $|\exp(\eta(z))| \leq 1$ for $z = i\xi$ in the proof of Theorem 3.2.2 and Proposition 4.1.1.

At the scale $o(t_n)$, our results — Theorems 4.3.1 and 3.3.1 — coincide with the ones of Hwang. Note, however, that our hypotheses are slightly different. We need $\eta(z)$ to be the Lévy exponent of an infinitely divisible distribution, while Hwang uses an hypothesis on the speed of convergence in Equation (24). In most examples, η is a Poisson or Gaussian Lévy exponent, so that our hypothesis is automatically verified. It

can thus be considered as a slight improvement that we do not require any hypothesis on the speed of convergence (but such an hypothesis allows us to refine our results at scale $O(t_n)$ in the lattice case, see Theorem 3.2.2).

REMARK 4.5.3 (The disk or the strip?). In the quasi-power framework, we assume convergence of the renormalised Laplace transform on a disk, while mod-ϕ convergence is defined as such convergence on the strip. It is thus natural to wonder which hypothesis is more natural. To this purpose, let us mention an old result of Lukacs and Szász [LS52, Theorem 2]: if X is a random variable with an analytic moment generating function $\mathbb{E}[e^{zX}]$ defined on the open disk $\mathcal{D}_{(0,c)}$, then this function is automatically defined and analytic on the strip $\mathcal{S}_{(-c,c)}$. This implies that the left-hand side of Eq. (2) is automatically defined on a strip, as soon as it is defined on a disk. Of course it could converge on a disk and not on a strip, but we shall see throughout this paper that, in many examples, the convergence on the strip indeed occurs. Actually, in most of our examples, $c = -\infty$ and $d = +\infty$, and the distinction between disk and strip disappears as $\mathcal{D}_{(0,+\infty)} = \mathcal{S}_{(-\infty,+\infty)} = \mathbb{C}$.

CHAPTER 5

An extended deviation result from bounds on cumulants

In this chapter, we discuss a particular case of mod-Gaussian variables, that arises from bounds on cumulants. We will see that in this case the deviation result given in Theorem 4.2.1 is still valid at a scale larger than t_n; see Proposition 5.2.1.

5.1. Bounds on cumulants and mod-Gaussian convergence

Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of real-valued centered random variables that admit moments of all order, and such that

$$|\kappa^{(r)}(S_n)| \leq (Cr)^r \, \alpha_n(\beta_n)^r \tag{25}$$

for all $r \geq 2$ and for some sequences $(\alpha_n)_{n\to\infty} \to +\infty$ and $(\beta_n)_{n\in\mathbb{N}}$ arbitrary. Assume moreover that there exists an integer $v \geq 3$ such that

(1) $\kappa^{(r)}(S_n) = 0$ for all $3 \leq r < v$ and all $n \in \mathbb{N}$;

(2) we have the following approximations for second and third cumulants:

$$\kappa^{(2)}(S_n) = \sigma^2 \, \alpha_n(\beta_n)^2 \left(1 + o\left((\alpha_n)^{-\frac{v-2}{v}}\right)\right);$$

$$\kappa^{(v)}(S_n) = L \, \alpha_n(\beta_n)^v \left(1 + o(1)\right). \tag{26}$$

Set $X_n = \frac{S_n}{(\alpha_n)^{\frac{1}{v}}\beta_n}$. The cumulant generating series of X_n is

$$\log \varphi_n(z) = \frac{\kappa^{(2)}(S_n)}{2\,(\alpha_n)^{\frac{2}{v}}\,(\beta_n)^2}\, z^2 + \frac{\kappa^{(v)}(S_n)}{v!\,\alpha_n(\beta_n)^v}\, z^v + \sum_{r=v+1}^{\infty} \frac{\kappa^{(r)}(S_n)}{r!\,(\alpha_n)^{\frac{r}{v}}\,(\beta_n)^r}\, z^r$$

$$= \frac{\sigma^2}{2}\,(\alpha_n)^{\frac{v-2}{v}} z^2 + \frac{L}{v!}\, z^v + \sum_{r=v+1}^{\infty} \frac{\kappa^{(r)}(S_n)}{r!\,(\alpha_n)^{\frac{r}{v}}\,(\beta_n)^r}\, z^r + o(1),$$

where the $o(1)$ is locally uniform. The remaining series is locally uniformly bounded in absolute value by

$$\sum_{r=v+1}^{\infty} C^r \frac{r^r}{r!}\, \frac{1}{(\alpha_n)^{\frac{r-v}{v}}}\, R^r \leq \alpha_n \sum_{r=v+1}^{\infty} \left(\frac{e\,CR}{(\alpha_n)^{\frac{1}{v}}}\right)^r = (\alpha_n)^{-\frac{1}{v}}\, \frac{(e\,CR)^{v+1}}{1 - e\,CR\,(\alpha_n)^{-\frac{1}{v}}} \to 0.$$

Hence,

$$\psi_n(z) = \exp\left(-(\alpha_n)^{\frac{v-2}{v}}\,\frac{\sigma^2 z^2}{2}\right)\, \varphi_n(z) \to \exp\left(\frac{L}{v!}\, z^v\right)$$

locally uniformly on $\mathbb{C}$, so one has again mod-Gaussian convergence, with parameters $t_n = \sigma^2\,(\alpha_n)^{\frac{v-2}{v}}$ and limiting function $\psi(z) = e^{\frac{L}{v!}\,z^v}$.

REMARK 5.1.1. The case of i.i.d. variables — Example 2.1.2 — fits in this framework, with $\alpha_n = n$ and $\beta_n = 1$. However, it includes many more examples than sums

of i.i.d. variables: in particular, in Chapter 9, we show that such bounds on cumulants typically occur in the framework of dependency graphs. Concrete examples are discussed in Chapters 10 and 11.

5.2. Precise deviations for random variables with control on cumulants

We use the same hypotheses as in the previous subsection, and without loss of generality, we suppose that $v = 3$. Then, the sequence of random variables

$$X_n = \frac{S_n}{\beta_n \, (\alpha_n)^{1/3}}$$

converges mod-Gaussian with parameters $t_n = (\alpha_n)^{1/3} \sigma^2$ and with limiting function $\psi(z) = \exp(Lz^3/6)$ (here, we may have $L = 0$). So, one can apply the previous theorems to estimate the tail of the distribution of S_n. In particular, $X_n/\sqrt{t_n} = S_n/(\beta_n \sigma \sqrt{\alpha_n})$ satisfies a central limit theorem with a normality zone of size $o(\sqrt{t_n}) = o((\alpha_n)^{1/6})$ (as for the sum of α_n i.i.d. variables) and one can describe the deviation probabilties at the edge of the normality zone — see Proposition 4.4.1.

We shall see now that, with stronger assumptions on the speed of convergence than Equation (26), we can extend these results to a larger scale. More precisely, we will assume:

$$\kappa^{(2)}(S_n) = \sigma^2 \, \alpha_n \, (\beta_n)^2 \, (1 + O((\alpha_n)^{-1/2}));$$
$$\kappa^{(3)}(S_n) = L \, \alpha_n \, (\beta_n)^3 \, (1 + O((\alpha_n)^{-1/4})). \tag{27}$$

We then have the following result:

PROPOSITION 5.2.1. *Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of centered real-valued random variables. Assume that the bound on cumulants (25) and the asymptotics of second and third cumulants given by Equation (27) hold. If x_n is a positive sequence, bounded away from 0 with $x_n = o((\alpha_n)^{1/12})$, then*

$$\mathbb{P}\left[S_n \geq \beta_n \sigma^2 \alpha_n^{2/3} x_n\right] = \mathbb{P}[X_n \geq t_n x_n] = \frac{e^{-\frac{(x_n)^2 (\alpha_n)^{1/3} \sigma^2}{2}}}{x_n (\alpha_n)^{1/6} \sigma \sqrt{2\pi}} \, e^{\frac{L(x_n)^3}{6}} \, (1 + o(1)).$$

$$\mathbb{P}\left[S_n \geq \beta_n \sigma^2 \alpha_n^{2/3} x_n\right] = \mathbb{P}[X_n \leq -t_n x_n] = \frac{e^{-\frac{(x_n)^2 (\alpha_n)^{1/3} \sigma^2}{2}}}{x_n (\alpha_n)^{1/6} \sigma \sqrt{2\pi}} \, e^{\frac{-L(x_n)^3}{6}} \, (1 + o(1)).$$

REMARK 5.2.2. The case where x_n is a constant sequence equal to x corresponds to Equations (22) and (23), which gives an equivalent for the deviation probability at the edge of the normality zone. Hence, the proposition asserts that, with appropriate assumptions on cumulants, this result is valid at a larger scale. Namely, we give an equivalent for the deviation probability of X_n of order up to $o((t_n)^{5/4})$, instead of the usual $O(t_n)$.

PROOF. Set $X_n = (\alpha_n)^{-1/3} S_n$; up to a renormalisation of the random variables, one can suppose $\beta_n = 1$, and also $\sigma^2 = 1$. Let z_n be a sequence of complex numbers with

$|z_n| = O((\alpha_n)^{1/12})$; we set $\eta_n = |z_n| (\alpha_n)^{-1/12}$. Then, following the computation of Section 5.1 with $v = 3$, we get:

$$
\log \varphi_{X_n}(z_n) = \frac{\kappa^{(2)}(S_n)}{2 (\alpha_n)^{\frac{2}{3}}} (z_n)^2 + \frac{\kappa^{(3)}(S_n)}{6 \alpha_n} (z_n)^3 + \sum_{r=4}^{\infty} \frac{\kappa^{(r)}(S_n)}{r! (\alpha_n)^{\frac{r}{3}}} (z_n)^r
$$

$$
= \frac{1}{2} (\alpha_n)^{\frac{1}{3}} (z_n)^2 + \frac{L}{6} (z_n)^3 + O((\eta_n)^2 + (\eta_n)^3 + (\eta_n)^4).
$$

If $|z_n| = o((\alpha_n)^{1/12})$, then $\eta_n \to 0$, so the remainder above is $o(1)$ and we have:

$$
\varphi_{X_n}(z_n) = \exp \left((\alpha_n)^{\frac{1}{3}} \frac{(z_n)^2}{2} + \frac{L(z_n)^3}{6} \right) (1 + o(1)). \tag{28}
$$

We make the change of probability measure

$$
\mathbb{P}[Y_n \in dy] = \frac{e^{x_n y}}{\varphi_{X_n}(x_n)} \mathbb{P}[X_n \in dy]
$$

with $x_n = o((\alpha_n)^{1/12})$; the generating function of Y_n is $\varphi_{Y_n}(z) = \frac{\varphi_{X_n}(x_n+z)}{\varphi_{X_n}(x_n)}$. So, using the inequality

$$
|(z + x_n)^r - (x_n)^r| \leq 2^r |z| \max(|z|, |x_n|)^{r-1},
$$

we get, setting $\eta_n^z = |z| (\alpha_n)^{-1/12}$ and $\eta_n^{x,z} = \max(|x_n|, |z|) (\alpha_n)^{-1/12}$,

$$
\log \varphi_{Y_n}(z) = (\alpha_n)^{1/3} \frac{(z + x_n)^2 - (x_n)^2}{2} + \frac{L((z + x_n)^3 - (x_n)^3)}{6}
$$

$$
+ O(\eta_n^z (\eta_n^{x,z} + (\eta_n^{x,z})^2 + (\eta_n^{x,z})^3))
$$

$$
= \left((\alpha_n)^{1/3} x_n + \frac{L (x_n)^2}{2} \right) z + \left(\frac{(\alpha_n)^{1/3} + L x_n}{2} \right) z^2 + \frac{L}{6} z^3
$$

$$
+ O(\eta_n^z (\eta_n^{x,z} + (\eta_n^{x,z})^2 + (\eta_n^{x,z})^3)) \tag{29}
$$

Thus, if

$$
Z_n = Y_n - (\alpha_n)^{1/3} x_n - \frac{L (x_n)^2}{2},
$$

then the sequence $(Z_n)_{n \in \mathbb{N}}$ converges in the mod-Gaussian sense, with parameters $t_n = (\alpha_n)^{1/3} + L x_n$ and limiting function $\exp(\frac{L z^3}{6})$. Moreover, in Equation (29), the approximation is valid for any z such that $|z| \leq \Delta (\alpha_n)^{1/12}$, for some constant Δ depending only on the constant C in the bound (25).

Besides, for $x_n = o((\alpha_n)^{1/12})$, one has

$$
\mathbb{P}\left[S_n \geq x_n (\alpha_n)^{\frac{2}{3}} \right] = \mathbb{P}\left[X_n \geq x_n (\alpha_n)^{\frac{1}{3}} \right] = \varphi_{X_n}(x_n) \int_{y=x_n (\alpha_n)^{1/3}}^{\infty} e^{-x_n y} \, \mathbb{P}[Y_n \in dy]
$$

$$
= \varphi_{X_n}(x_n) \, e^{-\left((\alpha_n)^{1/3} (x_n)^2 + \frac{L (x_n)^3}{2} \right)} \int_{z=-\frac{L (x_n)^2}{2}}^{\infty} e^{-x_n z} \, \mathbb{P}[Z_n \in dz]
$$

$$
= \exp \left(-\frac{(\alpha_n)^{1/3} (x_n)^2}{2} - \frac{L(x_n)^3}{3} \right) R_n (1 + o(1)),
$$

by replacing $\varphi_{X_n}(x_n)$ by its estimate (28), which holds since $x_n = o((\alpha_n)^{1/12})$; R_n is the integral of the second line.

To estimate the integral R_n, we shall adapt the proof of Proposition 4.1.1 to the special case of a sequence $(Z_n)_{n\in\mathbb{N}}$ that converges in the mod-Gaussian sense, with parameters t_n, limit function $\exp(Kz^3)$, and with the approximation

$$\log \varphi_{Z_n}(z) = \frac{t_n z^2}{2} + K z^3 + O\left(\frac{z}{(t_n)^{1/4}}\right) \tag{30}$$

that is valid for every $|z| \leq \Delta(t_n)^{1/4}$ with $\Delta > 0$. Notice that the sequence $(Z_n)_{n\in\mathbb{N}}$ previously constructed satisfies these hypotheses with $t_n = (\alpha_n)^{1/3} + Lx_n \simeq (\alpha_n)^{1/3}$. If one applies Proposition 4.1.1 to the case of mod-Gaussian convergence with a limit $\exp(Kz^3)$, then $\eta'''(0) = \psi'(0) = 0$, so the approximation of the law $dF_n(w)$ of $Z_n/\sqrt{t_n}$ is simply the Gaussian law $dG(w) = (2\pi)^{-1/2}\, e^{-w^2/2}\, dw$, and the Kolmogorov distance between F_n and G is a $o((t_n)^{-1/2})$. However, by using the validity of the approximation (30) on a larger scale than $z = O(1)$, it is possible to obtain a better Berry-Esseen bound, namely, $O((t_n)^{-3/4})$.

Recall that for any $T > 0$ and any $w \in \mathbb{R}$, the distance between cumulative distribution functions is smaller than

$$|F_n(w) - G(w)| \leq \frac{1}{\pi} \int_{-T}^{T} \left|\frac{f_n^*(\zeta) - g^*(\zeta)}{\zeta}\right| d\zeta + \frac{24m}{\pi T}.$$

However, for any $\zeta = O((t_n)^{3/4})$, one has

$$\left|\frac{f_n^*(\zeta) - g^*(\zeta)}{\zeta}\right| = e^{-\frac{\zeta^2}{2}} \left|\frac{\exp\left(\frac{K\zeta^3}{(t_n)^{3/2}} + O\left(\frac{\zeta}{(t_n)^{3/4}}\right)\right) - 1}{\zeta}\right|$$

$$e^{\frac{\zeta^2}{2}} \left|\frac{f_n^*(\zeta) - g^*(\zeta)}{\zeta}\right| \leq \left|e^{\frac{K\zeta^3}{(t_n)^{3/2}}}\right| \left|\frac{\exp\left(O\left(\frac{\zeta}{(t_n)^{3/4}}\right)\right) - 1}{\zeta}\right| + \left|\frac{\exp\left(\frac{K\zeta^3}{(t_n)^{3/2}}\right) - 1}{\zeta}\right|$$

$$\leq e^{\frac{K|\zeta|^3}{(t_n)^{3/2}}} O\left(\frac{1}{(t_n)^{3/4}} + \frac{\zeta^2}{(t_n)^{3/2}}\right).$$

In these inequalities, the constant hidden in the big O can be chosen uniform if $\frac{\zeta}{(t_n)^{3/4}}$ stays in a bounded, sufficiently small interval $[-\Delta, \Delta]$. As a consequence, setting $T = \Delta\,(t_n)^{3/4}$, one obtains from Feller's lemma:

$$|F_n(w) - G(w)| \leq O\left(\int_{-\Delta\,(t_n)^{3/4}}^{\Delta\,(t_n)^{3/4}} \left(\frac{1}{(t_n)^{3/4}} + \frac{\zeta^2}{(t_n)^{3/2}}\right) e^{-\frac{\zeta^2}{2} + \frac{K|\zeta|^3}{(t_n)^{3/2}}} d\zeta + \frac{1}{\Delta\,(t_n)^{3/4}}\right)$$

uniformly in w. Since $|\zeta|$ stays smaller than $\Delta\,(t_n)^{3/4}$, in this integral,

$$-\frac{\zeta^2}{2} + \frac{K|\zeta|^3}{(t_n)^{3/2}} \leq -\frac{\zeta^2}{2}\left(1 - \frac{K\Delta}{(t_n)^{3/4}}\right) \leq -\frac{\zeta^2}{4}$$

for t_n large enough. As claimed before, it follows that

$$\sup_{w\in\mathbb{R}} |F_n(w) - G(w)| = O\left(\frac{1}{(t_n)^{3/4}}\right).$$

We can now compute the asymptotics of the integral R_n. Set

$$\varepsilon_n = -\frac{L(x_n)^2}{2\sqrt{(\alpha_n)^{1/3} + Lx_n}} = -\frac{L(x_n)^2}{2\sqrt{t_n}};$$

since $x_n = o((\alpha_n)^{1/12})$, $\varepsilon_n \to 0$. Now,

$$\begin{aligned}
R_n &= \int_{w=\varepsilon_n}^{\infty} \exp\left(-w\, x_n \sqrt{t_n}\right)\, dF_n(w) \\
&= x_n \sqrt{t_n} \int_{\varepsilon_n}^{\infty} \exp\left(-w\, x_n \sqrt{t_n}\right)\, (F_n(w) - F_n(\varepsilon_n))\, dw \\
&= x_n \sqrt{t_n} \int_{\varepsilon_n}^{\infty} \exp\left(-w\, x_n \sqrt{t_n}\right)\, \left(G_n(w) - G_n(\varepsilon_n) + O\left(\frac{1}{(t_n)^{3/4}}\right)\right)\, dw \\
&= \frac{1}{\sqrt{2\pi}} \int_{\varepsilon_n}^{\infty} \exp\left(-w\, x_n \sqrt{t_n} - \frac{w^2}{2}\right)\, dw + O\left(\frac{e^{\frac{L(x_n)^3}{2}}}{(t_n)^{3/4}}\right).
\end{aligned}$$

The last Gaussian integral is given by Lemma 2.3.1, (3):

$$\begin{aligned}
\int_{\varepsilon_n}^{\infty} \exp\left(-w\, x_n \sqrt{t_n} - \frac{w^2}{2}\right)\, dw &= e^{\frac{t_n(x_n)^2}{2}} \int_0^{\infty} \exp\left(-\frac{(y + \varepsilon_n + x_n\sqrt{t_n})^2}{2}\right)\, dy \\
&= \frac{e^{\frac{L(x_n)^3}{2}}}{\varepsilon_n + x_n\sqrt{t_n}}(1 + o(1)) = \frac{e^{\frac{L(x_n)^3}{2}}}{x_n\sqrt{t_n}}(1 + o(1))
\end{aligned}$$

since $\varepsilon_n \to 0$. Since $x_n = o((\alpha_n)^{1/12}) = o((t_n)^{1/4})$, $\frac{1}{x_n\sqrt{t_n}}$ becomes much larger than $O(\frac{1}{(t_n)^{3/4}})$ as t_n goes to infinity, so finally:

$$R_n = \frac{1}{\sqrt{2\pi}} \frac{e^{\frac{L(x_n)^3}{2}}}{x_n(\alpha_n)^{1/6}}(1 + o(1))$$

as $t_n \simeq (\alpha_n)^{1/3}$. Gathering everything, we get

$$\mathbb{P}\left[S_n \geq x_n\,(\alpha_n)^{\frac{2}{3}}\right] = \frac{e^{-\frac{(x_n)^2(\alpha_n)^{1/3}}{2}}}{x_n(\alpha_n)^{1/6}\sqrt{2\pi}}\, e^{\frac{L(x_n)^3}{6}}(1 + o(1)),$$

and this ends the proof if $\beta_n = \sigma^2 = 1$ (set $T = x_n\,(\alpha_n)^{2/3}$ in the statement of the Proposition). In the general case, it suffices to replace S_n by $\frac{S_n}{\sigma \beta_n}$, which changes L into $\frac{L}{\sigma^3}$ in the previous computations. $\qquad \square$

REMARK 5.2.3. The argument which allows one to get a better Berry-Esseen estimate than in Proposition 4.1.1 can be used in a very general setting of mod-stable convergence, in order to get sharp bounds on the Kolmogorov distance. This will be the main topic of the forthcoming paper [FMN16].

5.3. Link with the Cramér-Petrov expansion

Proposition 5.2.1 hints at a possible expansion of the fluctuations up to any order $T = o((\alpha_n)^{1-\varepsilon})$, and indeed, it is a particular case of the results given by Rudzkis, Saulis and Statulevičius in [RSS78, SS91], see in particular [SS91, Lemma 2.3]. Suppose that

$$|\kappa^{(r)}(S_n)| \leq (Cr)^r \, \alpha_n \, (\beta_n)^r \quad ; \quad \kappa^{(r)}(S_n) = K(r) \, \alpha_n \, (\beta_n)^r \, (1 + O((\alpha_n)^{-1}))$$

the second estimate holding for any $r \leq v$; we denote $\sigma^2 = K(2)$. In this setting, one can push the expansion up to order $o((\alpha_n)^{1-1/v})$. Indeed, define recursively for a sequence of cumulants $(\kappa^{(r)})_{r \geq 2}$ the coefficients of the Cramér-Petrov series $\lambda^{(r)} = -b_{r-1}/r$, with

$$\sum_{r=1}^{j} \frac{\kappa^{(r+1)}}{r!} \left(\sum_{\substack{j_1 + \cdots + j_r = j \\ j_i \geq 1}} b_{j_1} b_{j_2} \cdots b_{j_r} \right) = \mathbb{1}_{j=1}.$$

For instance, $\lambda^{(2)} = -\frac{1}{2}$, $\lambda^{(3)} = \frac{\kappa^{(3)}}{6}$, $\lambda^{(4)} = \frac{\kappa^{(4)} - 3(\kappa^{(3)})^2}{24}$, *etc.* The appearance of these coefficients can be guessed by trying to push the previous technique to higher order; in particular, the simple form of $\lambda^{(3)}$ is related to the fact that the only term in z^3 in the expansion (29) is $\frac{\kappa^{(3)}}{6}$. If for the cumulants $\kappa^{(r)}$'s one has estimates of order $(\alpha_n)^{1-r/2}(1 + O((\alpha_n)^{-1}))$, then one has the same estimates for the $\lambda^{(r)}$'s, so there exists coefficients $L(r)$ such that

$$\lambda^{(r)} \left(\frac{S_n}{\sigma \beta_n (\alpha_n)^{\frac{1}{2}}} \right) = L(r) \, (\alpha_n)^{1-r/2} \, (1 + O((\alpha_n)^{-1})).$$

Take then $T = x_n \, (\alpha_n)^{\frac{v-1}{v}}$ with $x_n = O(1)$; Lemma 2.3 of [SS91] ensures that

$$\mathbb{P}\left[\frac{S_n}{\sigma \beta_n} \geq T \right] = \frac{e^{-\frac{T^2}{2\alpha_n}}}{\sqrt{2\pi \frac{T^2}{\alpha_n}}} \exp\left(\sum_{r=3}^{v} \lambda^{(r)} \left(\frac{T}{\sigma(\alpha_n)^{1/2}} \right)^r \right) (1 + o(1))$$

$$= \frac{e^{-\frac{T^2}{2\alpha_n}}}{\sqrt{2\pi \frac{T^2}{\alpha_n}}} \exp\left(\sum_{r=3}^{v} \frac{L(r) \, T^r}{\sigma^r \, (\alpha_n)^{r-1}} \right) (1 + o(1)).$$

Thus, the method of cumulants of Rudzkis, Saulis and Statulevičius can be thought of as a particular case (and refinement in this setting) of the notion of mod-ϕ convergence. However, their works do not yield a bound

$$|\kappa^{(r)}(S_n)| \leq (Cr)^r \, \alpha_n \, (\beta_n)^r$$

but for simple cases, such as sums of i.i.d. random variables. In Chapter 9, we show that dependency graphs are an adequate framework to provide such bounds.

CHAPTER 6

A precise version of the Ellis-Gärtner theorem

In the classical theory of large deviations, asymptotic results are formulated not only for the probabilities of tails $\mathbb{P}[X_n \geq t_n x]$, but more generally for probabilities

$$\mathbb{P}[X_n \in t_n B] \quad \text{with } B \text{ arbitrary Borelian subset of } \mathbb{R}.$$

In particular, under some technical assumptions on the generating series (that look like, but are somehow weaker than mod-convergence), Ellis-Gärtner theorem provides some asymptotic upper and lower bounds for $\log(\mathbb{P}[X_n \in t_n B])$, these bounds relying on a limiting condition on $(t_n)^{-1} \log \varphi_n(\cdot)$. When the topology of B is *nice* enough, these bounds coincide (see *e.g.* [DZ98, Theorem 2.3.6]). This generalises Cramér's large deviations for sums of i.i.d. random variables.

Our Theorems 3.2.2 and 4.2.1 give estimates for the probabilities $\mathbb{P}[X_n \geq t_n x]$ themselves, instead of their logarithm). Therefore, it is natural to establish in the framework of mod-convergence a precise version of Ellis-Gärtner theorem. In this chapter, we shall give some asymptotic upper and lower bounds for the probabilities $\mathbb{P}[X_n \in t_n B]$ itself instead of their logarithms. Once again, the upper and lower bounds coincide for *nice* borelian sets B.

REMARK 6.0.1. In [FMN15], we shall prove similar estimates of $\mathbb{P}[\mathbf{X}_n \in t_n B]$ in the setting of sequences of random *vectors* that converge in the multi-dimensional mod-Gaussian sense.

6.1. Technical preliminaries

In this section, we make the following assumptions:

(1) The random variables X_n satisfy the hypotheses of Definition 1.1.1 with $c = +\infty$ (in particular, ψ is entire on $\mathbb{C}$).

(2) The Legendre-Fenchel transform F is essentially smooth, that is to say that it takes finite values on a non-empty closed interval I_F and that $\lim F'(x) = \lim h = \pm\infty$ when x goes to a bound of the interval I_F (*cf.* [DZ98, Definition 2.3.5]).

The latter point is verified if ϕ is a Gaussian or Poisson law, which are the most important examples.

LEMMA 6.1.1. *Let C be a closed subset of $\mathbb{R}$. Either $\inf_{u \in C} F(u) = +\infty$, or $\inf_{u \in C} F(u) = m$ is attained and $\{x \in C \mid F(x) = \min_{u \in C} F(u)\}$ consists of one or two real numbers $a \leq b$, with $a < \eta'(0) < b$ if $a \neq b$.*
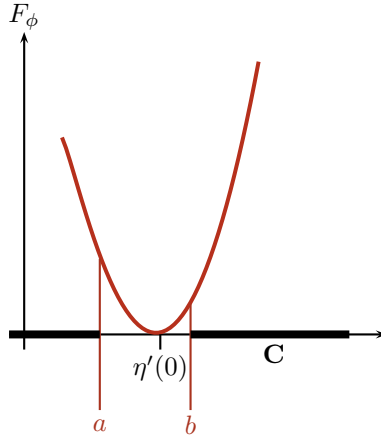
FIGURE 1. The infimum of F on an admissible closed set C is attained either at $a = \sup(C \cap (-\infty, \eta'(0)])$, or at $b = \inf(C \cap [\eta'(0), +\infty))$, or at both if $F(a) = F(b)$.

PROOF. Recall that F is strictly convex, since its second derivative is $1/\eta''(h)$, which is the inverse of the variance of a non-constant random variable. Also, $\eta'(0)$ is the point where F attains its global minimum, and it is the expectation of the law ϕ. If $C \cap I_F = \varnothing$, then $F_{|C} = +\infty$ and we are in the first situation. Otherwise, $F_{|C}$ is finite at some points, so there exists $M \in \mathbb{R}_+$ such that $C \cap \{x \in \mathbb{R} \mid F(x) \leq M\} \neq \varnothing$. However, the set $\{x \in \mathbb{R} \mid F(x) \leq M\}$ is compact by the hypothesis of essential smoothness: it is closed as the reciprocal image of an interval $]-\infty, M]$ by a lower semi-continuous function, and bounded since $\lim_{x \to (I_F)^c} |F'(x)| = +\infty$. So, $C \cap \{x \in \mathbb{R} \mid F(x) \leq M\}$ is a non-empty compact set, and the lower semi-continuous F attains its minimum on it, which is also $\min_{u \in C} F(u)$. Then, if $a \leq b$ are two points in C such that $F(a) = F(b) = \min_{u \in C} F(u)$, then by strict convexity of F, $F(x) < F(a)$ for all $x \in (a, b)$, hence, $(a, b) \subset C^c$. Also, $F(x) > F(a)$ if $a \neq b$ and $x \notin [a, b]$, so either $a = b$, or $\eta'(0) \in (a, b)$. □

We take the usual notations B^o and $\overline{B}$ for the interior and the closure of a subset $B \subset \mathbb{R}$. Call *admissible* a (Borelian) subset $B \subset \mathbb{R}$ such that there exists $b \in B$ with $F(b) < +\infty$, and denote then

$$F(B) = \inf_{u \in B} F(u) = \min_{u \in \overline{B}} F(u),$$

and $B_{\min} = \{a \in \overline{B} \mid F(a) = F(B)\}$; according to the previous discussion, $B_{\min}$ consists of one or two elements.

