Abstract. In this paper, we use the framework of mod-$\phi$ convergence to prove precise large or moderate deviations for quite general sequences of random variables $(X_n)_{n \in \mathbb{N}}$. The random variables considered can be lattice or non-lattice distributed, and single or multi-dimensional; and one obtains precise estimates of the fluctuations $\mathbb{P}[X_n \in t_n B]$, instead of the usual estimates for the rate of exponential decay $\log(\mathbb{P}[X_n \in t_n B])$. In the special setting of mod-Gaussian convergence, we shall see that our approach allows us to identify the scale at which the central limit theorem ceases to hold and we are able to quantify the "breaking of symmetry" at this critical scale thanks to the residue or limiting function occurring in mod-$\phi$ convergence.

The first sections of the article are devoted to a proof of these abstract results. We then propose new examples covered by this theory and coming from various areas of mathematics: classical probability theory (multi-dimensional random walks, random point processes), 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). 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. Although the latter methods can only be applied in the more restrictive setting of mod-Gaussian convergence, the large number as well as the variety of examples which are covered there hint at a universality class for second order fluctuations.
10.3. Moderate deviations for subgraph count statistics
11. Random character values from central measures on partitions
11.1. Preliminaries
11.2. Bounds and limits of the cumulants
11.3. Asymptotics of the random character values and partitions
12. Appendix: Berry-Esseen estimates for mod-convergence in $\mathbb{R}^d$
12.1. Properties of certain smoothing kernels
12.2. Estimates for test functions
12.3. Gaussian regular domains and convex bodies
12.4. Estimates for Gaussian regular domains
References
1. Introduction

1.1. Mod-$\phi$ convergence. The notion of mod-$\phi$ convergence has been studied in the articles [JKN11, DKN11, KN10, KN12, BKN13], in connection with problems from number theory, random matrix theory and probability theory. The main idea was to look for a natural renormalization of the characteristic functions of random variables which do not converge in law (instead of a renormalization of the random variables themselves). After this renormalization, the sequence of characteristic functions converges to some non-trivial limiting function. Here is the definition of mod-$\phi$ convergence that we will use throughout this article (see Section 1.4 for a discussion on the different part of this definition).

Definition 1.1. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of random variables, and $\varphi_n(z) = \mathbb{E}[e^{zX_n}]$ be their moment generating functions, which we assume to all exist in a strip

$$S_c = \{z, -c < \text{Re}z < c\},$$

with $c$ positive real number. We assume that there exists a non-constant infinitely divisible distribution $\phi$ with moment generating function $\int_{\mathbb{R}} e^{zx} \phi(dx) = \exp(\eta(z))$ that is convergent on $S_c$, and an analytic function $\psi(z)$ that does not vanish on the real part of $S_c$, such that locally uniformly in $z \in S_c$,

$$\exp(-t_n \eta(z)) \varphi_n(z) \to \psi(z),$$

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-$\phi$, with parameters $(t_n)_{n \in \mathbb{N}}$ and limiting function $\psi$. In the following we denote $\psi_n(z)$ the left-hand side of (1).

When $\phi$ is the standard Gaussian (resp. Poisson) distribution, we will speak of mod-Gaussian (resp. mod-Poisson) convergence.

It is immediate to see that mod-$\phi$ convergence implies a central limit theorem if the sequence of parameters $t_n$ goes to infinity (see the remark after Theorem 3.9). But in fact there is much more information encoded in mod-$\phi$ convergence than merely the central limit theorem. Indeed the works [JKN11, DKN11, KN10, KN12, BKN13] tend to illustrate the fact that mod-$\phi$ convergence appears as a natural extension of the framework of sums of independent random variables, in the sense that 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-$\phi$ sense. For instance, under some general extra assumptions on the convergence in Equation (1), it is proved in [DKN11, KN12, FMN14] 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 in the framework of mod-$\phi$ convergence. On the other hand, it applies also 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 [BKN13] that mod-Poisson convergence (in fact mod-$\phi$ convergence for $\phi$ 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-$\phi$ 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 $\psi$ accounts for the fact that prime numbers do not behave independently of each other. More precisely, the limiting function $\psi$ 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 Erdős-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-$\phi$ convergence by a factor which is contained in $\psi$. More generally, it seems that the limiting function encodes information about the dependence between the different components of a random vector: indeed it is noted in [KN12] that 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 normalized it converges to a Gaussian vector with independent components.

1.2. Our results. The goal of this paper is to prove that the framework of mod-$\phi$ convergence as described in Definition 1.1 is suitable to obtain precise (i.e., estimates of the probabilities without the log) large and moderate deviation results for the sequence $(X_n)_{n \in \mathbb{N}}$. 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 3.3 for a lattice distribution $\phi$ and Theorem 4.3 for a non-lattice distribution). This can be viewed as an analog (or an extension, see Remark 4.9) 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 nice Borelian $B$, see Theorem 6.3. This result is an analog 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 bound involves 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 3.9 in the lattice case, and Theorem 4.8 in the non-lattice case).

An interesting fact in our deviation results is the appearance of the limiting function $\psi$ in deviations of scale $t_n$. This means that, at smaller scales, a sequence $X_n$ converging mod-$\phi$ behaves exactly as a sum of $t_n$ i.i.d. variables with distribution $\phi$. However, at scale $t_n$, this is not true any more and the limiting function $\psi$ gives us exactly the correcting factor (cf. §5.1).

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 $\psi$ is the equivalent; see Remark 3.4.

A special case of mod-Gaussian deviations is the case where the convergence 1 is proved using bounds on the cumulants of $X_n$ — see Example 2.3. This case is particularly interesting as:

- it contains a large class of examples, see below;
- in this setting, one can obtain deviation results at a scale bigger that $t_n$ (typically, $o((t_n)^{5/4})$, see Proposition 5.4).

Besides, in the mod-Gaussian case, we are able to establish precise deviation results at the scale $t_n$ in a multidimensional setting (Theorem 7.6). 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. This Theorem seems to be new, even in the case of i.i.d. random variables.

The arguments involved in the proofs of our large deviation results in dimension one are standard, but they nonetheless need to be carefully adapted to the framework of Definition 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. However, it should be noted that in the multidimensional case the arguments are more involved, and we provide new Berry-Esseen type estimates for Gaussian regular domains or polytopes which may have an interest in their own, and which extend recent results from [BR10].

**Remark 1.2.** Our Berry-Esseen estimates (Proposition 4.1) shall hold in the general setting of mod-$\phi$ convergence with respect to a non-lattice infinitely divisible law. In return, they are not at all optimal: the correct order of decay of

$$d_{\text{Kol}} \left( \frac{X_n - t_n \eta'(0)}{\sqrt{t_n \eta''(0)}}, \mathcal{N}(0, 1) \right)$$

is usually better than $O((t_n)^{-1/2})$, as can be seen for instance in the case of sum of i.i.d. random variables: there, the Kolmogorov distance is an $O(n^{-1/2}) = O((t_n)^{-3/2}). It is possible to use the framework of mod-$\phi$ convergence in order to get such optimal bounds on the Kolmogorov distance, see [FMN14]. One then needs to control precisely the zone on which the mod-convergence happens, and make it grow with $n$ in an appropriate way. Besides, these estimates on the speed of convergence require the use of a class of test functions, which is also involved in the proof of local limit theorems ([DKN11, KN12]). We refer to our forthcoming article [FMN14] for details on this approach to Berry-Esseen estimates, which provides a mod-$\phi$ alternative to Stein’s method.

**Remark 1.3.** We should here mention some work of Hwang [Hwa96], with some similarities with ours.

Hwang works with some hypothesis similar to Definition 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 1.6 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 \( \psi \) 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 last four sections of the paper to examples, from a variety of different areas. Section 8 contains what one would be tempted to call “known examples”, in the sense that most of these examples were proved to satisfy mod-\( \phi \) convergence in earlier works. But, also in these cases, the precise moderate or large deviation results we obtain are new.

For instance, one can obtain a surprisingly new result on the simple random walk on \( \mathbb{Z}^d \) with \( n \) steps conditioned to end at distance \( r_n \) from the origin. When \( r_n = o(n^{3/4}) \), the distribution of the angle \( \theta \) of the end point of the walk is uniform (for a non-conditioned walk, this is the case because of the conformal invariance of Brownian motion). But when, \( r_n = cn^{3/4} \), the end point of the walk has higher chance to be near the axis, and we give explicitly the distribution of \( \theta \), see Section 8.1.

In §8.5, we compute deviation probabilities the characteristic polynomial of random unitary matrices of different types. This completes previous results by Hughes, Keating and O’Connell [HKO01] on large deviations for the characteristic polynomial. We also recover results of Radziwill [Rad09] on precise large deviations for additive arithmetic functions, by carefully recalling the principle of the Selberg-Delange method as well as results by Nikeghbali and Zeindler [NZ13] on precise large deviations for the total number of cycles for random permutations under the general weighted probability measure as an illustration of the singularity analysis method.

Our next examples lie in the framework in which mod-Gaussian convergence is obtained via bounds on cumulants (Example 2.3). In Section 9, we show that such bounds on cumulants typically arise in the context of sparse dependency graphs, that is for sums of partially dependent random variables (references and details are provided in Section 9). This allows us to provide new examples of variables converging in the mod-Gaussian sense:

- renormalized subgraph count statistics in Erdős-Rényi random graph \( G(n, p) \) (Theorem 10.1) for a fixed \( p \) between 0 and 1. The moderate deviation probabilities in this case are given and compared with the literature on the subject in Section 10. Our proof of the mod-Gaussian convergence also imply a local limit theorem and a Berry-Esseen estimate, see [FMN14].

- in our last application in §11, we use the machinery of dependency graphs in non-commutative probability spaces, namely, the algebras \( C\mathfrak{S}(n) \) 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 1.4. The idea of using bounds on cumulants to show precise deviations for a family of random variables with some given dependency graph is not new — see in particular [DE12]. Nevertheless, the bounds we obtain in Theorem 9.2 (and also in Theorem 9.3) are stronger than those which were previously known and, as a consequence, we obtain deviation results at larger scale.

1.4. Discussion on our hypotheses. The following remarks explain the role of each hypothesis of Definition 1.1. As we shall see later, some hypotheses can be removed in some of our results (e.g., the infinite-divisibility of the reference law), but Definition 1.1 provides a coherent setting where all the techniques presented in the paper do apply without further verification.

Remark 1.5 (Analyticity). The existence of the relevant moment generating function on a strip is crucial in our proof, as we consider in Section 4 the Fourier transform of $\tilde{X}_n$, obtained from $X_n$ by an exponential change of measure. We also use respectively the existence of continuous derivatives up to order 3 for $\eta$ and $\psi$ on the strip $S_c$, and the local uniform convergence of $\psi_n$ and its first derivatives (say, up to order 3) toward those of $\psi$. By Cauchy 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 hypotheses of analyticity are a bit restrictive, as they imply that the $X_n$’s and $\phi$ have moments of all order; in particular, $\phi$ cannot be any infinitely divisible distribution (for instance the Cauchy distribution is excluded). That explains that the theory of mod-$\phi$ 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. [DKN11, KNN13]), while the concept of mod-Cauchy convergence does not even make sense in the sense of Definition 1.1. We are unfortunately not able to give precise deviation results in this framework.

Remark 1.6 (The disk or the strip?). As mentioned in Remark 1.3, Hwang defined some framework, now called quasi-powers, which is similar to ours except that the convergence takes place in a disk rather than in a 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 $D_{(0,c)}$, then this function is automatically defined and analytic on the strip $S_c$. This implies that the left-hand side of Eq. (1) 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 and the distinction between disk and strip disappears as $D_{(0,c)} = S_c = C$.

Besides, we shall see that assuming a convergence on a strip is a good setting in order to obtain deviation results at order $t_n$ (which Hwang can not achieve).
Remark 1.7 (Infinite divisibility and non-vanishing of the terms of mod-$\phi$ convergence). The non-vanishing of $\psi$ is a natural hypothesis since evaluations of $\psi$ appear in many estimates of non-zero probabilities, and also in denominators in fractions, see for instance Lemma 4.7. However, the assumption that $\phi$ is an infinite divisibility distribution can be relaxed, and we shall only need good estimates on ratios
\[
\left( \int e^{(h+iu)x} \phi(dx) \right) / \left( \int e^{hx} \phi(dx) \right).
\]
These estimates will be provided by Proposition 3.11 in the infinitely divisible case.

In particular, it is very tempting to say that if $S_n = Y_1 + \cdots + Y_n$ is a sum of independent and identically distributed random variables, then $S_n$ converges mod-$Y$ with parameters $t_n = n$ and limiting function $\psi = 1$, even when $Y$ is not infinitely divisible: indeed, $E[e^{zS_n}] = (E[e^{zY}])^n \times 1$. This makes sense under quite weak hypothesis (see Remark 4.9 for a precise statement) and, with this viewpoint, the precise large deviation result of Bahadur-Rao [BR60] fits in our framework.

ACKNOWLEDGEMENTS

The authors would like to thank Martin Wahl for his input at the beginning of this project and for sharing with us his ideas. We would also like to address special thanks to Andrew Barbour, Reda Chhaibi and Kenny Maples for many fruitful discussions which helped us improve some of our arguments.

2. Preliminaries

2.1. Basic examples of mod-convergence. Let us give a few examples of mod-$\phi$ 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.

Definition 2.1. We say that the sequence of random variables $(X_n)_{n \in \mathbb{N}}$ converges mod-$\phi$ at speed $O((t_n)^{-v})$ if the difference of the two sides of Equation (1) can be bounded by $C_K(t_n)^{-v}$ for any $z$ in a given compact subset $K$ of $S_c$. We use the analogue definition with the $o(\cdot)$ notation.

Example 2.2. Let $(Y_n)_{n \in \mathbb{N}}$ be a sequence of centered, independent and identically distributed real-valued random variables, with $E[e^{zY}] = E[e^{zY_1}]$ analytic and non-vanishing on a strip $S_c$, possibly with $c = +\infty$. Set $S_n = Y_1 + \cdots + Y_n$. If the distribution of $Y$ is infinitely divisible (as mentioned earlier, this hypothesis could be weakened — see Remark 4.9), 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
\[
\log E[e^{zS_n}] = n \log E[e^{zY}] = n \sum_{r=2}^{\infty} \frac{\kappa(r)(Y)}{r!} z^r,
\]
which is also analytic on $S_c$ — the coefficients $\kappa(r)(Y)$ are the cumulants of the variable $Y$. 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 variable, and has therefore its odd moments and cumulants that vanish. One has

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

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

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

that is, $(X_n)_{n \in \mathbb{N}}$ converges in the mod-Gaussian sense with parameters $t_n = \sigma^2 n^{v-2}$, 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 [KNN13] to characterize the set of limiting functions in the setting of mod-$\phi$ 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-$\phi$ 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 |
|---------|-----------------|------------------------|
| $S_n$   | $n$             | $\mathcal{N}(0, 1)$    |
| $\frac{S_n}{n^{1/3}}$ | $\sigma^2 n^{1-\frac{2}{3}}$ | large deviations (cf. [BR60, DZ98]) |
| $\frac{S_n}{n^{1/(v+1)}}$ | $\mathcal{N}(0, 1)$ | moderate deviations |
| $\frac{S_n}{n^{1/2}}$ | $\mathcal{N}(0, 1)$ | central limit theorem |
| $\frac{S_n}{n^{1/2}}$ | $\mathcal{N}(0, 1)$ | local limit theorem (cf. [DKN11, KN12, FMN14]) |

**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 Section 5.

**Example 2.3.** More generally, 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 $|\kappa(r)(S_n)| \leq (Cr)^r \alpha_n(\beta_n)^r$ 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 $3 \leq r < \nu$ and all $n \in \mathbb{N}$;

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

$$
\kappa^{(2)}(S_n) = \sigma^2 \alpha_n(\beta_n)^2 \left(1 + o\left((\alpha_n)^{-\frac{\nu-2}{\nu}}\right)\right) \quad ; \quad \kappa^{(\nu)}(S_n) = L \alpha_n(\beta_n)^{\nu} (1 + o(1)).
$$

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

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

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

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

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

Hence,

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

locally uniformly on $\mathbb{C}$, so one has again mod-Gaussian convergence, with parameters $t_n = \sigma^2 (\alpha_n)^{\frac{\nu-2}{\nu}}$ and limiting function $\psi(z) = e^{\frac{L}{\nu!} z^\nu}$. The case of sums of i.i.d. variables 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 Section 9, we show that such bounds on cumulants can be obtained for sums of partially dependent random variables. This leads to a powerful method of cumulants to prove the mod-Gaussian convergence of a sequence of random variables; see Sections 5 and 9-11.