6.2. A precise upper bound

THEOREM 6.2.1. *Let B be a Borelian subset of $\mathbb{R}$.*

(1) If B is admissible, then

$$\limsup_{n \to \infty} \left(\sqrt{2\pi t_n} \, \exp(t_n F(B)) \, \mathbb{P}[X_n \in t_n B] \right) \leq \begin{cases} \sum_{a \in B_{\min}} \dfrac{\psi(h(a))}{(1 - e^{-|h(a)|}) \sqrt{\eta''(h(a))}} \\[2ex] \sum_{a \in B_{\min}} \dfrac{\psi(h(a))}{|h(a)| \sqrt{\eta''(h(a))}} \end{cases}$$

the distinction of cases corresponding to ϕ lattice or non-lattice distributed. The sum on the right-hand side consists in one or two terms — it is considered infinite if $a = \eta'(0) \in B_{\min}$.

(2) *If B is not admissible, then for any positive real number M,*

$$\lim_{n \to \infty} \left(\exp(t_n M) \, \mathbb{P}[X_n \in t_n B] \right) = 0.$$

PROOF. For the second part, one knows that $\varphi_n(x) \exp(-t_n \eta(x))$ converges to $\psi(x)$ which does not vanish on the real line, so by taking the logarithms,

$$\lim_{n \to \infty} \frac{\log \varphi_n(x)}{t_n} = \eta(x).$$

Then, Ellis-Gärtner theorem holds since F is supposed essentially smooth. So,

$$\limsup_{n \to \infty} \frac{\log \mathbb{P}[X_n \in t_n B]}{t_n} \leq -F(B),$$

and if B is not admissible, then the right-hand side is $-\infty$ and (2) follows immediately.

For the first part, suppose for instance ϕ non-lattice distributed. Take C a closed admissible subset, and assume $\eta'(0) \notin C$ — otherwise the upper bound in (1) is $+\infty$ and the inequality is trivially satisfied. Since C^c is an open set, there is an open interval $(a, b) \subset C^c$ containing $\eta'(0)$, and which we can suppose maximal. Then a and b are in C as soon as they are finite, and $C \subset (-\infty, a] \sqcup [b, +\infty)$. Moreover, by strict convexity of F, the minimal value $F(C)$ is necessarily attained at a or b. Suppose for instance $F(a) = F(b) = F(C)$ — the other situations are entirely similar. Then,

$$\mathbb{P}[X_n \in t_n C] \leq \mathbb{P}[X_n \leq t_n a] + \mathbb{P}[X_n \geq t_n b]$$

$$\lesssim \exp(-t_n F(C)) \left(\frac{\psi(h(a))}{-h(a)\sqrt{2\pi t_n \eta''(h(a))}} + \frac{\psi(h(b))}{h(b)\sqrt{2\pi t_n \eta''(h(b))}} \right)$$

by using Theorem 4.2.1 for $\mathbb{P}[X_n \geq t_n b]$, and also for $\mathbb{P}[X_n \leq t_n a] = \mathbb{P}[-X_n \geq -t_n a]$ — the random variables $-X_n$ satisfy the same hypotheses as the X_n's with $\eta(x)$ replaced by $\eta(-x)$, $\psi(x)$ replaced by $\psi(-x)$, *etc.* This proves the upper bound when B is closed, and since $F(B) = F(\overline{B})$ by lower semi-continuity of F and $B_{\min} = (\overline{B})_{\min}$, the result extends immediately to arbitrary admissible Borelian subsets. $\qquad\square$

6.3. A precise lower bound

One can then ask for an asymptotic lower bound on $\mathbb{P}[X_n \in t_n B]$, and in view of the classical theory of large deviations, this lower bound should be related to open sets and to the exponent $F(B^\circ)$. Unfortunately, the result takes a less interesting form than Theorem 6.2.1. If B is a Borelian subset of $\mathbb{R}$, denote B^δ the union of the open intervals $(x, x + \kappa)$ of width $\kappa \geq \delta$ that are included into B. The interior $O = B^\circ$ is a disjoint union of a countable collection of open intervals, and also the increasing union $\bigcup_{\delta > 0} B^\delta$.
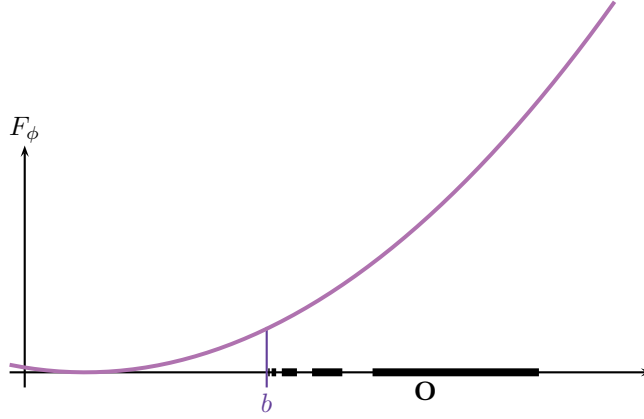
F_ϕ

b

O

FIGURE 2. In some problematic situations, one is only able to prove a non-precise lower bound for large deviations.

However, the topology of B^o may be quite intricate in comparison to the one of the B^δ's, as some points can be points of accumulation of open intervals included in B and of width going to zero (see Figure 2). This phenomenon prevents us to state a precise lower bound when one of this point of accumulation is $a = \sup(B^o \cap (-\infty, \eta'(0)])$ or $b = \inf(B^o \cap [\eta'(0), +\infty))$. Nonetheless, the following is true:

THEOREM 6.3.1. *For an admissible Borelian set B,*

$$\liminf_{\delta \to 0} \liminf_{n \to \infty} \left(\sqrt{2\pi t_n} \, \exp(t_n F(B^\delta)) \, \mathbb{P}[X_n \in t_n B] \right) \geq \begin{cases} \sum_{a \in (B^o)_{\min}} \dfrac{\psi(h(a))}{(1-e^{-|h(a)|}) \sqrt{\eta''(h(a))}} \\ \sum_{a \in (B^o)_{\min}} \dfrac{\psi(h(a))}{|h(a)| \sqrt{\eta''(h(a))}}, \end{cases}$$

with the usual distinction of lattice/non-lattice cases. In particular, the right-hand side in Theorem 6.2.1 is the limit of $\sqrt{2\pi t_n} \, \exp(t_n F(B)) \, \mathbb{P}[X_n \in t_n B]$ as soon as $F(B^\delta) = F(B)$ for some $\delta > 0$.

PROOF. Again we deal with the non-lattice case, and we suppose for instance that the set $(B^o)_{\min}$ consists of one point $b = \inf(B^o \cap [\eta'(0), +\infty))$, the other situations being entirely similar. As δ goes to 0, B^δ increases towards $B^o = \bigcup_{\delta > 0} B^\delta$, so the infimum $F(B^\delta)$ decreases and the quantity

$$L(\delta) = \liminf_{n \to \infty} \left(\sqrt{2\pi t_n} \, \exp(t_n F(B^\delta)) \, \mathbb{P}[X_n \in t_n B] \right)$$

is decreasing in δ. Actually, if $b^\delta = \inf(B^\delta \cap [\eta'(0), +\infty))$, then for δ small enough $F(B^\delta) = F(b^\delta)$, so $\lim_{\delta \to 0} F(B^\delta) = F(B^o)$ by continuity of F. On the other hand,

$$R(\delta) = \frac{\psi(h(b^\delta))}{h(b^\delta) \sqrt{\eta''(h(b^\delta))}}$$

tends to the same quantity with b instead of b^δ. Hence, it suffices to show that for δ small enough, $L(\delta) \geq R(\delta)$. However, by definition of B^δ, the open interval $(b^\delta, b^\delta + \delta)$

is included into B, so

$$\mathbb{P}[X_n \in t_n B] \geq \mathbb{P}[X_n \in t_n B^\delta] \geq \mathbb{P}[X_n > t_n b^\delta] - \mathbb{P}[X_n \geq t_n (b^\delta + \delta)]$$

$$\geq \left(\frac{\psi(h(b^\delta)) \, e^{-t_n F(b^\delta)}}{h(b^\delta) \sqrt{2\pi t_n \eta''(h(b^\delta))}} - \frac{\psi(h(b^\delta + \delta)) \, e^{-t_n F(b^\delta + \delta)}}{h(b^\delta + \delta) \sqrt{2\pi t_n \eta''(h(b^\delta + \delta))}} \right) (1 + o(1))$$

$$\geq \frac{\psi(h(b^\delta)) \, e^{-t_n F(b^\delta)}}{h(b^\delta) \sqrt{2\pi t_n \eta''(h(b^\delta))}} (1 + o(1))$$

since the second term on the second line is negligible in comparison to the first term —
$F(b^\delta + \delta) > F(b^\delta)$. This ends the proof. $\qquad\square$

CHAPTER 7

Examples with an explicit generating function

The general results of Chapters 3 and 6 can be applied in many contexts, and the main difficulty is then to prove for each case that one has indeed the estimate on the Laplace transform given by Definition 1.1.1. Therefore, the development of techniques to obtain mod-ϕ estimates is an important part of the work. Such an estimate can sometimes be established from an explicit expression of the Laplace transform (hence of the characteristic function); we give several examples of this kind in Section 7.1. But there also exist numerous techniques to study sequences of random variables without explicit expression for the characteristic function: complex analysis methods in number theory (Section 7.2) and in combinatorics (Section 7.3), localisation of zeros (Chapter 8) and dependency graphs (Chapters 9, 10 and 11) to name a few. These methods are known to yield central limit theorems and we show how they can be adapted to prove mod-convergence. We illustrate each case with one or several example(s).

In this chapter, we detail examples for which the mod-ϕ convergence has already been proved before (*cf.* [JKN11, DKN15]) or follows easily from formulas in the literature.

7.1. Mod-convergence from an explicit formula for the Laplace transform

Examples where mod-convergence is proved by using an explicit formula for the Laplace transform were already given in Section 2.1. We provide here two additional examples of mod-Gaussian convergence that can be obtained by this method.

EXAMPLE 7.1.1. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be a random analytic function, where the coefficients a_n are independent standard complex Gaussian variables. The random function f has almost surely its radius of convergence equal to 1, and its set of zeroes $\mathcal{Z}(f) = \{z \in \mathcal{D}_{(0,1)} \mid f(z) = 0\}$ is a determinantal point process on the unit disk, with kernel

$$\mathcal{K}(w,z) = \frac{1}{\pi(1 - w\bar{z})^2} = \frac{1}{\pi} \sum_{k=1}^{\infty} k(w\bar{z})^{k-1}.$$

We refer to [HKPV09] for precisions on these results. It follows then from the general theory of determinantal point processes, and the radial invariance of the kernel, that the number N_r of points of $\mathcal{Z}(f) \cap B_{(0,r)}$ can be represented in law as a sum of independent Bernoulli variables of parameters $\{r^{2k}\}_{k \geq 1}$:

$$N_r = \text{card}\{z \in \mathcal{Z}(f) \mid |z| \leq r\} =_{\text{law}} \sum_{k=1}^{\infty} \mathcal{B}(r^{2k}).$$

This representation as a sum of independent variables allows one to estimate the moment generating function of N_r under various renormalisations. Let us introduce the

hyperbolic area

$$h = \frac{4\pi\, r^2}{1 - r^2}$$

of $\mathcal{D}_{(0,r)}$, and denote $N_r = N^h$; we are interested in the asymptotic behavior of N^h as h goes to infinity, or equivalently as r goes to 1. Since $\mathbb{E}[e^{zN_r}] = \prod_{k=1}^{\infty}(1 + r^{2k}(e^z - 1))$, one has

$$\log\left(\mathbb{E}\left[e^{\frac{zN^h}{h^{1/3}}}\right]\right) = \sum_{k=1}^{\infty} \log\left(1 + r^{2k}\left(e^{\frac{z}{h^{1/3}}} - 1\right)\right) = \frac{h^{2/3}z}{4\pi} + \frac{h^{1/3}z^2}{16\pi} + \frac{z^3}{144\pi} + o(1)$$

with a remainder that is uniform when z stays in a compact domain of $\mathbb{C}$. Therefore,

$$\frac{N^h - \frac{h}{4\pi}}{h^{1/3}} \quad \text{converges mod-Gaussian with} \quad \begin{cases} \text{parameters } t_h = \frac{h^{1/3}}{8\pi}, \\ \text{limiting function } \psi(z) = \exp\left(\frac{z^3}{144\pi}\right). \end{cases}$$

Again, the limiting function is the exponential of a simple monomial. By Proposition 4.4.1,

$$X^h = \frac{N^h - \frac{h}{4\pi}}{\sqrt{\frac{h}{8\pi}}}$$

converges as $h \to \infty$ to a Gaussian law, with normality zone $o(h^{1/6})$. Moreover, by Theorem 4.2.1, at the edge of this normality zone,

$$\mathbb{P}\left[N^h - \frac{h}{4\pi} \geq \frac{h^{2/3}}{4\pi}\, x\right] = \frac{e^{-\frac{h^{1/3}x^2}{4\pi}}}{h^{1/6}\, x}\, \exp\left(\frac{x^3}{18\pi}\right)(1 + o(1))$$

for any $x > 0$.

EXAMPLE 7.1.2. Consider the Ising model on the discrete torus $\mathbb{Z}/n\mathbb{Z}$. Thus, we give to each spin configuration $\sigma : \mathbb{Z}/n\mathbb{Z} \to \{\pm 1\}$ a probability proportional to the factor $\exp(-\beta \sum_{i\sim j} \mathbb{1}_{\sigma(i)\neq\sigma(j)})$, the sum running over neighbours in the circular graph $\mathbb{Z}/n\mathbb{Z}$. The technique of the transfer matrix (see [Bax82, Chapter 2]) ensures that if $M_n = \sum_{i=1}^{n} \sigma(i)$ is the total magnetisation of the model, then

$$\mathbb{E}[e^{zM_n}] = \frac{\text{tr}\,(T(z))^n}{\text{tr}\,(T(0))^n}, \qquad \text{where } T(z) = \begin{pmatrix} e^{-z} & e^{z-\beta} \\ e^{-z-\beta} & e^z \end{pmatrix}.$$

The two eigenvalues of $T(z)$ are $\cosh z \pm \sqrt{(\sinh z)^2 + e^{-2\beta}}$, and their Taylor expansion shows that

$$\log\left(\mathbb{E}\left[e^{\frac{zM_n}{n^{1/4}}}\right]\right) = \frac{e^{\beta}\, n^{1/2}\, z^2}{2} - \frac{(3e^{3\beta} - e^{\beta})\, z^4}{24} + o(1).$$

So, one has mod-Gaussian convergence for $n^{-1/4}\, M_n$, and the estimates

$$\mathbb{P}[M_n \geq n^{3/4}x] = \frac{e^{-\frac{n^{1/2}\, e^{\beta}\, x^2}{2}}}{x\, \sqrt{2\pi n^{1/2}}}\, \exp\left(-\frac{(3e^{3\beta} - e^{\beta})\, x^4}{24}\right)(1 + o(1)).$$

In particular, $\frac{M_n}{n^{1/2}\, e^{\beta/2}}$ satisfies a central limit theorem with normality zone $o(n^{1/4})$.

REMARK 7.1.3. If instead of $\mathbb{Z}/n\mathbb{Z}$ we consider the graph $1 \leftrightarrow 2 \leftrightarrow \cdots \leftrightarrow n$ (*i.e.*, one removes the link $n \leftrightarrow 1$), then one can realise the spins $\sigma(i)$ of the Ising model as the n first states of a Markov chain with space of states $\{\pm 1\}$. The magnetisation M_n appears then as a linear functional of the empirical measure of this Markov chain. More generally, if $S_n = \sum_{i=1}^{n} f(X_i)$ is a linear functional of a Markov chain on a finite space, then under mild hypotheses this sum satisfies the Markov chain central limit theorem (*cf.* for instance [Cog72]). In [FMN16], we shall use arguments of the perturbation theory of operators in order to prove that one also has mod-Gaussian convergence for such linear functionals of Markov chains.

7.2. Additive arithmetic functions of random integers

In this paragraph, we sometimes write $\log(\log n) = \log_2 n$, which is negligible in comparison to $\log n$ as n goes to infinity.

7.2.1. Number of prime divisors, counted without multiplicities. Denote $\mathbb{P}$ the set of prime numbers, $\omega(k)$ the number of distinct prime divisors of an integer k, and ω_n the random variable $\omega(k)$ with k random integer uniformly chosen in $[n] := \{1, 2, \ldots, n\}$. The random variable ω_n satisfies the Erdös-Kac central limit theorem (*cf.* [EK40]):

$$\frac{\omega_n - \log\log n}{\sqrt{\log\log n}} \to \mathcal{N}_{\mathbb{R}}(0, 1).$$

In this section, we show that ω_n converges mod-Poisson and present precise deviation results for it. Such a result was established in [KN10, Section 4]. But in the latter article, mod-ϕ convergence is defined via convergence of the renormalised Fourier transform, while here we work with Laplace transform. Therefore, we need to justify that the convergence also holds for Laplace transforms.

We start from the *Dirichlet series* of $y^{\omega(k)}$, which is:

$$\sum_{k \geq 1} \frac{y^{\omega(k)}}{k^s} = \prod_{p \in \mathbb{P}} \left(1 + \frac{y}{p^s - 1}\right),$$

well-defined and absolutely convergent if $\mathrm{Re}(s) > 1$. The Selberg-Delange method allows to extract from this formula precise estimates for the generating function of ω_n; see [Ten95] and references therein.

PROPOSITION 7.2.1. *[Ten95, Section II.6, Theorem 1] For any $A > 0$, we have, for any y in $\mathbb{C}$ with $|y| \leq A$*

$$\sum_{k \leq n} y^{\omega(k)} = n (\log n)^{y-1} (\lambda_0(y) + O(1/\log n)),$$

where

$$\lambda_0(y) = \frac{1}{\Gamma(y)} \prod_{p \in \mathbb{P}} \left(1 + \frac{y}{p-1}\right) \left(1 - \frac{1}{p}\right)^y$$

and the constant hidden in the O symbol depends only on A.

REMARK 7.2.2. In fact, [Ten95, Section II.6, Theorem 1] gives a complete asymptotic expansion of $\sum_{k \leq n} y^{\omega(k)}$ in terms of powers of $1/\log(n)$. For our purpose, the first term is enough: we will see that it implies mod-Poisson convergence with a speed of convergence as precise as wanted.

Setting $y = e^z$, this can be rewritten as an asymptotic formula for the Laplace transform of ω_n.

$$\mathbb{E}[e^{z\,\omega_n}] = \frac{1}{n} \sum_{k=1}^{n} e^{z\,\omega(k)} = (\log n)^{e^z-1} \lambda_0(e^z)\, (1 + O(1/\log n))$$

$$= e^{(e^z-1)\log\log n} \lambda_0(e^z)\, (1 + O(1/\log n)).$$

Recall that the constant in the O symbol is uniform on sets $\{w, |w| \leq A\}$, that is on bands $\{z, \mathrm{Re}(z) \leq \log(A)\}$. In particular the convergence is uniform on compact sets. Therefore, one has the following result.

PROPOSITION 7.2.3. *The sequence of random variables* $(\omega_n)_{n \geq 1}$ *converges mod-Poisson with parameter* $t_n = \log_2 n$ *and limiting function* $\psi(z) = \lambda_0(e^z)$ *on the whole complex plane. This takes place with speed of convergence* $O(1/\log n)$, *that is* $O((t_n)^{-\nu})$ *for all* $\nu > 0$.

Using Theorem 3.2.2, we get immediately the following deviation result.

THEOREM 7.2.4. *Let* $x > 0$. *Assume* $x \log_2 n \in \mathbb{N}$. *Then*

$$\mathbb{P}[\omega_n = x \log_2 n] = \frac{\lambda_0(x)\,(1 + O(1/\log_2 n))}{(\log n)^{x\log(x)-x+1} \sqrt{2\pi x \log_2 n}}. \tag{31}$$

Furthermore, if $x > 1$, *then*

$$\mathbb{P}[\omega_n \geq x \log_2 n] = \frac{\lambda_0(x)\,(1 + O(1/\log_2 n))}{(\log n)^{x\log(x)-x+1} \sqrt{2\pi x \log_2 n}} \frac{1}{1 - \frac{1}{x}}. \tag{32}$$

REMARK 7.2.5. The first equation (31) is not new: it is due to Selberg [Sel54] and presented in a slightly different form than here in [Ten95, Section II.6, Theorem 4]. Note also that, as the speed of mod-Poisson convergence of ω_n is $O((t_n)^{-\nu})$ for all $\nu > 0$, Theorem 3.2.2 gives asymptotic expansions of the above probabilities, up to an arbitrarily large power of $1/\log\log n$. Theorem 4 in [Ten95, Section II.6] also gives such estimates. The second statement (32) follows from [Rad09, Theorem 2.8]; it is a nice feature of the theory of mod-ϕ convergence to allow to recover quickly such deep arithmetic results (though we still need Selberg-Delange asymptotics).

7.2.2. Number of prime divisors, counted with multiplicities. In this section, we give similar results for the number of prime divisors Ω_n of an integer chosen randomly in $\{1, 2, \ldots, n\}$, but this time *counted with multiplicities*. An important difference is that, here, the mod-Poisson convergence occurs only on a band and not on the whole complex plane $\mathbb{C}$. In this case the Dirichlet series is given by:

$$\sum_{k \geq 1} \frac{y^{\Omega(k)}}{k^s} = \prod_{p \in \mathbb{P}} \left(1 - \frac{y}{p^s}\right)^{-1},$$

which again is well-defined and absolutely convergent for $\text{Re}(s) > 1$. Note that, unlike in the case of $\omega(k)$, the right-hand side has some poles, the smallest in modulus being for $y = 2^s$. Again, a precise estimate for the generating function follows from the work of Selberg and Delange; see [Ten95] and references therein.

PROPOSITION 7.2.6. *[Ten95, Section II.6, Theorem 2] For any δ with $0 < \delta < 1$, we have, for any y in $\mathbb{C}$ with $|y| \leq 2 - \delta$*

$$\sum_{k \leq n} y^{\Omega(k)} = n \, (\log n)^{y-1} \, (\nu_0(y) + O(1/\log n)),$$

where

$$\nu_0(y) = \frac{1}{\Gamma(y)} \prod_{p \in \mathbb{P}} \left(1 - \frac{y}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^{y}$$

and the constant hidden in the O symbol depends only on δ.

Note the difference with Proposition 7.2.1: the function ν_0 has a simple pole for $y = 2$ (while λ_0 is an entire function) and the estimate in Proposition 7.2.6 holds (uniformly on compacts) *only for $|y| < 2$.*

Setting again $y = e^z$, this can be rewritten as an asymptotic formula for the Laplace transform of Ω_n on the band $\mathcal{S}_{(-\infty, \log 2)} = \{z, \, \text{Re}(z) < \log 2\}$.

$$\mathbb{E}[e^{z\, \Omega_n}] = \frac{1}{n} \sum_{k=1}^{n} e^{z\, \Omega(k)} = e^{(e^z - 1) \log \log n} \, \nu_0(e^z) \, (1 + O(1/\log n)).$$

Recall that the constant in the O symbol is uniform on sets $\{y, |y| \leq 2 - \delta\}$, that is on bands $\{z, \, \text{Re}(z) < \log(2 - \delta)\}$. In particular the convergence is uniform on compact sets. Therefore, one has the following result.

PROPOSITION 7.2.7. *The sequence of random variables $(\Omega_n)_{n \geq 1}$ converges mod-Poisson with parameter $t_n = \log_2 n$ and limiting function $\psi(z) = \nu_0(e^z)$, on the band $\mathcal{S}_{(-\infty, \log 2)}$. This takes place with speed of convergence $O(1/\log(n))$, that is $O((t_n)^{-\nu})$, for all $\nu > 0$.*

As for ω_n, this implies precise deviation results. However, as the convergence only takes place on a band, the range of these results is limited (note the condition $x < 2$ in the theorem below).

THEOREM 7.2.8. *Fix x with $0 < x < 2$. Assume $x \log_2 n \in \mathbb{N}$. Then*

$$\mathbb{P}[\Omega_n = x \log_2 n] = \frac{\nu_0(x) \, (1 + O(1/\log_2 n))}{(\log n)^{x \log(x) - x + 1} \sqrt{2\pi x \log_2 n}}. \tag{33}$$

Furthermore, if $1 < x < 2$,

$$\mathbb{P}[\Omega_n \geq x \log_2 n] = \frac{\nu_0(x) \, (1 + O(1/\log_2 n))}{(\log n)^{x \log(x) - x + 1} \sqrt{2\pi x \log_2 n}} \frac{1}{1 - \frac{1}{x}}. \tag{34}$$

Again, the first equation was discovered by Selberg [Sel54] and can be found in a slightly different form in [Ten95, Section II.6, Theorem 5].

REMARK 7.2.9. An extension of Equation (33) for $x > 2$ is given in [Ten95, Section II.6, Theorem 6]. This involves the type of singularity of the limiting function $v_0(e^z)$ at the edge of the convergence domain. Here, $v_0(y)$ has only a simple pole in $y = 2$, and the residue appears in the deviation results. It would be interesting to see if this kind of idea can be used in the general framework of mod-ϕ convergence, but this is outside the scope of this already quite long paper.

7.2.3. Other arithmetic functions. In this paragraph, we generalise the two examples ω_n and Ω_n above. The goal is to understand the phase transition between the convergence on a band and the convergence on the complex plane.

Let $f : \mathbb{N} \to \mathbb{Z}$ be a function with the following properties:

(i) f is additive, that is $f(mn) = f(m) + f(n)$ if $m \wedge n = 1$;

(ii) for every prime p, one has $f(p) = 1$.

We necessarily have $f(1) = 0$. For $C > 0$, we say that such a function *has a C-linear growth* if there exists B such that

$$|f(p^k)| \le B + C k \quad \text{for any } p \in \mathbb{P} \text{ and any } k > 0.$$

If f has a C-linear growth for any $C > 0$, we say that f has a *sublinear growth*. In particular, since $\Omega(p^k) = k$ (for all p and k), the function Ω has a 1-linear growth, while ω has a sublinear growth (for all p and k, $\omega(p^k) = 1$). We are interested in the random variable f_n, which is equal to the value $f(k)$ of f on a random integer k chosen uniformly between 1 and n. Note that the additivity condition ensures that the following Dirichlet series factorises:

$$F(y, s) := \sum_{k=1}^{\infty} \frac{y^{f(k)}}{k^s} = \prod_{p \in \mathbb{P}} \left(1 + \frac{y^{f(p)}}{p^s} + \frac{y^{f(p^2)}}{p^{2s}} + \cdots \right).$$

Again, one can use the Selberg-Delange method to get precise estimates for the Laplace transform.

PROPOSITION 7.2.10. *Suppose that f fulfills* (i) *and* (ii) *and has a C-linear groth for some $C > 0$. For any δ with $0 < \delta < \frac{1}{2}$, we have, for any y in $\mathbb{C}$ with $(2 - 2\delta)^{-1/C} \le |y| \le (2 - 2\delta)^{1/C}$,*

$$\sum_{k \le n} y^{f(k)} = n (\log n)^{y-1} (\tau_0(y) + O(1/\log n)), \tag{35}$$

where the constant hidden in the O symbol depends on δ, B and C (but not on y), and

$$\tau_0(y) = \frac{1}{\Gamma(y)} \prod_{p \in \mathbb{P}} \left(1 + \frac{y^{f(p)}}{p} + \frac{y^{f(p^2)}}{p^2} + \cdots \right) \left(1 - \frac{1}{p} \right)^y.$$

PROOF. The proof is an easy adaptation of the proof of [Ten95, Section II.6, Theorems 1 and 2]. We will only indicate the necessary modifications, assuming that the reader can consult Tenenbaum's book.

Fix $\delta \in (0, \frac{1}{2})$, and y in $\mathbb{C}$ with $(2 - 2\delta)^{-1/C} \le |y| \le (2 - 2\delta)^{1/C}$. We set $Y = \max(|y|, |y|^{-1})$; by assumption, $1 \le Y \le (2 - 2\delta)^{1/C}$, and on the other hand,

$$|y^{f(p^k)}| \le Y^{B + Ck}$$

as f has a C-linear growth. We have, for any prime p and $|\xi| < Y^{-C}$, the bound

$$\sum_{k \geq 0} |y^{f(p^k)} \xi^k| < 1 + \sum_{k \geq 1} Y^{B+Ck}|\xi|^k \leq 1 + Y^B \frac{Y^C|\xi|}{1 - Y^C|\xi|}.$$

Therefore, for each prime p, the following formula defines a holomorphic function for $|\xi| < Y^{-C} \leq 1$:

$$h_{y,p}(\xi) = \left(1 + y^{f(p)}\xi + y^{f(p^2)}\xi^2 + \dots\right)(1 - \xi)^y.$$

Moreover, we have

$$|h_{y,p}(\xi)| < \left(1 + Y^B \frac{Y^C|\xi|}{1 - Y^C|\xi|}\right) e^{Y|\log(1-\xi)|}.$$

Consider now the coefficients of the power series expansion of $h_{y,p}$ around the origin, i.e. the numbers $b_y(p^k)$ such that

$$h_{y,p}(\xi) = 1 + \sum_{k \geq 1} b_y(p^k)\xi^k.$$

Note that $b_y(p) = 0$ for any $p \in \mathbb{P}$. With the previous hypotheses,

$$\frac{1}{2 - \delta} < Y^{-C} \leq |y|^{-C}.$$

By Cauchy's formula applied to the circle of radius $1/(2 - \delta)$, one has, for any prime p, any complex number y such that $(2 - 2\delta)^{1/C} \leq |y| \leq (2 - 2\delta)^{1/C}$, and any $k \geq 1$:

$$|b_w(p^k)| \leq M(\delta, B, C)(2 - \delta)^k,$$

where

$$M(\delta, B, C) = \sup_{\substack{|\xi|=1/(2-\delta) \\ 1 \leq Y \leq (2-2\delta)^{1/C}}} \left(1 + Y^B \frac{Y^C|\xi|}{1 - Y^C|\xi|}\right) e^{Y|\log(1-\xi)|} < +\infty.$$

Denote ζ the Riemann zeta function. The same arguments as in the proof of [Ten95, Section II.6, Theorem 1] shows that

$$G(y,s) := F(y,s)\,\zeta(s)^{-y} = \prod_{p \in \mathbb{P}} \left(1 + \sum_{k \geq 2} \frac{b_y(p^k)}{p^{ks}}\right)$$

is convergent and uniformly bounded for $\mathrm{Re}(s) > 3/4$ and y such that $(2 - 2\delta)^{-1/C} \leq |y| \leq (2 - 2\delta)^{1/C}$. Therefore, we can apply [Ten95, Section II.5, Theorem 3], which gives the asymptotic formula (35). $\qquad \square$

Setting $y = e^z$, Proposition 7.2.10 can be rewritten as an asymptotic formula for the Laplace transform of f_n:

$$\mathbb{E}[e^{z f_n}] = \frac{1}{n} \sum_{k=1}^{n} e^{z f(k)} = e^{(e^z - 1) \log \log n} \left(\tau_0(e^z) + O(1/\log n)\right),$$

where the constant hidden in O depends on δ, B and C, but is uniform on $|\mathrm{Re}(z)| \leq (\log(2 - 2\delta))/C$. This yields the following mod-convergence result:

PROPOSITION 7.2.11. *Assume f is an arithmetic function that fulfills* (i) *and* (ii) *and has a C-linear growth for some constant $C > 0$. Then, the sequence of random variables $(f_n)_{n \geq 1}$ converges mod-Poisson with parameter $t_n = \log_2 n$ and limiting function $\psi(z) = \tau_0(e^z)$ on the band*

$$\mathcal{S}_{(-\frac{\log 2}{C}, \frac{\log 2}{C})}.$$

This takes place with speed of convergence $O(1/\log n)$, that is $O((t_n)^{-\nu})$ for all $\nu > 0$.

As a consequence, if f has a sublinear growth, then the convergence takes place on the whole complex plane.

It is then straightforward to write deviation results, analoguous to Equations (32) and (34), for a generic function f as above.

REMARK 7.2.12. If the additive function f is non-negative, as is the case for $\omega(n)$ and $\Omega(n)$, then it is easily seen that one can perform the whole proof of Proposition 7.2.10 assuming only $|y| \leq (2 - 2\delta)^{1/C}$ (but not $|y| \geq (2 - 2\delta)^{-1/C}$). Therefore, one has in this case a mod-Poisson convergence on the band

$$\mathcal{S}_{(-\infty, \frac{\log 2}{C})} \quad \text{instead of} \quad \mathcal{S}_{(-\frac{\log 2}{C}, \frac{\log 2}{C})}.$$

7.3. Number of cycles in weighted probability measure

We consider here the number of cycles of random permutations under the so-called weighted probability measure. This example was already considered in [NZ13] and we follow the presentation of this article.

Denote $X_n(\sigma)$ the number of disjoint cycles (including fixed points) of a permutation σ in the symmetric group $\mathfrak{S}(n)$. We write $C_j(\sigma)$ for the number of cycles of length j in the decomposition of σ as a product of disjoint cycles; thus, $X_n(\sigma) = \sum_{j=1}^{n} C_j(\sigma)$ and $n = \sum_{j=1}^{n} j\, C_j(\sigma)$. Let $\Theta = (\theta_m)_{m \geq 1}$ be given with $\theta_m \geq 0$. The generalised weighted measure is defined as the probability measure $\mathbb{P}_\Theta$ on the symmetric group $\mathfrak{S}(n)$:

$$\mathbb{P}_\Theta[\sigma] = \frac{1}{h_n\, n!} \prod_{m=1}^{n} (\theta_m)^{C_m(\sigma)}$$

with h_n a normalisation constant (or a partition function) and $h_0 = 1$. This model is coming from statistical mechanics and the study of Bose quantum gases (see [NZ13] for more references and details). It generalises the classical cases of the uniform measure (corresponding to $\theta_m \equiv 1$) and the Ewens measure (corresponding to the case $\theta_m \equiv \theta > 0$). It has been an open question to prove a central limit theorem for the total number of cycles X_n under such measures (or more precisely under some specific regimes related to the asymptotic behavior of the θ_m's; such a central limit theorem was already known for Ewens measure). We refer to [EU14] for a nice survey of this question. The difficulty comes from the fact the there is nothing such as the Feller coupling anymore, and we cannot apply the same method as in Example 2.1.3. We now show how singularity analysis allows us to prove mod-Poisson convergence, and hence the central limit theorem, but also distributional approximations and precise large deviations.

We consider the generating series

$$g_\Theta(t) = \sum_{n=1}^{\infty} \frac{\theta_n}{n} t^n.$$

It is well known that

$$\sum_{n=0}^{\infty} h_n t^n = \exp(g_\Theta(t)). \tag{36}$$

Set $F(t, w) = \exp(w g_\Theta(t))$. Using elementary combinatorial arguments (which are detailed in [NZ13]), one can prove for each $z \in \mathbb{C}$ the following identity as formal power series:

$$\sum_{n=0}^{\infty} h_n \, \mathbb{E}_\Theta[\exp(z X_n)] \, t^n = \exp(e^z g_\Theta(t)) = F(t, e^z). \tag{37}$$

Our goal is to obtain an asymptotic for h_n and for the moment generating function of X_n. Note that in general (and unlike the case of Ewens measure), neither h_n nor $\mathbb{E}_\Theta[\exp(z X_n)]$ have a closed expression. Nevertheless, they correspond to the coefficient of t^n in $F(t, 1)$ (respectively $F(t, e^z)$) and, thus, using Cauchy formula, they can be expressed as a contour integral. The idea of singularity analysis is to choose the contour in a clever way such that, asymptotically, the main part of this contour integral comes from the integral near the singularity. In this way, the integral depends asymptotically only of the type of singularity of $F(t, 1)$ (resp. $F(t, e^z)$).

We note r the radius of convergence of $g_\Theta(t)$. We need suitable assumptions on the analyticity properties of g together with assumptions on the nature of its singularity at the point r on the circle of convergence. This motivates the next definition:

DEFINITION 7.3.1. *Let $0 < r < R$ and $0 < \phi < \pi/2$ be given. We then define*

$$\Delta_0 = \Delta_0(r, R, \phi) = \{z \in \mathbb{C}; \, |z| < R, z \neq r, |\arg(z - r)| > \phi\},$$

see Figure 1. Assume we are further given $g(t)$, $\theta \geq 0$ and $r > 0$. We then say that $g(t)$ is in the class $\mathcal{F}(r, \theta)$ if

(i) *there exists $R > r$ and $0 < \phi < \pi/2$ such that $g(t)$ is holomorphic in $\Delta_0(r, R, \phi)$;*

(ii) *there exists a constant K such that*

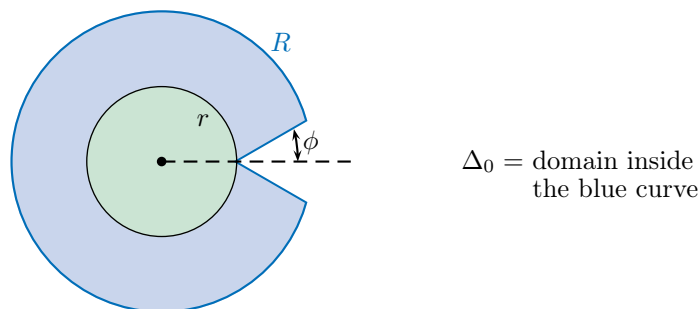$$g(t) = \theta \log\left(\frac{1}{1 - t/r}\right) + K + O(t - r) \quad \text{as } t \to r.$$

FIGURE 1. Domain $\Delta_0(r, R, \phi)$.

One readily notes that the generating series corresponding to the Ewens measure (i.e. $\theta_m \equiv \theta$) is of class $\mathcal{F}(1, \theta)$ since in this case

$$g_\Theta(t) \equiv \theta \log \left(\frac{1}{1-t} \right).$$

Consequently our results will provide alternative proofs to this case as well. The next theorem due to Hwang plays a key role in our example (we use the following notation: if $G(t) = \sum_{n=0}^\infty g_n t^n$, we denote $[t^n][G] \equiv g_n$ the coefficient of t^n in $G(t)$).

THEOREM 7.3.2 (Hwang, [Hwa94]). *Let $F(t, w) = \exp(wg(t)) S(t, w)$ be given. Suppose*

 (i) *$g(t)$ is of class $\mathcal{F}(r, \theta)$,*

 (ii) *$S(t, w)$ is holomorphic in a domain containing $\{(t, w) \in \mathbb{C}^2;\ |t| \leq r, |w| \leq \hat{r}\}$, where $\hat{r} > 0$ is some positive number.*

Then

$$[t^n][F(t, w)] = \frac{e^{Kw} n^{w\theta - 1}}{r^n} \left(\frac{S(r, w)}{\Gamma(\theta w)} + O\left(\frac{1}{n} \right) \right)$$

uniformly for $|w| \leq \hat{r}$ and with the same K as in the definition above.

The idea of the proof consists in taking a suitable Hankel contour and to estimate the integral over each piece. There exist several other versions of this theorem where one can replace $\log(1 - t/r)$ by other functions and we refer the reader to the monograph [FS09], chapter VI.5. As an application of Theorem 7.3.2, we obtain an asymptotic for h_n.

LEMMA 7.3.3. *Let $\theta > 0$ and assume that $g_\Theta(t)$ is of class $\mathcal{F}(r, \theta)$. We then have*

$$h_n = \frac{e^K n^{\theta - 1}}{r^n \, \Gamma(\theta)} \left(1 + O\left(\frac{1}{n} \right) \right).$$

PROOF. We have already noted that $\sum_{n=1}^\infty h_n t^n = \exp(g_\Theta(t))$. We can apply Theorem 7.3.2 with $g(t) = g_\Theta(t)$, $w = 1$ and $S(t, w) = 1$. $\qquad \square$

Using identity (37) and Theorem 7.3.2, we can also obtain an asymptotic for the Laplace transform $\mathbb{E}_\Theta[\exp(wX_n)]$:

THEOREM 7.3.4 (Nikeghbali-Zeindler, [NZ13]). *If $g_\Theta(t)$ is of class $\mathcal{F}(r, \theta)$, then*

$$\mathbb{E}_\Theta[\exp(zX_n)] = n^{\theta(e^z - 1)} e^{K(e^z - 1)} \left(\frac{\Gamma(\theta)}{\Gamma(\theta e^z)} + O\left(\frac{1}{n} \right) \right).$$

Consequently, the sequence $(X_n)_{n \in \mathbb{N}}$ converges in the mod-Poisson sense, with parameters $t_n = K + \theta \log n$ and limiting function $\psi(z) = \frac{\Gamma(\theta)}{\Gamma(\theta e^z)}$. The convergence takes place of the whole complex plane with speed $O(1/n)$.

PROOF. An application of Theorem 7.3.2 yields

$$[t^n][\exp(e^z g_\Theta(t))] = \frac{e^{Ke^z} n^{e^z \theta - 1}}{r^n} \left(\frac{1}{\Gamma(\theta e^z)} + O\left(\frac{1}{n} \right) \right),$$

with $O(\cdot)$ uniform for bounded $z \in \mathbb{C}$. Now a combination of identity (37) and Lemma 7.3.3 gives the desired result. $\qquad \square$

The above theorem not only implies the central limit theorem, but also Poisson approximations and precise large deviations. We only state here the precise large deviation result which extends earlier work of Hwang in the case $\theta = 1$ as a consequence of Theorem 3.2.2 and refer to [NZ13] for the distributional approximations results.

PROPOSITION 7.3.5 (Nikeghbali-Zeindler, [NZ13]). *Let* $Y_n = X_n - 1$ *and let* $x \in \mathbb{R}$ *such that* $t_n x \in \mathbb{N}$ *with* $t_n = K + \theta \log n$. *We note* $k = t_n x$. *Then*

$$\mathbb{P}[Y_n = x t_n] = \mathrm{e}^{-t_n} \frac{(t_n)^k}{k!} \left(\frac{\Gamma(\theta)}{x \, \Gamma(\theta x)} + O\left(\frac{1}{t_n}\right) \right).$$

In fact, an application of Theorem 3.2.2 would immediately yield an arbitrary long expansion for $\mathbb{P}[Y_n = x t_n]$ and also for $\mathbb{P}[Y_n \geq x t_n]$, since the speed of convergence is fast enough.

REMARK 7.3.6. Equations (36) and (37) fit naturally in the framework of *labeled combinatorial classes* [FS09, Chapter II]: a permutation is a *set* of cycles and the weighted exponential generating series of cycles is $g_\Theta(t)$. Using this framework, one could give a more general statement with the same proof. Let $\mathcal{A}$ be a (weighted) labeled combinatorial class and consider the class $\mathcal{B} = \mathrm{SET}(\mathcal{A})$. We denote X_n the number of components in a random element of size n of $\mathcal{B}$ (the probability of taking an element of $\mathcal{B}$ being proportional to its weight). Assume that the generating series of $\mathcal{A}$ is in the class $\mathcal{F}(r, \theta)$. Then $(X_n)_{n \in \mathbb{N}}$ converges in the mod-Poisson sense as in Theorem 7.3.4. We refer to [Hwa96, Section 4.1] for a similar discussion.

However, the hypothesis that the generating series of $\mathcal{A}$ is in the class $\mathcal{F}(r, \theta)$ is mainly natural in the case of cycles and permutations, which explains why we focused on this example here.

REMARK 7.3.7. The singularity analysis of generating functions is closely related to the Selberg-Delange method used in previous section. In both cases we do not have a closed expression for the Laplace transform of our variable, but for an appropriate sum of them — Dirichlet series for arithmetic functions and exponential generating functions for cycles in weighted permutations. From Perron's or Cauchy's formula, we get an integral formula for the Laplace transform. Then choosing an appropriate integral contour allows to find the asymptotics. It is quite striking that these closely related methods both yield mod-Poisson convergence. However, we are not aware of a way to bring both under one roof.

7.4. Rises in random permutations

As illustrated in the previous section, the methods of singularity analysis and the *transfer theorems* between the singularities of a generating series and the asymptotic behavior of its coefficients (see *e.g.* [FS09, Chapter VI]) can be used to prove mod-ϕ convergence (this is not surprising as they can be used to establish the related concept of *quasi-powers*, see Section 4.5.3). We examine here another example where we do not have strictly speaking mod-ϕ convergence since the reference law is not infinitely divisible, but where our methods can nevertheless be used.

If $\sigma \in \mathfrak{S}(n)$ is a permutation, a *rise* of σ is an integer $i \in [\![1, n-1]\!]$ such that $\sigma(i) < \sigma(i+1)$. The number of rises $R(\sigma)$ is thus a quantity between 0 and $n-1$, the

two extremal values corresponding to the permutations $n(n-1)\ldots 21$ and $12\ldots(n-1)n$. We denote $R_n = R(\sigma_n)$ the random variable which counts the number of rises of a random permutation $\sigma_n \in \mathfrak{S}(n)$ chosen under the uniform measure. The double generating series

$$R(z,t) = \sum_{n=0}^{\infty} \mathbb{E}[e^{zR_n}]\, t^n$$

is computed in [FS09, III.7, p. 209]:

$$R(z,t) = \frac{e^z - 1}{e^z - e^{t(e^z-1)}}.$$

If $\mathrm{Re}(z) \le \log 2$, then $|e^z| \le 2$ and the smallest (in modulus) root of the denominator is

$$t = \rho(z) = \frac{z}{e^z - 1}.$$

We then have $R(z,t) = \frac{1}{\rho(z)-t} + (e^z - 1)\, S(z,t)$ with $S(z,t) = \frac{1}{e^z - e^{t(e^z-1)}} - \frac{1}{z - t(e^z-1)}$ analytic in t and on the band $z \in \mathcal{S}_{(-\infty,\log 2)}$. This implies:

THEOREM 7.4.1 ([FS09], p. 658). *For $z \in \mathcal{S}_{(-\infty,\log 2)}$, one has locally uniformly*

$$\mathbb{E}[e^{zR_n}] = \tau(z)^{n+1} + O(|z|\, 2^{-n}),$$

where $\tau(z) = \frac{e^z - 1}{z}$.

Notice that $\tau(z) = \frac{e^z - 1}{z} = \int_0^1 e^{zu}\, du$ is the Laplace transform of a uniform random variable U in the interval $[0,1]$. Informally, the previous should be thought as follows: $(R_n)_{n\in\mathbb{N}}$ converges mod-U on $\mathcal{S}_{(-\infty,\log 2)}$ with parameters $t_n = n+1$ and limit $\psi(z) = 1$. However, the uniform law U is not infinitely divisible, and $\tau(z)$ vanishes on any $z \in 2i\pi\mathbb{Z}$, therefore the error term $O(|z|\, 2^{-n})$ cannot be rewritten as a multiplicative error term, as we did before. Still, every computation of Chapter 4 can be performed in this case. Let us stress out the necessary adjustments:

- Berry-Esseen estimates. Set $F_n(x) = \mathbb{P}[R_n - \frac{n+1}{2} \ge \sqrt{\frac{n+1}{12}}\, x]$ and $G_n(x) = \int_{-\infty}^x g(y)\, dy$. We have by Feller's lemma

$$|F_n(x) - G_n(x)| \le \frac{24m}{\Delta\pi\sqrt{t_n}} + \frac{1}{\pi\sqrt{t_n}} \int_{[\pm\delta\sqrt{t_n}]} \left| \frac{f_n^*(\xi) - e^{-\frac{\xi^2}{2}}}{\xi} \right| d\xi$$

$$+ \frac{1}{\pi\delta\sqrt{t_n}} \int_{[\pm\Delta\sqrt{t_n}]\setminus[\pm\delta\sqrt{t_n}]} \left| f_n^*(\xi) - e^{-\frac{\xi^2}{2}} \right| d\xi,$$

where as in the proof of Proposition 4.1.1, $f_n^*(\xi)$ is the Fourier transform of $\frac{R_n - \frac{t_n}{2}}{\sqrt{\frac{t_n}{12}}}$. From Theorem 7.4.1 and a Taylor expansion of $\tau(i\xi)$, one has

$$f_n^*(\xi) = e^{-\frac{\xi^2}{2}\left(1 + O\left(\frac{\xi^2}{t_n}\right)\right)} + O(|\xi|\, 2^{-t_n}).$$

In the first integral of Feller's bound, we are integrating

$$\left| \frac{f_n^*(\xi) - g^*(\xi)}{\xi} \right| = O\left(\delta^2 e^{-\frac{\xi^2}{2}} \right) + O(2^{-t_n}),$$

so we get a $o(\frac{1}{\sqrt{t_n}})$ as δ goes to zero. Then, regarding the second integral, we are integrating a $O((q_\delta)^{t_n})$ for some $q_\delta \in [\frac{1}{2}, 1)$, so it does not contribute to the asymptotics. Thus, we get the uniform estimate $|F_n(x) - G_n(x)| = o(\frac{1}{\sqrt{t_n}})$ as in Proposition 4.1.1.

- Tilting method and large deviation estimates. When tilting by $h > 0$, the new random variable R_n^h with law $\mathbb{P}[R_n^h \in dy] = \frac{e^{hy}}{\mathbb{E}[e^{hR_n}]} \mathbb{P}[R_n \in dy]$ has its Fourier transform that has asymptotics

$$\frac{\tau(h + i\xi)^{n+1} + O(2^{-n})}{\tau(h)^{n+1} + O(2^{-n})} = \left(\frac{\tau(h + i\xi)}{\tau(h)}\right)^{n+1} (1 + o(1))$$

the second formula holding if h and ξ are small enough (more precisely, we want $0 < h < \log 2$ and then ξ small enough with bounds depending on h). Notice that the new exponent $\frac{\tau(h+i\xi)}{\tau(h)}$ does not vanish anymore, so we can now use the same argument as in Theorem 4.2.1. Thus, if F is the Legendre-Fenchel transform of the uniform law on $[0, 1]$ (not really explicit), and if h is the solution of $\frac{1+\varepsilon}{2} = \eta'(h)$ with $\eta(h) = \log\left(\frac{e^h - 1}{h}\right)$, then for

$$\varepsilon \in (2\eta'(0) - 1, 2\eta'(\log 2) - 1) = \left(0, 3 - \frac{2}{\log 2}\right) \supseteq (0, 0.114),$$

we obtain the large deviation estimate:

$$\mathbb{P}\left[R_n \geq \frac{(1 + \varepsilon)(n + 1)}{2}\right] = \frac{e^{-t_n F(\frac{1+\varepsilon}{2})}}{h\sqrt{2\pi t_n \eta''(h)}}(1 + o(1))$$

since we have mod-U convergence of $(R_n)_{n \in \mathbb{N}}$ with parameters $t_n = \frac{n+1}{2}$ and residue $\psi(x) = 1$. Similarly, for the negative deviations, the estimate of the Fourier transform of the tilted random variable R_n^h can be used if $h > h_0$, where $h_0 \approx -1.594$ is the solution of the equation $\tau(h) = \frac{1}{2}$. Therefore, for

$$\varepsilon \in \left(0, \frac{2}{h_0} - 1\right) \supseteq (0, 0.254),$$

we have

$$\mathbb{P}\left[R_n \leq \frac{(1 - \varepsilon)(n + 1)}{2}\right] = \frac{e^{-t_n F(\frac{1-\varepsilon}{2})}}{|h|\sqrt{2\pi t_n \eta''(h)}}(1 + o(1)),$$

where h is the solution of $\frac{1-\varepsilon}{2} = \eta'(h)$.

- Extended central limit theorem. The result of Lemma 4.3.2 holds, because

$$\left|\frac{\tau(h + iu)}{\tau(h)}\right|^2 = \left|\frac{e^{h+iu} - 1}{e^h - 1} \frac{h}{h + iu}\right|^2$$

$$= \left(1 + \frac{2e^h (1 - \cos u)}{(e^h - 1)^2}\right) \frac{h^2}{h^2 + u^2} = \frac{h^2 + 2t(h)(1 - \cos(u))}{h^2 + u^2},$$

where $t(h) = \frac{h^2 e^h}{(e^h - 1)^2} \leq 1$ for any h. From this inequality, one deduces that

$$q_\delta = \max_{u \in \mathbb{R} \setminus [-\delta, \delta]} \left| \frac{\tau(h + iu)}{\tau(h)} \right| \leq 1 - D\delta^2$$

for some constant $D > 0$ and any $h \in (-\varepsilon, \varepsilon)$, with ε small enough.

Since $\eta(x) = \log(\tau(x))$ has Taylor expansion $\frac{x}{2} + \frac{x^2}{24} - \frac{x^4}{2880} + o(x^5)$, we thus get from Equation (20):

$$\mathbb{P}\left[R_n \geq \frac{n+1}{2} + \sqrt{\frac{n+1}{12}}\, y \right] = \frac{(1 + o(1))}{y\sqrt{2\pi}} \exp\left(-\frac{y^2}{2} + \frac{y^4}{240 t_n} \right)$$

for any positive y with $y = o((t_n)^{\frac{1}{2} - \frac{1}{6}}) = o((t_n)^{\frac{5}{12}})$.

7.5. Characteristic polynomials of random matrices in a compact Lie group

Introduce the classical compact Lie groups of type A, C, D:

$$\begin{array}{ll}
U(n) = \{g \in GL(n, \mathbb{C}) \mid gg^\dagger = g^\dagger g = I_n\} & \text{(unitary group)} \\
USp(n) = \{g \in GL(n, \mathbb{H}) \mid gg^\star = g^\star g = I_n\} & \text{(compact symplectic group)} \\
SO(2n) = \{g \in GL(2n, \mathbb{R}) \mid gg^t = g^t g = I_{2n} \;;\; \det g = 1\} & \text{(special orthogonal group)}
\end{array}$$

where for compact symplectic groups $g^\star$ denotes the transpose conjugate of a quaternionic matrix, the conjugate of a quaternionic number $q = a + ib + jc + kd$ being $q^\star = a - ib - jb - kd$. In the following we shall consider quaternionic matrices as complex matrices of size $2n \times 2n$ by using the map

$$a + ib + jc + kd \mapsto \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}.$$

The eigenvalues of a matrix $g \in G = SO(2n)$ or $U(n)$ or $USp(n)$ are on the unit circle $\mathbb{S}^1$, and the value of the characteristic polynomial $\det(1 - g)$ factorises as

$$\det(1 - g) = \prod_{i=1}^{n} (1 - e^{i\theta_i})$$

if $G = U(n)$ and $Sp(g) = (e^{i\theta_1}, \dots, e^{i\theta_n})$, and

$$\det(1 - g) = \prod_{i=1}^{n} (1 - e^{i\theta_i})(1 - e^{-i\theta_i}) = \prod_{i=1}^{n} |1 - e^{i\theta_i}|^2$$

if $G = SO(2n)$ or $USp(n)$ and $Sp(g) = (e^{i\theta_1}, e^{-i\theta_1}, \dots, e^{i\theta_n}, e^{-i\theta_n})$. We define

$$Y_n^{A,C,D} = \begin{cases} \mathrm{Re}\left[\log \det(1 - g) \right] & \text{in type A;} \\ \log \det(1 - g) & \text{in types C and D.} \end{cases}$$

An exact formula for the Laplace transform of $\log \det(1 - g)$, and the corresponding asymptotics have been given in [KS00a, KS00b, HKO01], see also [KN12, Sections 3 and 4]. Hence, in type A,

$$\mathbb{E}\left[e^{z_1 \mathrm{Re}(\log \det(1-g)) + z_2 \mathrm{Im}(\log \det(1-g))} \right] = \prod_{j=1}^{n} \frac{\Gamma(j)\,\Gamma(j + z_1)}{\Gamma(j + \frac{z_1 + iz_2}{2})\,\Gamma(j + \frac{z_1 - iz_2}{2})}$$

for every z_1 with $\mathrm{Re}(z_1) > -1$, and every $z_2 \in \mathbb{C}$ (see *e.g.* [KS00b, Formula (71)]). In particular, if $\mathrm{Re}(z) > -1$, then

$$\mathbb{E}\left[e^{zY_n^A}\right] = \prod_{j=1}^{n} \frac{\Gamma(j)\,\Gamma(j+z)}{\Gamma(j+\frac{z}{2})^2}$$

$$= \frac{G(1+\frac{z}{2})^2}{G(1+z)}\, n^{\frac{z^2}{4}}\, (1+o(1)).$$

Here G denotes Barnes' G-function, which is the entire solution of the functional equation $G(z+1) = \Gamma(z)\,G(z)$ with $G(1) = 1$. To go from the exact formula with ratios of Gamma functions, to the asymptotic formula with Barnes' functions, one has to interpret the formula as a Toeplitz determinant with one Fisher-Hartwig singularity, see the details in the aforementioned paper [KN12] (the arguments therein apply *mutatis mutandi* to the case of a complex random variable in the Laplace transform z, since the exact formula holds as long as $\mathrm{Re}(z) > -1$, and the theory of Toeplitz determinants works from the outset with complex functions). Therefore, $(Y_n^A)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense with parameters $t_n = \frac{\log n}{2}$, limiting function

$$\psi^A(z) = \frac{G(1+\frac{z}{2})^2}{G(1+z)},$$

and on the band $\mathcal{S}_{(-1,+\infty)}$.

Similarly, in type C and D, one has the exacts formulas

$$\mathbb{E}\left[e^{zY_n^C}\right] = 2^{2nz} \prod_{j=1}^{n} \frac{\Gamma(j+n+1)\,\Gamma(z+j+\frac{1}{2})}{\Gamma(j+\frac{1}{2})\,\Gamma(z+j+n+1)} \quad \text{for } \mathrm{Re}(z) > -\frac{3}{2};$$

$$\mathbb{E}\left[e^{zY_n^D}\right] = 2^{2nz} \prod_{j=1}^{n} \frac{\Gamma(j+n-1)\,\Gamma(z+j-\frac{1}{2})}{\Gamma(j-\frac{1}{2})\,\Gamma(z+j+n-1)} \quad \text{for } \mathrm{Re}(z) > -\frac{1}{2}$$

as well as the asymptotic formulas

$$\mathbb{E}\left[e^{zY_n^C}\right] = \left(\frac{\pi n}{2}\right)^{\frac{z}{2}} \left(\frac{n}{2}\right)^{\frac{z^2}{2}} \frac{G(\frac{3}{2})}{G(\frac{3}{2}+z)}\, (1+o(1));$$

$$\mathbb{E}\left[e^{zY_n^D}\right] = \left(\frac{8\pi}{n}\right)^{\frac{z}{2}} \left(\frac{n}{2}\right)^{\frac{z^2}{2}} \frac{G(\frac{1}{2})}{G(\frac{1}{2}+z)}\, (1+o(1))$$

which hold in the same range for z, locally uniformly. Therefore, setting

$$X_n^C = Y_n^C - \frac{1}{2}\log\frac{\pi n}{2}$$

$$X_n^D = Y_n^D - \frac{1}{2}\log\frac{8\pi}{n}$$

then $(X_n^C)_{n\in\mathbb{N}}$ and $(X_n^D)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense, with parameters $t_n = \log\frac{n}{2}$, limiting functions

$$\psi^C(z) = \frac{G(\frac{3}{2})}{G(\frac{3}{2}+z)} \quad \text{and} \quad \psi^D(z) = \frac{G(\frac{1}{2})}{G(\frac{1}{2}+z)},$$

and respectively on $\mathcal{S}_{(-\frac{3}{2},+\infty)}$ and on $\mathcal{S}_{(-\frac{1}{2},+\infty)}$.

Now, our large deviation theorems apply and one obtains:

THEOREM 7.5.1. *Over* $\mathrm{U}(n)$, *one has:*

$$\forall x > 0, \quad \mathbb{P}_n\left[\,|\det(1-g)| \geq n^{\frac{x}{2}}\right] = \frac{G(1+\frac{x}{2})^2}{G(1+x)} \frac{1}{x\,n^{\frac{x^2}{4}}\,\sqrt{\pi \log n}}(1+o(1));$$

$$\forall x \in (0,1), \quad \mathbb{P}_n\left[\,|\det(1-g)| \leq n^{-\frac{x}{2}}\right] = \frac{G(1-\frac{x}{2})^2}{G(1-x)} \frac{1}{x\,n^{\frac{x^2}{4}}\,\sqrt{\pi \log n}}(1+o(1)).$$

Over $\mathrm{USp}(n)$, *one has:*

$$\forall x > 0, \quad \mathbb{P}_n\left[\frac{\det(1-g)}{\sqrt{\frac{\pi}{2}}} \geq n^{\frac{1}{2}+x}\right] = \frac{G(\frac{3}{2})}{G(\frac{3}{2}+x)} \frac{1}{x\,n^{\frac{x^2}{2}}\,2^{\frac{x^2+x}{2}}\,\sqrt{2\pi \log n}}(1+o(1));$$

$$\forall x \in \left(0,\frac{3}{2}\right), \quad \mathbb{P}_n\left[\frac{\det(1-g)}{\sqrt{\frac{\pi}{2}}} \geq n^{\frac{1}{2}-x}\right] = \frac{G(\frac{3}{2})}{G(\frac{3}{2}-x)} \frac{1}{x\,n^{\frac{x^2}{2}}\,2^{\frac{x^2-x}{2}}\,\sqrt{2\pi \log n}}(1+o(1)).$$

Finally, over $\mathrm{SO}(2n)$, *one has:*

$$\forall x > 0, \quad \mathbb{P}_n\left[\frac{\det(1-g)}{\sqrt{8\pi}} \geq n^{-\frac{1}{2}+x}\right] = \frac{G(\frac{1}{2})}{G(\frac{1}{2}+x)} \frac{1}{x\,n^{\frac{x^2}{2}}\,2^{\frac{x^2-x}{2}}\,\sqrt{2\pi \log n}}(1+o(1));$$

$$\forall x \in \left(0,\frac{1}{2}\right), \quad \mathbb{P}_n\left[\frac{\det(1-g)}{\sqrt{8\pi}} \leq n^{-\frac{1}{2}-x}\right] = \frac{G(\frac{1}{2})}{G(\frac{1}{2}-x)} \frac{1}{x\,n^{\frac{x^2}{2}}\,2^{\frac{x^2+x}{2}}\,\sqrt{2\pi \log n}}(1+o(1)).$$

PROOF. These are immediate computations by using Theorem 4.2.1. $\square$

REMARK 7.5.2. From Proposition 4.4.1, one also gets normality zones for the random variables $X_n/\sqrt{t_n}$. On the other hand, the computations performed in the unitary case hint at a phenomenon of 2-dimensional mod-Gaussian convergence for the complex numbers $\log \det(1-g)$, with $g \in \mathrm{U}(n)$. In [FMN15], we shall prove this rigorously, and compute various probabilistic consequences, for instance, estimates of large deviations for the random vectors $\log \det(1-g)$.