**Example 2.4.** Denote $X_n$ the number of disjoint cycles (including fixed points) of a random permutation chosen uniformly in the symmetric group $S_n$. Feller’s coupling (cf. [ABT03, Chapter 1]) shows that $X_n = \text{law} \sum_{i=1}^{n} B_{(1/i)}$, where $B_p$ 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,

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

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

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

locally uniformly, i.e., one has mod-Poisson convergence with parameters $t_n = \log n$ and limiting function $\frac{1}{\Gamma(e^\nu)}$. Moreover, the speed of convergence is a $O\left(\frac{1}{n}\right)$, hence, a $o(t_n^{-\nu})$ for any integer $\nu$. 

scaling $X_n$ mod-convergence $H_n$ regime of fluctuations $X_n - H_n (H_n)^{1/3}$ large deviations (cf. [Rad09])

$X_n - H_n (H_n)^{1/3}$ moderate deviations

$X_n - H_n (H_n)^{2/3}$ central limit theorem (cf. [EK40])

$X_n - H_n (H_n)^{1/2}$ 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

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

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

It is a feature of every sequence that converges mod-Poisson to yield after rescaling a sequence that converges in the mod-Gaussian sense. The two mod-convergence correspond to two regime of fluctuations, namely, the regime of large deviations for the mod-Poisson convergence, and the regime of the central limit theorem and of moderate deviations for the mod-Gaussian convergence.

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, §2.2]) that we shall use a lot in this paper. The Legendre-Fenchel transform is the following operation on (convex) functions:

**Definition 2.5.** The Legendre-Fenchel transform of a function $\eta$ 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 $\eta$ is the logarithm of the moment generating series of a random variable. In this case, $\eta$ 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 notations lighter). This implies the following useful identities:

$$F(x) = xh - \eta(h) ; \quad F'(x) = h ; \quad F''(x) = h'(x) = \frac{1}{\eta''(h)}.$$
Example 2.6. If $\eta(z) = mz + \frac{\sigma^2 z^2}{2}$ (Gaussian variable with mean $m$ and variance $\sigma^2$), then

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

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

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

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 2.7 (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

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

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

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

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

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

$$\int_0^\infty y e^{-\frac{(y+a)^2}{2}} dy = \frac{e^{-\frac{a^2}{2}}}{a^2} \left(1 + O\left(\frac{1}{a^2}\right)\right)$$
\[
\int_0^\infty y^2 e^{-\frac{(y^2)^2}{2}} \, dy = O\left(\frac{e^{-\frac{y^2}{2}}}{\sqrt{a}}\right)
\]
\[
\int_0^\infty y^3 e^{-\frac{(y^2)^2}{2}} \, dy = O\left(\frac{e^{-\frac{y^2}{2}}}{\sqrt{a}}\right)
\]

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

(4) complex transform: for \(\beta > 0\),
\[
\int_{\mathbb{R}} \frac{e^{-\frac{x^2}{2}}}{\beta + i\omega} \, dw = \int_{\beta}^\infty \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \, d\alpha = P[\mathcal{N}(0,1) \geq \beta].
\]

Proof. Recall that the generating series of Hermite polynomials ([Sze75, Chapter 5]) is
\[
\sum_{r=0}^{\infty} H_r(x) \frac{t^r}{r!} = e^{\frac{x^2}{2}} \sum_{r=0}^{\infty} \frac{(-t)^r}{r!} \frac{\partial^r}{\partial x^r} \left( e^{-\frac{x^2}{2}} \right) = e^{\frac{x^2}{2}} e^{-\frac{(x-i\xi)^2}{2}} \equiv e^{-\frac{\beta^2 t}{2}} e^{i\beta \xi t}.
\]

Integrating against \(g(x) e^{i\xi x} \, dx\) yields
\[
\sum_{r=0}^{\infty} (g(x) H_r(x))^\prime \left(\frac{t}{r!}\right) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{(x-i\xi)^2}{2}} \, dx
\]
\[
= \frac{e^{i\xi t}}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{x^2}{2} + i\xi y} \, dy = e^{i\xi \alpha^2} \sum_{r=0}^{\infty} (i\xi)^r e^{-\frac{\beta^2 t}{r!}}
\]
whence the identity (2) for Fourier transforms.

With \(r = 0\), one gets the Fourier transform of the Gaussian \(g^\prime(\xi) = e^{-\frac{\xi^2}{2}}\), hence the moments (1) by derivation at \(\xi = 0\). 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
\[
F(\beta) = \int_{\mathbb{R}} \frac{e^{-\frac{x^2}{2}}}{\beta + i\omega} \, dw = \frac{1}{2\pi} \int_{\Gamma = \beta + i\mathbb{R}} \frac{e^{\frac{(z-\beta)^2 - \beta^2}{2}}}{z} \, dz,
\]
the second integral being along the complex curve \(\Gamma = \beta + i\mathbb{R}\). Since \(\lim_{\beta \to +\infty} F(\beta) = 0\),
\[
F(\beta) = -\int_{\beta}^{\infty} F'(\alpha) \, d\alpha = \int_{\beta}^{\infty} \left(\frac{1}{2\pi} \int_{\Gamma = \alpha + i\mathbb{R}} \frac{e^{\frac{(z-\alpha)^2 - \alpha^2}{2}}}{z} \, dz\right) \, d\alpha = \int_{\beta}^{\infty} \frac{e^{-\frac{\alpha^2}{2}}}{\sqrt{2\pi}} \, d\alpha,
\]
which is the tail \(P[\mathcal{N}(0,1) \geq \beta]\) of a standard Gaussian law.

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.
3. Fluctuations in the case of lattice distributions

3.1. Lattice and non-lattice distributions. If \( \phi \) is an infinitely divisible distribution, recall that its characteristic function writes uniquely as

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

where \( \Pi \) is the Lévy measure of \( \phi \) and is required to integrate \( 1 \wedge x^2 \). If \( \sigma^2 > 0 \), then \( \phi \) has a normal component and its support \( \text{supp}(\phi) = \left( \text{smallest closed subset } S \text{ of } \mathbb{R} \text{ with } \phi(S) = 1 \right) \) is the whole real line, since \( \phi \) 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 \( \phi \). We refer to [SH04, Chapter 4, Theorem 8.4] for the following result:

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

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

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 \( \Pi \) is either \( \mathbb{R} \), or the half-line \([\gamma, +\infty)\), or the half-line \((-\infty, \gamma)\).

Recall on the other hand 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 \( \gamma \) is well-defined and finite, and if the subgroup \( \mathbb{Z}[\text{supp}(\Pi)] \) is discrete. Otherwise, we say that \( \phi \) is a non-lattice infinitely divisible distribution.

**Proposition 3.1.** An infinitely divisible distribution \( \phi \) 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 \( \gamma \) and \( \lambda > 0 \).

2. with \( \lambda \) taken maximal in the previous item, the characteristic function \( \phi(e^{iux}) \) has modulus \( \left| \phi(e^{iux}) \right| = 1 \) if and only if \( u \in \frac{2\pi}{\lambda} \mathbb{Z} \).

An infinitely divisible distribution \( \phi \) is of type non-lattice if and only if \( \left| \phi(e^{iux}) \right| < 1 \) for all \( u \neq 0 \).

For the convenience of the reader, we provide a full proof of this characterization, since we were not able to find an adequate reference.

**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, \( \phi \) is of type non-lattice and with support \( \mathbb{R} \), and the inequality \( \left| \phi(e^{iux}) \right| < 1 \) 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 $\phi$ of type lattice. Then, since $\mathbb{Z}[\text{supp}(\Pi)] = \lambda \mathbb{Z}$ for some $\lambda > 0$, the semigroup $\mathbb{N}[\text{supp}(\Pi)] \subset \lambda \mathbb{Z}$ is discrete hence closed, and by the previous discussion,

$$\text{supp}(\phi) = \gamma + \mathbb{N}[\text{supp}(\Pi)] \subset \gamma + \lambda \mathbb{Z}.$$ 

Conversely, if $\text{supp}(\phi)$ is included in a shifted lattice $\gamma + \lambda \mathbb{Z}$, then the second case of the previous discussion is excluded, so $\gamma$ is well-defined and finite, and then

$$\text{supp}(\phi) = \gamma + \mathbb{N}[\text{supp}(\Pi)] \subset \gamma + \lambda \mathbb{Z}.$$ 

This forces $\mathbb{N}[\text{supp}(\Pi)] \subset \lambda \mathbb{Z}$, and therefore $\mathbb{Z}[\text{supp}(\Pi)] \subset \lambda \mathbb{Z}$, so the first assertion is indeed equivalent to the definition of lattice infinitely divisible distribution. In the following, if $\phi$ is of type lattice, we call period of $\phi$ the smallest quotient $p = \frac{2\pi}{\lambda}$ with $\lambda > 0$ such that $\text{supp}(\phi) \subset \gamma + \lambda \mathbb{Z}$ (hence, with $\lambda$ as big as possible). Notice then that

$$f(u) = \frac{\phi(e^{iux})}{e^{iur\gamma}} = \sum_{n \in \mathbb{Z}} \phi(\{\gamma + n\lambda\}) e^{iun\lambda}$$

is a $p$-periodic function with $f(0) = 1$. Suppose that $|f(u)| = 1$, which is equivalent to $|\phi(e^{iux})| = 1$. Since $\sum_{n \in \mathbb{Z}} \phi(\{\gamma + n\lambda\}) = 1$, this is only possible if $e^{iun\lambda} = 1$ for all $n\lambda \in (\text{Supp}(\phi) - \gamma)$, or if $e^{iun\lambda} = -1$ for all $n\lambda \in (\text{Supp}(\phi) - \gamma)$. The second case is impossible, since $\gamma$ belongs to $\text{Supp}(\phi)$, and therefore $0 \lambda \in (\text{Supp}(\phi) - \gamma)$. So, $e^{iun\lambda} = 1$ for all $n \in \frac{1}{\lambda}(\text{Supp}(\phi) - \gamma)$, and $2\pi$ divides $u\lambda$ for all $n \in \frac{1}{\lambda}(\text{Supp}(\phi) - \gamma)$.

To take $\lambda$ maximal amounts to impose that the greatest common divisor of these integers $n$ is $1$; so, $2\pi$ divides $u\lambda$, and $u \in \frac{2\pi}{\lambda} \mathbb{Z}$. The converse implication is true since $f$ is $p = \frac{2\pi}{\lambda}$ periodic. Hence, if $\phi$ is of type lattice with period $p = \frac{2\pi}{\lambda}$, then $|\phi(e^{iu})| = 1$ if and only if $u \in p\mathbb{Z}$.

It remains to see that if $\phi$ is not of type lattice, then $|\phi(e^{iux})| < 1$ for all $u \neq 0$. Fix $u > 0$, and suppose first that we are in the second case of the classification from [SH04] previously described. Notice then that $\text{supp}(\phi)$ contains an half-line, and therefore an interval $[a, a + \frac{2\pi}{u}]$ of length $\frac{2\pi}{u}$. Then, then

$$|\phi(e^{iux})| \leq \int_{\mathbb{R}\setminus[a, a + \frac{2\pi}{u}]} \phi(dx) + \left| \int_a^{a + \frac{2\pi}{u}} e^{iux} \phi(dx) + \int_{a + \frac{2\pi}{u}}^{a + \frac{2\pi}{u}} e^{iux} \phi(dx) \right|.$$ 

In the right hand-side, the two terms $\int_a^{a + \frac{2\pi}{u}} e^{iux} \phi(dx)$ and $\int_{a + \frac{2\pi}{u}}^{a + \frac{2\pi}{u}} e^{iux} \phi(dx)$ have their complex arguments belonging to two opposite half-circles $(ua, ua + \pi)$ and $(ua + \pi, ua + 2\pi)$.

Therefore, one has a strict triangular inequality

$$|\phi(e^{iux})| < \int_{\mathbb{R}\setminus[a, a + \frac{2\pi}{u}]} \phi(dx) + \int_a^{a + \frac{2\pi}{u}} e^{iux} \phi(dx) + \int_{a + \frac{2\pi}{u}}^{a + \frac{2\pi}{u}} e^{iux} \phi(dx) \leq \int_{\mathbb{R}} \phi(dx) = 1.$$
Thus, this case is treated. Suppose finally that we are in the first case of the classification, but with \( \phi \) that is not of type lattice. Then,

\[
\mathbb{R} = \mathbb{Z} \supset \mathbb{N} \supset \mathbb{Z} - \mathbb{N} \supset \mathbb{N} \supset \emptyset \supset \emptyset
\]

So, for every \( u > 0 \), there exists sequences \( (a_n)_{n \in \mathbb{N}} \) and \( (b_n)_{n \in \mathbb{N}} \) in \( \supp(\phi) \) such that

\[
\frac{\pi}{u} = \lim_{n \to \infty} b_n - a_n
\]

Fix \( \varepsilon > 0 \) and \( n \in \mathbb{N} \), such that \( \varepsilon u \leq \frac{\pi}{3} \) and \( \frac{2\pi}{3} \leq (b_n - a_n)u \leq \frac{4\pi}{3} \). The part of the Fourier transform \( \phi(e^{iu}) \) that corresponds to the disjoint intervals \([a_n - \varepsilon, a_n + \varepsilon]\) and \([b_n - \varepsilon, b_n + \varepsilon]\) is

\[
\int_{[a_n - \varepsilon, a_n + \varepsilon]} e^{iu} \phi(dx) + \int_{[b_n - \varepsilon, b_n + \varepsilon]} e^{iu} \phi(dx),
\]

with the first part that has its argument in

\[
((a_n - \varepsilon)u, (a_n + \varepsilon)u) \subset \left( a_n u - \frac{\pi}{3}, a_n u + \frac{\pi}{3} \right),
\]

and the second part that has its argument in

\[
((b_n - a_n)u + (a_n - \varepsilon)u, (b_n - a_n)u + (a_n + \varepsilon)u) \subset \left( a_n u + \frac{\pi}{3}, a_n u + \frac{5\pi}{3} \right).
\]

Therefore, one can as before write a strict triangular inequality, which ensures that \( |\phi(e^{iu})| < 1 \).

\[
\square
\]

In the remaining of this section, we place ourselves in the setting of Definition 1.1, and we suppose that the \( X_n \)'s and the (non-constant) infinitely divisible distribution \( \phi \) both take values in the lattice \( \mathbb{Z} \), and furthermore, that \( \phi \) has period \( 2\pi \) (in other words, the lattice \( \mathbb{Z} \) is minimal for \( \phi \)). In particular, for every \( u \in (0, 2\pi) \), \( |\exp(\eta(\pi))| < 1 \), 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 \( |\phi(e^{iu})| = 1 \). For precisions on (discrete) infinitely-divisible distributions, we refer to the aforementioned textbook [SH04], and also to [Kat67] and [Fel71, Chapter XVII].

3.2. Large deviations in the scale \( O(t_n) \).

**Lemma 3.2.** Let \( X \) be a \( \mathbb{Z} \)-valued random variable whose generating function \( \varphi_X(z) = \mathbb{E}[e^{zX}] \) converges absolutely in the strip \( \mathcal{S}_c \). For \( k \in \mathbb{Z} \),

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

\[
\forall h \in (0, c), \quad \mathbb{P}[X \geq k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-k(h+iu)}}{1 - e^{-(h+iu)}} \varphi_X(h + iu) \, du.
\]

**Proof.** Since

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

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

\[
\mathbb{P}[X \geq k] = \sum_{l=k}^{\infty} \mathbb{P}[X = l] = \sum_{l=k}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-l(h+iu)} \varphi_X(h + iu) \, du,
\]
and the sum of the moduli of the functions on the right-hand side is dominated by the integrable function \( \frac{e^{-th}}{1-e^{-h}} \varphi_X(h) \); so by Lebesgue’s dominated convergence theorem, one can exchange the integral and the summation symbol, which yields the second equation.

\[ \square \]

We now work under the assumptions of Definition 1.1, and furthermore, we assume that the convergence is at speed \( O((t_n)^{-\varphi}) \). Note that necessarily \( \eta(0) = 0 \) and \( \varphi(0) = 1 \). A simple computation gives also the following approximation formulas:

\[
\begin{align*}
\mathbb{E}(X_n) &= \varphi'_n(0) = t_n\eta'_n(0) + \psi'(0) + t_n\eta'_n(0) = t_n\eta'_n(0) + O(1); \\
\text{Var}(X_n) &= \varphi''_n(0) - \varphi'_n(0)^2 = t_n\eta''_n(0) + O(1).
\end{align*}
\]

**Theorem 3.3.** Let \( x \) be a real number in the interval \((\eta'(-c), \eta'(c))\), and \( h \) defined by the implicit equation \( \eta'(h) = x \). We assume \( t_n x \in \mathbb{N} \).