REMARK 7.5.3. The analogue of Theorem 7.5.1 in the setting of random matrices in the β-ensembles, or of general Wigner matrices, has been studied in the recent paper [DE13a]. They can be easily restated in the mod-Gaussian language, since their proofs rely on the computation of the asymptotics of the cumulants of the random variables $X_n = \log|\det M_n|$, with for instance $(M_n)_{n\in\mathbb{N}}$ random matrices of the Gaussian unitary ensembles.

Mod-Gaussian convergence from a factorisation of the probability generating function

Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of bounded random variables with nonnegative integer values, and such that $\sigma_n^2 := \mathrm{Var}(X_n)$ tends to infinity. Denote $P_n(t) = \mathbb{E}[t^{X_n}]$ the *probability generating function* of X_n. Each $P_n(t)$ is a polynomial in t. Then, it is known that a sufficient condition so that X_n is asymptotically Gaussian is that $P_n(t)$ has negative real roots (see references below). In this chapter, we prove that if the third cumulant $L_n^3 := \kappa^{(3)}(X_n)$ also tends to infinity with light additional hypotheses, then a suitable renormalised version of X_n converges in the mod-Gaussian sense. We then give an application for the number of blocks in a uniform set-partition of $[n]$.

8.1. Background: central limit theorem from location of zeros

The idea of proving central limit theorem by looking at the zeros of the probability generating function originates from an article of Harper [Har67]. Harper was interested in the number of blocks of a random set-partition, which will be our main example below. The argument was then generalised by Haigh [Hai71]. Hwang and Steyaert [HS02, Lemma 4] refined Haigh's result by giving a bound on the speed of convergence towards the normal distribution. Finally, let us mention a recent work of Lebowitz, Pittel, Ruelle and Speer [LPRS14], where the authors prove central limit theorems and local limit laws for random variables under various assumptions on the location of the zeros of the probability generating function. Lebowitz, Pittel, Ruelle and Speer apply their theoretical results to graph counting polynomials and to the Ising model in $\mathbb{Z}^d$ (using Lee-Yang's theorem). We plan to address the problem of mod-Gaussian convergence for these models in the future. The presentation of the results here is inspired from the one in [HS02].

8.2. Mod-Gaussian convergence from non-negative factorisations

We will prove the following statement, which is stronger than what was announced.

THEOREM 8.2.1. *Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of bounded random variables with non-negative integer values, with mean μ_n, variance σ_n^2 and third cumulant L_n^3. Suppose that the probability generating function $P_n(t)$ of X_n can be factorised as*

$$P_n(t) = \prod_{1 \le j \le k_n} P_{n,j}(t), \tag{38}$$

where

- *$(k_n)_{n \in \mathbb{N}}$ is a sequence of positive integers;*
- *each $P_{n,j}$ is a polynomial with non-negative coefficients;*

- *we assume $L_n = o(\sigma_n)$ and $M_n = o((L_n)^2/\sigma_n)$, where $M_n = \max_{1 \leq j \leq k_n} \deg(P_{n,j})$.*

Then, the sequence

$$\widetilde{X_n} = \frac{X_n - \mu_n}{L_n}$$

converges in the mod-Gaussian sense with limiting function $\psi = \exp(z^3/6)$ and parameters $t_n = \sigma_n^2/L_n^2$. The convergence takes place of the whole complex plane, with speed of convergence $O(M_n/|L_n|)$.

REMARK 8.2.2. Before doing the proof, let us discuss briefly the hypothesis. First observe that if P_n has only real negative roots $-r_{n,j}$, then

$$P_n(t) = C \prod_{1 \leq j \leq n} (t + r_{n,j}),$$

so that our theorem includes this case, as claimed. In particular, this theorem contains the mod-Gaussian convergence of a sum of i.i.d. bounded variables with a non-zero third cumulant (see Example 2.1.2).

Conversely, the factorisation (38) and the fact that all $P_{n,j}$ have non-negative coefficients imply that X_n can be written as a sum of independent (but not necessarily identically distributed) random variables. However, an interesting feature is that this representation of X_n may not be explicit at all, as in the example of blocks in set-partitions in next section.

Finally, we point our that the hypotheses of Theorem 8.2.1 imply $M_n = o(|L_n|)$: indeed,

$$M_n \ll \frac{(L_n)^2}{\sigma_n} = |L_n| \frac{|L_n|}{\sigma_n} \ll |L_n|$$

since $L_n = o(\sigma_n)$.

PROOF. Since $P_n(1) = \prod_{1 \leq j \leq k_n} P_{n,j}(1) = 1$, we have

$$P_n(t) = \prod_{1 \leq j \leq k_n} \frac{P_{n,j}(t)}{P_{n,j}(1)}.$$

Therefore, we can assume without loss of generality that $P_{n,j}(1) = 1$ (otherwise replace $P_{n,j}(t)$ by $P_{n,j}(t)/P_{n,j}(1)$). Since $P_{n,j}(t)$ has non-negative coefficients, $P_{n,j}(t)$ is the probability generating function of a variable $X_{n,j}$ (for $1 \leq j \leq k_n$), defined by $\mathbb{P}[X_{n,j} = k] = [t^k]P_{n,j}(t)$.

The factorisation (38) implies that X_n can be represented as the sum of independent copies of the variables $X_{n,j}$. In particular,

$$\mu_n = \sum_{1 \leq j \leq k_n} \mu_{n,j};$$

$$\sigma_n^2 = \sum_{1 \leq j \leq k_n} \sigma_{n,j}^2;$$

$$L_n^3 = \sum_{1 \leq j \leq k_n} L_{n,j}^3$$

where $\mu_{n,j}$, $\sigma^2_{n,j}$ and $L^3_{n,j}$ are the three first cumulants of $X_{n,j}$. Notice that each $X_{n,j}$, and hence $|X_{n,j} - \mu_{n,j}|$ is bounded by M_n. We will use this repeatedly below. In particular one has: $\sigma^2_{n,j} \leq (M_n)^2$ and $|L^3_{n,j}| \leq (M_n)^3$.

Call $\widetilde{\varphi_n}$ the Laplace transform of $\widetilde{X_n}$. We have

$$\widetilde{\varphi_n}(z) = \mathbb{E}\left[\exp\left(\frac{z}{L_n}(X_n - \mu_n)\right)\right] = \prod_{j=1}^{k_n} \mathbb{E}\left[\exp\left(\frac{z}{L_n}(X_{n,j} - \mu_{n,j})\right)\right].$$

Fix $K > 0$ and assume that $\frac{M_n|z|}{|L_n|} \leq K$. From the Taylor expansion

$$e^w = 1 + w + \frac{w^2}{2} + \frac{w^3}{6} + O(w^4) \text{ (uniformly for } |w| \leq K),$$

we have that, uniformly for $\frac{M_n|z|}{|L_n|} \leq K$,

$$\mathbb{E}\left[\exp\left(\frac{z}{L_n}(X_{n,j} - \mu_{n,j})\right)\right] = 1 + \frac{\sigma^2_{n,j}z^2}{2(L_n)^2} + \frac{L^3_{n,j}z^3}{6(L_n)^3} + O\left(\mathbb{E}[(X_{n,j} - \mu_{n,j})^4]\left(\frac{|z|}{|L_n|}\right)^4\right).$$
(39)

Note that, since $M_n|z|/|L_n| \leq K$, the two first terms are bounded. Besides, we have the following bounds:

$$\left|\left(\frac{\sigma^2_{n,j}z^2}{2(L_n)^2}\right)^2\right| \leq \frac{\mathbb{E}[(X_{n,j} - \mu_{n,j})^4]}{4}\left(\frac{|z|}{|L_n|}\right)^4;$$

$$\left|\frac{\sigma^2_{n,j}z^2}{2(L_n)^2} \times \frac{L^3_{n,j}z^3}{6(L_n)^3}\right| \leq \frac{\mathbb{E}[|X_{n,j} - \mu_{n,j}|^5]}{12}\left(\frac{|z|}{|L_n|}\right)^5$$

$$\leq \frac{\mathbb{E}[(X_{n,j} - \mu_{n,j})^4]}{12}\left(\frac{|z|}{|L_n|}\right)^4\frac{M_n|z|}{|L_n|};$$

$$\left|\left(\frac{L^3_{n,j}z^3}{6(L_n)^3}\right)^2\right| \leq \frac{\mathbb{E}[(X_{n,j} - \mu_{n,j})^6]}{36}\left(\frac{|z|}{|L_n|}\right)^6$$

$$\leq \frac{\mathbb{E}[(X_{n,j} - \mu_{n,j})^4]}{36}\left(\frac{|z|}{|L_n|}\right)^4\left(\frac{M_n|z|}{|L_n|}\right)^2.$$

Taking the logarithm of Equation (39) and using $\log(1 + t) = t + O(t^2)$, we get, thanks to the above bounds, that

$$\log\mathbb{E}\left[\exp\left(\frac{z}{L_n}(X_{n,j} - \mu_{n,j})\right)\right] = \frac{\sigma^2_{n,j}z^2}{2(L_n)^2} + \frac{L^3_{n,j}z^3}{6(L_n)^3} + O\left(\mathbb{E}[(X_{n,j} - \mu_{n,j})^4]\left(\frac{|z|}{|L_n|}\right)^4\right).$$

Summing these identities, we obtain

$$\log\widetilde{\varphi_n}(z) = \left(\frac{\sigma_n}{L_n}\right)^2\frac{z^2}{2} + \frac{z^3}{6} + O\left(\left(\frac{|z|}{|L_n|}\right)^4\sum_{j=1}^{k_n}\mathbb{E}[(X_{n,j} - \mu_{n,j})^4]\right).$$

The error term can be bounded as follows:

$$\left(\frac{|z|}{|L_n|}\right)^4 \sum_{j=1}^{k_n} \mathbb{E}[(X_{n,j} - \mu_{n,j})^4] \leq \left(\frac{|z|}{|L_n|}\right)^4 (M_n)^2 \sum_{j=1}^{k_n} \mathbb{E}[(X_{n,j} - \mu_{n,j})^2] \leq |z|^4 \frac{(M_n)^2 (\sigma_n)^2}{|L_n|^4}.$$

By the assumption made on M_n, this is a $o(|z|^4)$. Finally, we get

$$\log \widetilde{\varphi_n}(z) = \left(\frac{\sigma_n}{L_n}\right)^2 \frac{z^2}{2} + \frac{z^3}{6} + o(|z|^4),$$

which can be rewritten as

$$\widetilde{\varphi_n}(z) \, e^{-\left(\frac{\sigma_n}{L_n}\right)^2 \frac{z^2}{2}} = \exp\left(\frac{z^3}{6}\right) (1 + o(|z|^4)),$$

this being uniform for $M_n |z| / |L_n| \leq K$, thus on compacts of $\mathbb{C}$. This ends the proof of the theorem. □

REMARK 8.2.3. Suppose for instance that $M_n = 1$, *i.e.*, the probability generating function of X_n has non-negative real roots. Then, the conditions on the first cumulants of X_n in order to apply Theorem 8.2.1 are

$$\sqrt{\sigma_n} \ll L_n \ll \sigma_n,$$

which are usually easy to check.

8.3. Two examples: uniform permutations and uniform set-partitions

The first example that fills in this framework is the number of disjoint cycles X_n of a uniform random permutation in $\mathfrak{S}(n)$. As mod-convergence of X_n has already been discussed in this article (Example 2.1.3), we will skip details. Using Feller's coupling, it is easily seen that

$$P_n(t) = \prod_{i=1}^{n} \frac{t + i - 1}{1 + i - 1}.$$

Moreover, a straight-forward computation yields $\mu_n = H_n + O(1)$, $\sigma_n^2 = H_n + O(1)$ and $L_n^3 = H_n + O(1)$, where H_n is the n-th harmonic number as in Example 2.1.3. Theorem 8.2.1 implies that $(X_n - H_n)/(H_n)^{1/3}$ converges in the mod-Gaussian sense, as established at the end of Example 2.1.3. Note however that Theorem 8.2.1 does not give the stronger mod-Poisson convergence of $(X_n)_{n \in \mathbb{N}}$ without renormalisation.

The second and more interesting example is the number of blocks in a random uniform set-partition of $[n]$. By definition, a *set-partition* of $[n]$ is a set of disjoint non-empty subsets of $[n]$, whose union is $[n]$. These subsets are called *blocks* or *parts* of the set-partition. For instance, $\{\{2,4\}, \{1\}, \{3\}\}$ is a set-partition of $[4]$ with 3 blocks. We denote $\mathfrak{Q}(n)$ the set of all set-partitions of $[n]$. For each integer $n \geq 0$, we then consider a random uniform set-partition in $\mathfrak{Q}(n)$, and denote X_n its number of blocks.

It was proved by Harper [Har67, Lemma 1] that the probability generating function of X_n has only real non-negative roots. Moreover, the asymptotic behaviour of $\mu_n = \mathbb{E}[X_n]$ and $\sigma_n^2 = \text{Var}(X_n)$ are known — see *e.g.* [CDKR15, Theorem 2.1] —

$$\mu_n = \frac{n}{\log n} (1 + o(1)), \quad \sigma_n^2 = \frac{n}{(\log n)^2} (1 + o(1)).$$

We will prove in next subsection that

$$L_n^3 = \frac{2n}{(\log n)^3} \, (1 + o(1)).$$

Since $M_n = 1$ and

$$\sqrt{\sigma_n} = O\left(\frac{n^{1/4}}{(\log n)^{1/2}}\right) \ll L_n = O\left(\frac{n^{1/3}}{\log n}\right) \ll \sigma_n = O\left(\frac{n^{1/2}}{\log n}\right),$$

we can apply Theorem 8.2.1: the variables $(X_n - \mu_n)/L_n$ converge in the mod-Gaussian sense with parameter $t_n = \sigma_n^2/L_n^2 = (\frac{n}{4})^{1/3} (1 + o(1))$ and limiting function $\psi(z) = \exp(z^3/6)$. As a corollary, applying Theorem 4.2.1 and Proposition 4.4.1 yields the following precise deviation result.

PROPOSITION 8.3.1. *Let X_n be the number of blocks in a uniform set-partition of $[n]$. Define μ_n, σ_n, L_n and t_n as above. Then the random variable $\frac{X_n - \mu_n}{\sigma_n}$ converges towards a Gaussian law, with a normality zone of size $o(n^{1/6})$. Moreover, at the edge of this normality zone, the deviation probabilities are given by: for any fixed $x > 0$,*

$$\mathbb{P}\left[\frac{X_n - \mu_n}{L_n} \geq t_n x\right] = \frac{\exp(-t_n \frac{x^2}{2})}{x\sqrt{2\pi t_n}} \exp\left(\frac{x^3}{6}\right) (1 + o(1));$$

$$\mathbb{P}\left[\frac{X_n - \mu_n}{L_n} \leq -t_n x\right] = \frac{\exp(-t_n \frac{x^2}{2})}{x\sqrt{2\pi t_n}} \exp\left(-\frac{x^3}{6}\right) (1 + o(1)).$$

8.4. Third cumulant of the number of blocks in uniform set-partitions

Fix $n \geq 0$. Let B_n be the number of set-partitions of $[n]$, known as the n-th Bell number. Dobinski's formula states that $B_n = e^{-1} \sum_{k=0}^{\infty} k^n/k!$, which allows us to consider a random variable M with the following distribution:

$$\mathbb{P}[M = k] = \frac{1}{eB_n} \frac{k^n}{k!}.$$

An easy observation, useful below, is that $\mathbb{E}[M^r] = B_{n+r}/B_n$. As above, we denote X_n the number of blocks in a uniform set-partition of size $[n]$. We also consider a Poisson variable P of parameter 1, independent from X_n. Stam [Sta83] proved the following relation (with a very nice probabilistic explanation).

LEMMA 8.4.1. *[Sta83, Theorem 2] We have the following equality of random variables in law:*

$$M \overset{law}{=} X_n + P.$$

Therefore, $\kappa^{(3)}(M) = \kappa^{(3)}(X_n) + \kappa^{(3)}(P)$. But $\kappa^{(3)}(P) = 1$ is a constant (independent of n), whose value will not be relevant for the asymptotic of $\kappa^{(3)}(X_n)$. Let us consider $\kappa^{(3)}(M)$. It is given by:

$$\kappa^{(3)}(M) = \mathbb{E}[M^3] - 3\,\mathbb{E}[M^2]\mathbb{E}[M] + 2(\mathbb{E}[M])^3 = \frac{B_{n+3}}{B_n} - 3\frac{B_{n+2}}{(B_n)^2}\frac{B_{n+1}}{} + 2\frac{(B_{n+1})^3}{(B_n)^3}.$$

In order to find the asymptotic of the above formula, we use the following estimate for Bell numbers. This is a variant of the Moser-Wyman formula [MW55], that can be found in an unpublished note of Mohr [Moh95].

LEMMA 8.4.2. *Let α_n be the positive real number defined by the equation $\alpha_n \, e^{\alpha_n} = n$; in particular $\alpha_n = (\log n)\,(1 - o(1))$. Then, one has, for any integer r,*

$$B_{n+r} = \frac{(n+r)!}{(\alpha_n)^{n+r}} \frac{\exp(e^{\alpha_n} - 1)}{\sqrt{2\pi\beta_n}}$$
$$\times \left(1 + \frac{P_0 + rP_1 + r^2 P_2}{e^{\alpha_n}} + \frac{Q_0 + rQ_1 + r^2 Q_2 + r^3 Q_3 + r^4 Q_4}{e^{2\alpha_n}} + O(e^{-3\alpha_n}) \right),$$

where $\beta_n = ((\alpha_n)^2 + \alpha_n)\, e^{\alpha_n}$, and the P_i's and Q_i's are explicit rational functions of α_n, with $P_i = O((\alpha_n)^{-i})$ and $Q_i = O((\alpha_n)^{-i})$.

This lemma yields an estimate for quotients of Bell numbers (we substitute e^{α_n} by n/α_n):

$$(\alpha_n)^r \frac{B_{n+r}}{B_n} = (n+1) \cdots (n+r) \left(1 + \alpha_n \frac{rP_1 + r^2 P_2}{n} \right.$$
$$\left. + (\alpha_n)^2 \frac{-P_0(rP_1 + r^2 P_2) + rQ_1 + r^2 Q_2 + r^3 Q_3 + r^4 Q_4}{n^2} + O(\alpha_n^3/n^3) \right)$$

Consider the expression

$$k^3 := \alpha_n^3 \left(\frac{B_{n+3}}{B_n} - 3\frac{B_{n+2}\, B_{n+1}}{B_n^2} + 2\frac{(B_{n+1})^3}{(B_n)^3} \right).$$

With the help of a computer algebra program, we find

$$k^3 = n \left((-6\,P_2\,(P_1 + P_2) + 6\,Q_3 + 36\,Q_4)(\alpha_n)^2 + 12P_2\,\alpha_n + 2 \right) + O((\alpha_n)^3).$$

From the estimate $P_i = O((\alpha_n)^{-i})$ and $Q_i = O((\alpha_n)^{-i})$, we see that the dominant term in k^3 comes from the constant 2 in the above equation. Namely,

$$k^3 = 2n + O\left(\frac{n}{\alpha_n} \right), \quad \text{that is } \kappa^{(3)}(X_n) = \frac{2n}{(\alpha_n)^3} + O\left(\frac{n}{(\alpha_n)^4} \right),$$

as claimed in the previous subsection.

CHAPTER 9

Dependency graphs and mod-Gaussian convergence

The theory of dependency graphs is a powerful toolbox to prove asymptotic normality. A dependency graph encodes the dependency structure in a family of random variables: roughly we take a vertex for each variable in the family and connect dependent random variables by edges. The idea is that, if the degrees in a sequence of dependency graphs do not grow too fast, then the corresponding variables behave as if independent and the sum of the corresponding variables is asymptotically normal. Precise normality criteria using dependency graphs have been given by Petrovskaya/Leontovich, Janson, Baldi/Rinott and Mikhailov [PL83, Jan88, BR89, Mik91]. These results are black boxes to prove asymptotic normality of sums of *partially dependent* variables and can be applied in many different contexts. The original motivation of Petrovskaya and Leontovich comes from the mathematical modelisation of cell populations [PL83]. On the contrary, Janson was interested in random graph theory: dependency graphs are used to prove central limit theorems for some statistics, such as subgraph counts, in $G(n, p)$ [BR89, Jan88, JŁR00]; see also [Pen02] for applications to *geometric* random graphs. The theory has then found a field of application in geometric probability [AB93, PY05, BV07]. More recently it has been used to prove asymptotic normality of pattern counts in random permutations [Bón10, HJ10]. Dependency graphs also generalise the notion of *m*-dependence [HR48, Ber73], widely used in statistics [Das08].

Beyond asymptotic normality, dependency graphs may be used for other kind of results. The first family of such results consists in refinements of central limit theorems.

- In their original paper [BR89], Baldi and Rinott combined dependency graphs with Stein's method. In addition to give a central limit theorem, this approach provides precise estimates for the Kolmogorov distance between a renormalised version of X_n and the Gaussian distribution. For more general and in some cases sharper bounds, we also refer the reader to the work of Chen and Shao [CS04].

- More recently, H. Döring and P. Eichelsbacher have shown how dependency graphs can be used to obtain some moderate deviation principles, see [DE13b, Section 2].

The Gaussian law is not the only limit law that is accessible with the dependency graph approach. Convergence to Poisson distribution can also be proved this way, as demonstrated in [AGG89]; again, this result has found applications, *e.g.*, in the theory of random geometric graphs (*cf.* [Pen02]).

We now leave convergence in distribution to discuss probabilities of rare events:

- In [Jan04], S. Janson has established some large deviation result involving the fractional chromatic number of the dependency graph.

- Another important, historically first use of dependency graphs is the *Lovász local lemma* [EL75, Spe77]. The goal here is to find a lower bound for the probability that $X_n = 0$ when the $Y_{n,i}$'s are indicator random variables, that is the probability that none of the $Y_{n,i}$'s is equal to 1. This inequality has found a large range of application to prove by probabilistic arguments the existence of an object (often a graph) with given properties: this method is known as the *probabilistic method*, see [AS08, Chapter 5].

Here, we shall see that dependency graphs can be used to prove mod-Gaussian convergence. This gives us a large collection of examples, for which the material of this article gives automatically some *precise* moderate deviation and normality zone results. Our deviation result has a larger domain of validity than the one of Döring and Eichelsbacher — see below.

In this chapter, we establish a general result involving dependency graphs (Theorem 9.1.7). In the next two chapters, we focus on examples and derive the mod-Gaussian convergence of the following renormalised statistics:

- subgraph count statistics in Erdös-Rényi random graphs (Chapter 10);
- random character values from central measures on partitions (Chapter 11).

Dependency graphs can also be used to cover the framework of *m*-dependence, widely used in statistics [Das08]. We shall develop this application in a forthcoming work [FMN16]. Unfortunately, we haven't been able to obtain similar result for mod-Poisson convergence, using dependency graphs.

9.1. The theory of dependency graphs

Let us consider a variable X, which writes as a sum

$$X = \sum_{\alpha \in V} Y_\alpha$$

of random variables Y_α indexed by a set V.

DEFINITION 9.1.1. *A graph G with vertex set V is called a* dependency graph *for the family of random variables* $\{Y_\alpha, \ \alpha \in V\}$ *if the following property is satisfied:*

> *If V_1 and V_2 are disjoint subsets of V such that there are no edges in G with one extremity in V_1 and one in V_2, then the sets of random variables $\{Y_\alpha\}_{\alpha \in V_1}$ and $\{Y_\alpha\}_{\alpha \in V_2}$ are independent (i.e., the σ-algebras generated by these sets are independent).*

EXAMPLE 9.1.2. Let $(Y_1, \ldots, Y_7)$ be a family with dependency graph
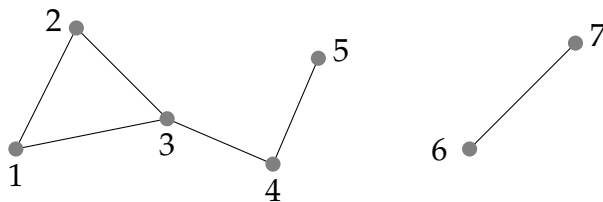
Then, $(Y_1, Y_2, Y_3, Y_4, Y_5)$ and (Y_6, Y_7) are independent (obvious), but the vectors (Y_1, Y_2) and (Y_4, Y_5) are also independent: although they are in the same connected component of the graph G, they are not directly connected by an edge $e \in E$.

REMARK 9.1.3. Note that a family of random variables may admit several dependency graphs. In particular, the complete graph with vertex set V is always a dependency graph. We are interested in dependency graphs with as few edges as possible. Note that a family of random variables does not always have a unique minimal dependency graph (minimal for edge-set inclusion), as shown by the following example.

EXAMPLE 9.1.4. Consider three independent Bernoulli random variables X_1, X_2, X_3, and $Y_1 = \mathbb{1}_{(X_2 = X_3)}$, $Y_2 = \mathbb{1}_{(X_1 = X_3)}$ and $Y_3 = \mathbb{1}_{(X_1 = X_2)}$. Then, the following graphs are minimal dependency graphs for (Y_1, Y_2, Y_3):
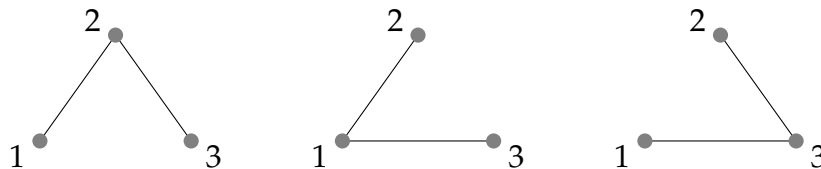
EXAMPLE 9.1.5. Fix a finite graph $G = (V, E)$. Take a family of independent non-constant random variables $(Y_e)_{e \in E}$ indexed by the edge-set E of G. For a vertex $v \in V$, define $X_v = \sum_e Y_e$ where the sum runs over incident edges to v. Then G is a dependency graph for the family $(X_v)_{v \in V}$. Moreover, it is minimal for edge-set inclusion and unique with this property.

The following bound on cumulants of sums of random variables has been established by S. Janson [Jan88, Lemma 4].

THEOREM 9.1.6. *For any integer $r \geq 1$, there exists a constant C_r with the following property. Let $\{Y_\alpha\}_{\alpha \in V}$ be a family of random variables with dependency graph G. We denote $N = |V|$ the number of vertices of G and D the maximal degree of G. Assume that the variables Y_α are uniformly bounded by a constant A. Then, if $X = \sum_\alpha Y_\alpha$, one has:*

$$|\kappa^{(r)}(X)| \leq C_r N (D+1)^{r-1} A^r.$$

In most applications for counting substructures in random objects, the Y_α are indicator variables, so that the *uniformly bounded* assumption is not restrictive. This theorem is often used to prove some central limit theorem. In [DE13b], Döring and Eichelsbacher analysed Janson's original proof and established that the theorem holded with $C_r = (2e)^r (r!)^3$. Then they used this new bound to obtain some moderate deviation results. Here, we will give a new proof of Janson's result, with a smaller value of the constant C_r. Namely, we will prove:

THEOREM 9.1.7. *Theorem 9.1.6 holds with $C_r = 2^{r-1} r^{r-2}$.*

We shall see at the end of this section, and in the next sections that this stronger version can be used to establish mod-Gaussian convergence and, thus, precise moderate deviation results. In fact, we prove a slightly stronger bound.

THEOREM 9.1.8. *With the same assumptions as above, one has:*

$$|\kappa^{(r)}(X)| \leq 2^{r-1}\, r^{r-2} \left(\sum_{\alpha} \mathbb{E}[|Y_\alpha|] \right) (D+1)^{r-1} A^{r-1}.$$

The next subsections are devoted to the proof of Theorem 9.1.8. Note that $\sum_\alpha \mathbb{E}[|Y_\alpha|] \leq N\,A$, so that Theorem 9.1.8 is indeed stronger than Theorem 9.1.7. On the other hand, if Y_α are Bernoulli variables indexed by $\alpha \in \{1,\dots,N\}$ with parameters $1/\alpha$, then $\sum_\alpha \mathbb{E}[|Y_\alpha|] \sim \log N$. In this case, when D is constant, the bound given in Theorem 9.1.8 is exponentially smaller than the one from Theorem 9.1.7.

REMARK 9.1.9. The hypothesis of bounded random variables can sometimes be lifted by mean of truncation methods. Indeed, if $(Y_\alpha)_{\alpha \in V}$ is a family of unbounded random variables with dependency graph G, then any truncated family $(Y_\alpha\, \mathbb{1}_{|Y_\alpha| \leq L_\alpha})_{\alpha \in V}$ with fixed levels of truncation L_α has the same dependency graph G. Thus, in many situations, one can prove the mod-Gaussian convergence of (an adequate renormalisation of) the truncated sum $S_{\text{truncated}} = \sum_{\alpha \in V} Y_\alpha\, \mathbb{1}_{|Y_\alpha| \leq L_\alpha}$, and then to use *ad-hoc* arguments in order to control the remainder $S_{\text{remainder}} = \sum_{\alpha \in V} Y_\alpha\, \mathbb{1}_{|Y_\alpha| > L_\alpha}$, such as moments inequalities (Bienaymé-Chebyshev). We shall develop these arguments in details in the forthcoming paper [FMN16].

9.2. Joint cumulants

There exists a multivariate version of cumulants, called joint cumulants, that we shall use to prove Theorem 9.1.7. We present in this paragraph its definition and basic properties. Most of this material can be found in Leonov's and Shiryaev's paper [LS59] (see also [JŁR00, Proposition 6.16]).

9.2.1. Preliminaries: set-partitions. We denote by $[n]$ the set $\{1,\dots,n\}$. Recall from Section 8.3 that a *set partition* of $[n]$ is a (non-ordered) family of non-empty disjoint subsets of S (called parts of the partition), whose union is $[n]$. For instance,

$$\{\{1,3,8\},\{4,6,7\},\{2,5\}\}$$

is a set partition of $[8]$. Denote $\mathfrak{Q}(n)$ the set of set partitions of $[n]$. Then $\mathfrak{Q}(n)$ may be endowed with a natural partial order: the *refinement* order. We say that π is *finer* than π' or π' *coarser* than π (and denote $\pi \leq \pi'$) if every part of π is included in a part of π'.

Lastly, denote μ the Möbius function of the poset $\mathfrak{Q}(n)$. In this paper, we only use evaluations of μ at pairs $(\pi, \{[n]\})$ (the second argument is the partition of $[n]$ in only one part, which is the maximum element of $\mathfrak{Q}(n)$), so we shall use abusively the notation $\mu(\pi)$ for $\mu(\pi, \{[n]\})$. In this case, the value of the Möbius function is given by:

$$\mu(\pi) = \mu(\pi, \{[n]\}) = (-1)^{\#(\pi)-1}(\#(\pi)-1)!\,. \tag{40}$$

9.2.2. Definition and properties of joint cumulants. If $X_1, \ldots, X_r$ are random variables with finite moments on the same probability space (denote $\mathbb{E}$ the expectation on this space), we define their joint cumulant by

$$\kappa(X_1, \ldots, X_r) = [t_1 \cdots t_r] \, \log \left(\mathbb{E}\left[e^{t_1 X_1 + \cdots + t_r X_r} \right] \right). \tag{41}$$

As usual, $[t_1 \ldots t_r]F$ stands for the coefficient of $t_1 \cdots t_r$ in the series expansion of F in positive powers of $t_1, \ldots, t_r$. Note that joint cumulants are multilinear functions. In the case where all the X_i's are equal, we recover the r-th cumulant $\kappa^{(r)}(X)$ of a single variable. Using set-partitions, joint cumulants can be expressed in terms of joint moments, and *vice-versa*:

$$\mathbb{E}[X_1 \cdots X_r] = \sum_{\pi \in \mathfrak{Q}(r)} \prod_{C \in \pi} \kappa(X_i \, ; \, i \in C); \tag{42}$$

$$\kappa(X_1, \ldots, X_r) = \sum_{\pi \in \mathfrak{Q}(r)} \mu(\pi) \prod_{C \in \pi} \mathbb{E}\left[\prod_{i \in C} X_i \right]. \tag{43}$$

In these equations, $C \in \pi$ shall be understood as "C is a part of the set partition π". Recall that $\mu(\pi)$ has an explicit expression given by Equation (40). For example the joint cumulants of one or two variables are simply the mean of a single random variable and the covariance of a couple of random variables:

$$\kappa(X_1) = \mathbb{E}[X_1] \qquad ; \qquad \kappa(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1] \, \mathbb{E}[X_2].$$

For three variables, one has

$$\kappa(X_1, X_2, X_3) = \mathbb{E}[X_1 X_2 X_3] - \mathbb{E}[X_1 X_2] \, \mathbb{E}[X_3] - \mathbb{E}[X_1 X_3] \, \mathbb{E}[X_2]$$
$$- \mathbb{E}[X_2 X_3] \, \mathbb{E}[X_1] + 2 \, \mathbb{E}[X_1] \, \mathbb{E}[X_2] \, \mathbb{E}[X_3].$$

REMARK 9.2.1. The most important property of cumulants is their relation with independence: if the variables $X_1, \ldots, X_r$ can be split in two non-empty sets of variables which are independent with each other, then $\kappa(X_1, \ldots, X_r)$ vanishes [JŁR00, Proposition 6.16 (v)]. We will not need this property here. In fact, we will prove a more precise version of it, see Equation (48).

9.2.3. Statement with joint cumulants. Theorems 9.1.6 and 9.1.7 have some analogues with joint cumulants. Let $\{Y_\alpha\}_{\alpha \in V}$ be a family of random variables with dependency graph G. As in Theorem 9.1.6, we assume that the variables Y_α are uniformly bounded by a constant A: *i.e.*, for all $\alpha \in V$,

$$\|Y_\alpha\|_\infty \leq A.$$

Consider r subsets $V_1, V_2, \ldots, V_r$ of V, non necessarily distinct and set $X_i = \sum_{\alpha \in V_i} Y_\alpha$ (for $i \in [r]$). We denote D_i the maximal number of vertices in V_i adjacent to a given vertex (not necessarily in V_i). Then one has the following result.

THEOREM 9.2.2. *With the notation above,*

$$|\kappa(X_1, \ldots, X_r)| \leq 2^{r-1} \, r^{r-2} \, |V_1| \, (D_2 + 1) \cdots (D_r + 1) \, A^r.$$

The proof of this theorem is very similar to the one of Theorem 9.1.7. However, to simplify notation, we only prove the latter here.

9.3. Useful combinatorial lemmas

We start our proof of Theorem 9.1.7 by stating a few lemmas on graphs and spanning trees.

9.3.1. A functional on graphs. In this section, we consider graphs H with multiple edges and loops. We use the standard notation $V(H)$ and $E(H)$ for their vertex and edge sets. For a graph H and a set partition π of $V(H)$, we denote $\pi \perp H$ when the following holds: for any edge $\{i, j\} \in E(H)$, the elements i and j lie in different parts of π (in this case we use the notation $i \not\sim_\pi j$). We introduce the following functional on graphs H:

$$F_H = (-1)^{|V(H)|-1} \sum_{\pi \perp H} \mu(\pi).$$

LEMMA 9.3.1. *For any graph H, one has*

$$F_H = \sum_{\substack{E \subset E(H) \\ (V(H),E) \text{ connected}}} (-1)^{|E|-|V(H)|+1}.$$

PROOF. To simplify notation, suppose $V(H) = [r]$. We denote $\mathbb{1}_{(P)}$ the characteristic function of the property (P). By inclusion-exclusion,

$$(-1)^{|V(H)|-1} F_H = \sum_{\pi \in \mathfrak{Q}(r)} \left(\prod_{(i,j) \in E(H)} \mathbb{1}_{i \not\sim_\pi j} \right) \mu(\pi) = \sum_{\pi \in \mathfrak{Q}(r)} \left(\prod_{(i,j) \in E(H)} (1 - \mathbb{1}_{i \sim_\pi j}) \right) \mu(\pi)$$

$$= \sum_{E \subset E(H)} \sum_{\pi \in \mathfrak{Q}(r)} (-1)^{|E|} \left(\prod_{(i,j) \in E} \mathbb{1}_{i \sim_\pi j} \right) \mu(\pi)$$

$$= \sum_{E \subset E(H)} (-1)^{|E|} \left[\sum_{\substack{\pi \text{ such that} \\ \forall (i,j) \in E, \, i \sim_\pi j}} \mu(\pi) \right].$$

But the quantity in the bracket is 0 unless the only partition in the sum is the maximal partition $\{[r]\}$, in which case it is 1. This corresponds to the case where the edges in E form a connected subgraph of H. □

COROLLARY 9.3.2. *The functional F_H fulfills the deletion-contraction induction, i.e., if e is an edge of H which is not a loop, then*

$$F_H = F_{H/e} + F_{H \setminus e},$$

where $H \setminus e$ (respectively H/e) are the graphs obtained from H by deleting (resp. contracting) the edge e.

PROOF. The first term corresponds to sets of edges containing e, and the second to those that do not contain e. □

This induction (over-)determines F_H together with the initial conditions:

$$\begin{cases} F_\bullet = 1, \\ F_{\underset{\bullet}{\bigcirc}} = F_{\underset{\bullet}{\bigvee}} = \cdots = 0, \\ F_H = 0 \quad \text{if } H \text{ is disconnected.} \end{cases}$$

COROLLARY 9.3.3. *For any graph H, the quantity F_H is nonnegative and less or equal than the number ST_H of spanning trees in H.*

PROOF. The quantity ST_H fulfills the same induction as F_H with initial conditions:

$$\begin{cases} ST_\bullet = 1, \\ ST_{\underset{\bullet}{\bigcirc}} = ST_{\underset{\bullet}{\bigvee}} = \cdots = 1, \\ ST_H = 0 \quad \text{if } H \text{ is disconnected.} \end{cases} \qquad \square$$

REMARK 9.3.4. If H is connected, both F_H and ST_H are actually specialisations of the bivariate *Tutte polynomial* $T_H(x, y)$ of H (*cf.* [Bol98, Chapter X]):

$$F_H = T_H(1, 0) \qquad ; \qquad ST_H = T_H(1, 1).$$

This explains the deletion-contraction relation. As the bivariate Tutte polynomials has non-negative coefficients, it also explains the inequality $0 \leq F_H \leq ST_H$.

9.3.2. Induced graphs containing spanning trees. Fix a graph G (typically the dependency graphs of our family of variables). For a list $(v_1, \ldots, v_r)$ of r vertices of G, we define the induced graph $G[v_1, \ldots, v_r]$ as follows:

- its vertex set is $[r]$;
- there is an edge between i and j if and only if $v_i = v_j$ or v_i and v_j are linked in G.

We will be interested in spanning trees of induced graphs. As the vertex set is $[r]$, these spanning trees may be seen as *Cayley trees*. Recall that a Cayley tree of size r is by definition a tree with vertex set $[r]$ (Cayley trees are neither rooted, nor embedded in the plane, they are only specified by an adjacency matrix). These objects are enumerated by the well-known Cayley formula established by C. Borchardt in [Bor60]: there are exactly r^{r-2} Cayley trees of size r.

LEMMA 9.3.5. *Fix a Cayley tree T of size r and a graph G with N vertices and maximal degree D. Fix a vertex v_1 of G. The number of lists $(v_1, \ldots, v_r)$ of r vertices of G such that T is contained in the induced subgraph $G[v_1, \ldots, v_r]$ is bounded from above by*

$$(D + 1)^{r-1}.$$

PROOF. Lists $(v_1, \ldots, v_r)$ as in the lemma are constructed as follows. First consider a neighbour j of 1 in T. As we require $G[v_1, \ldots, v_r]$ to contain T, the vertices 1 and j must also be neighbours in $G[v_1, \ldots, v_r]$, *i.e.* $v_j = v_1$ or v_j is a neighbour of v_1 in T. Thus, once v_1 is fixed, there are at most $D + 1$ possible values for v_j. The same is true for all neighbours of 1 and then for all neighbours of neighbours of 1 and so on. $\square$

We have the following immediate consequence.

COROLLARY 9.3.6. *Let G be a graph with n vertices and maximal degree D and $r \geq 1$. Fix a vertex v_1 of G. The number of couples $((v_1, \ldots, v_r), T)$ where each v_i is a vertex of V and T a spanning tree of the induced subgraph $G[v_1, \ldots, v_r]$ is bounded above by*

$$r^{r-2} (D+1)^{r-1}.$$

9.3.3. Spanning trees and set partitions of vertices. Recall that ST_H denotes the number of spanning trees of a graph H. Consider now a graph H with vertex set $[r]$ and a set partition $\pi = (\pi_1, \ldots, \pi_t)$ of $[r]$. For each i, we denote $ST^{\pi_i}(H) = ST_{H[\pi_i]}$ the number of spanning trees of the graph induced by H on the vertex set π_i. We also use the multiplicative notation

$$ST^{\pi}(H) = \prod_{j=1}^{t} ST^{\pi_j}(H).$$

We can also consider the contraction H/π of H with respect to π. By definition, it is the multigraph (*i.e.* graph with multiple edges, but no loops) defined as follows. Its vertex set is the index set $[t]$ of the parts of π and, for $i \neq j$, there are as many edges between i and j as edges between a vertex of π_i and a vertex of π_j in H. Denote $ST_{\pi}(H) = ST_{H/\pi}$ the number of spanning trees of this contracted graph (multiple edges are here important). This should not be confused with $ST^{\pi}(H)$: in the latter, π is placed as an exponent because the quantity is multiplicative with respect to the part of π.

Note that the union of a spanning tree $\overline{T}$ of H/π and of spanning trees T_i of $H[\pi_i]$ (one for each $1 \leq i \leq t$) gives a spanning tree T of H. Conversely, take a spanning tree T on H and a bicolouration of its edges. Edges of colour 1 can be seen as a subgraph of H with the same vertex set $[r]$. This graph is of course acyclic. Its connected components define a partition $\pi = \{\pi_1, \ldots, \pi_t\}$ of $[r]$ and edges of colour 1 correspond to a collection of spanning trees T_i of $H[\pi_i]$ (for $1 \leq i \leq t$). Besides, edges of colour 2 define a spanning tree $\overline{T}$ on H/π.

Therefore, we have described a bijection between spanning trees T_0 of H with a bicolouration of their edges and triples $(\pi, \overline{T}, (T_i)_{1 \leq i \leq t})$ where:

- π is a set partition of the vertex set $[r]$ of H (we denote t its number of parts);
- $\overline{T}$ is a spanning tree of the contracted graph H/π;
- for each $1 \leq i \leq t$, T_i is a spanning tree of the induced graph $H[\pi_i]$.

Before giving a detailed example, let us state the numerical corollary of this bijection:

$$2^{r-1} ST_H = \sum_{\pi} ST_{\pi}(H) ST^{\pi}(H), \tag{44}$$

where the sum runs over all set partitions π of $[r]$.

Our bijection is illustrated on Figure 1, with the following conventions.

- On the left, blue plain edges are edges of colour 1 in the tree T_0; on the right, these blue plain edges are the edges of the spanning trees T_i.
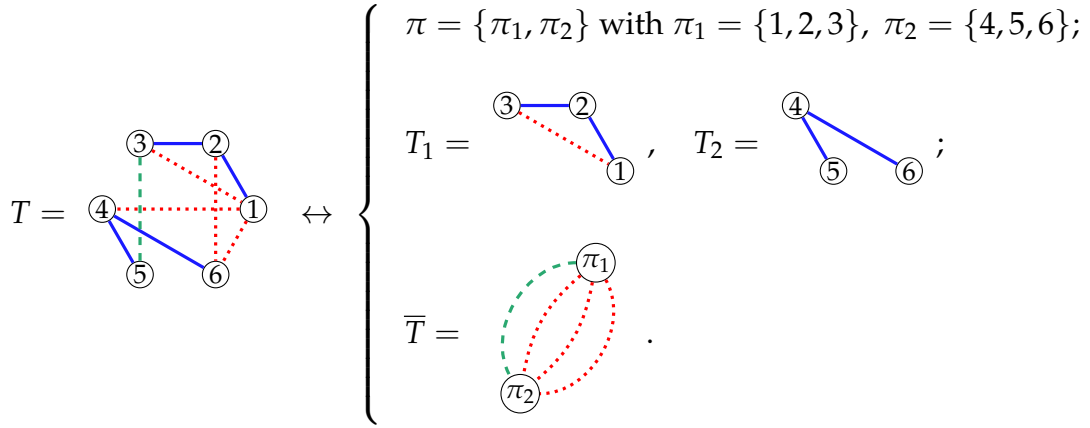
FIGURE 1. Bijection explaining Identity (44).

- On the left, green dashed edges are edges of colour 2 in the tree T_0; on the right, these green dashed edges are edges of the spanning tree $\overline{T}$.
- Red dotted edges belong to the graphs H, H/π, $H[\pi_i]$ but not to their spanning tree $T_0, \overline{T}, T_i$.

Note that in this example the graph H/π has two vertices linked by four edges. These four edges correspond to the edges $(1,4)$, $(1,6)$, $(2,6)$ and $(3,5)$ of H. In the example, the spanning tree $\overline{T}$ is the edge $\{(3,5)\}$. If we had chosen another edge, the tree T on the left-hand side would have been different. Hence, in the Equality (44), the multiple edges of H/π must be taken into account: in our example, $\mathrm{ST}_\pi(H) = 4$.

9.4. Proof of the bound on cumulants

Recall that we want to find a bound for $\kappa^{(r)}(X)$. As X writes $X = \sum_{\alpha \in V} Y_\alpha$, we may use joint cumulants and expand by multilinearity:

$$\kappa^{(r)}(X) = \sum_{\alpha_1,\ldots,\alpha_r} \kappa(Y_{\alpha_1},\ldots,Y_{\alpha_r}). \tag{45}$$

The sum runs over lists of r elements in V, that is vertices of the dependency graph G. The proof consists in bounding each summand $\kappa(Y_{\alpha_1},\ldots,Y_{\alpha_r})$, with a bound depending on the induced subgraph $G[\alpha_1,\ldots,\alpha_r]$.

9.4.1. Bringing terms together in joint cumulants. Recall the moment-cumulant formula (43) which states $\kappa(Y_{\alpha_1},\ldots,Y_{\alpha_r}) = \sum_\pi \mu(\pi) M_\pi$, where

$$M_\pi = \prod_{B \in \pi} \mathbb{E}\left[\prod_{i \in B} Y_{\alpha_i}\right].$$

We warn the reader that the notation M_π is a little bit abusive as this quantity depends also on the list $(\alpha_1,\ldots,\alpha_r)$. By hypothesis, G is a dependency graph for the family $\{Y_\alpha\}_{\alpha \in V}$. Hence if some block B of a partition π can be split into two sub-blocks B_1 and B_2 such that the sets of vertices $\{\alpha_i\}_{i \in B_1}$ and $\{\alpha_i\}_{i \in B_2}$ are disjoint and do not share an edge, then

$$\mathbb{E}\left[\prod_{i \in B} Y_{\alpha_i}\right] = \mathbb{E}\left[\prod_{i \in B_1} Y_{\alpha_i}\right] \times \mathbb{E}\left[\prod_{i \in B_2} Y_{\alpha_i}\right]. \tag{46}$$

Therefore, $M_\pi = M_{\phi_H(\pi)}$, where $H = G[\alpha_1, \ldots, \alpha_r]$ and $\phi_H(\pi)$ is the refinement of π obtained as follows: for each part π_i of π, consider the induced graph $H[\pi_i]$ and replace π_i by the collection of vertex sets of the connected components of $H[\pi_i]$. This construction is illustrated on Figure 2.
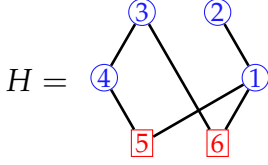
For example, consider the graph H here opposite and the partition $\pi = \{\pi_1, \pi_2\}$ with $\pi_1 = \{1, 2, 3, 4\}$ and $\pi_2 = \{5, 6\}$. Then $H[\pi_1]$ (respectively, $H[\pi_2]$) has two connected components with vertex sets $\{1, 2\}$ and $\{3, 4\}$ (resp. $\{5\}$ and $\{6\}$). Thus

$$\phi_H(\pi) = \{\{1, 2\}, \{3, 4\}, \{5\}, \{6\}\}.$$

FIGURE 2. Illustration of the definition of ϕ_H.

We can thus write

$$\kappa(G) = \sum_{\pi'} M_{\pi'} \left(\sum_{\pi \in \phi_H^{-1}(\pi')} \mu(\pi) \right).$$

Fix $\pi' = (\pi_1', \ldots, \pi_t')$ and let us have a closer look to the expression in the parentheses that we will call $\alpha_{\pi'}$. To compute it, it is convenient to consider the contraction H/π' of the graph H with respect to the partition π'.

LEMMA 9.4.1. *Let π' be a set partition of $[r]$. If one of the induced graph $H[\pi_i']$ is disconnected, then $\alpha_{\pi'} = 0$. Otherwise, $\alpha_{\pi'} = (-1)^{\ell(\pi')-1} F_{H/\pi'}$.*

PROOF. The first part is immediate, as $\phi_H^{-1}(\pi') = \varnothing$ in this case.

If all induced graphs are connected, let us try to describe $\phi_H^{-1}(\pi')$. All set partitions π of this set are coarser than π', so can be seen as set partitions of the index set $[r]$ of the parts of π'. This identification does not change their Möbius functions, which depends only on the number of parts. Then, it is easy to see that π lies in $\phi_H^{-1}(\pi')$ if and only if π is coarser than π' and two elements in the same part of π never share an edge in H/π' (here, π is seen as a set partition of $[r]$). In other words, π lies in $\phi_H^{-1}(\pi')$ if and only if $\pi \perp (H/\pi')$. This implies the Lemma. $\square$

Consequently,

$$\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r}) = \sum_{\pi'} (-1)^{\ell(\pi')-1} M_{\pi'} F_{H/\pi'} \left(\prod_{i=1}^{t} \mathbb{1}_{H[\pi_i'] \text{ connected}} \right), \qquad (47)$$

where the sum runs over all set partitions π' of $[r]$.

9.4.2. Bounding all the relevant quantities. Using $|Y_{\alpha_i}| \le A$ for all i, we get the inequality:

$$|M_\pi| \le A^{r-1} \, \mathbb{E}[|Y_{\alpha_1}|].$$

Finally, to bound each summand $\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r})$, we shall use the following bounds:

$$|F_{H/\pi'}| \le ST_{H/\pi'} \qquad \text{by Corollary 9.3.3;}$$

$$\mathbb{1}_{H[\pi_i'] \text{ connected}} \le ST_{H[\pi_i']}.$$

Thus, Equation (47) gives

$$|\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r})| \leq A^{r-1} \, \mathbb{E}[|Y_{\alpha_1}|] \sum_{\pi'} \mathrm{ST}_{H/\pi'} \left(\prod_{i=1}^{t} \mathrm{ST}_{H[\pi_i']} \right) = A^{r-1} \, \mathbb{E}[|Y_{\alpha_1}|] \, 2^{r-1} \, \mathrm{ST}_H,$$

(48)

the last equality corresponding to Equation (44). Recall that $H = G[\alpha_1, \ldots, \alpha_r]$. Summing over $\alpha_1, \ldots, \alpha_r$ and using Equation (45), we get

$$|\kappa^{(r)}(X)| \leq (2A)^{r-1} \sum_{\alpha_1} \mathbb{E}[|Y_{\alpha_1}|] \left[\sum_{\alpha_2, \ldots, \alpha_r} \mathrm{ST}_{G[\alpha_1, \ldots, \alpha_r]} \right].$$

From Corollary 9.3.6, we know that the expression in the bracket is bounded by the quantity $r^{r-2}(D+1)^{r-1}$ (for any fixed α_1). This completes the proof of Theorem 9.1.8.

9.5. Sums of random variables with a sparse dependency graph

An immediate application of Theorem 9.1.7 is the following general result on sums of weakly dependent random variables, to be compared with [Pen02, §2.3]. Let $X_n = \sum_{i=1}^{N_n} Y_i$ be a sum of random variables, where the Y_i's have a dependency graph of maximal degree $D_n - 1$, and satisfy $\|Y_i\|_\infty \leq A$ for some $A \geq 0$ (independent of i). We also assume that X_n is not deterministic, so that its variance is non-zero.

THEOREM 9.5.1. *We assume that the dependency graph of $(Y_i)_{1 \leq i \leq N_n}$ is sparse, in the sense that $\lim_{n \to \infty} \frac{D_n}{N_n} = 0$.*

(1) There exists a positive constant C such that, for all $r \geq 2$, $\left| \kappa^{(r)} \left(\frac{X_n}{D_n} \right) \right| \leq (Cr)^r \frac{N_n}{D_n}$.

(2) Consider the bounded sequences

$$\sigma_n^2 = \frac{D_n}{N_n} \, \kappa^{(2)} \left(\frac{X_n}{D_n} \right) \qquad ; \qquad L_n = \frac{1}{\sigma_n^3} \frac{D_n}{N_n} \, \kappa^{(3)} \left(\frac{X_n}{D_n} \right).$$

If they have limits $\sigma^2 > 0$ and L, then

$$\frac{X_n - \mathbb{E}[X_n]}{D_n \, \sigma_n} \left(\frac{D_n}{N_n} \right)^{1/3}$$

converges in the mod-Gaussian sense with parameters $t_n = (N_n/D_n)^{1/3}$ and limiting function $\psi(z) = \exp(Lz^3/6)$.

(3) If furthermore,

$$\sigma_n^2 = \sigma^2 \left(1 + o\left((D_n/N_n)^{1/3} \right) \right),$$

then the variable

$$\frac{X_n - \mathbb{E}[X_n]}{D_n \, \sigma} \left(\frac{D_n}{N_n} \right)^{1/3}$$

converges in the mod-Gaussian sense with parameters $(N_n/D_n)^{1/3}$ and limiting function $\psi(z) = \exp(Lz^3/6)$ (the difference with the previous item is that σ_n has been replaced by the usually more explicit quantity σ).

Of course, if (2) holds, then the results of this paper imply that

- $X_n/(\sigma_n\sqrt{N_nD_n})$ satisfies a central limit theorem, with a normality zone of size $o((\frac{N_n}{D_n})^{1/6})$;

- furthermore, for $x > 0$,

$$\mathbb{P}[X_n - \mathbb{E}[X_n] \geq (N_n)^{2/3}(D_n)^{1/3}\sigma_n x] = \frac{e^{-\left(\frac{N_n}{D_n}\right)^{1/3}\frac{x^2}{2}}}{x\sqrt{2\pi(N_n/D_n)^{1/3}}} \exp\left(\frac{Lx^3}{6}\right)(1+o(1))$$

(49)

and a similar result holds for negative deviations.

The same holds replacing σ_n by σ if one has the assumption on the speed of convergence of σ_n towards σ, given in the third item of Theorem 9.5.1. If, moreover, the cumulants of X_n satisfy (27), then (49) still holds when we replace x by a positive sequence x_n tending to infinity with $x_n = o((N_n/D_n)^{1/12})$ — see Proposition 5.2.1.
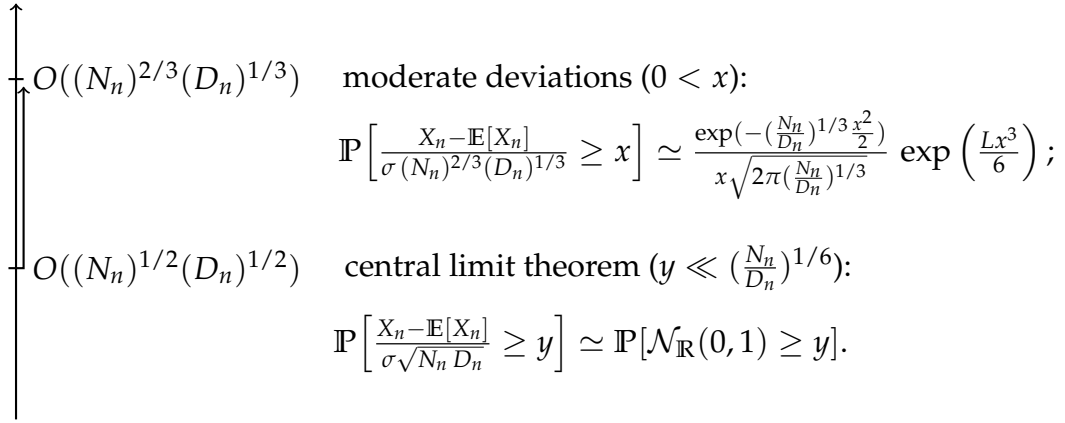
order of fluctuations

$O((N_n)^{2/3}(D_n)^{1/3})$ moderate deviations ($0 < x$):

$$\mathbb{P}\left[\frac{X_n - \mathbb{E}[X_n]}{\sigma(N_n)^{2/3}(D_n)^{1/3}} \geq x\right] \simeq \frac{\exp(-(\frac{N_n}{D_n})^{1/3}\frac{x^2}{2})}{x\sqrt{2\pi(\frac{N_n}{D_n})^{1/3}}} \exp\left(\frac{Lx^3}{6}\right);$$

$O((N_n)^{1/2}(D_n)^{1/2})$ central limit theorem ($y \ll (\frac{N_n}{D_n})^{1/6}$):

$$\mathbb{P}\left[\frac{X_n - \mathbb{E}[X_n]}{\sigma\sqrt{N_n D_n}} \geq y\right] \simeq \mathbb{P}[\mathcal{N}_\mathbb{R}(0,1) \geq y].$$

FIGURE 3. Panorama of the fluctuations of a sum of N_n random variables that have a sparse dependency graph, of degree $D_n - 1$ with $\lim_{n\to\infty}\frac{D_n}{N_n} = 0$.

Notice that in the case of independent random variables, $D_n = 1$ and we recover the mod-Gaussian convergence of Example 2.1.2.

In the next chapters, we shall study some particular cases of Theorem 9.5.1, and see that most computations are related to the asymptotic analysis of the first cumulants of X_n.

Subgraph count statistics in Erdös-Rényi random graphs

In this chapter, we consider Erdös-Rényi model $\Gamma(n, p_n)$ of random graphs. A random graph Γ with this distribution is described as follows. Its vertex set is $[n]$ and for each pair $\{i, j\} \subset [n]$ with $i \neq j$, there is an edge between i and j with probability p_n. Moreover, all these events are independent. We are then interested in the following random variables, called *subgraph count statistics*. If γ is a fixed graph of size k, then $X_\gamma^{(n)}$ is the number of copies of γ *contained* in the graph $\Gamma(n, p_n)$ (a more formal definition is given in the next paragraph). This is a classical parameter in random graph theory; see, *e.g.*, the book of S. Janson, T. Łuczak and A. Ruciński [JŁR00].

The first result on this parameter was obtained by P. Erdös and A. Rényi, *cf.* [ER60]. They proved that, if γ belongs to some particular family of graphs (called *balanced*), then one has a threshold: namely,

$$\lim_{n \to \infty} \mathbb{P}[X_\gamma^{(n)} > 0] = \begin{cases} 0 & \text{if } p_n = o\left(n^{-1/m(\gamma)}\right); \\ 1 & \text{if } n^{-1/m(\gamma)} = o(p_n), \end{cases}$$

where $m(\gamma) = |E(\gamma)| / |V(\gamma)|$. This result was then generalised to all graphs by B. Bollobás [Bol01], but the parameter $m(\gamma)$ is in general more complicated than the quotient above. Consider the case $n^{-1/m(\gamma)} = o(p_n)$, when the graph $\Gamma(n, p_n)$ contains with high probability a copy of γ. It was then proved by A. Ruciński (see [Ruc88]) that, under the additional assumption $n^2(1 - p_n) \to \infty$, the fluctuations of $X_\gamma^{(n)}$ are Gaussian. This result can be obtained using dependency graphs; see *e.g.* [JŁR00, pages 147-152].