1. The following expansion holds:

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

for some numbers \( a_k \).

2. Similarly, if \( x \) is a real number in the range of \( \eta'(0, c) \), then

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

for some numbers \( b_k \).

Both \( a_k \) and \( b_k \) are rational fractions in the derivatives of \( \eta \) and \( \psi \) at \( h \), that can be computed explicitly — see Remark 3.6.

**Proof.** With the notations of Definition 1.1, the first equation of Lemma 3.2 becomes

\[
\begin{align*}
\mathbb{P}[X_n = t_n x] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-t_n x (h+iu)} \varphi_n(h+iu) \, du \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-t_n xh} e^{t_n(\eta(h+iu) - iux)} \varphi_n(h+iu) \, du \\
&= \frac{e^{-t_n F(x)}}{2\pi} \int_{-\pi}^{\pi} e^{t_n(\eta(h+iu) - \eta(h) - iux)} \varphi_n(h+iu) \, du.
\end{align*}
\]

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

\[
\exp(\eta(h+iu) - \eta(h)) = \frac{\mathbb{E}[e^{(h+iu)X}]}{\mathbb{E}[e^{hX}]} = \mathbb{E}_Q[e^{iuX}]
\]
is the characteristic function of $X$ under the new probability $dQ(\omega) = \frac{e^{\phi X(\omega)}}{E[e^{\phi X}]} dP(\omega)$ (and $X$ has minimum lattice $\mathbb{Z}$). Note that Proposition 3.11 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

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

for $n$ big enough, since $\psi_n$ converges uniformly towards $\psi$ on the compact set $K = h + i[-\pi, \pi]$. Since $q_\delta < 1$, for any $\delta > 0$ fixed, $I_{(-\delta,\delta)^c} e^{t_nF(x)}$ 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 $\psi_n$ by $\psi$ up to a $(1 + O((t_n)^{-v}))$, since the integral is taken on a compact subset of $S_c$. We then set $u = \frac{w}{\sqrt{t_n\eta''(h)}}$:

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

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 $\psi$ 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 $\alpha_k$ 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 $\psi$ and $\eta$ at point $h$. So,

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

For any power $w^m$,

$$\left|\int_{-\infty}^{\infty} w^m e^{-\frac{w^2}{2}} dw - \int_{-\delta}^{\delta} w^m e^{-\frac{w^2}{2}} dw\right|$$

is smaller than any negative power of $t_n$ as $n$ goes to infinity (see Lemma 2.7, (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 $m, h, \delta$ 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

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

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

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

where $(2m - 1)!!$ is the $2m$-th moment of the Gaussian distribution (cf. Lemma 2.7, (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

$$\frac{1}{1 - \exp(-h - iu)} = \frac{1}{1 - \exp(-h)} \left(\frac{1 - \exp(-h)}{1 - \exp(-h - \frac{iu}{\sqrt{\eta''(h)}})}\right)$$

in the integrals. \hfill \Box

**Remark 3.4.** 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 $\psi$ measures the difference between the distribution of $X_n$ and the distribution of $Y_{t_n}$ in the interval $(t_n \eta'(0), t_n \eta'(c))$.

**Remark 3.5.** 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 $t_n$,

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

i.e., the expansions of both sides up to any given power $O\left(\frac{1}{(t_n)^\nu}\right)$ agree.
Remark 3.6. 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

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

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

the last expansion holding in a neighborhood of zero. The coefficient \( a_{2k}(w) \) is an even polynomial in \( w \) with valuation \( 2k \) and coefficients which are polynomials in the derivatives of \( \psi \) and \( \eta \) at \( h \), and in \( \frac{1}{\eta''(h)} \). Then,

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

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

\[
g_n(w) = \frac{1 - \exp(-h)}{1 - \exp(h - \frac{iw}{\sqrt{t_n \eta''(h)}})} \cdot f_n(w).
\]

Remark 3.7. The infinite divisibility of the reference law \( \phi \) is only used for the existence of bounds \( q_k \) for ratios of values of the moment generating function. Therefore, one can easily adapt the proof to the case where such bounds exist. Consider for instance a sum \( S_n \) of \( n \) i.i.d. Bernoulli variables with \( \mathbb{P}[B = 1] = \mathbb{P}[B = 0] = \frac{1}{2} \). The moment generating function of \( B \) is \( \frac{1+e^x}{2} \), so

\[
\mathbb{E}[e^{xS_n}] = \left( \frac{1+e^x}{2} \right)^n,
\]

and informally speaking one has mod-\( B \) convergence with parameter \( n \) and limiting function \( \psi(z) = 1 \). Fix \( x \in (0,1) \), and \( h = \log \left( \frac{1-x}{x} \right) \). One has indeed

\[
\left| \frac{1+e^{h+iu}}{1+e^h} \right| < 1
\]

for all \( u \in (-\pi, \pi) \setminus \{0\} \), with uniform bound on complementary sets of neighborhoods of 0. Therefore, the proof adapts readily and if \( nx \in \mathbb{N} \), then

\[
\mathbb{P}[S_n = nx] = \frac{e^{-n(\log 2 + x \log x + (1-x) \log(1-x))}}{\sqrt{2\pi nx(1-x)}} (1 + o(1)),
\]

as can be seen more directly from Stirling’s estimates of factorials, or as a particular case of Bahadur-Rao’s estimates ([BR60]). We refer to Remark 4.9 for a more general statement, which enables one to recover the Bahadur-Rao theorem for general sums of i.i.d. random variables (discrete or continuous).
Example 3.8. 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 3.3 reads then as follows:

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

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 2.4 shows that for \(x > 0\) such that \(x \log n \in \mathbb{N}\),

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

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

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

As the speed of convergence is very good in this case, precise expansions in \(1/\log n\) to any order could be also 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 \(X_n\) 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 4.

| Order of Fluctuations | Large Deviations \((\eta'(0) < x)\) | Extended Central Limit Theorem \((t_n)^{1/6} \lesssim y \ll (t_n)^{1/2}\) | Central Limit Theorem \(y \ll (t_n)^{1/6}\) |
|------------------------|-------------------------------------|-----------------------------------------------|-----------------------------------------------|
| \(O(t_n)\)            | \(P \left[ \frac{X_n - t_n F(x)}{\sqrt{t_n \eta'(x)}} \right] \geq x \) \simeq \frac{\exp\left(-t_n F(x)\right)}{\sqrt{2\pi t_n \eta'(x)}} \frac{1}{1 - e^{-F(x)}} \psi(F'(x))\); | \(P \left[ \frac{X_n - t_n \eta'(0) \sqrt{t_n \eta''(0)}}{\eta''(0)} \right] \geq y \) \simeq \frac{\exp\left(-t_n F(x)\right)}{F'(x) \sqrt{2\pi t_n \eta'(x)}}\); | \(P \left[ \frac{X_n - t_n \eta'(0) \sqrt{t_n \eta''(0)}}{\eta''(0)} \right] \geq y \) \simeq \(P[\mathcal{N}(0,1) \geq y]\). |
| \(O((t_n)^{2/3})\)    | \(P \left[ \frac{X_n - t_n \eta'(0) \sqrt{t_n \eta''(0)}}{\eta''(0)} \right] \geq y \) \simeq \(P[\mathcal{N}(0,1) \geq y]\). |
| \(O((t_n)^{1/2})\)    |                                                                                     |                                                                                          |                                                                                          |

Figure 4. Panorama of the fluctuations of a sequence of random variables \((X_n)_{n \in \mathbb{N}}\) that converges modulo a lattice distribution (with \(x = \eta'(0) + \sqrt{\eta''(0)/t_n} y\)).
Theorem 3.9. We consider as in Theorem 3.3 a sequence \((X_n \in \mathbb{N})\) that converges mod-\(\phi\), with a reference infinitely divisible law \(\phi\) that is a lattice distribution. Assume \(y = o((t_n)^{1/6})\). Then,
\[
\mathbb{P}\left[X_n \geq t_n\eta'(0) + \sqrt{t_n\eta''(0)} y\right] = \mathbb{P}[\mathcal{N}(0,1) \geq y] (1 + o(1)).
\]
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
\[
\mathbb{P}\left[X_n \geq t_n\eta'(0) + \sqrt{t_n\eta''(0)} y\right] = e^{-t_n F(x)} h \sqrt{2\pi t_n \eta''(h)} (1 + o(1)).
\]

Remark 3.10. The case \(y = O(1)\), which is the classical central limit theorem, follows immediately from the assumptions of Definition 1.1, since by a Taylor expansion around 0 of \(\eta\) 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 \(e^{-\xi^2/2}\), 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 \(\phi\). Consider a non-constant infinitely divisible law \(\phi\), which has as in Definition 1.1 a convergent moment generating function \(\int e^{zx} \varphi(dx) = e^{\eta(z)}\) on a strip \(S_c\) with \(c > 0\). We assume that \(\phi\) is:

- either of type lattice, in which case one can assume without loss of generality that \(Z\) is the minimal lattice;
- or, of type non-lattice and absolutely continuous w.r.t. Lebesgue measure. This is a strictly stronger hypothesis than being non-lattice, but one has a sufficient condition on the Lévy-Khintchine representation of \(\phi\) for this to happen (cf. [SH04, Chapter 4, Theorem 4.23]): it is the case if \(\sigma^2 > 0\), or if \(\sigma^2 = 0\) and if the absolutely continuous part of the Lévy measure \(\Pi\) has infinite mass.

Lemma 3.11. Under the previous assumptions, there exists a constant \(D > 0\) only depending on \(\phi\), and an interval \((-\varepsilon, \varepsilon) \subset (-c, c)\), such that for all \(h \in (-\varepsilon, \varepsilon)\) and all \(\delta\) small enough,
\[
q_\delta = \max\left\{\max_{u \in [-\pi, \pi]\setminus(-\delta, \delta)} |\exp(\eta(h + iu) - \eta(h))|, \max_{u \in \mathbb{R}\setminus(-\delta, \delta)} |\exp(\eta(h + iu) - \eta(h))|\right\}
\]
is smaller than \(1 - D \delta^2\).

Remark 3.12. The difficult part of Lemma 3.11 is the non-lattice case. The hypotheses on the infinitely divisible law \(\phi\) 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 $-\frac{\sigma^2}{2}$. As a consequence,

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

This expression can be expanded in series of $u$ as

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

Therefore, in the non-lattice case, Lemma 3.11 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 neighborhood of zero. Unfortunately, for general probability measures, Riemann-Lebesgue lemma does not apply, and even for $h = 0$, it is unclear whether for a general exponent $\eta$ one has

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

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 oneselfs to absolutely continuous measures in the non-lattice case, in order to use Riemann-Lebesgue lemma. In the following we provide ad hoc proofs of Lemma 3.11 in the lattice and absolutely continuous cases, that do not rely on the Kolmogorov representation.

**Proof of Lemma 3.11.** Denote $X$ a random variable under the infinitely divisible distribution $\phi$. Suppose first that $\phi$ is of type lattice. We then claim that there exists 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 $\Pi$ is the Lévy measure of $\phi$, then

$$Z = Z[\supp(\Pi)] - N[\supp(\Pi)],$$

so there exists $a$ and $b$ in $N[\supp(\Pi)]$ such that $b - a = 1$. However, $\supp(\phi) = \gamma + N[\supp(\Pi)]$ for some $\gamma \in Z$, 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 $dQ = \frac{e^{hx}}{E[e^{hx}]} d\mathbb{P}$. So, for any $u$,

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

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

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 (\( \mathcal{Q} \) tends to \( \mathbb{P} \) for \( h \to 0 \)). As \( \cos u \leq 1 - \frac{u^2}{2} \) for all \( u \in (-\pi, \pi) \),
\[
|\exp(\eta(h + iu) - \eta(h))|^2 \leq 1 + 15D (\cos u - 1) \leq 1 - 3D u^2;
\]
\[
q_\delta \leq \sqrt{1 - 3D \delta^2} \leq 1 - D \delta^2 \text{ for } \delta \text{ small enough.}
\]
This concludes the proof in the lattice case.

Suppose now that \( \phi \) is of type non-lattice, and absolutely continuous with respect to Lebesgue measure, with density \( f \). We shall adapt the previous arguments from the discrete to the continuous case. Though the density \( f \) cannot be supported on a compact segment (cf. the previously described classification of the possible supports of an infinitely divisible law), one can work as if it were the case, thanks to the following calculation:

\[
|\exp(\eta(h + iu) - \eta(h))| = \left| \frac{\phi(e^{(h+iu)x})}{\phi(e^{hx})} \right| = \frac{|\phi_{<a}(e^{(h+iu)x}) + \int_a^b e^{(h+iu)x} f(x) \, dx + \phi_{>b}(e^{(h+iu)x})|}{\phi_{<a}(e^{hx}) + \int_a^b e^{hx} f(x) \, dx + \phi_{>b}(e^{hx})} \leq \frac{\phi_{<a}(e^{hx}) + \phi_{>b}(e^{hx}) + \int_a^b e^{hx} f(x) \, dx}{\phi_{<a}(e^{hx}) + \phi_{>b}(e^{hx}) + \int_a^b e^{hx} f(x) \, dx}.
\]

Therefore, it suffices to show:

\[
\max_{u \in (-\delta, \delta)} \frac{\int_a^b e^{(h+iu)x} f(x) \, dx}{\int_a^b e^{hx} f(x) \, dx} \leq 1 - D \delta^2
\]
for \( \delta \) and \( h \) small enough. This reduction to a compact support will be convenient later in the computations.

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

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

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

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

The problem is to show that

\[
\sup_{h \in (-\epsilon, \epsilon)} \sup_{u \in (-\delta, \delta)} F(h, u) \leq 1 - D \delta^2
\]
for some constant \( D \). With \( h \) fixed, by 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 \( \frac{2\pi}{|a|} \mathbb{Z} \), which is impossible for a measure continuous with respect to the Lebesgue measure. Combining these two observations, one sees that for any \( \delta > 0 \),

\[
\sup_{u \in (-\delta, \delta)} F(h, u) \leq C(h, \delta) < 1
\]

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 \( C_\delta \):

\[
\forall \delta > 0, \exists C_\delta < 1 \text{ such that } \sup_{h \in (-e, e)} \sup_{u \in (-\delta, \delta)} F(h, u) \leq C_\delta. \tag{6}
\]

On the other hand, notice that

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

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

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

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

\[
\exists B > 0 \text{ such that for all } u \leq \frac{\pi}{2(b-a)}, \sup_{h \in (-e, e)} F(h, u) \leq 1 - B u^2. \tag{7}
\]

The two assertions (6) (with \( \delta = \frac{\pi}{2(b-a)} \)) and (7) enables one to conclude, with

\[
D = \inf \left( B, \frac{1 - C\delta}{\delta^2} \right), \text{ where } \delta = \frac{\pi}{2(b-a)}. \tag{8}
\]

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 bigger in absolute value than \( 1 - \delta^2 \).

**Proof of Theorem 3.9.** Notice that \( \eta''(0) \neq 0 \) since this is the variance of the law \( \phi \), assumed to be non-trivial. Set \( x = \eta'(0) + s \), and assume \( s = o(1) \). The analogue of Equation (4) reads in our setting

\[
\mathbb{P}[X_n \geq t_n x] = \frac{e^{-t_n F(x)}}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\eta(h+iu) - \eta(h) - i\eta'(h)}}{1 - e^{-h - iu}} \psi_n(h + iu) \, du. \tag{8}
\]

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 3.3 shows that the integral over \((-\delta, \delta)^c\) is bounded by \( C \delta (q_\delta)^n \), where \( q_\delta < 1 \), and \( C \delta \) (with \( C \) a constant independent from \( s \) and \( \delta \)) 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 $\delta$ 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 $I_{(-\delta, \delta)} e^{t_n F(x)}$ 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 3.11. The other part of (8) is

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

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

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

by continuity of $\psi$, so one can replace the term with $\psi$ 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 $e^{t_n \Delta_n(w)}$ by the constant 1.