Here, we consider the case where $p_n = p$ is a constant sequence ($0 < p < 1$). The possibility of relaxing this hypothesis is discussed in Subsection 10.3.3. Denote $\alpha_n = n^2$ and $\beta_n = n^{k-2}$, where k is the number of vertices of γ. It is easy to check that

$$\mathbb{E}[X_\gamma^{(n)}] = c \, \alpha_n \, \beta_n \quad ; \quad \mathrm{Var}(X_\gamma^{(n)}) = \sigma^2 \, \alpha_n \, (\beta_n)^2$$

for some positive constants c and σ — see, *e.g.*, [JŁR00, Lemma 3.5]. Hence, Ruciński's central limit theorem asserts that, if $T \sim x\sqrt{\alpha_n}$ for some fixed real x, then

$$\lim_{n \to \infty} \mathbb{P}\left[\frac{X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]}{\beta_n} \geq T\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/\sigma} e^{-u^2/2} du.$$

Using Theorem 9.1.7, we shall extend this result to a framework where x tends to infinity, but not to quickly: $(\alpha_n)^{1/2} \ll T \ll (\alpha_n)^{3/4}$.

THEOREM 10.0.1. *Let $0 < p < 1$ and γ be a graph with k vertices. We consider $X_\gamma^{(n)}$ the number of copies of γ contained in Erdös-Rényi random graph $\Gamma(n, p)$. Let α_n and β_n be defined as above.*

(1) *The renormalised variable* $(X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}])/((\alpha_n)^{1/3}\beta_n)$ *converges mod-Gaussian with parameters* $t_n = \sigma^2 \alpha_n^{1/3}$ *and limiting function* $\psi(z) = \exp\left(\frac{L}{6}z^3\right)$, *where* σ *and* L *are computed in Section 10.3.2. The convergence takes place on the whole complex plane with speed of convergence* $o(\alpha_n^{-1/3})$.

(2) *Therefore,* $(X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}])/((\alpha_n)^{1/2}\beta_n)$ *converges in distribution towards a Gaussian law, with normality zone* $o((\alpha_n)^{1/6})$. *Moreover, the deviation probabilities at the edge of the normality zone and at a slightly larger scale are given as follows: is* x_n *is a positive sequence bounded away from* 0 *with* $x_n = o((\alpha_n)^{1/12})$, *then*

$$\mathbb{P}\left[\frac{X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]}{(\alpha_n)^{1/2}\beta_n} \geq \sigma^2(\alpha_n)^{1/6}x_n\right] = \frac{e^{-\frac{(x_n)^2(\alpha_n)^{1/3}\sigma^2}{2}}}{x_n(\alpha_n)^{1/6}\sigma\sqrt{2\pi}} e^{\frac{L(x_n)^3}{6}}(1 + o(1)).$$

$$\mathbb{P}\left[\frac{X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]}{(\alpha_n)^{1/2}\beta_n} \leq -\sigma^2(\alpha_n)^{1/6}x_n\right] = \frac{e^{-\frac{(x_n)^2(\alpha_n)^{1/3}\sigma^2}{2}}}{x_n(\alpha_n)^{1/6}\sigma\sqrt{2\pi}} e^{\frac{-L(x_n)^3}{6}}(1 + o(1)).$$

A similar result has been obtained by H. Döring and P. Eichelsbacher in [DE13b, Theorem 2.3]. However,

- their result is less precise as they only obtain the equivalence of the logarithms of the relevant quantities (in particular, when we look at the logarithm, the second factor of the right-hand side is negligible);

- and they cover a smaller zone of deviation.

Unfortunately, we cannot get deviation results when $x_n \sim t(\alpha_n)^{1/3}$ for some real number t; this would amount to evaluate $\mathbb{P}[X_\gamma^{(n)} > (1 + \varepsilon)\mathbb{E}[X_\gamma^{(n)}]]$. For large deviations equivalents of

$$\log \mathbb{P}[X_\gamma^{(n)} > (1 + \varepsilon)\mathbb{E}[X_\gamma^{(n)}]],$$

there is a quite large literature, see [CV11, Theorem 4.1] and [Cha12] for recent results in this field. As we consider deviations of a different scale, our result is neither implied by, nor implies these results. Note, however, that their large deviation results are equivalents of the logarithm of the probability, while our statement is an equivalent for the probability itself.

10.1. A bound on cumulants

10.1.1. Subgraph count statistics. In the following we denote $\mathfrak{A}(n, k)$ the set of *arrangements* in $[n]$ of length k, *i.e.*, lists of k distinct elements in $[n]$. The cardinality of $\mathfrak{A}(n, k)$ is the falling factorial $n^{\downarrow k} = n(n-1)\cdots(n-k+1)$. Let $A = (a_1, \ldots, a_k)$ be an arrangement in $[n]$ of length k, and γ be a fixed graph with vertex set $[k]$. Recall that $\Gamma = \Gamma(n, p_n)$ is a random Erdös-Rényi graph on $[n]$. We denote $\delta_\gamma(A)$ the following random variable:

$$\delta_\gamma(A) = \begin{cases} 1 & \text{if } \gamma \subseteq \Gamma[a_1, \ldots, a_k]; \\ 0 & \text{else.} \end{cases} \tag{50}$$

Here $\Gamma[a_1, \ldots, a_k]$ denotes the graph induced by Γ on vertex set $\{a_1, \ldots, a_k\}$. As our data is an *ordered* list $(a_1, \ldots, a_k)$, this graph can canonically be seen as a graph on

vertex set $[k]$, that is the same vertex set as γ. Then the inclusion should be understood as inclusion of edge sets.

For any graph γ with k vertices and any integer $n \geq 1$, we then define the random variable $X_\gamma^{(n)}$ by

$$X_\gamma^{(n)} = \sum_{A \in \mathfrak{A}(n,k)} \delta_\gamma(A).$$

REMARK 10.1.1. It would also be natural to replace in Definition (50) the inclusion by an equality $\gamma = \Gamma[a_1, \ldots, a_k]$. This would lead to other random variables $Y_\gamma^{(n)}$, called *induced* subgraph count statistics. Their asymptotic behavior is harder to study (in particular, fluctuations are not always Gaussian; see [JŁR00, Theorem 6.52]). Notice that if γ is a complete graph, then both definitions coincide.

10.1.2. A dependency graph of the subgraph count statistics. Fix some graph γ with vertex set $[k]$. By definition, the variable we are interested in writes as a sum

$$X_\gamma^{(n)} = \sum_{A \in \mathfrak{A}(n,k)} \delta_\gamma(A).$$

We shall describe a dependency graph for the variables $\{\delta_\gamma(A)\}_{A \in \mathfrak{A}(n,k)}$.

For each pair $e = \{v, v'\} \subset [n]$, denote I_e the indicator function of the event: *e is in the graph* $\Gamma(n, p)$. By definition of the model $\Gamma(n, p)$, the random variables I_e are independent Bernouilli variables of parameter p. Then, for an arrangement A, denote $E(A)$ the set of pairs $\{v, v'\}$ where v and v' appear in the arrangement A. One has

$$\delta_\gamma(A) = \prod_{e \in E'(A)} I_e,$$

where $E'(A)$ is a subset of $E(A)$ determined by the graph γ. In particular, if, for two arrangements A and A', one has $|E(A) \cap E(A')| = \varnothing$ (equivalently, $|A \cap A'| \leq 1$), then the variables $\delta_\gamma(A)$ and $\delta_\gamma(A')$ are defined using different variables I_e (and, hence, are independent). This implies that the following graph denoted $\mathbb{B}$ is a dependency graph for the family of variables $\{\delta_\gamma(A)\}_{A \in \mathfrak{A}(n,k)}$:

- its vertex set is $\mathfrak{A}(n,k)$;
- there is an edge between A and A' if $|A \cap A'| \geq 2$.

Considering this dependency graph is quite classical — see, *e.g.*, [JŁR00, Example 1.6].

All variables in this graph are Bernoulli variables and, hence, bounded by 1. Besides the graph $\mathbb{B}$ has $N = n^{\downarrow k}$ vertices, and is regular of degree D smaller than

$$\binom{k}{2}^2 2 (n-2)(n-3) \cdots (n-k+1) < k^4 n^{k-2}.$$

Indeed, a neighbour A' of a fixed arrangement $A \in \mathfrak{A}(n,k)$ is given as follows:

- choose a pair $\{a_i, a_j\}$ in A that will appear in A';
- choose indices i' and j' such that $a'_{i'} = a_i$ and $a'_{j'} = a_j$ (these indices are different but their order matters);

- choose the other values in the arrangement A'.

So, we may apply Theorem 9.1.7 and we get:

PROPOSITION 10.1.2. *Fix a graph γ of vertex set $[k]$. For any $r \leq 1$, one has*

$$\left| \kappa^{(r)}(X_\gamma^{(n)}) \right| \leq 2^{r-1} r^{r-2} n^k (k^4 n^{k-2})^{r-1}.$$

10.2. Polynomiality of cumulants

10.2.1. Dealing with several arrangements. Consider a list $(A^1, \ldots, A^r)$ of arrangements. We associate to this data two graphs (unless said explicitly, we always consider loopless simple graphs, and $V(G)$ and $E(G)$ denote respectively the edge and vertex sets of a graph G):

- the graph $G_\mathbb{A}$ has vertex set

$$V_\mathbb{k} = V(G_\mathbb{A}) = \{(t,i) \mid 1 \leq i \leq r, 1 \leq t \leq k_i\}$$

  and an edge between (t,i) and (s,j) if and only if $a_t^i = a_s^j$. It is always a disjoint union of cliques. If the A^i's are arrangements, then the graph $G_\mathbb{A}$ is endowed with a natural proper r-colouring, (t,i) being of colour i.

- the graph $H_\mathbb{A}^m$ has vertex set $[r]$ and an edge between i and j if $|A^i \cap A^j| \geq m$.

Notice that $H_\mathbb{A}^1$ is the *contraction* of the graph $G_\mathbb{A}$ by the map $\varphi : (t,i) \mapsto i$ from the vertex set of $G_\mathbb{A}$ to the vertex of $H_\mathbb{A}^1$. Indeed,

$$(i,j) \in E(H_\mathbb{A}^1) \iff \exists v_i \in \varphi^{-1}(i), v_j \in \varphi^{-1}(j) \text{ such that } (v_i, v_j) \in E(G_\mathbb{A}).$$

An example of a graph $G_\mathbb{A}$ and of its 1- and 2-contractions $H_\mathbb{A}^1$ and $H_\mathbb{A}^2$ is drawn on Figure 1. For $m \geq 2$, the definition of $H_\mathbb{A}^m$ is less common. We call it the *m-contraction* of $G_\mathbb{A}$. The 2-contraction is interesting for us. It corresponds exactly to the graph induced by the dependency graph $\mathbb{B}$ on the list of arrangement $\mathbb{A}$, considered in the proof of Theorem 9.1.7.

REMARK 10.2.1. Graphs associated to families of arrangements are a practical way to encode some information and should not be confused with the random graphs or their induced subgraphs. Therefore we used greek letters for the latter and latin letter for graphs $G_\mathbb{A}$ and their contractions. The dependency graph will always be called $\mathbb{B}$ to avoid confusions.

10.2.2. Exploiting symmetries. The dependency graph of our model has much more structure than a general dependency graph. In particular, all variables $\delta_\gamma(A)$ are identically distributed. More generally, the joint distribution of

$$\left(\delta_\gamma(A^1), \ldots, \delta_\gamma(A^r) \right)$$

depends only on $G_\mathbb{A}$. Here, we state a few consequences of this invariance property that will be useful in the next section.
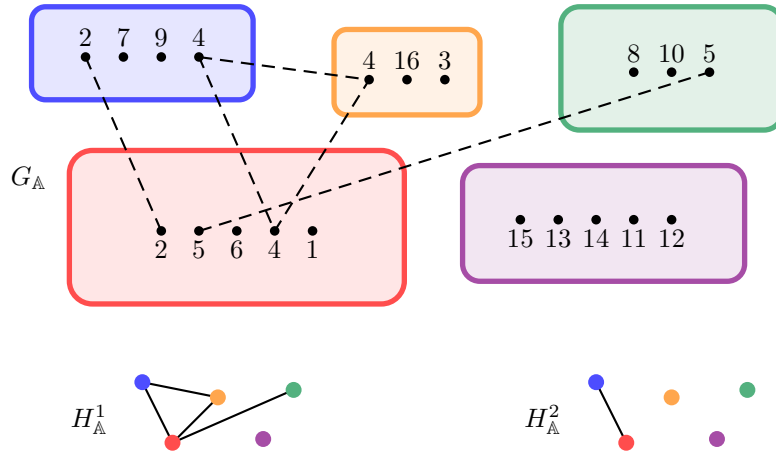
FIGURE 1. The graphs $G_{\mathbb{A}}$, $H_{\mathbb{A}}^1$ and $H_{\mathbb{A}}^2$ corresponding to the family of arrangements $(15, 13, 14, 11, 12)$, $(2, 5, 6, 4, 1)$, $(2, 7, 9, 4)$, $(4, 16, 3)$, and $(8, 10, 5)$.

LEMMA 10.2.2. *Fix a graph γ of size k, the quantity*

$$\mathbb{E}\left[\delta_\gamma(A^1)\cdots\delta_\gamma(A^r)\right]$$

depends only on the graph $G_{\mathbb{A}}$ associated to the family of arrangements $(A^1, \ldots, A^r)$. The same is true for the joint cumulant $\kappa(\delta_\gamma(A^1), \ldots, \delta_\gamma(A^r))$.

PROOF. The first statement follows immediately from the invariance of the model $\Gamma(n, p)$ by relabelling of the vertices. The second is a corollary, using the moment-cumulant relation (43). □

COROLLARY 10.2.3. *Fix some graph γ. Then the joint cumulant $\kappa(X_\gamma^{(n)}, \ldots, X_\gamma^{(n)})$ is a polynomial in n.*

PROOF. Using Lemma 10.2.2, we can rewrite the expansion (45) as

$$\kappa(X_\gamma^{(n)}, \ldots, X_\gamma^{(n)}) = \sum_G \kappa(G) N_G, \tag{51}$$

where:

- the sum runs over graphs G of vertex set $V_{\mathbb{k}}$ that correspond to some arrangements (that is G is a disjoint union of cliques and, for any s, t and i, there is no edge between (s, i) and (t, i));
- $\kappa(G)$ is the common value of $\kappa\left(\delta_\gamma(A^1), \ldots, \delta_\gamma(A^r)\right)$, where $(A^1, \ldots, A^r)$ is any list of arrangements with associated graph G;
- N_G is the number of lists of arrangements with associated graph G.

But it is clear that the sum index is finite and that neither the summation index nor the quantity $\kappa(G)$ depend on n. Besides, the number N_G is simply the falling factorial $n(n-1)\ldots(n-c_G+1)$, where c_G is the number of connected components of G. The corollary follows from these observations. □

10.3. Moderate deviations for subgraph count statistics

10.3.1. End of the proof of Theorem 10.0.1. We would like to apply Proposition 5.2.1 to the sequence $S_n = X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]$ with $\alpha_n = n^2$ and $\beta_n = n^{k-2}$. Let us check that S_n indeed fulfills the hypothesis.

(1) The uniform bound $|\kappa^{(r)}(S_n)| \leq (Cr)^r \alpha_n (\beta_n)^r$, where C does not depend on n, corresponds to Proposition 10.1.2; we may even choose $C = 2k^4$.

(2) We also have to check the speed of convergence:

$$\kappa^{(2)}(S_n) = \sigma^2 \, \alpha_n \, (\beta_n)^2 \, (1 + O(\alpha_n^{-1/2})) \quad ; \quad \kappa^{(3)}(S_n) = L \, \alpha_n \, (\beta_n)^3 \, (1 + O(\alpha_n^{-1/4})). \quad (52)$$

But these estimates follow directly from the bound above for $r = 2, 3$ and the fact that $\kappa^{(r)}(S_n)$ is always a polynomial in n — see Corollary 10.2.3.

Finally, the mod-Gaussian convergence follows from the observations in Chapter 5. The normality zone result follows from Proposition 4.4.1 and we can apply Proposition 5.2.1 to get the moderate deviation statement. This ends the proof of Theorem 10.0.1. $\qquad\square$

REMARK 10.3.1. Using Theorem 9.2.2, we could obtain a bound for joint cumulants of subgraph count statistics. Hence, it would be possible to derive mod-Gaussian convergence and moderate deviation results for *linear combinations* of subgraph count statistics. However, since we do not have a specific motivation for that and since the statement for a single subgraph count statistics is already quite technical, we have chosen not to present such a result.

10.3.2. Computing σ^2 and L. The proof above does not give an explicit value for σ^2 and L. Yet, these values can be obtained by analyzing the graphs G that contribute to the highest degree term of $\kappa^{(2)}$ and $\kappa^{(3)}$.

LEMMA 10.3.2. *Let γ be a graph with k vertices and h edges. Then the positive number σ appearing in Theorem 10.0.1 is given by*

$$\sigma^2 = 2h^2 p^{2h-1}(1 - p).$$

PROOF. By definition, σ^2 is the coefficient of n^{2k-2} in $\kappa^{(2)}(X_\gamma^{(n)})$. As seen in Equation (51), the quantity $\kappa^{(2)}(X_\gamma^{(n)})$ can be written as

$$\sum_G \kappa(G) N_G,$$

where the sum runs over some graphs G with vertex set $V \sqcup V$. However, we have seen that $\kappa(G) = 0$ unless the 2-contraction H^2 of G is connected — see Inequality (48) — and on the other hand, N_G is a polynomial in n, whose degree is the number c_G of connected component of G.

As we are interested in the coefficient of n^{2k-2}, we should consider only graphs G with at least $2k - 2$ connected components and a connected 2-contraction. These graphs are represented on Figure 2.
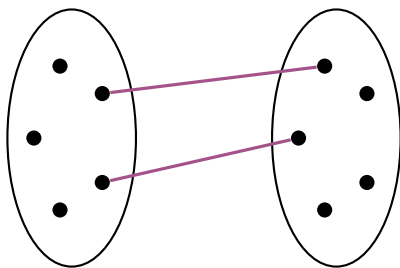
FIGURE 2. Graphs involved in the computation of the main term in $\kappa^{(2)}(X_\gamma^{(n)})$.

Namely, we have to choose a pair of vertices on each side and connect each of these vertices to one vertex of the other pair (there are 2 ways to make this connection, if both pairs are fixed). A quick computation shows that, for such a graph G,

$$\kappa(G) = \begin{cases} p^{2h-1}(1-p) & \text{if both pairs correspond to an edge of } \gamma; \\ 0 & \text{else.} \end{cases}$$

Finally there are $2h^2$ graphs with a non-zero contribution to the coefficient of n^{2k-2} in $\kappa^{(2)}(X_\gamma^{(n)})$. For each of these graphs, $N_G = n(n-1)\cdots(n-2k+3) = n^{2k-2}(1+o(1))$. Therefore, the coefficient of n^{2k-2} in $\kappa^{(2)}(X_\gamma^{(n)})$ is $2h^2 p^{2h-1}(1-p)$, as claimed. $\qquad\square$

The number L can be computed by the same method.

LEMMA 10.3.3. *Let γ be a graph with k vertices and h edges. Then the number L appearing in Theorem 10.0.1 is given by*

$$L = 12h^3(h-1)p^{3h-2}(1-p)^2 + 4h^3 p^{3h-2}(1-p)(1-2p).$$

PROOF. Here, we have to consider graphs G on vertex set $V \sqcup V \sqcup V$ with at least $3k-4$ connected components and with a connected 2-contraction. These graphs are of two kinds, see Figure 3.
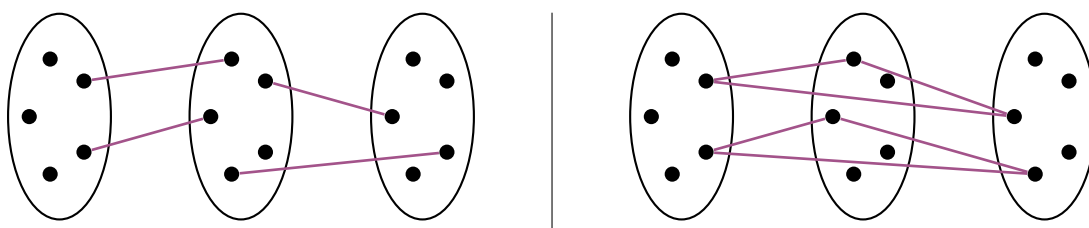
FIGURE 3. Graphs involved in the computation of the main term in $\kappa^{(3)}(X_\gamma^{(n)})$.

In the first case (left-hand side picture), an edge on the left can possibly have an extremity in common with an edge on the right. In this case, one has to add an edge to complete the triangle (indeed, all graphs G are disjoint unions of cliques). But this cannot happen for both edges on the left simultaneously, otherwise the graph belong to the second family.

The following is now easy to check. There are $12h^3(h-1)$ graphs of the first kind with a non-zero cumulant $\kappa(G)$ — 3 choices for which copy of V plays the central role, $h^3(h-1)$ for pairs of vertices and 4 ways to link the chosen vertices — and, for these graphs, the corresponding cumulant is always $\kappa(G) = p^{3h-2}(1-p)^2$. Similarly, there are $4h^3$ graphs of the second kind with a non-zero cumulant $\kappa(G)$. For these graphs, $\kappa(G) = p^{3h-2}(1-p)(1-2p)$. In both cases, $N_G = n^{3k-4}(1+o(1))$. This completes the proof. $\square$

EXAMPLE 10.3.4. Denote T_n the number of triangles in a random Erdös-Rényi graph $\Gamma(n,p)$ (each triangle being counted 6 times). According to the previous Lemmas, the parameters σ^2 and L are respectively

$$\sigma^2 = 18\, p^5\,(1-p) \quad \text{and} \quad L = 108\, p^7\,(1-p)(7-8p).$$

Moreover, $\mathbb{E}[T_n] = n^{\downarrow 3}\, p^3 = n^3\, p^3 - 3\, n^2\, p^3 + O(n)$. So,

$$\mathbb{P}\left[T_n \geq n^3 p^3 + n^2(v - 3p^3)\right] \simeq \sqrt{\frac{9p^5(1-p)}{\pi\, v^2}}\, \exp\left(-\frac{v^2}{36\, p^5(1-p)} + \frac{(7-8p)\, v^3}{324\, n\, p^8(1-p)^2}\right)$$

for $1 \ll v \ll n^{1/2}$.

10.3.3. Case of a non-constant sequence p_n. Proposition 10.1.2 still holds when p_n is a non-constant sequence (the particularly interesting case is $p_n \to 0$). Applying Theorem 9.1.8 instead of Theorem 9.1.7, one gets a slightly sharper bound

$$\left|\kappa^{(r)}(X_\gamma^{(n)})\right| \leq C^r\, r^{r-2}\, n^{r(k-2)+2}\, (p_n)^h.$$

But, unlike in the case $p_n = p$ constant, this bound is not always optimal for a fixed r (even up to a multiplicative constant) . Indeed, [JŁR00, Lemma 6.17] gives stronger bounds than ours (see also Example 6.19 in [JŁR00]). Finding a uniform bound for cumulants, whose dependence in r is of order $(Cr)^r$ (so that we have mod-Gaussian convergence), and which is optimal for fixed r is an open problem.

Yet, we can still give a lower bound on the normality zone. Let α_n and β_n be defined as follows:

$$\alpha_n = n^2\, (p_n)^{4h-3}\,(1-p_n)^3;$$
$$\beta_n = n^{k-2}\, (p_n)^{1-h}\,(1-p_n)^{-1}.$$

With these choices, one has

$$\left|\kappa^{(r)}(X_\gamma^{(n)})\right| \leq (Cr)^r\, \alpha_n(\beta_n)^r.$$

Notice that $\kappa^{(2)}(X_\gamma^{(n)})$ and $\kappa^{(3)}(X_\gamma^{(n)})$ are polynomials in n and p_n, of degree $2h$ and $3h$ in p_n. Thanks to this observation, if $p_n \to 0$, then

$$\kappa^{(2)}(X_\gamma^{(n)}) = 2h^2\, n^{2k-2}\, (p_n)^{2h-1}(1-p_n) + O(n^{2k-3})$$

$$= 2h^2\, \alpha_n\,(\beta_n)^2\left(1 + O\left(\frac{1}{n\,(p_n)^{2h-1}}\right)\right);$$

$$\kappa^{(3)}(X_\gamma^{(n)}) = h^3\, n^{3k-4}\, (p_n)^{3h-2}(1 - p_n)\, (12(h-1)(1 - p_n) + 4(1 - 2p_n)) + O(n^{3k-5})$$

$$= \alpha_n\, (\beta_n)^3\, \left(O((p_n)^{2h-2}) + O\left(\frac{1}{n\,(p_n)^h}\right)\right).$$

From the discussion of Section 5.1, we conclude that we have mod-Gaussian convergence if $(\alpha_n)^{1/3} = o(n\,(p_n)^{2h-1})$ and $n\,(p_n)^h \to +\infty$. The first condition is equivalent to $n\,(p_n)^{2h} \to +\infty$, so, if

$$n^{-\frac{1}{2h}} \ll p_n \ll 1,$$

then Conditions (26) are met, and one has a mod-Gaussian convergence of

$$\frac{X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]}{n^{k-\frac{4}{3}}\,(p_n)^{\frac{h}{3}}}$$

with parameters $t_n = 2h^2\, n^{\frac{2}{3}}\,(p_n)^{\frac{4h}{3}-1}\,(1 - p_n)$ and limit $\psi(z) = 1$. Therefore, in this setting,

$$\frac{X_\gamma^{(n)} - \mathbb{E}[X_\gamma^{(n)}]}{h\, n^{k-1}\,\sqrt{2(p_n)^{2h-1}(1 - p_n)}}$$

has normality zone at least $O(n^{\frac{1}{3}}\,(p_n)^{\frac{2h}{3}-\frac{1}{2}})$. Note that we only have a lower bound on the normality zone, since the limit in the mod-Gaussian convergence is trivial.

REMARK 10.3.5. Unfortunately the convergence speed hypotheses (52) are not satisfied, so one cannot apply Proposition 5.2.1.

CHAPTER 11

Random character values from central measures on partitions

11.1. Background

Our Theorem 9.1.7 can also be used to study certain models of random integer partitions, which encode the lengths of the longest increasing subsequences in random permutations or random words. Consider first a random permutation $\sigma = \sigma_n$ chosen uniformly among the $n!$ permutations of size n. An increasing subsequence of the permutation σ is a subword $\sigma(i_1)\sigma(i_2)\cdots\sigma(i_l)$ of the word $\sigma(1)\sigma(2)\cdots\sigma(n)$, with

$$i_1 < i_2 < \cdots < i_l \qquad ; \qquad \sigma(i_1) \leq \sigma(i_2) \leq \cdots \leq \sigma(i_l).$$

Denote $\ell_n = \ell(\sigma_n)$ the longest length of an increasing subsequence of σ_n; it is a random variable with values in $[\![1, n]\!]$. In the 60's, numerical experiments ([Ula61]) lead Ulam to conjecture the existence of a limit in probability

$$\lim_{n \to \infty} \frac{\ell_n}{\sqrt{n}} = c > 0$$

and this was proved later by Hammersley, see [Ham72]. The value of c was then computed by Logan and Shepp and separately by Kerov and Vershik to be equal to 2, *cf.* [LS77, KV77] for the inequality $c \geq 2$, and [KV86] for the converse inequality. More recently, the fluctuations $n^{-1/6}(\ell_n - 2\sqrt{n})$ of the length of the longest increasing subsequence were shown to converge to the Tracy-Widom distribution, see [BDJ99, BOO00, Oko00, Joh01].

The proof of the Logan-Shepp-Kerov-Vershik law of large numbers relies on the use of the Robinson-Schensted-Knuth algorithm ([Rob38, Sch61]), which associates to a permutation σ or more generally to a word w of length n two numberings (tableaux) of the cells of a Young diagram λ of size n. The first row of this Young diagram has length $\lambda_1 = \ell(w)$, and more generally, for any $k \geq 1$,

$\lambda_1 + \lambda_2 + \cdots + \lambda_k$

$= \max\{\text{sum of the lengths of } k \text{ disjoint weakly increasing subwords of } w\}.$

One can then study the random Young diagram λ under the law that is the image of the uniform probability measure on $\mathfrak{S}(n)$, or more generally under the image by RSK of any probability measure on words of length n. Denote $\mathfrak{P}(n)$ the set of Young diagrams of size n. If $\sigma = \sigma_n$ is a uniform random permutation of size n and λ is its image by RSK, then the probability measure on $\mathfrak{P}(n)$ that one obtains is the so-called *Plancherel measure*. It can be shown that it is given by

$$\mathbb{P}[\lambda] = \frac{(\dim S^\lambda)^2}{n!},$$

where S^λ is the Specht module of label λ, that is one of the irreducible complex linear representations of $\mathfrak{S}(n)$ (see [Sag01]). This algebraic point of view is particular useful, and its generalisation with spectral measures of traces of symmetric groups leads to

the study of certain random variables, which shall be shown in this chapter to be mod-Gaussian convergent.

11.2. Definitions and results

Recall that if G is a finite group and if τ is a function $G \to \mathbb{C}$ with $\tau(e_G) = 1$ and $\tau(gh) = \tau(hg)$ (such a function is called a *trace* on G), then τ can be expanded uniquely as a linear combination of normalised irreducible characters:

$$\tau = \sum_{\lambda \in \widehat{G}} \mathbb{P}_\tau[\lambda] \, \widehat{\chi}^\lambda,$$

where $\widehat{G}$ is the (finite) set of isomorphism classes of irreducible representations of G and $\widehat{\chi}^\lambda$ the normalised (*i.e.* divided by the dimension of the space) character of the irreducible representation associated to λ.

DEFINITION 11.2.1. *The map $\lambda \mapsto \mathbb{P}_\tau[\lambda]$ is called the spectral measure of the trace τ. It takes non-negative values if and only if, for every family $(g_1, \ldots, g_n)$ of elements of G, the matrix $(\tau(g_i(g_j)^{-1}))_{1 \leq i,j \leq n}$ is Hermitian non-negative definite. Then, the spectral measure is a probability measure on $\widehat{G}$.*

When $G = \mathfrak{S}(n)$ is the symmetric group of order n, the irreducible representations of G are indexed by the integer partitions of size n, so a spectral measure associated to a normalised trace on $\mathfrak{S}(n)$ is a (complex) measure on $\mathfrak{P}(n)$. In the sequel, we shall also consider the infinite symmetric group $\mathfrak{S}(\infty) = \bigcup_{n \geq 1} \mathfrak{S}(n)$, which is the group of permutations of the set of natural numbers that move a finite number of elements.

DEFINITION 11.2.2. *A central measure on partitions is a family $(\mathbb{P}_{\tau,n})_{n \in \mathbb{N}}$ of spectral measures on the sets $\mathfrak{P}(n)$ that come from the same trace of the infinite symmetric group $\mathfrak{S}(\infty)$. In other words $(\mathbb{P}_{\tau,n})_{n \in \mathbb{N}}$ is a central measure if there exists a trace $\tau : \mathfrak{S}(\infty) \to \mathbb{C}$ such that*

$$\tau_{|\mathfrak{S}(n)} = \sum_{\lambda \in \mathfrak{P}(n)} \mathbb{P}_{\tau,n}[\lambda] \, \widehat{\chi}^\lambda.$$

This definition of central measures comes from [KV81, Section 3], and it can be generalised to the setting of approximately finite algebras, see [SV75] and [Ker03].

EXAMPLE 11.2.3. Consider the regular trace $\tau(\sigma) = \mathbb{1}_{\sigma=\mathrm{id}}$ of the infinite symmetric group $\mathfrak{S}(\infty)$. The corresponding spectral measures are the Plancherel measures given by the formula $\mathbb{P}_n[\lambda] = (\dim S^\lambda)^2/n!$, see *e.g.* [Oko00, IO02]. Thus, the previous definition generalises Ulam's problem and the study of Plancherel measures of symmetric groups.

A central measure $(\mathbb{P}_{\tau,n})_{n \in \mathbb{N}}$ is non-negative if and only if $(\tau(\rho_i \rho_j^{-1}))_{1 \leq i,j \leq n}$ is Hermitian non-negative definite for any finite family of permutations $\rho_1, \ldots, \rho_n$. The set of non-negative central measures, *i.e.*, coherent systems of probability measures on partitions has been identified in [Tho64] and later studied in [KV81]; see also [Ols90, Oko99] for more general results on representations of $\mathfrak{S}(\infty)$, and [Buf12, Mél12] for a probabilistic study of the central measures. Call extremal a non-negative trace on $\mathfrak{S}(\infty)$ that

is not a positive linear combination of other non-negative traces. Then, extremal central measures are indexed by the infinite-dimensional *Thoma simplex*

$$\Omega = \left\{ \omega = (\alpha, \beta) = ((\alpha_1 \geq \alpha_2 \geq \cdots \geq 0), (\beta_1 \geq \beta_2 \geq \cdots \geq 0)) \mid \sum_{i=1}^{\infty} \alpha_i + \beta_i \leq 1 \right\}.$$

The trace on the infinite symmetric group corresponding to a parameter ω is given by

$$\tau_\omega(\sigma) = \prod_{c \in C(\sigma)} p_{|c|}(\omega) \quad \text{with } p_1(\omega) = 1, \ p_{k \geq 2}(\omega) = \sum_{i=1}^{\infty} (\alpha_i)^k + (-1)^{k-1}(\beta_i)^k, \quad (53)$$

$C(\sigma)$ denoting the set of cycles of σ.

We can finally introduce the random variables that we are interested in. In [KV81], Kerov and Vershik have shown that if $\omega \in \Omega$ and $\rho \in \mathfrak{S}(\infty)$ are fixed, then the random character value $\widehat{\chi}^\lambda(\rho)$ with λ chosen according to the central measure $\mathbb{P}_{\omega,n}$ converges in probability towards the trace $\tau_\omega(\rho)$. Informally, central measures on partitions and extremal traces of $\mathfrak{S}(\infty)$ are *concentrated*. More recently, it was shown by Féray and Méliot that this concentration is Gaussian, see [FM12, Mél12]. The aim of this chapter is to use the techniques of Sections 9-10 in order to prove the following:

THEOREM 11.2.4. *Fix a parameter $\omega \in \Omega$, and denote ρ a k-cycle in $\mathfrak{S}(\infty)$ (for $k \geq 2$). Assume $p_{2k-1}(\omega) - (p_k(\omega))^2 \neq 0$. Denote then $X_\rho^{(n)}$ the random character value $\widehat{\chi}^\lambda(\rho)$, where:*

- *$\lambda \in \mathfrak{P}(n)$ is picked randomly according to the central measure $\mathbb{P}_{\omega,n}$,*
- *and $\widehat{\chi}^\lambda$ denotes the normalised irreducible character indexed by λ of $\mathfrak{S}(n)$, that is*

$$\widehat{\chi}^\lambda(\rho) = \frac{\operatorname{tr} \Pi^\lambda(\rho)}{\dim V^\lambda} \qquad \text{with } (V^\lambda, \Pi^\lambda) \text{ irreducible representation of } \mathfrak{S}(n).$$

The rescaled random variable $n^{2/3} \left[X_{(k)}^{(n)} - p_k(\omega) \right]$ converges in the mod-Gaussian sense with parameters $t_n = n^{1/3} \sigma^2$ and limiting function $\psi(z) = \exp(L \frac{z^3}{6})$, where

$$\sigma^2 = k^2 \left(p_{2k-1}(\omega) - p_k(\omega)^2 \right);$$
$$L = k^3 \left((3k-2) \, p_{3k-2}(\omega) - (6k-3) \, p_{2k-1}(\omega) \, p_k(\omega) + (3k-1) \, p_k(\omega)^3 \right).$$

As in Chapter 10, the mod-Gaussian convergence is proved by bounds on cumulants of type (25). Furthermore, the hypothesis (27) on the third and second cumulant are ensured by the polynomiality of joint cumulants — Lemma 11.3.3. Therefore, we can apply Propositions 5.2.1 and 4.4.1.

COROLLARY 11.2.5. *Let $X_k^{(n)} = X_{(k)}^{(n)}$ be the random character value on a k-cycle as defined above. Then $n^{1/2} (X_k^{(n)} - p_k(\omega))$ satisfies a central limit theorem with normality zone $o(n^{1/6})$. Moreover, at the edge of this normality zone and at a slightly larger scale, the deviation*

probabilities are given by:

$$\mathbb{P}\left[n^{1/2}(X_k^{(n)} - p_k(\omega)) \geq n^{1/6}\sigma^2 x_n\right] = \frac{e^{-\frac{(x_n)^2 n^{1/3}\sigma^2}{2}}}{x_n \, n^{1/6} \, \sigma\sqrt{2\pi}} \, e^{\frac{L(x_n)^3}{6}} \, (1 + o(1));$$

$$\mathbb{P}\left[n^{1/2}(X_k^{(n)} - p_k(\omega)) \leq -n^{1/6}\sigma^2 x_n\right] = \frac{e^{-\frac{(x_n)^2 n^{1/3}\sigma^2}{2}}}{x_n \, n^{1/6} \, \sigma\sqrt{2\pi}} \, e^{\frac{-L(x_n)^3}{6}} \, (1 + o(1)).$$

for any sequence $(x_n)_{n \in \mathbb{N}}$ of positive numbers bounded away from 0 with $x_n = o(n^{1/12})$, and where σ^2 and L as in Theorem 11.2.4.

REMARK 11.2.6. The condition $\sigma^2 > 0$ is satisfied as soon as the sequence $\alpha \sqcup \beta$ associated to the Thoma parameter ω contains two different non-zero coordinates. Indeed, for any summable non-increasing non-negative sequence $\gamma = (\gamma_1, \gamma_2, \dots)$ with $\gamma_i > \gamma_{i+1}$ for some i, one has

$$\left(\sum_{i=1}^{\infty}(\gamma_i)^{k-\varepsilon}\right)\left(\sum_{i=1}^{\infty}(\gamma_i)^{k+\varepsilon}\right) \geq \left(\sum_{i=1}^{\infty}(\gamma_i)^k\right)^2$$

with equality if and only if $\varepsilon = 0$. Indeed, the derivative of the function of ε on the left-hand side is

$$\sum_{i<j}(\gamma_i\gamma_j)^k \log\left(\frac{\gamma_i}{\gamma_j}\right)\left(\left(\frac{\gamma_i}{\gamma_j}\right)^{\varepsilon} - \left(\frac{\gamma_j}{\gamma_i}\right)^{\varepsilon}\right).$$

Applying the result to $\gamma = \alpha \sqcup \beta$ and $\varepsilon = k - 1$, we obtain on the left-hand side

$$\left(\sum_{i=1}^{\infty}\gamma_i\right)\left(\sum_{i=1}^{\infty}(\gamma_i)^{2k-1}\right) = \left(\sum_{i=1}^{\infty}\alpha_i + \beta_i\right) p_{2k-1}(\omega) \leq p_{2k-1}(\omega)$$

and on the right-hand side

$$\left(\sum_{i=1}^{\infty}(\gamma_i)^k\right)^2 \geq \left(\sum_{i=1}^{\infty}(\alpha_i)^k + (-1)^{k-1}(\beta_i)^k\right)^2 = (p_k(\omega))^2.$$

This proof shows that the condition $\sigma^2 > 0$ is also satisfied if $0 < \sum_{i=1}^{\infty}\alpha_i + \beta_i < 1$.

This chapter is organised as follows. In Section 11.3, we present the necessary material. In Section 11.4, we then prove bounds on these cumulants similar to those of Proposition 10.1.2, and we compute the limits of the second and third cumulants. This will allow us to use in Section 11.5 the framework of Section 5.2 in order to prove the results stated above. We shall also detail some consequences of these results for the shapes of the random partitions $\lambda \sim \mathbb{P}_{n,\omega}$ viewed as Young diagrams, in the spirit of [FM12, Mél12].

REMARK 11.2.7. By using the Hopf algebra of free quasi-symmetric functions, one can show that the central measures $\mathbb{P}_{\omega,n}$ associated to parameters ω of the Thoma simplex are the images by the Robinson-Schensted-Knuth algorithm of certain *non-uniform* probability measures on permutations, that are constructed by random shuffles of cards; see [DHT02] and [Sta01, Ful04]. Therefore, the study of central measures is indeed a generalisation of Ulam's problem of the longest increasing subsequence in a random *uniform* permutation.

As this combinatorial interpretation of the central measures is not at all required to understand the results of this chapter, we choose to skip it; and everything hereafter can be followed starting only from the algebraic definition 11.2.2.

11.3. Technical preliminaries

11.3.1. Non-commutative probability theory. The originality of this last chapter is that, although the problem is formulated in a classical probability space, it is natural to work in the setting of non-commutative probability theory.

DEFINITION 11.3.1. *A non-commutative probability space is a complex unital algebra $\mathscr{A}$ endowed with an adjunction $\star$, and with some linear functional $\varphi : \mathscr{A} \to \mathbb{C}$ such that $\varphi(1) = 1$ and, for any $a \in \mathscr{A}$, $\varphi(a^\star a) \geq 0$.*

This generalises the notion of probability space: elements of $\mathscr{A}$ should be thought as random variables, while φ should be thought as the expectation. The difference is that, unlike random variables in usual probability theory, elements of $\mathscr{A}$ are not assumed to commute. Usually one also assumes that φ is tracial, *i.e.,* $\varphi(ab) = \varphi(ba)$ for any $a, b \in \mathscr{A}$.

There are five natural analogues of the notion of independence in non-commutative probability theory, see [Mur03]. In our context, the relevant one is the following one, sometimes called *tensor independence* [HS10, Definition 1.1].

DEFINITION 11.3.2. *Two subalgebras $\mathscr{A}_1$ and $\mathscr{A}_2$ of $\mathscr{A}$ are* tensor independent *if and only if, for any sequence $a_1, \ldots, a_r$ with, for each i, $a_i \in \mathscr{A}_1$ or $a_i \in \mathscr{A}_2$, one has*

$$\varphi(a_1 \ldots a_r) = \varphi\left(\overrightarrow{\prod_{\substack{1 \leq i \leq r \\ a_i \in \mathscr{A}_1}}} a_i \right) \varphi\left(\overrightarrow{\prod_{\substack{1 \leq i \leq r \\ a_i \in \mathscr{A}_2}}} a_i \right).$$

The arrow on the product sign means that the a_i in the product appear in the same order as in $a_1, \ldots, a_r$.

With this definition of independence, the notion of dependency graph presented in Section 9.1 is immediately extended to the non-commutative framework. One can also define joint cumulants as follows: if $a_1, \ldots, a_r$ are elements in $\mathscr{A}$, we set

$$\kappa(a_1, \ldots, a_r) = \sum_{\pi} \mu(\pi) \prod_{B \in \pi} \varphi\left(\overrightarrow{\prod_{i \in B}} a_i \right).$$

Note that, in the proof of Theorem 9.1.7, independence is only used in equation (46). By definition of tensor independence, equation (46) also holds in the non-commutative setting and hence, so does Theorem 9.1.7.

11.3.2. Two probability spaces. From now on, we fix an element ω in the Thomas simplex Ω. Denote $\mathbb{C}\mathfrak{S}(n)$ the group algebra of $\mathfrak{S}(n)$. The function τ_ω, defined by Equation (53), can be linearly extended to $\mathbb{C}\mathfrak{S}(n)$. Then $(\mathbb{C}\mathfrak{S}(n), \tau_\omega)$ is a non-commutative probability space.

Note that we are now working with two probability spaces: the non-commutative probability space $(\mathbb{C}\mathfrak{S}(n), \tau_\omega)$ and the usual probability space that we want to study (the set of Young diagrams of size n with the probability measure $\mathbb{P}_{\omega,n}$). They are related as follows. Consider an element y in $\mathbb{C}\mathfrak{S}(n)$ and define the random variable (for the usual probability space) X_y by $X_y(\lambda) = \widehat{\chi}(y)$. Then one has

$$\mathbb{E}_{\mathbb{P}_{\omega,n}}(X_y) = \sum_{\lambda \vdash n} \mathbb{P}_{\omega,n}(\lambda)\widehat{\chi}(y) = \tau_\omega(y).$$

In other words, the *expectations* of y (in the non-commutative probability space) and of X_y (in the usual probability space) coincide (recall that the trace of a non-commutative probability space, here τ_ω, is considered as an expectation).

Besides, if we restrict to the *center* of $(\mathbb{C}\mathfrak{S}(n), \tau_\omega)$, then the map $y \mapsto X_y$ is an *algebra morphism*, called *discrete Fourier transform*. As a consequence, if $y_1, \ldots, y_r$ lie in the center of $(\mathbb{C}\mathfrak{S}(n), \tau_\omega)$, their joint moments (and joint cumulants) are the same as those of $X_{y_1}, \ldots, X_{y_r}$.

11.3.3. Renormalised conjugacy classes and polynomiality of cumulants. Given a partition $\mu = (\mu_1, \ldots, \mu_\ell)$ of size $|\mu| = \sum_{i=1}^\ell \mu_i = k$, we denote

$$\Sigma_{\mu,n} = \sum_A \rho_\mu(A) \text{ where } \rho_\mu(A) = (a_1, \ldots, a_{\mu_1})(a_{\mu_1+1}, \ldots, a_{\mu_1+\mu_2}) \cdots .$$

In the equation above, the formal sum is taken over arrangements in $\mathfrak{A}(n,k)$ and is considered as an element of the group algebra $\mathbb{C}\mathfrak{S}(n)$. It clearly lies in its center. Besides, $\Sigma_{\mu,n}$ is the sum of $n^{\downarrow k}$ elements of cycle-type μ, hence, if ρ is a fixed permutation of type μ, one has:

$$X_{\Sigma_{\mu,n}} = n^{\downarrow k} X_\rho.$$

Note that considering these elements $\Sigma_{\mu,n}$ and their normalised characters is a classical trick in the study of central measures on Young diagrams; see [IO02, Śni06b, FM12].

Fix some permutations $\rho_1, \ldots, \rho_r$ of respective size $k_1, \ldots, k_r$. Denote $\mu^1, \ldots, \mu^r$ their cycle-types.

$$n^{\downarrow k_1} \cdots n^{\downarrow k_r} \, \kappa\left(X_{\rho_{\mu^1}}^{(n)}, \ldots, X_{\rho_{\mu^r}}^{(n)}\right) = \kappa(\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r})$$

$$= \sum_{\substack{A^1 \in \mathfrak{A}(n,k_1) \\ \vdots \\ A^r \in \mathfrak{A}(n,k_r)}} \kappa\left(\rho_{\mu^1}(A^1), \ldots, \rho_{\mu^r}(A^r)\right). \tag{54}$$

As in the framework of subgraph count statistics, the invariance of τ_ω by conjugacy of its argument implies that the joint cumulant $\kappa\left(\rho_{\mu^1}(A^1), \ldots, \rho_{\mu^r}(A^r)\right)$ depends only on the graph $G_\mathbb{A}$ associated to the family of arrangement $\mathbb{A} = (A^1, \ldots, A^r)$. Copying the proof of Corollary 10.2.3, we get:

LEMMA 11.3.3. *Fix some integer partitions $\mu^1, \ldots, \mu^r$. Then the rescaled joint cumulant*

$$\kappa(\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r}) = n^{\downarrow k_1} \cdots n^{\downarrow k_r} \, \kappa\left(X^{(n)}_{\rho_{\mu^1}}, \ldots, X^{(n)}_{\rho_{\mu^r}}\right)$$

is a polynomial in n.

11.4. Bounds and limits of the cumulants

11.4.1. The dependency graph for random character values. If $k = |\mu|$, we are interested in the non-commutative random variables

$$\Sigma_{\mu,n} = \sum_{A \in \mathfrak{A}(n,k)} \rho_\mu(A) \in \mathbb{C}\mathfrak{S}(n).$$

Again, to control the cumulants, we shall exhibit a dependency graph for the families of random variables $\{\rho_\mu(A)\}_{A \in \mathfrak{A}(n,k)}$.

Due to the multiplicative form of Equation (53), if I and J have disjoint subsets of $[n]$, the subalgebras $\mathbb{C}\mathfrak{S}(I)$ and $\mathbb{C}\mathfrak{S}(J)$ are tensor independent (here, $\mathfrak{S}(I)$ denotes the group of permutations of I, canonically included in $\mathfrak{S}(n)$). Therefore, one can associate to $\{\rho_\mu(A)\}_{A \in \mathfrak{A}(n,k)}$ the dependency graph $\mathbb{B}$ defined by:

- its vertex set is $\mathfrak{A}(n,k)$;
- there is an edge between A and A' if $|A \cap A'| \geq 1$.

The graph $\mathbb{B}$ is obviously regular with degree strictly smaller than $k^2 \, n^{\downarrow k-1}$. On the other hand, all joint moments of the family $(\rho_\mu(A))_{A \in \mathfrak{A}(n,k)}$ are normalised characters of single permutations and hence bounded by 1 in absolute value. So one can once again apply Theorem 9.1.7 and we get:

PROPOSITION 11.4.1. *Fix a partition μ of size k. For any $r \leq 1$, one has*

$$\left|\kappa^{(r)}(\Sigma_\mu)\right| \leq 2^{r-1} r^{r-2} \, n^{\downarrow k} \, (k^2 n^{\downarrow k-1})^{r-1};$$

$$\left|\kappa^{(r)}(X^{(n)}_{\rho_\mu})\right| \leq r^{r-2} \left(\frac{2k^2}{n}\right)^{r-1}.$$

REMARK 11.4.2. Using Theorem 9.2.2, one can also obtain a bound for joint cumulants of the $X^{(n)}_{\rho_\mu}$, namely,

$$\left|\kappa\left(X^{(n)}_{\rho_{\mu^1}}, \ldots, X^{(n)}_{\rho_{\mu^r}}\right)\right| \leq k_1 \cdots k_r \left(r \cdot \max_{1 \leq i \leq r} k_i\right)^{r-2} \left(\frac{2}{n}\right)^{r-1}$$

for integer partitions $\mu^1, \ldots, \mu^r$ of sizes $k_1, \ldots, k_r$. In the following, we shall focus on the case of simple random variables $X^{(n)}_{\rho_\mu}$, though most results also hold in the multi-dimensional setting. Actually, in order to compute the asymptotics of the first cumulants of X_{ρ_μ}, it will be a bit clearer to manipulate joint cumulants of variables $\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r}$ with arbitrary integer partitions.

11.4.2. Limits of the second and third cumulants. Because of Lemma 11.3.3 and Proposition 11.4.1, for any fixed integer partitions,

$$\kappa\left(X^{(n)}_{\rho_{\mu^1}},\ldots,X^{(n)}_{\rho_{\mu^r}}\right) n^{1-r} \simeq \kappa\left(\Sigma_{\mu^1},\ldots,\Sigma_{\mu^r}\right) n^{k_1+\cdots+k_r-(r-1)}$$

converges to a constant. Let us compute this limit when $r = 2$ or 3; we use the same reasoning as in Section 10.3.2.

As κ is invariant by simultaneous conjugacy of its arguments, the summand in Equation (54) depends only on the graph $G = G_{\mathbb{A}}$ associated to the collection $\mathbb{A} = (A^1,\ldots,A^r)$, and we shall denote it $\kappa(G)$. We fix partitions $\mu^1,\ldots,\mu^r$ of respective sizes $k_1,\ldots,k_r$, and write

$$\kappa(\Sigma_{\mu^1},\ldots,\Sigma_{\mu^r}) = \sum_G \kappa(G)\, N_G.$$

Here, as in Chapter 10, N_G denotes the number of list $\mathbb{A}$ of arrangements with associated graph G.

When $r = 2$, we have to look for graphs G on vertex set $V_{\mathbb{k}} = [k_1] \sqcup [k_2]$ with 1-contraction connected and at least $k_1 + k_2 - 1$ connected components, because these are the ones that will give a contribution for the coefficient of $n^{k_1+k_2-1}$. For $i \in [\ell(\mu^1)]$ and $j \in [\ell(\mu^2)]$, denote

$$(\mu^1 \bowtie \mu^2)(i,j) = (\mu^1 \setminus \mu^1_i) \sqcup (\mu^2 \setminus \mu^2_j) \sqcup \{\mu^1_i + \mu^2_j - 1\}.$$

This is the cycle type of a permutation $\rho_{\mu^1}(A^1)\,\rho_{\mu^2}(A^2)$, where $G_{\mathbb{A}}$ is the graph with one edge joining an element of A^1 in the cycle of length μ^1_i with an element of A^2 in the cycle of length μ^2_j. These graphs are the only ones involved in our computation, and they yield

$$\kappa(G) = p_{(\mu^1 \bowtie \mu^2)(i,j)}(\omega) - p_{\mu^1 \sqcup \mu^2}(\omega),$$

where for a partition μ we denote $p_\mu(\omega)$ the product $\prod_{i=1}^{\ell(\mu)} p_{\mu_i}(\omega)$. So,

PROPOSITION 11.4.3. *For any partitions μ and ν, the limit of $n\,\kappa(X^{(n)}_{\rho_\mu}, X^{(n)}_{\rho_\nu})$ is*

$$\sum_{i=1}^{\ell(\mu)}\sum_{j=1}^{\ell(\nu)} \mu_i\,\nu_j\left(p_{(\mu \bowtie \nu)(i,j)}(\omega) - p_{\mu \sqcup \nu}(\omega)\right).$$

In particular, for cycles $\mu = (k)$ and $\nu = (l)$,

$$\lim_{n\to\infty} n\,\kappa\left(X^{(n)}_k, X^{(n)}_l\right) = kl\left(p_{k+l-1}(\omega) - p_{k,l}(\omega)\right).$$

On the other hand, if $\mu = \nu$, then the limit of $n\,\kappa^{(2)}(X^{(n)}_{\rho_\mu})$ is

$$(p_\mu(\omega))^2 \sum_{1\le i,j\le \ell(\mu)} \mu_i\,\mu_j\left(\frac{p_{(\mu_i\mu_j-1)}(\omega)}{p_{\mu_i}(\omega)\,p_{\mu_j}(\omega)} - 1\right).$$

When $r = 3$, we look for graphs G on vertex set $V_{\mathbb{k}} = [k_1] \sqcup [k_2] \sqcup [k_3]$ with 1-contraction connected and at least $k_1 + k_2 + k_3 - 2$ connected components. They are of three kinds:

(1) One cycle in $\rho_{\mu^2}(A^2)$ is connected to two cycles in $\rho_{\mu^1}(A^1)$ and $\rho_{\mu^3}(A^3)$, but not by the same point in this cycle of $\rho_{\mu^2}(A^2)$. This gives for the product $\rho_{\mu^1}(A^1)\,\rho_{\mu^2}(A^2)\,\rho_{\mu^3}(A^3)$ a permutation of cycle type

$$(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k) = (\mu^1 \setminus \mu_i^1) \sqcup (\mu^2 \setminus \mu_j^2) \sqcup (\mu^3 \setminus \mu_k^3) \sqcup \{\mu_i^1 + \mu_j^2 + \mu_k^3 - 2\};$$

and the corresponding cumulant is

$$\kappa(G) = p_{\mu^1 \sqcup \mu^2 \sqcup \mu^3}(\omega) + p_{(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k)}(\omega)$$
$$- p_{((\mu^1 \bowtie \mu^2)(i,j)) \sqcup \mu^3}(\omega) - p_{((\mu^2 \bowtie \mu^3)(j,k)) \sqcup \mu^1}(\omega).$$

In this description, one can permute cyclically the indices $1, 2, 3$, and this gives 3 different graphs.

(2) One cycle in $\rho_{\mu^2}(A^2)$ is connected to two cycles in $\rho_{\mu^1}(A^1)$ and $\rho_{\mu^3}(A^3)$, and by the same point in this cycle of $\rho_{\mu^2}(A^2)$. In other words, there is an identity $a_s^1 = a_t^2 = a_u^3$. This gives again for $\rho_{\mu^1}(A^1)\,\rho_{\mu^2}(A^2)\,\rho_{\mu^3}(A^3)$ a permutation of cycle type $(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k)$, but the corresponding cumulant takes now the form

$$\kappa(G) = 2\,p_{\mu^1 \sqcup \mu^2 \sqcup \mu^3}(\omega) + p_{(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j,k)}(\omega) - p_{((\mu^1 \bowtie \mu^2)(i,j)) \sqcup \mu^3}(\omega)$$
$$- p_{((\mu^2 \bowtie \mu^3)(j,k)) \sqcup \mu^1}(\omega) - p_{((\mu^1 \bowtie \mu^3)(i,k)) \sqcup \mu^2}(\omega).$$

Here, there is no need to permute cyclically the indices in the enumeration for N_G.

(3) Two distinct cycles in $\rho_{\mu^2}(A^2)$ are connected to a cycle of $\rho_{\mu^1}(A^1)$ and to a cycle of $\rho_{\mu^3}(A^3)$, which gives a permutation of cycle type

$$(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j;k,l)$$
$$= (\mu^1 \setminus \mu_i^1) \sqcup (\mu^2 \setminus \{\mu_j^2, \mu_k^2\}) \sqcup (\mu^3 \setminus \mu_l^3) \sqcup \{\mu_i^1 + \mu_j^2 - 1, \mu_k^2 + \mu_l^3 - 1\}.$$

The cumulant corresponding to this last case is

$$\kappa(G) = p_{\mu^1 \sqcup \mu^2 \sqcup \mu^3}(\omega) + p_{(\mu^1 \bowtie \mu^2 \bowtie \mu^3)(i,j;k,l)}(\omega)$$
$$- p_{((\mu^1 \bowtie \mu^2)(i,j)) \sqcup \mu^3}(\omega) - p_{((\mu^2 \bowtie \mu^3)(k,l)) \sqcup \mu^1}(\omega),$$

and again one can permute cyclically the indices $1, 2, 3$ to get 3 different graphs.

Consequently:

PROPOSITION 11.4.4. *For any partitions μ, ν and δ, the limit of $n^2\,\kappa(X_{\rho_\mu}^{(n)}, X_{\rho_\nu}^{(n)}, X_{\rho_\delta}^{(n)})$ is*

$$\sum_{\mathbb{Z}/3\mathbb{Z}} \left(\sum_{i=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \sum_{k=1}^{\ell(\delta)} \mu_i\, \nu_j\, (\nu_j - 1)\, \delta_k \left(\begin{matrix} p_{\mu \sqcup \nu \sqcup \delta}(\omega) + p_{(\mu \bowtie \nu \bowtie \delta)(i,j,k)}(\omega) \\ -p_{((\mu \bowtie \nu)(i,j)) \sqcup \delta}(\omega) - p_{((\nu \bowtie \delta)(j,k)) \sqcup \mu}(\omega) \end{matrix} \right) \right.$$

$$+ \sum_{i=1}^{\ell(\mu)} \sum_{(j \neq k)=1}^{\ell(\nu)} \sum_{l=1}^{\ell(\delta)} \mu_i\, \nu_j\, \nu_k\, \delta_l \left(\begin{matrix} p_{\mu \sqcup \nu \sqcup \delta}(\omega) + p_{(\mu \bowtie \nu \bowtie \delta)(i,j;k,l)}(\omega) \\ -p_{((\mu \bowtie \nu)(i,j)) \sqcup \delta}(\omega) - p_{((\nu \bowtie \delta)(k,l)) \sqcup \mu}(\omega) \end{matrix} \right) \left. \vphantom{\sum} \right)$$

$$+ \sum_{i=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \sum_{k=1}^{\ell(\delta)} \mu_i\, \nu_j\, \delta_k \left(\begin{matrix} 2\,p_{\mu \sqcup \nu \sqcup \delta}(\omega) + p_{(\mu \bowtie \nu \bowtie \delta)(i,j,k)}(\omega) - p_{((\mu \bowtie \nu)(i,j)) \sqcup \delta}(\omega) \\ -p_{((\nu \bowtie \delta)(j,k)) \sqcup \mu}(\omega) - p_{((\mu \bowtie \delta)(i,k)) \sqcup \nu}(\omega) \end{matrix} \right),$$

where $\sum_{\mathbb{Z}/3\mathbb{Z}}$ means that one permutes cyclically the partitions μ, ν, δ.