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

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

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

Indeed, setting $\beta = h \sqrt{t_n \eta''(h)}$, this leads directly to the computation done in Lemma 2.7, (4).

Hence, we have shown so far:

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

(9)

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 2.7, (3) to replace in Equation (9) the function of \( \beta \) by the tail-estimate of the Gaussian:

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

which is the content of the second part of our 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 \( h = 0 \))

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

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

- On the other hand,

\[
\beta = h \sqrt{t_n \eta''(h)} = \frac{s}{\eta''(0)} (1 + O(s)) \sqrt{t_n (\eta''(0) + O(s))}
\]

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 \( F_y(\lambda) = \mathbb{P}[N(0,1) \geq \lambda y] \). Then, for \( |\lambda| \) say between \( \frac{1}{2} \) and 2,

\[
|F_y'(\lambda)| = \left| \frac{y}{\sqrt{2\pi}} e^{-\frac{\lambda^2 y^2}{2}} \right| \leq \max_{y \in \mathbb{R}} \left| \frac{y}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \right| = C < +\infty;
\]

\[
|\mathbb{P}[N(0,1) \geq \beta] - \mathbb{P}[N(0,1) \geq y]| = \left| F_y(1 + o((t_n)^{-1/3})) - F_y(1) \right| \leq \frac{C}{(t_n)^{1/3}} = o(1).
\]

This ends the proof of Theorem 3.9.

\[\square\]

**Remark 3.13.** Equation (9) is the probabilistic counterpart of the number-theoretic results of [Kub72, Rad09], see in particular Theorems 2.1 and 2.2 in [Rad09]. In §8.2, 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.

To summarize, in the lattice case, mod-\( \phi \) convergence implies a large deviation principle (Theorem 3.3) and a precised central limit theorem (Theorem 3.9), and these two results cover a whole interval of possible scalings for the fluctuations of the sequence \((X_n)_{n \in \mathbb{N}}\). This phenomenon is also going to be true for non-lattice reference distributions, and we shall say more about it in Section 5.
4. FLUCTUATIONS IN THE NON-LATTICE CASE

In this section we prove the analogues of Theorems 3.3 and 3.9 when $\phi$ is not lattice-distributed; hence, by Proposition 3.1, $|e^{iu\eta}| < 1$ for any $u \neq 0$. In this setting, assuming $\phi$ absolutely continuous w.r.t. the Lebesgue measure, there is a formula equivalent to the one given in Lemma 3.2, namely,

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

(10)

if $\phi_X(h) = \mathbb{E}[e^{hX}] < +\infty$ for $h > 0$ (see [Fel71, Chapter XV, §3]). However, in order to manipulate this formula as in Section 3, one would need strong additional assumptions of integrability on the characteristic functions of the random variables $X_n$. Thus, instead of Equation (10), our main tool will be a Berry-Esseen estimate (see Proposition 4.1 hereafter), which we shall then combine with techniques of tilting of measures (Lemma 4.7) 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-$\phi$ convergence.

**Proposition 4.1** (Berry-Esseen expansion). Denote $g(y) = (2\pi)^{-1/2} e^{-y^2/2}$ the density of a standard Gaussian variable, and $F_n(x) = \mathbb{P}[X_n \leq t_n\eta'(0) + \sqrt{t_n\eta''(0)} x]$. Under the assumptions of Definition 1.1, with $\phi$ non-lattice,

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

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. 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 $T > 0$,

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

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
\[ F_n(x) = \text{cumulative distribution function of } Y_n; \]
\[ G_n(x) = \int_{-\infty}^{x} \left( 1 + \frac{\psi'(0) y}{\sqrt{t_n \eta''(0)}} + \frac{\eta'''(0)}{6 \sqrt{t_n \eta''(0)}} (y^3 - 3y) \right) g(y) \, dy. \]

Note that each \( G_n \) is clearly a Lipschitz function (with uniform Lipschitz constant, i.e. that do not depend on \( n \)). Besides, by Lemma 2.7, (2), the Fourier transform corresponding to the distribution function \( G_n \) is, setting \( z = i \zeta \),
\[ g_n^*(\zeta) = e^{\frac{\zeta^2}{2}} \left( 1 + \frac{\psi'(0) z}{\sqrt{t_n \eta''(0)}} + \frac{\eta'''(0) z^3}{6 \sqrt{t_n \eta''(0)}} \right). \] (11)

Consider now \( f_n^*(\zeta) \): if \( z = i \zeta \), then
\[ f_n^*(\zeta) = \mathbb{E} \left[ e^{\left( \frac{X_n - t_n \eta'(0)}{\sqrt{t_n \eta''(0)}} \right)} \right] = \exp \left( -z \sqrt{\frac{t_n}{\eta''(0)}} \eta'(0) \right) \times \varphi_n \left( \frac{z}{\sqrt{t_n \eta''(0)}} \right) \]
\[ = \exp \left( t_n \left( \eta \left( \frac{z}{\sqrt{t_n \eta''(0)}} \right) - \frac{\eta'(0) z}{\sqrt{t_n \eta''(0)}} \right) \right) \times \varphi_n \left( \frac{z}{\sqrt{t_n \eta''(0)}} \right) \]
But
\[ \varphi_n \left( \frac{z}{\sqrt{t_n \eta''(0)}} \right) = \left( 1 + \frac{\psi'(0) z}{\sqrt{t_n \eta''(0)}} + o \left( \frac{|z|}{\sqrt{t_n}} \right) \right) \left( 1 + \frac{\psi'(0) z}{\sqrt{t_n \eta''(0)}} + o \left( \frac{|z|}{\sqrt{t_n}} \right) \right) \]
where the \( o \) is uniform in \( n \) because of the local uniform convergence of the analytic functions \( \varphi_n \) to \( \psi \) (and hence, of \( \psi'_n \) and \( \psi''_n \) to \( \psi' \) and \( \psi \)). Thus
\[ f_n^*(\zeta) = \exp \left( \frac{z^2}{2} + \eta'''(0) \frac{z^3}{6 \sqrt{t_n \eta''(0)}} + |z|^2 o \left( \frac{z}{\sqrt{t_n}} \right) \right) \times \left( 1 + \frac{\psi'(0) z}{\sqrt{t_n \eta''(0)}} + o \left( \frac{|z|}{\sqrt{t_n}} \right) \right) \]
\[ = e^{\frac{\zeta^2}{2}} \left( 1 + \frac{\psi'(0) z}{\sqrt{t_n \eta''(0)}} + \frac{\eta'''(0) z^3}{6 \sqrt{t_n \eta''(0)}} + (1 + |z|^2) o \left( \frac{z}{\sqrt{t_n}} \right) \right). \] (12)

Beware that in the previous expansions, the \( o(\cdot) \) is
\[ o \left( \frac{z}{\sqrt{t_n}} \right) = \frac{|z|}{\sqrt{t_n}} \varepsilon \left( \frac{z}{\sqrt{t_n}} \right) \text{ with } \lim_{t \to 0} \varepsilon(t) = 0. \]

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 (11) and (12) and using Feller’s lemma, we get:
\[ |F_n(x) - G_n(x)| \leq \frac{1}{\pi} \int_{-\Delta \sqrt{t_n}}^{\Delta \sqrt{t_n}} \left| f_n^*(\zeta) - g_n^*(\zeta) \right| d\zeta + \frac{24m}{\Delta \pi \sqrt{t_n}} \]
\[ \leq \frac{1}{\pi \sqrt{t_n}} \int_{-\delta \sqrt{t_n}}^{\delta \sqrt{t_n}} e^{-\frac{\zeta^2}{2}} \left( 1 + |\zeta|^2 \right) \varepsilon \left( \frac{\zeta}{\sqrt{t_n}} \right) d\zeta + \frac{24m}{\Delta \pi \sqrt{t_n}} \]
\[ + \frac{1}{\pi \delta \sqrt{t_n}} \int_{[-\Delta \sqrt{t_n}, \Delta \sqrt{t_n}] \setminus [-\delta \sqrt{t_n}, \delta \sqrt{t_n}]} |f_n^*(\zeta) - g_n^*(\zeta)| d\zeta. \] (13)
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 \( |\zeta| \in \left[ \delta \sqrt{t_n}, \Delta \sqrt{t_n} \right] \),

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

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

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

but the maximum is a constant \( q_{\delta,\Delta} \) strictly smaller than 1, because \( \eta \) is not lattice distributed. This implies that on the 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 (11) shows that the same kind of bound holds for \( |g_n^*(\zeta)| \). We shall use the notations \( \tilde{K}(\Delta) \) and \( \bar{q}_{\delta,\Delta} \) for constants valid for both \( |f_n^*(\zeta)| \) and \( |g_n^*(\zeta)| \). Thus the third summand in the bound (13) is bounded by

\[
\frac{4\Delta}{\pi \delta} \tilde{K}(\Delta) (\bar{q}_{\delta,\Delta}) t_n.
\]

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

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

for \( t_n \) big enough. This ends the proof of the proposition. \( \square \)

**Remark 4.2.** Proposition 4.1 gives an approximation for the Kolmogorov distance between the law \( \mu_n \) and the normal law. Indeed, assume to simplify that the reference law \( \phi \) is the Gaussian law. Then, \( \eta''(0) = 1 \) and \( \eta'''(0) = 0 \), and one computes

\[
d_{\text{Kolmogorov}}(\mu_n, \mathcal{N}(0, 1)) = \frac{1}{\sqrt{t_n}} \sup_{x \in \mathbb{R}} \int_{-\infty}^{x} \psi'(0) y g(y) \, dy + o \left( \frac{1}{\sqrt{t_n}} \right)
\]

\[
= \frac{|\psi'(0)|}{\sqrt{2\pi t_n}} + o \left( \frac{1}{\sqrt{t_n}} \right).
\]

This makes explicit the bound given by Theorem 1 in [Hwa98]. However, if \( \psi'(0) = 0 \), then the estimate \( d_{\text{Kol}} = o(1/\sqrt{t_n}) \) is not at all 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

\[
d_{\text{Kol}} \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i, \mathcal{N}(0, 1) \right) = o \left( \frac{1}{n^{1/6}} \right),
\]

which is not as good as the classical Berry-Esseen estimate \( O \left( \frac{1}{n^{1/2}} \right) \). There is a way to modify the arguments in order to get such optimal estimates, under an assumption
over the zone of mod-convergence. However, in this setting, it is more convenient to relate the Kolmogorov distance to calculations with test functions (this is an alternative to Feller’s lemma). We refer to [FMN14], where such "optimal" computations of Kolmogorov distances will be performed.

4.2. Large deviations in the scale $O(t_n)$.

**Theorem 4.3.** Suppose $\varphi$ non-lattice. If $x \in (\eta'(0), \eta'(c))$, 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 4.4.** 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 4.5.** Theorem 4.3 should be compared with [Hwa96, Theorem 1], which studies another regime of large deviations in the mod-$\varphi$ setting, namely, when $h$ goes to zero (or equivalently, $x \to \eta'(0)$). We shall also look at this regime in our Theorem 4.8.

**Remark 4.6.** The main difference between Theorems 3.3 and 4.3 is the replacement of the factor $\psi(h)/(1 - e^{-h})$ by $\psi(h)/h$; the same happens with Bahadur-Rao’s estimates when going from lattice distributions to non-lattice distributions.

**Lemma 4.7.** Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of random variables that converges mod-$\varphi$ with parameters $(t_n)_{n \in \mathbb{N}}$ and limiting function $\psi$. For $h \in (0, c)$, we make the exponential change of measure

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

and denote $\tilde{X}_n$ a random variable following this law. The sequence $(\tilde{X}_n)_{n \in \mathbb{N}}$ converges mod-$\varphi'$, where $\varphi'$ 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

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

**Proof.** Obvious since $\varphi_{\tilde{X}_n}(z) = \varphi_{X_n}(z + h)/\varphi_{X_n}(h)$. 

**Proof of Theorem 4.3.** Consider the sequence $(\tilde{X}_n)_{n \in \mathbb{N}}$ of Lemma 4.7. All the assumptions of Proposition 4.1 are satisfied, so, the distribution function $F_n(u)$ of $\frac{\tilde{X}_n - t_n \eta'(h)}{\sqrt{t_n \eta''(h)}}$ is

$$G_n(u) = \int_{-\infty}^{u} \left(1 + \frac{\psi'(h)}{\psi(h) \sqrt{t_n \eta''(h)}} \frac{\eta'''(h)}{\sqrt{t_n(\eta''(h))^3}} (y^3 - 3y)\right) g(y) dy$$
up to a uniform $o(1/\sqrt{t_n})$. Then,

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

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

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

To compute the integral $I$, we choose the primitive of $dF_n(u)$ that vanishes at $u = 0$, and we make an integration by parts:

$$I = \left[ e^{-h \sqrt{t_n \eta''(h)} u} F_n(u) \right]_{u=0}^{t_n \eta'(h) + \sqrt{t_n \eta''(h)}} + h \sqrt{t_n \eta''(h)} \int_{u=0}^{t_n \eta'(h) + \sqrt{t_n \eta''(h)}} e^{-h \sqrt{t_n \eta''(h)} u} (G_n(u) - G_n(0) + o\left(\frac{1}{\sqrt{t_n}}\right)) du$$

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

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

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

where on the three last lines the symbol $\simeq$ means that the remainder is a $o((t_n)^{-1/2})$. By Lemma 2.7, (3), the only contribution in the integral that is not a $o((t_n)^{-1/2})$ is

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

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

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

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

$$\mathbb{P}\left[X_n \geq t_n \eta'(0) + \sqrt{t_n \eta''(0)} y\right] = \frac{e^{-t_n F(x)}}{h \sqrt{2\pi t_n \eta''(h)}} \left(1 + o(1)\right).$$
Proof. The proof is exactly the same as in the lattice case (Theorem 3.9). Indeed, the conclusions of the technical Lemma 3.11 hold, and on the other hand, the equivalents for $P[X_n \geq t_n x]$ in the lattice and non-lattice cases (Theorems 3.3 and 4.3) differ only by the fact that $1 - e^{-h}$ is replaced by $h$. But in the proof of Theorem 3.9, the quantity $1 - e^{-h}$ is approximated by $h$, so everything works the same as in the non-lattice case. □

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 5.

| Order of Fluctuations | Large Deviations ($\eta'(0) < x$): $P[X_n/t_n \geq x] \simeq \frac{\exp(-t_n F(x))}{F'(x) \sqrt{2\pi n \eta'(x)}} \psi(F'(x))$; | 
| --- | --- | 
| $O(t_n)$ | Extended Central Limit Theorem ($(t_n)^{1/6} \lesssim y \ll (t_n)^{1/2}$): $P\left[ X_n - t_n \eta'(0) \sqrt{t_n \eta''(0)} \geq y \right] \simeq \frac{\exp(-t_n F(x))}{F'(x) \sqrt{2\pi n \eta'(x)}}$; | 
| $O((t_n)^{2/3})$ | Central Limit Theorem ($y \ll (t_n)^{1/6}$): $P\left[ \frac{X_n - t_n \eta'(0)}{\sqrt{t_n \eta''(0)}} \geq y \right] \simeq P[N(0,1) \geq y]$. |

Figure 5. Panorama of the fluctuations of a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ that converges modulo an absolutely continuous distribution (with $x = \eta'(0) + \sqrt{\eta''(0)/t_n y}$).

Remark 4.9. By using the arguments of Lemma 3.11, one can get in a quite general setting a control on ratios of moment generating functions, and therefore extend Theorems 3.3, 3.9, 4.3 and 4.8 to the case of a non-infinitely divisible reference law. In particular, consider a sum $S_n = Y_1 + \cdots + Y_n$ of i.i.d. random variables admitting an analytic moment generating function. We assume that:

- either $Y$ takes its values in $\mathbb{Z}$, and there are two consecutive integers $n, m = n - 1$ such that $P[Y = n] \neq 0$ and $P[Y = m] \neq 0$;
- or, $Y$ has a component absolutely continuous w.r.t. Lebesgue measure.

Then, the proofs and conclusions of Lemma 3.11 hold, so one has a good control on the modulus of the moment generating function of $Y$ and one can apply Theorems 3.3 and 3.9 or Theorems 4.3 and 4.8 to the mod-$Y$ convergence of $S_n$ with parameters $n$ and limiting function $\psi = 1$. This leads in particular to the Bahadur-Rao estimates: for $x > \mathbb{E}[Y]$,

$$P[S_n \geq n x] \simeq \begin{cases} \frac{\exp(-n F(x))}{(1-e^{-h})\sqrt{2\pi n \eta''(h)}} & \text{in the first case (with } nx \in \mathbb{Z}) \\ \frac{\exp(-n F(x))}{h \sqrt{2\pi n \eta''(h)}} & \text{in the second case} \end{cases}$$
where \( \eta(h) = \log \mathbb{E}[e^{hY}] \) and \( F \) is the Legendre-Fenchel transform of \( \eta \); see Theorem 3.7.4 in [DZ98], and also the papers [BR60, Ney83, Ilt95]. It should be noticed that the assumption on discrete random variables with two consecutive integers charging the measure is necessary. For instance, if one considers a sum \( S_n \) of \( n \) independent Bernoulli random variables with \( \mathbb{P}[B = 1] = \mathbb{P}[B = -1] = 0 \), 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 \( z = i\pi \).

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 \( X_n \) converging mod-\( \phi \) with limiting distribution \( \psi \), with parameters \( t_n \) such that \( t_n \gg (\log n)^{3/2} \). We also assume that the random variables \( X_n \) are defined on the same probability space, and we look for sequences \( \gamma_n \) such that almost surely,

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

The main difference between the following proposition and the usual law of the iterated logarithm is that one does 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]. We obtain a less precise result (we only obtain an upper bound), but which does not depend at all on the way one realizes the random variables \( X_n \). In other words, for every possible coupling of the variables \( X_n \), the following holds:

**Proposition 4.10.** Under the previous assumptions,

\[
\limsup_{n \to \infty} \frac{X_n - t_n \eta'(0)}{\gamma_n} \leq 1 \quad \text{almost surely.}
\]

**Proof.** Notice the term \( \log n \) instead of \( \log \log n \) for the usual law of iterated logarithm. One uses of course Borel-Cantelli lemma, and computes

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

Set \( y = \sqrt{2(1 + \epsilon) \log n} \). Due to the hypotheses on \( t_n \), one has \( y = o\left((t_n)^{1/6}\right) \) and one can apply Theorem 4.8: using the classical equivalent

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

we get

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

for \( n \) big enough. For any \( \epsilon > 0 \), this is summable, so almost surely \( X_n - t_n \eta'(0) < \sqrt{2(1 + \epsilon) \eta''(0) t_n \log n} \). Since this is true for every \( \epsilon \), one has (almost surely):

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

\( \square \)
5. Panorama of the Fluctuations and the Method of Cumulants

In Sections 6 and 7, we shall precise the topology underlying Theorems 3.3 and 4.3, and describe a multi-dimensional generalization of Theorem 4.3 which relies on such topological analysis. Before doing that, we want:

- to compare the results previously shown to classical results on the fluctuations of sequences of random variables;
- and to give a concrete method that allows one to prove the asymptotics of Definition 1.1, and then to obtain the estimates given by Theorems 3.3 and 4.3.

Most of what we are going to say in this Section holds in the restricted setting of mod-Gaussian convergence; however, many interesting examples fall in this framework, up to the point that it can be considered as a universality class for second order fluctuations.

5.1. The residue $\psi$ as a measure of breaking of symmetry. To start our discussion, consider the simple case of a sum $S_n = Y_1 + \cdots + Y_n$ of i.i.d. random variables (with convergent moment generating function), for instance non-lattice distributed, and with $E[Y] = 0$ and $E[Y^2] = 1$. There are classically three regimes of fluctuations of $S_n$ that can be studied:

- central limit theorem: at scale $\sqrt{n}$, for any $x$, $P[S_n \geq \sqrt{n} x]$ is given by the normal law
  \[ \left( \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy \right) (1 + o(1)). \]

  Notice that the tail of the Gaussian law is given by $\frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-\frac{y^2}{2}} dy \simeq \frac{e^{-\frac{x^2}{2}}}{a\sqrt{2\pi}}$ (cf. Lemma 2.7, (3)), but one cannot use it directly in the CLT, which only holds a priori for $x$ fixed.

- moderate deviations: by Ellis-Gärtner theory, at scale $\sqrt{n} \ll s_n \ll n$, if $x \geq 0$, then
  \[ \log P[S_n \geq s_n x] = -\frac{(s_n x)^2}{2n} (1 + o(1)), \]
  see [DZ98, Theorem 3.7.1].

- large deviations: at scale $n$, if $x > 0$, then $P[S_n \geq n x]$ is given by Bahadur-Rao estimates
  \[ \frac{e^{-nF(x)}}{h \sqrt{2\pi n \eta''(h)}} (1 + o(1)). \]

This panorama leads to the following intuition: there should be a particular scale $s_n$ between $\sqrt{n}$ and $n$ such that:

- before $s_n$, the estimate $P[S_n \geq s_n x] = P[N(0,n) \geq s_n x] (1 + o(1))$ still holds (extended central limit theorem).
- at scale $s_n$ and after, the probability $P[S_n \geq s_n x]$ “transforms” into the large deviation estimate, so in particular, $P[S_n \geq s_n x] = P[N(0,n) \geq s_n x] (1 + o(1))$ is no more true.
The identification of this scale, and the transformation of the form of the probability of tails, is exactly the outcome of our results. Indeed, in the case of sums of centered i.i.d., suppose $E[Y^3] = L \neq 0$. Then, a Taylor expansion of $E[e^{zS_n/n^{1/3}}]$ shows that

$$n^{-1/3} S_n \text{ converges mod-Gaussian with } \left\{ \begin{array}{l} \text{parameters } n^{1/3}, \\ \text{limiting function } \psi(z) = \exp \left( \frac{Lz^3}{6} \right). \end{array} \right.$$}

Therefore, for $x > 0$,

$$\Pr[S_n \geq n^{2/3}x] = \Pr[N(0, n^{1/3}) \geq n^{1/3}x] \psi(x) (1 + o(1)) = \frac{e^{-n^{1/3}x^2}}{x\sqrt{2\pi n^{1/3}}} \exp \left( \frac{Lx^3}{6} \right) (1 + o(1)). \quad (14)$$

Hence, the extended central limit theorem holds up to scale $n^{2/3}$, and then its breaking is measured by the residue $\psi(x) = e^{Lx^3/6}$. It should be noticed that by looking at $-S_n$, one also has

$$\Pr[S_n \leq -n^{2/3}x] = \frac{e^{-n^{1/3}x^2}}{x\sqrt{2\pi n^{1/3}}} \exp \left( -\frac{Lx^3}{6} \right) (1 + o(1)), \quad (15)$$

so assuming for instance $L > 0$, starting at scale $n^{2/3}$, the positive fluctuations (Equation (14)) become more probable than the negative fluctuations (Equation (15)). Therefore, the residue $\psi$ measures also a breaking of symmetry that does not occur at scale $s_n \ll n^{2/3}$, and starts precisely at this scale. We shall also shed light on this breaking of symmetry in a striking 2-dimensional example in §8.1.

5.2. Mod-Gaussian convergence and typical limiting functions. It now turns out that the phenomenon observed for sums of i.i.d. is shared by many other important examples of sequences of random variables. More precisely, one often deals with sequences $(X_n)_{n \in \mathbb{N}}$ for which there exists 3 distinct scales $c_n \ll s_n \ll d_n$ such that:

1. at scale $c_n$, one has a central limit theorem $\frac{X_n}{c_n} \rightarrow \mathcal{N}(0, 1)$,
2. at scale $d_n$, one has large deviations with some rate function $F$,
3. at scale $s_n$, there is an exponent $0 < \alpha < 1$ such that one has mod-Gaussian convergence for $\frac{X_n}{(s_n)^\alpha}$ with parameters $(s_n)^{1-\alpha}$, and therefore the “breaking of symmetry and of the central limit theorem” evoked in the previous paragraph.

Moreover, the limiting function $\psi$ of the mod-Gaussian convergence is typically the exponential of a monomial of degree $r \geq 3$. In the following we give three examples of such behavior.

Example 5.1. Consider a sum $S_n = Y_1 + \cdots + Y_n$ of symmetric random variables, independent and identically distributed, with $E[Y] = 0$ and $E[Y^2] = 1$. Since $E[Y^3] = 0$ by symmetry, at scale $n^{2/3}$, one has mod-Gaussian convergence of $n^{-1/3} S_n$ with parameters $n^{1/3}$, but limiting function

$$\psi(z) = \exp \left( \frac{E[Y^3] z^3}{6} \right) = 1.$$
Thus, the central limit theorem still holds at scale $n^{2/3}$ in this case. However, if one looks at $n^{-1/4} S_n$, then a Taylor expansion of the moment generating function leads to:

$$n^{-1/4} S_n \text{ converges mod-Gaussian with } \left\{ \begin{array}{l} \text{parameters } n^{1/2}, \\
\text{limiting function } \psi(z) = \exp \left( \frac{Lz^4}{24} \right), 
\end{array} \right.$$ 

where $L = \mathbb{E}[Y^4] - 3$. The quantity $L$ is the so-called kurtosis of $Y$, and it measures the deviation of the tails $\mathbb{P}[S_n \geq n^{3/4}x]$ from the Gaussian distribution: the tails at scale $n^{3/4}$ are smaller than Gaussian tails if the kurtosis is negative, and larger if the kurtosis is positive. If the kurtosis vanishes, then the extended central limit theorem still holds at scale $n^{3/4}$, and one can push it further, up to some scale $n^{v-1}$, where $v \geq 3$ is the smallest integer such that the $v$-th moment of $Y$ does not agree with the $v$-th moment of a Gaussian. At this precise scale, the central limit theorem breaks and this is measured by:

$$n^{-1/v} S_n \text{ converges mod-Gaussian with } \left\{ \begin{array}{l} \text{parameters } n^{(v-2)/v}, \\
\text{limiting function } \psi(z) = \exp \left( \frac{k^{(v)}(Y) z^v}{v!} \right), 
\end{array} \right.$$ 

as explained in Example 2.2. With this result, we have fully explained the content of Figure 1.

**Example 5.2.** 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 $\mathcal{Z}(f) = \{ z \in D_{(0,1)} \mid f(z) = 0 \}$ is a determinantal point process on the unit disk, with kernel

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

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 $\{ r^{2k} \}_{k \geq 1}$:

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

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

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

of $D_{(0,r)}$, 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 $\mathbb{E}[e^{zN^h}] = \prod_{k=1}^{\infty} (1 + r^{2k}(e^z - 1))$, one has

$$\log \left( \mathbb{E} \left[ e^{zN^h} \right] \right) = \sum_{k=1}^{\infty} \log \left( 1 + r^{2k} \left( e^{\frac{z^2}{h^{2/3}}} - 1 \right) \right) = \frac{h^{2/3}z}{4\pi} + \frac{h^{1/3}z^2}{16\pi} + \frac{z^3}{144\pi} + o(1)$$
with a remainder that is uniform when \( z \) stays in a compact domain of \( C \). Therefore, 
\[
\frac{N_h - \frac{h}{N}}{\frac{h^{1/3}}{N}} \text{ converges mod-Gaussian with } \left\{ \begin{array}{l}
\text{parameters } \alpha_{1/3}, \\
\text{limiting function } \psi(z) = \exp\left(\frac{z^3}{144\pi}\right).
\end{array} \right.
\]
Again, the limiting function is the exponential of a simple monomial.

**Example 5.3.** Consider the Ising model on the discrete torus \( \mathbb{Z}/n\mathbb{Z} \). Thus, we give to each spin configuration \( \sigma : \mathbb{Z}/n\mathbb{Z} \rightarrow \{\pm 1\} \) a probability proportional to the factor \( \exp(-\beta \sum_{i,j=1}^{n} \delta_{\sigma(i) \neq \sigma(j)}) \), the sum running over neighbors in the circular graph \( \mathbb{Z}/n\mathbb{Z} \). The technique of the transfer matrix ensures that if \( M_n = \sum_{i=1}^{n} \sigma(i) \) is the total magnetization of the model, then
\[
\mathbb{E}[e^{zM_n}] = \frac{\text{tr}(T(z))}{\text{tr}(T(0))} = \begin{pmatrix} e^{-z} & e^{z-\beta} \\ e^{-z-\beta} & e^{z} \end{pmatrix}.
\]
The two eigenvalues of \( T(z) \) are \( \cosh z \pm \sqrt{(\sinh z)^2 + e^{-2\beta}} \), and a Taylor expansion of them shows that
\[
\log \left( \mathbb{E}[e^{zM_n}] \right) = \frac{e^\beta n^{1/2} z^2}{2} - \frac{(3e^{3\beta} - e^\beta) z^4}{24} + o(1).
\]
So, one has mod-Gaussian convergence for \( n^{-1/4} M_n \), and the estimates
\[
\mathbb{P}[M_n \geq n^{3/4}x] = \frac{e^{-\frac{n^{1/2}e^\beta x^2}{2}}}{\sqrt{2\pi n^{1/2}}} \exp\left(-\frac{(3e^{3\beta} - e^\beta) x^4}{24}\right) (1 + o(1)).
\]

### 5.3. Precise deviations for random variables with control on cumulants.

Since many sequences of random variables \((X_n)_{n \in \mathbb{N}}\) exhibit mod-Gaussian convergence with a limiting function \( \exp(Lz^2) \), it becomes crucial to give concrete criterions that ensure this asymptotic behavior. The method of cumulants yields such a sufficient criterion. Consider as in the Example 2.3 a sequence \((S_n)_{n \in \mathbb{N}}\) of centered real-valued random variables such that
\[
|\kappa^{(r)}(S_n)| \leq (Cr)^r \alpha_n (\beta_n)^r
\]
with \( \alpha_n \rightarrow \infty \) and
\[
\kappa^{(2)}(S_n) = \sigma^2 \alpha_n (\beta_n)^2 \left(1 + o\left(\left(\alpha_n\right)^{-5/12}\right)\right) \\
\kappa^{(3)}(S_n) = L \alpha_n (\beta_n)^3 \left(1 + o\left(\left(\alpha_n\right)^{-1/6}\right)\right).
\]
Then, as explained in Example 2.3, the sequence \( X_n := S_n / (\beta_n (\alpha_n)^{1/3}) \) converges mod-Gaussian with parameters \((\alpha_n)^{1/3} \sigma^2 \) and limiting function \( \psi(z) = \exp\left(\frac{Lz^2}{\beta}\right) \). So, one can apply the previously established theorems to estimate the tail of the distribution of \( S_n \). By replacing \( S_n \) by \( S_n / \beta_n \), one can assume without loss of generality that \( \beta_n = 1 \). Then:

1. For any sequence \( s_n = o((\alpha_n)^{2/3}) \),
\[
\mathbb{P}[S_n \geq s_n] = \mathbb{P}\left[\mathcal{N}(0, \sigma^2) \geq (\alpha_n)^{-1/6} s_n\right] (1 + o(1)).
\]
that is to say that the central limit theorem holds (extended up to the scale \( o((\alpha_n)^{2/3}) \)). This can be compared with the results of [Hiwa96] for lattice-distributed random variables.
(2) At the next level, the function \( \psi(x) = \exp\left(\frac{Lx^3}{6}\right) \) comes into play, and for any sequence \( x_n = O(1) \), by Theorem 4.3,
\[
\mathbb{P}\left[S_n \geq x_n (\alpha_n)^{\frac{3}{2}}\right] = \frac{\exp\left(-\left(\frac{\alpha_n}{2}\right)^{\frac{3}{2}}(x_n)^2\right)}{x_n (\alpha_n)^{\frac{3}{2}} \sqrt{2\pi \sigma^2}} \psi(x_n) (1 + o(1)).
\]