In particular, for cycles $\mu = (k)$, $\nu = (l)$ and $\delta = (m)$, $\lim_{n\to\infty} n^2 \kappa(X_k^{(n)}, X_l^{(n)}, X_m^{(n)})$ is equal to

$$klm\big((k+l+m-1)\, p_{k,l,m}(\omega) + (k+l+m-2)\, p_{k+l+m-2}(\omega)$$
$$- (k+l-1)\, p_{k+l-1,m}(\omega) - (l+m-1)\, p_{l+m-1,k}(\omega) - (l+m-1)\, p_{k+m-1,l}(\omega)\big).$$

One recovers for $k = l = m$ the values of σ^2 and L announced in Theorem 11.2.4. On the other hand, if $\mu = \nu = \delta$, then the limit of $\dfrac{n^2\, \kappa^{(3)}(X_{\rho\mu}^{(n)})}{(p_\mu(\omega))^3}$ is

$$\sum_{1\leq i,j,k\leq\ell(\mu)} 3\,\mu_i\,\mu_j^2\,\mu_k \left(1 + \frac{p_{\mu_i+\mu_j+\mu_k-2}(\omega)}{p_{\mu_i}(\omega)\,p_{\mu_j}(\omega)\,p_{\mu_k}(\omega)} - \frac{p_{\mu_i+\mu_j-1}(\omega)}{p_{\mu_i}(\omega)\,p_{\mu_j}(\omega)} - \frac{p_{\mu_j+\mu_k-1}(\omega)}{p_{\mu_j}(\omega)\,p_{\mu_k}(\omega)}\right)$$

$$+ \sum_{1\leq i,j\neq k,l\leq\ell(\mu)} 3\,\mu_i\,\mu_j\,\mu_k\,\mu_l \left(1 - \frac{p_{\mu_i+\mu_j-1}(\omega)}{p_{\mu_i}(\omega)\,p_{\mu_j}(\omega)}\right)\left(1 - \frac{p_{\mu_k+\mu_l-1}(\omega)}{p_{\mu_k}(\omega)\,p_{\mu_l}(\omega)}\right)$$

$$+ \sum_{1\leq i,j,k\leq\ell(\mu)} \mu_i\,\mu_j\,\mu_k \left(3\,\frac{p_{\mu_i+\mu_j-1}(\omega)}{p_{\mu_i}(\omega)\,p_{\mu_j}(\omega)} - 2\,\frac{p_{\mu_i+\mu_j+\mu_k-2}(\omega)}{p_{\mu_i}(\omega)\,p_{\mu_j}(\omega)\,p_{\mu_k}(\omega)} - 1\right).$$

11.5. Asymptotics of the random character values and partitions

Fix an integer $k \geq 2$, and consider the random variable

$$V_{n,k} = n^{2/3}\,(X_\rho^{(n)} - p_k(\omega)),$$

where ρ is a k-cycle. If $p_{2k-1}(\omega) - (p_k(\omega))^2 > 0$, then the previous results show that

$$\mathbb{E}[e^{zV_{n,k}}] = \exp\left(\frac{n^{1/3}}{2}\,\sigma^2\,z^2 + \frac{1}{6}\,L\,z^3\right)(1 + o(1)),$$

where σ and L are the quantities given in the statement of Theorem 11.2.4. This ends the proof of Theorem 11.2.4 and Corollary 11.2.5. Note that the speed of convergence of the cumulants is each time a $O((\alpha_n)^{-1})$ because of the polynomial behavior established in Lemma 11.3.3; therefore, one can indeed apply Proposition 5.2.1 with $\alpha_n = n$ and $\beta_n = n^{-1}$. The theorem can be extended to other permutations ρ of $\mathfrak{S}(\infty)$ than cycles: if ρ has cycle-type μ, then define

$$V_{n,\mu} = n^{2/3}\,(X_\rho^{(n)} - p_\mu(\omega)).$$

The generating series of $V_{n,\mu}$ is asymptotically given by:

$$\mathbb{E}[e^{zV_{n,\mu}}] = \exp\left(\frac{n^{1/3}}{2}\,\sigma^2(\mu)\,z^2 + \frac{1}{6}\,L(\mu)\,z^3\right),$$

where $\sigma^2(\mu) = \lim_{n\to\infty} n\,\kappa^{(2)}(X_{\rho\mu}^{(n)})$ and $L(\mu) = \lim_{n\to\infty} n^2\,\kappa^{(3)}(X_{\rho\mu}^{(n)})$ are the limiting quantities given in Section 11.4.2. Hence, *provided that $\sigma^2(\mu) > 0$*, one has mod-Gaussian convergence of $V_{n,\mu}$, and the limiting variance $\sigma^2(\mu)$ is non-zero under the same conditions as those given in Remark 11.2.6. Under these conditions, one can also

easily establish mod-Gaussian convergence for every vector of renormalised random character values $(V_{n,\mu_1}, \ldots, V_{n,\mu_\ell})$

REMARK 11.5.1. There is one case which is not covered by our theorem, but is of particular interest: the case $\omega = ((0,0,\ldots),(0,0,\ldots))$. This parameter of the Thoma simplex corresponds to the Plancherel measures of the symmetric groups, and in this case, since $p_2(\omega) = p_3(\omega) = \cdots = 0$, the parameters of the mod-Gaussian convergence are all equal to 0. Indeed, the random character values under Plancherel measures do not have fluctuations of order $n^{-1/2}$. For instance, Kerov's central limit theorem (*cf.* [Hor98, IO02]) ensures that the random character values

$$\frac{n^{k/2}\,\widehat{\chi}^\lambda(c_k)}{\sqrt{k}}$$

on cycles c_k of lengths $k \geq 2$ converges in law towards independent Gaussian variables; so the fluctuations are of order $n^{-k/2}$ instead of $n^{-1/2}$. One still expects a mod-Gaussian convergence for adequate renormalisations of the random character values; however, the combinatorics underlying the asymptotics of Plancherel measures are much more complex than those of general central measures — see [Śni06a] — and we have not been able to prove mod-Gaussian convergence here.

From the estimates on the laws of the random character values, one can prove many estimates for the parts $\lambda_1, \lambda_2, \ldots$ of the random partitions taken under central measures. The arguments of algebraic combinatorics involved in these deductions are detailed in [FM12, Mél12], so here we shall only state results. Given a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ of size n, the Frobenius coordinates of λ are the two sequences

$$\left(\lambda_1 - \frac{1}{2}, \lambda_2 - \frac{3}{2}, \ldots, \lambda_d - d + \frac{1}{2}\right), \left(\lambda_1' - \frac{1}{2}, \lambda_2' - \frac{3}{2}, \ldots, \lambda_d' - d + \frac{1}{2}\right)$$

where λ_1', λ_2', *etc.* are the sizes of the columns of the Young diagram of λ, and d is the size of the diagonal of the Young diagram. Denote $(a_1, \ldots, a_d), (b_1, \ldots, b_d)$ these coordinates, and

$$X_\lambda = \sum_{i=1}^{d} \frac{a_i}{n}\, \delta_{\left(\frac{a_i}{n}\right)} + \sum_{i=1}^{d} \frac{b_i}{n}\, \delta_{\left(-\frac{b_i}{n}\right)}.$$

This is a (random) discrete probability measure on $[-1, 1]$ whose moments

$$p_k(\lambda) = n^k \int_{-1}^{1} x^{k-1}\, X_\lambda(dx)$$

are also the moments of the Frobenius coordinates, so X_λ encodes the geometry of the Young diagram λ. We shall also need

$$X_\omega = \sum_{i=1}^{d} \alpha_i\, \delta_{(\alpha_i)} + \sum_{i=1}^{d} \beta_i\, \delta_{(-\beta_i)} + \gamma\, \delta_{(0)},$$

which will appear in a moment as the limit of the random measures X_λ. Here, $\gamma = 1 - \sum_{i=1}^{\infty} \alpha_i - \sum_{i=1}^{\infty} \beta_i$. Notice that $\mathbb{E}_{n,\omega}[X_k^{(n)}] = \tau_\omega(c_k) = X_\omega(x^{k-1})$.

It is shown in [IO02] that for any partition λ of size n and for any k,

$$p_k(\lambda) = \Sigma_k(\lambda) + \text{remainder},$$

where the remainder is a linear combination of symbols Σ_μ with $|\mu| < k$. It follows that the cumulants of the p_k's satisfy the same estimates as the cumulants of the Σ_k's. Therefore, the rescaled random variable $\nabla(x^{k-1}) = n^{2/3}(X_\lambda(x^{k-1}) - X_\omega(x^{k-1}))$ converges in the mod-Gaussian sense with parameters $n^{1/3}\sigma^2$ and limiting function $\psi(z) = \exp(L\frac{z^3}{6})$, where σ^2 and L are given by the same formula as in the case of the random character value $X_k^{(n)}$, that is to say

$$\sigma^2 = k^2 \left(p_{2k-1}(\omega) - p_k(\omega)^2\right);$$
$$L = k^3 \left((3k-2)\, p_{3k-2}(\omega) - (6k-3)\, p_{2k-1}(\omega)\, p_k(\omega) + (3k-1)\, p_k(\omega)^3\right).$$

Actually, one has mod-Gaussian convergence for any finite vector of random variables $\nabla(x^{k-1})$, $k \geq 1$.

We don't know how to obtain from there a moderate deviation result for the parts $\lambda_1, \lambda_2, \ldots$ of the partition; but one has at least a central limit theorem when one has strict inequalities $\alpha_1 > \alpha_2 > \cdots > \alpha_i > \cdots$ and $\beta_1 > \beta_2 > \cdots > \beta_i > \cdots$ (see [Mél12], and also [Buf12]). Indeed, for any smooth test function ψ_i equal to 1 around α_i and to 0 outside a neighbourhood of this point, one has

$$X_\lambda(\psi_i) - X_\omega(\psi_i) = \frac{a_i}{n} - \alpha_i \tag{55}$$

with probability going to 1, and then the quantities in the left-hand side renomalised by $\sqrt{n}$ converge jointly towards a Gaussian vector with covariance

$$\kappa(i,j) = \delta_{ij}\, \alpha_i - \alpha_i \alpha_j. \tag{56}$$

So, the fluctuations

$$\sqrt{n}\left(\frac{\lambda_i}{n} - \alpha_i\right)$$

of the rows of the random partitions taken under central measures $\mathbb{P}_{n,\omega}$ converge jointly towards a Gaussian vector with covariances given by Equation (56), and one can include in this result the fluctuations

$$\sqrt{n}\left(\frac{\lambda_j'}{n} - \beta_j\right)$$

of the columns of the random partitions, with a similar formula for their covariances.

The reason why it becomes difficult to get by the same technique the moderate deviations of the rows and columns is that in Equation (55), one throws away an event of probability going to zero (because of the law of large numbers satisfied by the rows and the columns, see *e.g.* [KV81]). However, one cannot *a priori* neglect this event in comparison to rare events such as $\{a_i - n\alpha_i \geq n^{2/3}x\}$; indeed, these rare events are themselves of probability going exponentially fast to zero. Also, there is the problem of approximation of smooth test functions by polynomials, which one has to control precisely when doing these computations. One still conjectures these moderate deviations to hold, and $n^{2/3}\left(\frac{a_i}{n} - \alpha_i\right)$ to converge in the mod-Gaussian sense with parameters $n^{1/3}(\alpha_i - \alpha_i^2)$ and limiting function

$$\psi(z) = \exp\left(\frac{\alpha_i - 3\alpha_i^2 + 2\alpha_i^3}{6}z^3\right)$$

— this is what one obtains if we ignore the previous caveats, and still suppose the α_i and β_j all distinct. As explained in [Mél12] (see also [KV86]), this would give moderate deviations for the lengths of the longest increasing subsequences in a random permutation obtained by generalised riffle shuffle.

Bibliography

[AB93]   F. Avram and D. Bertsimas. On central limit theorems in geometrical probability. *Ann. Appl. Prob.*, pages 1033–1046, 1993.

[ABT03] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic Combinatorial Structures: a Probabilistic Approach*. EMS Monographs in Mathematics. European Mathematical Society, 2003.

[AGG89] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Prob.*, pages 9–25, 1989.

[AS08]   N. Alon and J. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, 3rd edition, 2008.

[Bax82] R. J. Baxter. *Exactly solved models in statistical mechanics*. Academic Press, 1982.

[BC05]   A. Barbour and L. H. Y. Chen. *An introduction to Stein's method*, volume 4 of *Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore*. Singapore University Press, 2005.

[BDJ99] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12:1119–1178, 1999.

[Ber41] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49(1):122–136, 1941.

[Ber73] K. Berk. A central limit theorem for m-dependent random variables with unbounded m. *Ann. Probab.*, 1(2):352–354, 1973.

[Bin86] N. H. Bingham. Variants on the Law of the iterated logarithm. *Bull. London Math. Soc.*, 18(5):433–467, 1986.

[BKN14] A. D. Barbour, E. Kowalski, and A. Nikeghbali. Mod-discrete expansions. *Probab. Th. Rel. Fields*, 158(3):859–893, 2014.

[Bol98] B. Bollobás. *Modern graph theory*, volume 184 of *Graduate texts in mathematics*. Springer-Verlag, 1998.

[Bol01] B. Bollobás. *Random graphs*, volume 73 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2001. Second edition.

[Bón10] M. Bóna. On three different notions of monotone subsequences. In *Permutation Patterns*, volume 376 of *London Math. Soc. Lecture Note Series*, pages 89–113. Cambridge University Press, 2010.

[BOO00] A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.*, 13:491–515, 2000.

[Bor60]   C. W. Borchardt. Über eine der Interpolation entsprechende Darstellung der Eliminations-Resultante. *Journal für die Reine und angewandte Mathematik*, 1860(57):111–121, 1860.

[BR60]   R. R. Bahadur and R. R. Rao. On deviations of the sample mean. *Ann. Math. Statis.*, 31:1015–1027, 1960.

[BR89]   P. Baldi and Y. Rinott. On normal approximations of distributions in terms of dependency graphs. *Ann. Probab.*, 17(4):1646–1650, 1989.

[Bra05]   R. C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys*, 2:107–144, 2005.

[Buf12]   A. I. Bufetov. A central limit theorem for extremal characters of the infinite symmetric group. *Functional Analysis and Its Applications*, 46(2):83–93, 2012.

[BV07]   I. Bárány and V. Vu. Central limit theorems for Gaussian polytopes. *Ann. Probab.*, 35(4):1593–1621, 2007.

[CDKR15] B. Chern, P. Diaconis, D. M. Kane, and R. C. Rhoades. Central Limit Theorems for some Set Partition Statistics. *Advances in Applied Mathematics*, 70:92–105, 2015.

[CGS11]   L. H. Y. Chen, L. Goldstein, and Q. M. Shao. *Normal approximation by Stein's method*. Probability and Its Applications. Springer-Verlag, 2011.

[Cha12]   S. Chatterjee. The missing log in large deviations for triangle counts. *Random Structures & Algorithms*, 40(4):437–451, 2012.

[Cog72]   R. Cogburn. The central limit theorem for Markov processes. In *Proceedings of the Sixth Annual Berkeley Symposium on Mathematical Statistics and Probability*, volume 2, pages 485–512. University of California Press, 1972.

[Cra38]   H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Sci. Indust.*, 736:5–23, 1938.

[CS04]   L. Chen and Q.-M. Shao. Normal approximation under local dependence. *The Annals of Probability*, 32(3):1985–2028, 2004.

[CV11]   S. Chatterjee and S. R. S. Varadhan. The large deviation principle for the Erdös-Rényi random graph. *Eur. J. Comb.*, 32(7):1000–1017, 2011.

[Das08]   A. DasGupta. *Asymptotic Theory of Statistics and Probability*. Springer Texts in Statistics. Springer-Verlag New York, 2008.

[Dav73]   Y. A. Davydov. Mixing conditions for Markov chains. *Theory of Probability and its Applications*, XVIII(2):312–328, 1973.

[DE13a]   H. Döring and P. Eichelsbacher. Moderate Deviations for the Determinant of Wigner Matrices. In *Limit Theorems in Probability, Statistics and Number Theory*, volume 42 of *Springer Proceedings in Mathematics & Statistics*, pages 253–275. Springer, 2013.

[DE13b]   H. Döring and P. Eichelsbacher. Moderate Deviations via Cumulants. *Journal of Theoretical Probability*, 26(2):360–385, 2013.

[DHT02]   G. Duchamp, F. Hivert, and J.-Y. Thibon. Noncommutative symmetric functions VI: free quasi-symmetric functions and related algebras. *International Journal of Algebra and Computation*, 12:671–717, 2002.

[Dia55]   P. H. Diananda. The central limit theorem for m-dependent variables. *Proc. Comb. Phil. Soc.*, 51:92–95, 1955.

[DKN15]   F. Delbaen, E. Kowalski, and A. Nikeghbali. Mod-ϕ convergence, 2015.

[dM56]   A. de Moivre. *The Doctrine of Chances, or a Method of Calculating the Probability of Events in Play*. A. Millar, 3rd edition, 1756.

[DZ98]   A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2nd edition, 1998.

[EK40]   P. Erdös and M. Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *American Journal of Mathematics*, 62:738–742, 1940.

[EL75]   P. Erdös and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. *Infinite and finite sets*, 10(2):609–627, 1975.

[Ell84]   R. S. Ellis. Large deviations for a general class of random vectors. *Ann. Probab.*, 12:1–12, 1984.

[EN78]   R. S. Ellis and C. M. Newman. Fluctuationes in Curie Weiss exemplis. In *Mathematical problems in theoretical physics*. Springer-Verlag, 1978.

[ER60]   P. Erdös and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[Ess45]   C.-G. Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gauss law. *Acta Math.*, 77(1):1–125, 1945.

[EU14]   N. M. Ercolani and D. Ueltschi. Cycle structure of random permutations with cycle weights. *Random Structures & Algorithms*, 44:109–133, 2014.

[Fel71]   W. Feller. *An Introduction to Probability Theory and Its Applications, Volume II*. Wiley Series in Probability and Mathematical Statistics. Wiley, 2nd edition, 1971.

[Fis11]   H. Fischer. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer-Verlag, 2011.

[FM12]   V. Féray and P.-L. Méliot. Asymptotics of q-Plancherel measures. *Probab. Th. Rel. Fields*, 152(3-4):589–624, 2012.

[FMN15]   V. Féray, P.-L. Méliot, and A. Nikeghbali. Mod-ϕ convergence, III: Multidimensional mod-Gaussian convergence and related estimates of probabilities. In preparation, 2015.

[FMN16]   V. Féray, P.-L. Méliot, and A. Nikeghbali. Mod-ϕ convergence, II: Estimates of the speed of convergence and local limit theorems. In preparation, 2016.

[FS09]   P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[Ful04]   J. Fulman. A card shuffling analysis of deformations of the Plancherel measure of the symmetric group. *Electronic Journal of Combinatorics*, 11(1):R21, 2004.

[FW32]   R. A. Fisher and J. Wishart. The derivation of the pattern formulae of two-way partitions from those of simpler patterns. *Proceedings of the London Mathematical Society*, s2-33(1):195–208, 1932.

[Gär77]   J. Gärtner. On large deviations from the invariant measure. *Th. Probab. Appl.*, 22:24–39, 1977.

[GJL75]   G. Gallavotti and G. Jona-Lasinio. Limit theorems for multidimensional Markov processes. *Comm. Math. Phys.*, 41:301–307, 1975.

[GL78]   M. Gordin and B. A. Lifšic. Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR*, 239:766–767, 1978.

[Gne48]   B. V. Gnedenko. On the local limit theorem in the theory of probability. *Uspekhi Mat. Nauk*, 3:187–194, 1948.

[Hai71]   J. Haigh. A neat way to prove asymptotic normality. *Biometrika*, 58(3):677–678, 1971.

[Ham72]   J. M. Hammersley. A few seedlings of research. In *Proc. Sixth Berkeley Symp. Math. Statist. and Probability*, volume 1, pages 345–394. University of California Press, 1972.

[Har67]   L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 38(2):410–414, 1967.

[HH80]   P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application*. Probability and Mathematical Statistics. Academic Press, 1980.

[HJ10]   P. Hitczenko and S. Janson. Asymptotic normality of statistics on permutation tableaux. *Contemporary Math.*, 520:83–104, 2010.

[HKO01]   C.P. Hughes, J.P. Keating, and N. O'Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.*, 220(2):429–451, 2001.

[HKPV09]   J. Ben Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, volume 51 of *University Lecture Series*. AMS, 2009.

[Hor98]   A. Hora. Central Limit Theorem for the Adjacency Operators on the Infinite Symmetric Group. *Comm. Math. Phys.*, 195(2):405–416, 1998.

[HR48]   W. Hoeffding and H. Robbins. The central limit theorem for dependent random variables. *Duke Math. J.*, 15(3):773–780, 1948.

[HS02]   H.-K. Hwang and J.-M. Steyaert. On the number of heaps and the cost of heap construction. In *Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, pages 294–310, 2002.

[HS10]   T. Hasebe and H. Saigo. Joint cumulants for natural independence. *Electronic Communications in Probability*, 16:491–506, 2010.

[Hwa94]   H.-K. Hwang. *Théoremes limites pour les structures combinatoires et les fonctions arithmétiques, phd thesis, Ecole Polytechnique*. PhD thesis, Ecole Polytechnique, 1994.

[Hwa96]   H.-K. Hwang. Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.*, 6(1):297–319, 1996.

[Hwa98] H.-K. Hwang. On convergence rates in the central limit theorems for combinatorial structures. *Europ. J. Combin.*, 19(3):329–343, 1998.

[IL71] I. A. Ibragimov and Y. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff, 1971.

[Ilt95] M. Iltis. Sharp asymptotics of large deviations in $\mathbb{R}^d$. *J. Theoretical Prob.*, 8:501–524, 1995.

[IO02] V. Ivanov and G. Olshanski. Kerov's central limit theorem for the Plancherel measure on Young diagrams. In *Symmetric Functions 2001*, volume 74 of *NATO Science Series II. Mathematics, Physics and Chemistry*, pages 93–151, 2002.

[Jan88] S. Janson. Normal Convergence by Higher Semiinvariants with Applications to Sums of Dependent Random Variables and Random Graphs. *Ann. Probab.*, 16(1):305–312, 1988.

[Jan04] S. Janson. Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms*, 24(3):234–248, 2004.

[JKN11] J. Jacod, E. Kowalski, and A. Nikeghbali. Mod-Gaussian convergence: new limit theorems in Probability and Number Theory. *Forum Mathematicum*, 23(4):835–873, 2011.

[JŁR00] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*, volume 45 of *Wiley Series in Discrete Mathematics and Optimization*. Wiley-Interscience, 2000.

[Joh01] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. Math.*, 153:259–296, 2001.

[Jon04] G. L. Jones. On the Markov chain central limit theorem. *Probability Surveys*, 1:299–320, 2004.

[Kal97] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer-Verlag, 1997.

[Kat67] S. K. Katti. Infinite divisibility of integer valued random variables. *Ann. Math. Statist.*, 38:1306–1308, 1967.

[Ker03] S. V. Kerov. *Asymptotic Representation Theory of the Symmetric Group and Its Applications in Analysis*, volume 219 of *Translations of Mathematical Monographs*. AMS, 2003.

[KN10] E. Kowalski and A. Nikeghbali. Mod-Poisson convergence in probability and number theory. *Intern. Math. Res. Not.*, 18:3549–3587, 2010.

[KN12] E. Kowalski and A. Nikeghbali. Mod-Gaussian distribution and the value distribution of $\zeta(1/2 + it)$ and related quantities. *J. London Math. Soc.*, 86(2):291–319, 2012.

[KNN15] E. Kowalski, J. Najnudel, and A. Nikeghbali. A characterization of limiting functions arising in mod-* convergence. *Electronic Communications in Probability*, 20(79):1–11, 2015.

[KS00a] J. P. Keating and N. C. Snaith. Random matrix theory and L-functions at $s = 1/2$. *Comm. Math. Phys.*, 214(1):91–110, 2000.

[KS00b] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[Kub72] J. P. Kubilius. On large deviations of additive arithmetic functions. *Trudy Mat. Inst. Steklov.*, 128:163–171, 1972.

[KV77] S. V. Kerov and A. M. Vershik. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. *Soviet Mathematics Doklady*, 18:527–531, 1977.

[KV81] S. V. Kerov and A. M. Vershik. Asymptotic theory of characters of the symmetric group. *Funct. Anal. Appl.*, 15(4):246–255, 1981.

[KV86] S.V. Kerov and A. M. Vershik. The Characters of the Infinite Symmetric Group and Probability Properties of the Robinson-Schensted-Knuth Algorithm. *SIAM J. Algebraic Discrete Methods*, 7(1):116–124, January 1986.

[Lap40] P.-S. Laplace. *Essai philosophique sur les probabilités*. Bachelier, 6th edition, 1840.

[Lév25] P. Lévy. *Calcul des probabilités*. Gauthier-Villars, 1925.

[Lin22] J. W. Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Zeitschrift*, 15(1):211–225, 1922.

[LL96] Z. Lin and C. Lu. *Limit Theory for Mixing Dependent Random Variables*. Kluwer Academic Publishers, 1996.

[LPRS14] J.L. Lebowitz, B. Pittel, D. Ruelle, and E.R. Speer. Central limit theorems, Lee-Yang zeros, and graph-counting polynomials. arXiv preprint 1408.4153, 2014.

[LS52] E. Lukacs and O. Szász. On analytic characteristic functions. *Pacific Journal of Mathematics*, 2(4):615–625, 1952.

[LS59] V.P. Leonov and A.N. Shiryaev. On a method of calculation of semi-invariants. *Theory Prob. Appl.*, 4:319–329, 1959.

[LS77] B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Advances in Math.*, 26:206–222, 1977.

[Lya00] A. M. Lyapunov. Sur une proposition de la théorie des probabilités. *Bulletin de l'Académie Impériale des Sciences de St.-Pétersbourg (5)*, 13:359–386, 1900.

[Lya01] A. M. Lyapunov. Sur un théorème du calcul des probabilités. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences de Paris*, 132:126–128, 1901.

[Mél12] P.-L. Méliot. Fluctuations of central measures on partitions. In *Proceedings of the 24th International Conference on Formal Power Series and Algebraic Combinatorics (Nagoya, Japan)*, pages 387–398, 2012.

[Mik91] V. G. Mikhailov. On a Theorem of Janson. *Theory of Probability & Its Applications*, 36(1):173–176, 1991.

[Moh95] A. Mohr. The Moser-Wyman expansion of the Bell numbers. Unpublished note, available on the web, 1995.

[Mur03] N. Muraki. The five independences as natural products. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 6(3):337–371, 2003.

[MW55] L. Moser and M. Wyman. An Asymptotic Formula for the Bell Numbers. *Trans. Royal Soc. Canada*, 49:49–54, 1955.

[New80] C. M. Newman. Normal fluctuations and the FKG inequalities. *Comm. Math. Phys.*, 74(2):119–128, 1980.

[Ney83] P. Ney. Dominating points and the asymptotics of large deviations for random walk on $\mathbb{R}^d$. *Ann. Probab.*, 11:158–167, 1983.

[NZ13] A. Nikeghbali and D. Zeindler. The generalized weighted probability measure on the symmetric group and the asymptotic behavior of the cycles. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(4):961–981, 2013.

[Oko99] A. Okounkov. On representations of the infinite symmetric group. *Journal of Mathematical Sciences*, 96(5):3550–3589, 1999.

[Oko00] A. Okounkov. Random matrices and random permutations. *Internat. Math. Res. Notices*, 20:1043–1095, 2000.

[Ols90] G. Olshanski. Unitary representations of (G, K)-pairs connected with the infinite symmetric group $S(\infty)$. *Leningrad Math. J.*, 1:983–1014, 1990.

[Pen02] M. Penrose. *Random Geometric graphs*, volume 5 of *Oxford studies in probability*. Oxford University Press, 2002.

[Pet54] V. V. Petrov. A generalization of Cramér's limit theorem. *Uspehi Mat. Nauk.*, 9(4):195–202, 1954.

[Pet75] V. V. Petrov. *Sums of independent random variables*, volume 82 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1975.

[Pet95] V. Petrov. *Limit theorems of probability theory*. Oxford Science Publications, 1995.

[Phi69] W. Philipp. The central limit problem for mixing sequences of random variables. *Z. Wahrsch. verw. Gebiete*, 12:155–171, 1969.

[PL83] M.B. Petrovskaya and A. M. Leontovich. The central limit theorem for a sequence of random variables with a slowly growing number of dependences. *Th. Prob. Appl.*, 27(4):815–825, 1983.

[PY05] M. Penrose and JE Yukich. Normal approximation in geometric probability. *Stein's Method and Applications, Lecture Note Series, Institute for Mathematical Sciences, National University of Singapore*, 5:37–58, 2005.

[Rad09] M. Radziwill. On large deviations of additive functions. B.Sc thesis, `arXiv:09095274v4 [math.NT]`, 2009.

[RH48] H. Robbins and W. Hoeffding. The central limit theorem for dependent random variables. 15:773–780, 1948.

[Rob38] G. Robinson. On the representations of the symmetric group. *Amer. J. Math.*, 60(3):745–760, 1938.

[Ros56] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA*, 42:43–47, 1956.

[Ros11] N. Ross. Fundamentals of Stein's method. *Probability Surveys*, 8:210–293, 2011.

[RSS78] R. Rudzkis, L. Saulis, and V. A. Statulevičius. A general lemma for large deviations. *Lithuanian Math. J.*, 18:99–116, 1978.

[Ruc88] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probab. Th. Rel. Fields*, 78(1):1–10, 1988.

[Sag01] B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms and Symmetric Functions*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, 2nd edition, 2001.

[Sch61] C. Schensted. Longest increasing and decreasing subsequences. *Canadian Journal of Mathematics*, 13:179–191, 1961.

[Sel54] A. Selberg. Note on a paper by LG Sathe. *J. Indian Math. Soc*, 18:83–87, 1954.

[SH04] F. W. Steutel and K. Van Harn. *Infinite divisibility of probability distributions on the real line*, volume 259 of *Monographs and textbooks in pure and applied mathematics*. Marcel Dekker, 2004.

[Śni06a] P. Śniady. Asymptotics of characters of symmetric groups, genus expansion and free probability. *Discrete Math.*, 306(7):624–665, 2006.

[Śni06b] P. Śniady. Gaussian fluctuations of characters of symmetric groups and of Young diagrams. *Probab. Th. Rel. Fields*, 136(2):263–297, 2006.

[Spe77] J. Spencer. Asymptotic lower bounds for ramsey functions. *Discrete Mathematics*, 20:69–76, 1977.

[SS91] L. Saulis and V. A. Statulevičius. *Limit Theorems for Large Deviations*. Kluwer Academic Publications, 1991.

[Sta83] A. J. Stam. Generation of a random partition of a finite set by an urn model. *J. Comb. Th., Series A*, 35(2):231–240, 1983.

[Sta01] R. P. Stanley. Generalized riffle shuffles and quasisymmetric functions. *Annals of Combinatorics*, 5(3):479–491, 2001.

[Str11] D. W. Stroock. *Probability Theory: An analytic view*. Cambridge University Press, 2nd edition, 2011.

[SV75] S.-V. Strâtilâ and D.-V. Voiculescu. *Representations of AF-Algebras and of the Group $U(\infty)$*, volume 486 of *Lecture Notes in Mathematics*. Springer-Verlag, 1975.

[Sze75] G. Szegö. *Orthogonal polynomials*, volume XXIII of *AMS Colloquium Publications*. American Mathematical Society, 4th edition, 1975.

[Ten95] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 46 of *Cambridge Studies in Advanced Mathematics*. Cambridge university press, 1995.

[Tho64] E. Thoma. Die unzerlegbaren, positive-definiten Klassenfunktionen der abzählbar unendlichen symmetrischen Gruppe. *Math. Zeitschrift*, 85:40–61, 1964.

[Ula61] S. M. Ulam. Monte-Carlo calculations in problems of mathematical physics. In *Modern Mathematics for the Engineers*, pages 261–281. McGraw-Hill, 1961.

[Wol83] S. J. Wolfe. Continuity properties of decomposable probability measures on Euclidean spaces. *J. Multivariate Analysis*, 13:534–538, 1983.

[WW27] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, 4th edition, 1927.