Moreover, as we shall see now, one can use the estimates on cumulants (16) and (17) to obtain more precise results than in the general setting of mod-Gaussian convergence. In particular, the estimate of moderate deviations can be pushed with sequences \( x_n \) growing to infinity not too fast. Indeed, set \( X_n = (\alpha_n)^{-1/3} S_n \), up to a renormalization of the random variables, one can suppose \( \beta_n = 1 \), and also \( \sigma^2 = 1 \). Notice then that as long as \( z_n = o((\alpha_n)^{1/12}) \), the estimate
\[
\varphi_{X_n}(z_n) = \exp\left(\left(\frac{\alpha_n}{2}\right)^{\frac{3}{2}}\frac{(z_n)^2}{2} + \frac{L(z_n)^3}{6}\right) (1 + o(1))
\]
holds, because the terms of order \( r \geq 4 \) in the series expansion of \( \log \varphi_{X_n}(z_n) \) all go to zero — this is the same computation as in Example 2.3, but noticing that the remaining series is a \( O((z_n)^4 (\alpha_n)^{-1/3}) \), that is, a \( o(1) \) under the previous assumptions. As a consequence, if one makes 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 \( x_n = o((\alpha_n)^{1/12}) \), then the generating function of \( Y_n \) is \( \varphi_{Y_n}(z) = \frac{\varphi_{X_n}(x_n + z)}{\varphi_{X_n}(x_n)} \), so
\[
\log \varphi_{Y_n}(z) = \left(\frac{\alpha_n}{2}\right)^{\frac{3}{2}}\frac{(z + x_n)^2 - (x_n)^2}{2} \left(1 + o(\left(\frac{\alpha_n}{n}\right))\right)
+ \frac{L((z + x_n)^3 - (x_n)^3)}{6} \left(1 + o\left(\left(\frac{\alpha_n}{n}\right)^{-\frac{1}{6}}\right)\right) + o(1)
= \left(\frac{\alpha_n}{2}\right)^{\frac{3}{2}}\frac{z^2 + 2z x_n}{2} + \frac{L}{6} \left(z^3 + 3z^2 x_n + 3z (x_n)^2\right) + o(1)
= \left(\frac{\alpha_n}{2}\right)^{\frac{3}{2}} x_n + \frac{L (x_n)^2}{2} z + \left(\frac{\alpha_n}{2}\right)^{\frac{3}{2}} L \left(x_n \right)^2 z^2 + \frac{L}{6} z^3 + o(1)
(18)
\]

Thus, if \( Z_n = Y_n - (\alpha_n)^{1/3} x_n - \frac{L (x_n)^2}{2} \), then the \( Z_n \)'s converge in the mod-Gaussian sense with parameter \( (\alpha_n)^{\frac{3}{2}} + L x_n \) and the same limiting function \( \exp(\frac{Lz^3}{6}) \). One can therefore apply Proposition 4.1, and this leads to:

**Proposition 5.4.** Let \( (S_n)_{n \in \mathbb{N}} \) be a sequence of centered real-valued random variables with the bound on cumulants (16), and the asymptotics of second and third cumulants given by Equation (17). If \( \langle \alpha_n \rangle^{1/2} \ll T \ll \langle \alpha_n \rangle^{3/4} \), then
\[
\mathbb{P}\left[\frac{S_n}{\beta n \sigma} \geq T\right] = \frac{e^{-\frac{T^2}{2 \sigma^2}}}{\sqrt{2 \pi \frac{T^3}{\sigma^2}} \alpha_n^{3/2}} \exp\left(\frac{L T^3}{6 \sigma^3 (\alpha_n)^{3/2}}\right) (1 + o(1)),
\]
the second term in the expansion being different from 1 for \( \langle \alpha_n \rangle^{2/3} \ll T \ll \langle \alpha_n \rangle^{3/4} \).
Proof. Pursuing the previous computations (with $\beta_n = 1$ and $\sigma^2 = 1$), one has for $x_n = o((\alpha_n)^{1/12})$:

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

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

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

by replacing $\varphi_{X_n}(x_n)$ by its estimate, which holds since $x_n = o((\alpha_n)^{1/12})$; $R_n$ is the integral of the second line. To compute this integral, one applies Proposition 4.1 to $R_n$, thus, if $F_n(dw)$ is the law of $\frac{Z_n}{\sqrt{(\alpha_n)^{1/3} + Lx_n}}$, then

$$R_n = \int_{w=-\frac{L(\alpha_n)^2}{2\sqrt{(\alpha_n)^{1/3} + Lx_n}}}^{\infty} \exp \left( -wx_n \sqrt{(\alpha_n)^{1/3} + Lx_n} \right) F_n(dw)$$

$$\approx \frac{1}{\sqrt{2\pi}} \int_{w=-\frac{L(\alpha_n)^2}{2\sqrt{(\alpha_n)^{1/3} + Lx_n}}}^{\infty} \exp \left( -wx_n \sqrt{(\alpha_n)^{1/3} + Lx_n} - \frac{w^2}{2} \right) dw$$

$$\approx \frac{1}{\sqrt{2\pi}} \exp \left( \frac{\alpha_n \left( x_n \right)^2 + L(\alpha_n) x_n^3}{2} \right) \int_{w=-\frac{L(\alpha_n)^2}{2\sqrt{(\alpha_n)^{1/3} + Lx_n}}}^{\infty} e^{-\frac{w^2}{2}} dt,$$

by using the same argument as in the proof of Proposition 4.1 to replace $F_n(dw)$ by the Gaussian distribution. So, gathering everything, we get

$$P \left[ S_n \geq x_n \left( \alpha_n \right)^{3/2} \right] = \exp \left( \frac{L(x_n)^3}{6} \right) T_a \left( 1 + o(1) \right)$$

where $T_a$ is the tail of the standard Gaussian distribution, starting from

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

Since $a = x_n \left( \alpha_n \right)^{1/6} (1 + o(1))$ and $a^2 = \left( x_n \right)^2 \left( \alpha_n \right)^{1/3} + o(1)$, by Lemma 2.7, (3),

$$T_a = \frac{e^{-\frac{a^2}{2\pi}}}{a \sqrt{2\pi}} \left( 1 + o(1) \right) = \frac{e^{-\frac{(x_n \left( \alpha_n \right)^{1/3})^2}{2\pi}}}{x_n \sqrt{2\pi} \left( \alpha_n \right)^{1/3}} \left( 1 + o(1) \right),$$

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

Proposition 5.4 hints at a possible expansion of the fluctuations up to any order $T = o((\alpha_n)^{1-\epsilon})$, 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

$$|\kappa^{(r)}(S_n)| \leq (Cr)^r \alpha_n (\beta_n)^r \quad ; \quad \kappa^{(r)}(S_n) = K(r) \alpha_n (\beta_n)^r \left( 1 + O((\alpha_n)^{-1}) \right)$$
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 Petrov-Cramér series \( \lambda^{(r)} = -b_{r-1}/r \), with

\[
\sum_{r=1}^{j} \frac{\kappa^{(r+1)}}{r!} \left( \sum_{j_1 + \cdots + j_r = j} b_{j_1} b_{j_2} \cdots b_{j_r} \right) = \delta_{j,1},
\]

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 \( (18) \) is \( \lambda^{(3)} \). 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

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

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

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

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

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-\( \phi \) convergence. However, their works do not provide concrete tools to prove the estimates \( (17) \), and above all the bounds \( (16) \), that are usually the most difficult to obtain. In Sections 9-11, we shall present a set of combinatorial tools to prove such bounds in some random discrete models, where the moment generating function is intractable. We will then use extensively the estimates of Proposition 5.4.

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 \( P[X_n \geq t_n x_n] \), but more generally for probabilities

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

In particular, under some technical assumptions on the generating series (that look like, but are somehow weaker than mod-convergence), the Ellis-Gärtner theorem provides some asymptotic upper and lower bounds for \( \log(P[X_n \in t_n B]) \), 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 generalizes Cramér’s large deviations for sums of i.i.d. random variables. Meanwhile, our Theorems 3.3 and
4.3 give estimates for the probabilities \( P[X_n \geq t_n x] \) themselves (instead of their logarithms), thereby generalizing Bahadur-Rao precise estimates for sums of i.i.d. variables. Therefore, it is natural to establish in the framework of mod-convergence a precise version of the Ellis-Gärtner theorem. Thus, in this section, we shall give some asymptotic upper and lower bounds for the probabilities \( P[X_n \in t_n B] \) itself instead of their logarithms. Once again, the upper and lower bounds coincide for nice borelian sets \( B \). Though our results (Theorems 6.2 and 6.3) are minor improvements over those of the previous sections, they also prepare the discussion of the multi-dimensional case, see Section 7.

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 with \( c = +\infty \) (in particular, \( \psi \) is entire on \( C \)).

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 \( \lim F'(x) = \pm \infty \) when \( x \) goes to a bound of the interval \( I_F \) (cf. [DZ98, Definition 2.3.5]).

The later point is verified if \( \phi \) is a Gaussian or Poisson law, which are the most important examples.

Lemma 6.1. 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 6. 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 \( F(a) = F(b) \).](image)

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 \( \phi \). If \( C \cap I_F = \emptyset \), 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 \( C \cap \{ x \in \mathbb{R} \mid F(x) \leq M \} \neq \emptyset \). 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 $B^\circ$ 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.** Let $B$ be a Borelian subset of $\mathbb{R}$.

1. If $B$ is admissible, then

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

the distinction of cases corresponding to $\phi$ 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 $\phi$ 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 $C \subset (-\infty, a] \cup [b, +\infty)$. 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 4.3 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.

### 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 $F(B^o)$. Unfortunately, the result takes a less interesting form than Theorem 6.2. If $B$ is a Borelian subset of $\mathbb{R}$, denote $B^\delta$ the union of the open intervals $(x, x + \kappa)$ of width $\kappa \geq \delta$ that are included into $B$. The interior $O = B^o$ is a disjoint union of a countable collection of open intervals, and also the increasing union $\bigcup_{\delta > 0} B^\delta$.

![Figure 7](image)

**Figure 7.** 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 $B^\delta$s, as some points can be points of accumulation of open intervals included in $B$ and of width going to zero (see Figure 7). 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.** For an admissible Borelian set $B$,

$$
\liminf_{\delta \to 0} \liminf_{n \to \infty} \left( \sqrt{2\pi t_n} \exp(t_n F(B^\delta)) \right) \mathbb{P}[X_n \in t_n B] \geq \left\{ \begin{array}{ll}
\sum_{a \in (B^o)_{\min}} \frac{\psi(h(a))}{(1-e^{-|h(a)|}) \sqrt{\eta''(h(a))}} & \\
\sum_{a \in (B^o)_{\min}} \frac{\psi(h(a))}{|h(a)| \sqrt{\eta''(h(a))}} & \\
\sum_{a \in (B^o)_{\min}} \frac{\psi(h(a))}{\sqrt{\eta''(h(a))}} &
\end{array} \right.
$$
with again the distinction of cases lattice/non-lattice. In particular, the right-hand side in Theorem 6.2 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 \((B^\delta)_{\min}\) consists of one point \( b = \inf(B^\delta \cap [\eta'(0), +\infty)) \), the other situations being entirely similar. As \( \delta \) goes to 0, \( B^\delta \) increases towards \( B^\circ = \bigcup_{\delta > 0} B^\delta \), 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 \( \delta \). Actually, if \( b^\delta = \inf(B^\delta \cap [\eta'(0), +\infty)) \), then for \( \delta \) small enough \( F(B^\delta) = F(b^\delta) \), so \( \lim_{\delta \to 0} F(B^\delta) = F(B^\circ) \) 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 \( b^\delta \). Hence, it suffices to show that for \( \delta \) small enough, \( L(\delta) \geq R(\delta) \). However, by definition of \( B^\delta \), 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. \( \square \)

### 7. Multi-dimensional extensions

In the following, vectors in \( \mathbb{R}^d \) are denoted by bold letters \( \mathbf{V}, \mathbf{W}, \ldots \) and their coordinates are indexed by exponents, so \( \mathbf{V} = (\mathbf{V}^{(1)}, \ldots, \mathbf{V}^{(d)}) \). This convention enables us to manipulate sequences \( (\mathbf{V}_n)_{n \in \mathbb{N}} \) of (random) vectors in \( \mathbb{R}^d \) without ambiguity.

#### 7.1. Sum of a Gaussian noise and an independent random variable.

The multi-dimensional extension of the previous results is geometrically non-trivial, and we shall restrict ourselves to the following situation, which is the mod-Gaussian setting:

1. A sequence of \( \mathbb{R}^d \)-valued random variables \( (\mathbf{X}_n)_{n \in \mathbb{N}} \) is said to converge in the mod-Gaussian sense with parameters \( A_n \in \mathbb{GL}(\mathbb{R}^d) \) and limiting function \( \psi \) if locally uniformly on \( \mathbb{C}^d \),

\[
\exp \left( -\frac{1}{2} \sum_{i=1}^d (A_n^i z^{(i)})^2 \right) \mathbb{E} \left[ e^{\sum_{i=1}^d A_n^i z^{(i)}} \right] \to \psi(z),
\]

with \( \psi \) analytic in \( d \) variables and non-vanishing on its real domain, and \( \Sigma_n = (A_n)^{-1} \) going to zero.
(b) By replacing $A_n$ by $\sqrt{A_n} A_n$, it does not cost anything in our definition to suppose that $A_n$ is symmetric definite positive.

(c) Our goal is then to find the precise asymptotics of the probabilities $P[X_n \in (A_n)^2 B]$, where $B$ is a Borelian subset of $\mathbb{R}^d$. In the one-dimensional case, $A_n = \sqrt{n}$ and we can rewrite the estimate of Theorem 6.2 as

$$\limsup_{n \to \infty} \left( \frac{2\pi}{t_n} (t_n b)^2 \exp \left( \frac{t_n b^2}{2} \right) P[X_n \in t_n B] \right) \leq \int_{B_{\min}} \psi(x) N(dx)$$

where $b = \inf \{ |x| : x \in B \}$, and $N$ is the counting measure on $B_{\min}$, which consists of one or two points. Moreover this estimate is sharp as soon as $B$ contains small open intervals $[b, b + \delta]$ or $(-b - \delta, -b]$ (according to the type of $B_{\min}$).

Example 7.1. In order to make a correct conjecture, let us consider a toy model. We suppose here that $X_n = Y + A_n G$, where $G$ is a $d$-dimensional standard Gaussian variable, the $A_n$’s are symmetric and positive definite, and $Y$ is a fixed random variable independent of the Gaussian $G$ and with generating series $\psi(z)$. A simple computation yields

$$\mathbb{E} \left[ e^{\sum_{i=1}^{d} X_n^{(i)} z^{(i)}} \right] = \mathbb{E} [e^{\langle A_n G \mid z \rangle}] \psi(z) = \mathbb{E} [e^{\langle G \mid A_n^* z \rangle}] \psi(z) = \exp \left( \frac{1}{2} \sum_{i=1}^{d} (A_n^* z)^{(i)} \right) \psi(z),$$

so we are typically in the mod-Gaussian setting. We assume in the following that:

(d) The random variable $Y$ has a density $\nu(y)$ with respect to the Lebesgue measure, and is a bounded random variable. In particular, $\psi(b) \leq e^{C \|Ab\|}$ for some constant $C$.

The probabilities that we want to estimate then as:

$$P[X_n \in (A_n)^2 B] = \frac{1}{(2\pi)^{d/2}} \int_{x \in \mathbb{R}^d} \int_{y \in \mathbb{R}^d} e^{-x^2/2} \chi_{(y + A_n x \in (A_n)^2 B)} \nu(y) \, dx \, dy$$

$$= \det A_n \frac{1}{(2\pi)^{d/2}} \int_{u \in \mathbb{R}^d} \int_{y \in \mathbb{R}^d} e^{-u^2/2} \chi_{(\Sigma_n)^2 y + u \in B} \nu(y) \, du \, dy$$

$$= \det A_n \frac{1}{(2\pi)^{d/2}} \int_{b \in B} e^{-b^2/2} \left( \int_{y \in \mathbb{R}^d} e^{(b \mid y)} e^{-y^2/2} \nu(y) \, dy \right).$$

Since $\Sigma_n$ goes to zero, the term in parentheses goes by Lebesgue dominated convergence theorem to $\int_{\mathbb{R}^d} e^{(b \mid y)} \nu(y) \, dy = \psi(b)$. To simplify a bit the discussion, we suppose:

(e) For some fixed symmetric positive definite matrix $A$, $A_n = \sqrt{t_n} A$ with $t_n$ increasing to infinity.

Then, in particular, the term in parentheses goes to $\psi(b)$ in an increasing way, and the convergence is locally uniform in $b$. One can even suppose $A = I_d$, the general situation following by obvious changes of coordinates. We can then perform a Laplace method on the previous expression, but in a quite unusual multi-dimensional setting. Set

$$b = \inf \left\{ \sqrt{(b \mid b)} \mid b \in B \right\},$$
and suppose \( b \neq 0 \). For any \( \varepsilon > 0 \), we split \( B \) into \( B_{< b+\varepsilon} \), the set of points of \( B \) that are of norm \( \|b\| \) smaller than \( b+\varepsilon \), and \( B_{\geq b+\varepsilon} \). We also denote \( B_{\min} = \{x \in \mathbb{B} \mid \sqrt{(x \cdot x)} = b\} \).

The integral over \( B_{\geq b+\varepsilon} \) is smaller than the integral on the whole complementary of the ball of radius \( b+\varepsilon \) and center the origin. Recall that the measure of the \( (d-1) \)-dimensional regular sphere is \( \text{vol}(S^{d-1}) = \frac{\pi^{d/2}}{\Gamma(d/2)} \). So, if \( I_{\geq b+\varepsilon} \) denotes the part of \( \mathbb{P}[X_n \in t_n B] \) corresponding to the integral over \( B_{\geq b+\varepsilon} \), then

\[
I_{\geq b+\varepsilon} \leq \left( \frac{t_n}{2\pi} \right)^{d/2} \int_{\|x\| \geq b+\varepsilon} e^{-\frac{t_n|x|^2}{2}} e^C|x| \, dx \leq \frac{2}{\Gamma(d/2)} \left( \frac{t_n}{2} \right)^{d/2} \int_{b+\varepsilon}^{\infty} w^{d-1} e^{-\frac{t_n w^2}{2}} e^{Cw} \, dw
\]

\[
\leq \frac{2 e^{C^2}}{\Gamma(d/2)} \left( \frac{t_n}{2} \right)^{d/2} \int_{b+\varepsilon}^{\infty} \left( x + \frac{C}{t_n} \right)^{d-1} e^{-\frac{t_n x^2}{2}} \, dx
\]

\[
\leq \frac{2 e^{C^2}}{\Gamma(d/2)} \left( \frac{t_n}{2} \right)^{d/2} \left( \frac{1 + C/t_n b}{t_n b} \right)^{d-1} \int_{b+\varepsilon}^{\infty} x^{d-1} e^{-\frac{t_n x^2}{2}} \, dx
\]

\[
\leq \frac{1}{\Gamma(d/2)} \left( \frac{t_n}{2} \right)^{d/2-1} \left( 1 + C/t_n b \right)^{d-1} \left( 1 - \frac{d-2}{t_n b^2} \right)^{-1} e^{-\frac{t_n(b+\varepsilon)^2}{2}} e^{C(b+\varepsilon)} (b+\varepsilon)^{d-2}
\]

assuming \( C/t_n \leq \varepsilon \), and using integration by parts at the end to estimate the Gaussian integral. To fix the ideas, suppose that \( t_n b \geq C \) and \( t_n b^2 \geq 2(d-2) \), which is surely the case for \( n \) big enough. Then,

\[
I_{\geq b+\varepsilon} \leq \frac{2}{\Gamma(d/2)} (2t_n)^{d/2} e^{-\frac{t_n(b+\varepsilon)^2}{2}} e^{C(b+\varepsilon)} (b+\varepsilon)^{d-2}.
\]

(19)

As for the integral over \( B_{< b+\varepsilon} \), we have to use more precisely the topology of \( B \) in the neighborhood of the sphere of radius \( b \). Recall that for any non-negative measurable function over \( \mathbb{R}^d \), one can make the polar change of coordinates

\[
\int_{\mathbb{R}^d} f(x) \, dx = \int_{r=0}^{\infty} \left( \int_{S^{d-1}} f(r x) \, d\mu_{S^{d-1}}(x) \right) r^{d-1} \, dr,
\]

where \( \mu_{S^{d-1}} \) is up to a scalar the unique SO(\( \mathbb{R}^d \))-invariant measure on the sphere — the “spherical” Lebesgue measure. Consequently, we can write

\[
(1 + o(1)) I_{< b+\varepsilon} = \left( \frac{t_n}{2\pi} \right)^{d/2} \int_{\|b\| < b+\varepsilon} e^{-\frac{t_n \|b\|^2}{2}} 1_B(b) \psi(b) \, db
\]

\[
= \left( \frac{t_n}{2\pi} \right)^{d/2} e^{-\frac{t_n b^2}{2}} \int_{\|b\| < b+\varepsilon} \frac{t_n (2-\|x\|^2)}{2} 1_B(c) \psi(c) \, dc
\]

\[
= \left( \frac{t_n}{2\pi} \right)^{d/2} e^{-\frac{t_n b^2}{2}} \int_{r=b}^{b+\varepsilon} t_n (2-\frac{r^2}{2}) r^{d-1} \left( \int_{S^{d-1}} 1_B(rc) \psi(rc) \, d\mu_{S^{d-1}}(c) \right) dr,
\]

the multiplicative factor \((1 + o(1))\) corresponding to the replacement of the integral

\[
\int_{y \in \mathbb{R}^d} e^{(b \cdot y)} e^{-\frac{y^T S^{-1} y}{2}} v(y) \, dy
\]
by $\psi(b)$. Suppose for a moment that the integral in parentheses is continuous with respect to $r$, or at least continuous at $r = b$. Then, up to a multiplicative factor $(1 + o(1))$, we can replace it by the constant

$$\frac{1}{b^{d-1}} \int_{B_{\min}} \psi(b) \, d\mu_{S^{d-1}(0,b)}(b)$$

where $\mu_{S^{d-1}(0,b)} = \mu_{\text{surface}}$ is the spherical measure on the sphere of radius $b$ that gives for total weight $b^{d-1} \text{vol}(S^{d-1})$. Hence,

$$I_{<b+\epsilon} \simeq \left( \frac{t_n}{2\pi} \right)^{d/2} e^{-\frac{t_n b^2}{2}} \left( \int_{B_{\min}} \psi(b) \, d\mu_{\text{surface}}(b) \right) \left( \int_{r=b}^{b+\epsilon} e^{\frac{t_n}{b} \left( \frac{r^2}{2} - \frac{1}{3} \right)} r^{d-1} dr \right). \quad (20)$$

Set $\epsilon = (t_n)^{-2/3}$. In (19), the expansion of $e^{-\frac{t_n(b+\epsilon)^2}{2}}$ gives $e^{-\frac{t_n b^2}{2}} e^{-b(t_n)^{1/3}(1 + o(1))}$, so

$$I_{\geq b+\epsilon} \lesssim 2e^{\frac{2b}{3}} \left( \frac{2t_n}{d/2} \right)^{d/2} e^{-\frac{t_n b^2}{2}} e^{-b(t_n)^{1/3}(1 + o(1))}.$$ 

On the other hand, since $t_n \epsilon = (t_n)^{1/3}$ and $t_n \epsilon^2 = (t_n)^{-1/3}$, in Formula (20), one obtains

$$I_{<b+\epsilon} = \left( \frac{t_n}{2\pi} \right)^{d/2} e^{-\frac{t_n b^2}{2}} \left( \int_{B_{\min}} \psi(b) \, d\mu_{\text{surface}}(b) \right) \left( \int_{s=0}^{\epsilon} e^{-\frac{t_n b^2}{2}} e^{-t_n b s} ds \right) (1 + o(1))$$

$$= \left( \frac{t_n}{2\pi} \right)^{d/2} e^{-\frac{t_n b^2}{2}} \left( \int_{B_{\min}} \psi(b) \, d\mu_{\text{surface}}(b) \right) \left( \int_{t=0}^{(t_n)^{1/3}} e^{-bt} dt \right) (1 + o(1))$$

$$= \left( \frac{t_n}{2\pi} \right)^{d/2} e^{-\frac{t_n b^2}{2}} \left( \int_{B_{\min}} \frac{\psi(b)}{\|b\|} \, d\mu_{\text{surface}}(b) \right) (1 + o(1)).$$

The first term is negligible in comparison to the second term, so, one can reasonably conjecture that

$$\limsup_{n \to \infty} \left( \frac{2\pi}{t_n} \right)^{d/2} (t_n b) \exp \left( \frac{t_n b^2}{2} \right) \mathbb{P}[X_n \in t_n B] \leq \int_{B_{\min}} \psi(x) \, d\mu_{\text{surface}}(dx).$$

The reason why one has only a inequality with a lim sup and not a equality with the limit is that the spherical integral with $B$ a closed set (which does not cost anything since $B_{\min} = \overline{(B)}_{\min}$) is only upper semi-continuous, see the following Lemma. Therefore,

$$\limsup_{r \to b} \left( \int_{S^{d-1}} 1_B(rc) \psi(rc) \, d\mu_{S^{d-1}}(c) \right) \leq \int_{S^{d-1}} 1_B(bc) \psi(bc) \, d\mu_{S^{d-1}}(c),$$

and then the discussion is the same but with inequalities instead of equalities in the estimates of $I_{<b+\epsilon}$.

**Lemma 7.2.** Let $\psi$ be a non-negative continuous function, and $C$ a closed subset of $\mathbb{R}^d$. The function

$$F_C : r \mapsto \int_{S^{d-1}} \psi(r x) 1_C(rx) \, d\mu_{S^{d-1}}(x)$$

is upper semi-continuous on $(0, +\infty)$. 
Proof. If instead of $\mathbb{1}_C$ we had a continuous function, then by standard results on integrals of continuous functions in two parameters, the integral $F$ would be continuous in $r$. However, by using the distance function to $C$, one can write $\mathbb{1}_C$ as the infimum of continuous functions, and therefore, $F_C$ is itself the infimum of continuous functions, whence upper semi-continuous. □

Now working in the more general case of an arbitrary symmetric matrix $A_n = \sqrt{t_n} A$, we get the following upper bound:

$$
\limsup_{n \to \infty} \left( \frac{2\pi}{t_n} \right)^{d/2} (t_n b^2) \exp \left( \frac{t_n b^2}{2} \right) \mathbb{P}[X_n \in (A_n)^2 B] \leq \int_{B_{\text{min}}} \psi(x) \mu_{\text{surface}}(dx)
$$

where $b = \inf \{ \sqrt{\langle Ab, Ab \rangle} \mid b \in B \}$; $B_{\text{min}} = \{ x \in \overline{B} \mid \sqrt{\langle Ax, Ax \rangle} = b \}$; and $\mu_{\text{surface}}$ is the surface measure on the $A$-sphere of radius $b$ obtained by the linear automorphism $A$ from the surface measure on $S^{d-1}(0, b)$ — in particular it has the same mass, so beware that unless $A = I_d$, this is not the surface measure obtained by restricting the Riemannian structure of the euclidian space $\mathbb{R}^d$ to the $A$-sphere. The goal of this section is now to prove the estimate (21) assuming only hypotheses (a), (b) and (e); and to give a sufficient condition for the limsup to be a limit.

7.2. Berry-Esseen estimates for hypercubes and polytopes. In this section, we assume that 7.1.(a), 7.1.(b) and 7.1.(e) are fulfilled, and we set

$$Y_n = \frac{\Sigma(X_n)}{\sqrt{t_n}} = \Sigma_n X_n,$$

where $\Sigma = A^{-1}$. We then denote $\mu_n$ the law of $Y_n$. As in the one-dimensional case, the main tool for estimates of precise deviations will be a Berry-Esseen estimate, analogue to Proposition 4.1. In a multi-dimensional setting, there are several ways to tackle this problem:
(1) Denote \( g(x) = (2\pi)^{-d/2} e^{-||x||^2/2} \) the distribution of a standard \( d \)-dimensional Gaussian variable, and \( G[F] = \int_{\mathbb{R}^d} F(x) g(x) \, dx \) the expectation of a test function \( F \) against the Gaussian distribution. One can estimate the difference between 
\[ \mu_n(F) = \mathbb{E}[F(Y_n)] \quad \text{and} \quad G[F] \]
for \( F \) in a suitable class of test functions, say functions that are bounded and Lipschitz, or smooth functions.

(2) Alternatively, one can estimate of the difference between \( \mu_n(B) \) and \( G(B) \), where \( B \) runs over a class of regular domains, for instance, the class of all convex bodies with a sufficiently regular boundary, or the class of all polytopes.

In both cases, one expects to be able to write estimates with a uniform remainder \( o(1/\sqrt{n}) \). We were only able to do so for the probabilities of hypercubes \( \prod_{i=1}^d [a(i), b(i)] \) under the law \( \mu_n \) of \( Y_n \); and we refer to the additional Section 12 for other Berry-Esseen estimates, which are less sharp but hold in a more general setting. Though not needed for our results of large deviations in several dimensions, these new estimates seem of interest in their own right; they should be compared with the approximation results of [BR10], which are also proved in several dimensions, but only in the setting of sums of i.i.d. random variables.

**Theorem 7.3** (Berry-Esseen estimates for probabilities of polytopes). For every \( d \)-dimensional hypercube \( C = \{ \sum_{i=1}^d t(i)v_i, \ t(i) \in [a(i), b(i)] \} \), with the \( v_i \)’s not necessarily orthogonal, and possibly with infinite bounds \( a(i) \) or \( b(i) \),
\[
\mathbb{P}[X_n \in \sqrt{n} A(C)] = \mu_n(C) = \frac{1}{(2\pi)^{d/2}} \left( \int_C e^{-\frac{||x||^2}{2}} (1 + \langle \Sigma_n \psi'(0) \ | \ x \rangle) \, dx \right) + o \left( \frac{1}{\sqrt{n}} \right)
\]
with a remainder uniform on hypercubes. The same uniform estimate holds over convex polytopes with a bounded number of faces \( F \).

**Proof.** We claim that the proof of [Fel71, Lemma XVI.3.1] and of Proposition 4.1 adapts readily to the multi-dimensional setting. The main difference between the polytope case and the more general computations performed in Section 12 is the possibility to use the monotony of the cumulative distribution function
\[
F(x^{(1)}, \ldots, x^{(d)}) = \mathbb{P} \left[ Y_n = \sum_{i=1}^d y^{(i)}v_i, \ y^{(i)} \leq x^{(i)} \right] = \mu_n(C(x^{(1)}, \ldots, x^{(d)}))
\]
with respect to all of its parameters. Denote
\[
G(x^{(1)}, \ldots, x^{(d)}) = \frac{1}{(2\pi)^{d/2}} \int_{C(x^{(1)}, \ldots, x^{(d)})} e^{-\frac{||x||^2}{2}} (1 + \langle v \ | \ x \rangle) \, dx;
\]
\[
\delta(x^{(1)}, \ldots, x^{(d)}) = (F - G)(x^{(1)}, \ldots, x^{(d)}) \quad \text{and} \quad \eta = \sup_{x^{(1)}, \ldots, x^{(d)}} |\delta(x^{(1)}, \ldots, x^{(d)})|. \]
The \( d \)-dimensional analogue of [Fel71, Lemma XVI.3.1] yields
\[
\eta \leq 2\eta_T + \frac{M}{T},
\]
for
with
\[
\eta_T = \sup_{x^{(1)}, \ldots, x^{(d)}} |\delta_T(x^{(1)}, \ldots, x^{(d)})|
\]
\[
= \sup_{x^{(1)}, \ldots, x^{(d)}} |(\Delta_{T,R^d}^0 * F)(x^{(1)}, \ldots, x^{(d)}) - (\Delta_{T,R^d}^0 * G)(x^{(1)}, \ldots, x^{(d)})|,
\]
and where:

- \(M\) is a multiple of a Lipschitz bound for \(G\), that is to say a constant depending only on \(v\).
- \(\Delta_{T,R^d}^0\) is a \(d\)-dimensional probability measure with Fourier transform supported on \([-T, T]^d\). We refer to §12.1 for precisions on the properties of these kernels: they enable one to smoothen distributions and to cut the support of their Fourier transforms.

The monotony of \(F\) plays here a crucial role: indeed, Feller’s proof uses the fact that if \((x^{(1)}, \ldots, x^{(d)})\) are parameters such that \(\eta = \delta(x^{(1)}, \ldots, x^{(d)})\), then for a certain vector \(m \in \mathbb{R}^d\),
\[
\delta(\tilde{x}^{(1)} + e^{(1)}, \ldots, \tilde{x}^{(d)} + e^{(d)}) \geq F(x^{(1)}, \ldots, x^{(d)}) - G(\tilde{x}^{(1)} + e^{(1)}, \ldots, \tilde{x}^{(d)} + e^{(d)}) \\
\geq \frac{\eta}{2} + \langle m \mid e \rangle
\]
in a neighborhood \(\prod_{i=1}^d [\tilde{x}^{(i)} - e^{(i)}, \tilde{x}^{(i)} + e^{(i)}]\) of the vector \((\tilde{x}^{(1)}, \ldots, \tilde{x}^{(d)}) = (x^{(1)} + e^{(1)}, \ldots, x^{(d)} + e^{(d)})\).

Now, the same arguments as in Proposition 4.1 yield the uniform bound of the theorem for infinite hypercubes \(C(x^{(1)}, \ldots, x^{(d)})\) — notice that the term \(\eta'''(0)\) disappears here since \(\eta''' = 0\). Then, for any convex polytope \(C\) bounded by \(F\) faces (including bounded hypercubes with \(F = 2d\)), one can use inclusion-exclusion to write \(1_C\) as a linear combination of \(2^F\) cumulative distribution functions:
\[
1_C = \prod_{f \in \text{faces}(C)} 1_{f^+} = \prod_{f \in F(C)} (1 - 1_{f^-}) = \sum_{A \subseteq F(C)} \pm 1_{C(A)},
\]
where \(1_{f^+}\) and \(1_{f^-}\) denote the indicator functions of the two half-spaces determined by the affine hyperplane \(f\), with the + side containing \(C\); and for \(A\) subset of the set of faces \(F(C)\), \(C(A)\) is the corresponding infinite hypercube \(\bigcap_{f \in A} f^-\). Since the bound \(o((\mu_n)^{-1/2})\) for \(|\mu_n(C) - G(C)|\) is uniform over infinite hypercubes, we get from (22) the same uniform bound for convex polytopes with a bounded number of faces. \qed

### 7.3. Estimates of probabilities of Borel sets in a multi-dimensional setting

We are now ready to prove the large deviations principle (21); again, we suppose \(A = \Sigma = I_d\), as this only amounts to easy changes of variables. Fix \(x \neq 0\), and as before, introduce a random variable \(X_n\) with probability
\[
Q_n[dy] = \frac{\exp((y \mid x) \cdot n)}{\varphi_n(x)} \mathbb{P}_n[dy],
\]
where \( P_n \) is the law of \( X_n \). The characteristic function of \( \tilde{X}_n \) is 
\[ \phi_n(z+x) = \phi_n(x) \exp \left( -\frac{t_n \| z \|^2}{2} + \langle t_n x \mid z \rangle \right), \]
which is the Laplace transform of a Gaussian random variable with mean \( t_n x \) and 
covariance matrix \( t_n I_d \), it converges locally uniformly to \( \psi(z+x)/\psi(x) \); this is the 
multi-dimensional version of Lemma 4.7. Notice that if \( x \) stays in a compact subset 
of \( \mathbb{R}^d \), then the bounds measuring this convergence can be made independent of \( x \), so 
the previous theory applies locally uniformly with respect to \( x \). In particular, for any 
convex polytope \( C \),
\[ \left| \mathbb{Q} \left[ \frac{\tilde{X}_n - t_n x}{\sqrt{t_n}} \in C \right] - G_{\psi'(x)}(C) \right| = o \left( \frac{1}{(t_n)^{1/2}} \right), \tag{23} \]
where \( G_{\psi'(x)} \) is the signed measure with density
\[ \frac{1}{(2\pi)^{d/2}} \left( 1 + \frac{\langle \psi'(x) \mid y \rangle}{\sqrt{t_n} \psi(x)} \right) e^{-\| y \|^2/2} \text{d}y. \]

We fix a polytope neighborhood \( B \) of \( 0 \) in the \((d-1)\) vector space orthogonal to \( x \), 
and we consider the convex polytopes
\[ C(0,B) = \{(\lambda-1)x + \lambda b, \ b \in B, \ \lambda \geq 1 \}; \]
\[ C(x,B) = x + C(0,B) = \{\lambda(x+b), \ b \in B, \ \lambda \geq 1 \}, \]
see the following Figure.

\[ \text{FIGURE 9. For cones } C(x,B) \text{ with polytope basis } B, \text{ one can approximate } \]
the probability \( \mathbb{P}[X_n \in t_n C] \) by the conjectured upper bound, and this 
locally uniformly in \( x \).

Denote \( \tilde{\mu}_n \) the law of \( \frac{\tilde{X}_n - t_n x}{\sqrt{t_n}} \), and suppose \( \| x \| = x \). One has
\[ \mathbb{P}[X_n \in t_n C(x,B)] = \phi_n(x) \int_{\mathbb{R}^d} 1_{y \in t_n C(x,B)} \exp(-\langle y \mid x \rangle) \mathbb{Q}_n(\text{d}y). \]
Setting \( u = (y - t_n x) / \sqrt{t_n} \) in the integral, we get
\[ \mathbb{P}[X_n \in t_n C(x,B)] = \psi_n(x) \exp \left( -\frac{t_n x^2}{2} \right) \int_{\mathbb{R}^d} 1_{u \in \sqrt{t_n} C(0,B)} \exp(-\sqrt{t_n} \langle u \mid x \rangle) \tilde{\mu}_n(\text{d}u). \]

Denote
\[ C_{n,\lambda} = \{ \sqrt{t_n}(\nu x + (\nu + 1)b), \ 0 \leq \nu \leq \lambda, \ b \in B \}, \tag{24} \]
so that in particular $C_{n,\infty} = \sqrt{t_n} C(0, B)$. By slicing this cone $C_{n,\infty}$ orthogonally to the direction $x$, we see that for every distribution $F(u) = F(\sqrt{t_n}(\nu x + (v+1)b)) = F(v)$ only depending on $\nu$,

$$
\int_{C_{n,\infty}} F(v) \overline{\mu}_{n}(du) = \int_0^\infty -F'(\lambda) \overline{\mu}_{n}(C_{n,\lambda}) \, d\lambda,
$$

so in particular the integral $I$ in (24) is also equal to

$$
I = \int_0^\infty t_n x^2 \exp(-\lambda t_n x^2) \overline{\mu}_{n}(C_{n,\lambda}) \, d\lambda.
$$

We can now replace $\overline{\mu}_{n}(C_{n,\lambda})$ by its approximation (23), and as in dimension 1, it suffices to take the Gaussian mass, as this is the only part that will give something bigger than $o(1/\sqrt{t_n})$. So,

$$
I \simeq \int_0^\infty t_n x^2 \exp(-\lambda t_n x^2) G(C_{n,\lambda}) \, d\lambda
$$

$$
\simeq \int_0^\infty \int_B (t_n)^{d/2} x e^{-t_n x^2} g(\sqrt{t_n}(\nu x + (v+1)b)) \, db \, dv
$$

$$
\simeq \frac{\chi(t_n)^{d/2}}{(2\pi)^{d/2}} \int_0^\infty e^{-t_n x^2 (v + \frac{x^2}{\pi})} \left( \int_B e^{-t_n(\frac{(x+1)^2||b||^2}{2})} \, db \right) \, dv
$$

where the symbol $\simeq$ means up to a uniform remainder $o((t_n)^{-1/2})$.

Let $\epsilon$ be a function going to 0 as its argument goes to zero, and such that the uniform remainder in the previous formula always satisfies

$$
0\left(\frac{1}{\sqrt{t_n}}\right) \leq \frac{1}{\sqrt{t_n}} \epsilon^3 \left(\frac{1}{\sqrt{t_n}}\right).
$$

We suppose that $B = \frac{S}{\sqrt{t_n}}$ with $S$ polytope of surface measure of order

$$
\mu(S) \asymp \epsilon \left(\frac{1}{\sqrt{t_n}}\right);
$$

here by “of order” we mean equal up to a positive multiplicative constant bounded from below and from above. Under this assumption, the asymptotics of the integral are

$$
\left| I - \frac{(t_n)^{d/2-1}}{(2\pi)^{d/2}} x \mu_{\text{surface}}(B) \right| \leq \frac{1}{\sqrt{t_n}} \epsilon^2 \left(\frac{1}{\sqrt{t_n}}\right),
$$

and the estimate is of order

$$
(t_n)^{d/2-1-(d-1)/2} \mu(S) \simeq \frac{1}{\sqrt{t_n}} \epsilon \left(\frac{1}{\sqrt{t_n}}\right),
$$

so larger than the remainder. Indeed, in the integral over $B$, $t_n ||b||^2$ is always a $o(1)$ by assumption on the size of the surface, and then one only has to compute a one dimensional Gaussian integral, which is done by mean of Lemma 2.7. So:

**Lemma 7.4.** Fix $x \neq 0$ and a $(d-1)$-dimensional polytope $S$ orthogonal to $x$ and of surface measure of order $\epsilon \left(\frac{1}{\sqrt{t_n}}\right)$, with $\epsilon$ chosen as before. One has

$$
\exp\left(\frac{t_n x^2}{2}\right) \mathbb{P}\left[ X_n \in t_n C\left( x, \frac{S}{\sqrt{t_n}} \right) \right] = \frac{(t_n)^{d/2-1} \psi_n(x)}{(2\pi)^{d/2} x} \mu_{\text{surface}}\left( \frac{S}{\sqrt{t_n}} \right)
$$
up to a remainder of order smaller than $\frac{1}{\sqrt{n}} \epsilon^2 \left( \frac{1}{\sqrt{n}} \right)$.

This leads finally to the analogue of Theorems 4.3 and 6.2 in several dimensions:

**Theorem 7.5.** Fix a Jordan measurable part $S$ of the A-sphere $\{x, \|Ax\| = b\}$, $b > 0$; and denote $S^+ = [1, +\infty) S$. Locally uniformly in $b$,

$$P[X_n \in (A_n)^2 S^+] = \left( \frac{t_n}{2\pi} \right)^d \exp \left( -\frac{t_n b^2}{2} \right) \left( \int_S \frac{\psi(x)}{t_n b} d\mu_{\text{surface}}(dx) \right) \left( 1 + o(1) \right).$$

**Proof.** Supposing again $A = \Sigma = I_d$, it suffices to approximate the surface $S$ by polytopes $S/\sqrt{t_n}$ with the hypotheses of the previous Lemma (see the next Figure). The hypothesis of Jordan measurability will ensure the existence of such approximations (see [Spi65, p. 56] for details on the notion of Jordan measurability). More precisely, we shall take two kinds of approximations:

- outer approximations $S_{\text{ext}}$, such that the projection of $S_{\text{ext}}$ onto the sphere is included into $S$ and converges in measure to it; the disjoint union of polytopes $S_{\text{ext}}^+$ is always included into $S^+$, and

$$\lim \int_{S_{\text{ext}}} \psi(b) d\mu_{\text{surface}}(b) = \int_S \psi(b) d\mu_{\text{surface}}(b).$$

- inner approximations $S_{\text{int}}$, such that the projection of $S_{\text{int}}$ onto the sphere eventually contains $S$ and converges in measure to it; the disjoint union of polytopes $S_{\text{int}}^+$ eventually contains $S^+$, and

$$\lim \int_{S_{\text{int}}} \psi(b) d\mu_{\text{surface}}(b) = \int_S \psi(b) d\mu_{\text{surface}}(b).$$

*Figure 10. The two approximations of a part of the sphere by polytopes.*
In each case, one wants the polytopes of the approximation to be of surface measure of order $\varepsilon \left(\frac{1}{t_n}\right)^{\frac{d-1}{2}}$, so let us take an approximation with a number of polytopes of order $\varepsilon^{-1} \left(\frac{1}{t_n}\right)^{\frac{d}{2}-1} = o\left((\frac{1}{t_n})^{d/2-1}\right)$.

Then, for the outer approximations,

$$
\mathbb{P}[X_n \in (A_n)^2 S^+] \geq \mathbb{P}[X_n \in (A_n)^2 S_{\text{ext}}^+] \geq \left(\frac{t_n}{2\pi}\right)^{d/2} \exp\left(-\frac{t_n b_+^2}{2}\right) \left(\int_{S_{\text{ext}}} \psi(x) \mu_{\text{surface}}(dx)\right) \left(1 + o(1)\right)
$$

where $b_+$ is the maximum of the norms of the points that are the centers of the polytope faces of the approximation. Since the polytopes are of width $o\left(\frac{1}{\sqrt{t_n}}\right)$, one can replace $b_+$ by $b$ without changing the previous formula, which proves the lower bound

$$
\liminf_{n \to \infty} \left(\frac{2\pi}{t_n}\right)^{\frac{d}{2}} (t_n b) \mathbb{P}[X_n \in (A_n)^2 B] \geq \int_{S} \psi(x) \mu_{\text{surface}}(dx).
$$

Taking inner approximations, one obtains the same upper bound for the limsup.

\begin{proof}
These are now the same arguments as for the toy-model, cf. §7.1.
\end{proof}

\textbf{Theorem 7.6.} Let $B$ be a Borelian subset of $\mathbb{R}^d$, such that $B_{\text{min}}$ is a Jordan measurable part of the sphere of radius $b$. With the notations of Formula (21),

$$
\limsup_{n \to \infty} \left(\frac{2\pi}{t_n}\right)^{d/2} (t_n b) \exp\left(\frac{t_n b^2}{2}\right) \mathbb{P}[X_n \in (A_n)^2 B] \leq \int_{B_{\text{min}}} \psi(x) \mu_{\text{surface}}(dx)
$$

with equality if for instance $B = C$ is closed and the function $F_C$ of Lemma 7.2 is continuous at $r = b$.

\begin{proof}
These are now the same arguments as for the toy-model, cf. §7.1.
\end{proof}

\section{First Examples}

The general results of Sections 3-7 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. An explicit formula for this characteristic function, which is mainly accessible for sums of i.i.d. variables (§8.1), is fortunately not required. Therefore, the development of techniques to obtain mod-$\phi$ estimates becomes an interesting part of the work. In probabilistic number theory, this will usually be related to the Selberg-Delange method (§8.2), whereas for random combinatorial objects, we will have to combine the methods of Section 5 with some new tools (Sections 9-11). In this section, we detail examples for which the mod-$\phi$ convergence has already been proved before (cf. [JKN11, DKN11]).
8.1. **Sums of independent random variables.** Though the theory of mod-\(\phi\) convergence is meant to be used with complicate random variables (statistics of random combinatorial objects, arithmetic properties of random integers, sums of dependent random variables, etc.), it already gives interesting results for sums of independent and identically distributed random variables. Let \(X\) be a random variable in \(\mathbb{R}^d\) with entire characteristic function, and \((X_n)_{n \in \mathbb{N}}\) be independent copies of it. The sum \(S_n = \sum_{i=1}^{n} X_i\) has characteristic function \(e^{n \langle X|z\rangle}\), so \(Z_n = n^{-1/3} S_n\) converges modulo a Gaussian of mean \(n^{2/3}\mathbb{E}[X]\) and covariance matrix \(n^{1/3} \text{Cov}[X] = n^{1/3} (\mathbb{E}[X^{(i)}X^{(j)}])_{1 \leq i,j \leq n}\), with limiting function

\[
\psi(z) = \exp \left( \frac{\kappa^{(3)}(X)(z^{\otimes 3})}{6} \right).
\]

Here, \(\kappa^{(3)}(X)\) is the trilinear form

\[
u \otimes v \otimes w \mapsto \sum_{1 \leq i,j,k \leq n} \kappa^{(3)}(X^{(i)}, X^{(j)}, X^{(k)}) \nu^{(i)} v^{(j)} w^{(k)}.
\]

In particular, the following moderate deviation principle holds: assuming \(\mathbb{E}[X] = 0\) and \(X\) truly \(d\)-dimensional, i.e., \(A = \text{Cov}[X] = \mathbb{E}[X^{\otimes 2}]\) is symmetric positive definite, one has

\[
\mathbb{P}\left[ \|\Sigma S_n\| \geq n^{2/3}b \right] \approx \frac{(n^{1/6} b)^{d-2}}{(2\pi)^{d/2}} \exp \left( \frac{\kappa^{(3)}(X)(z^{\otimes 3}) b^3}{6} \right) \mu_{\text{surface}}(dz)
\]

where \(\Sigma = A^{-1}\). This estimate holds for \(b\) up to order \(o(n^{1/12})\) according to the discussion of §5.3. If \(X\) is symmetric in law (\(X\) and \(-X\) have same law), then the third cumulants vanishes and one has to look for the fourth cumulant: hence, the random walk \(S_n\) satisfies in this case the moderate deviation principle

\[
\mathbb{P}\left[ \|\Sigma S_n\| \geq n^{3/4}b \right] \approx \frac{(n^{1/4} b)^{d-2}}{(2\pi)^{d/2}} \exp \left( \frac{\kappa^{(4)}(X)(z^{\otimes 4}) b^4}{24} \right) \mu_{\text{surface}}(dz)
\]

for \(b\) up to order \(o(n^{1/20})\).

A simple consequence of these multi-dimensional results is the loss of symmetry of the random walks on \(\mathbb{Z}^d\) conditioned to be far away from the origin; this loss of symmetry has also been brought out in dimension 2 in the recent paper [Ben13]. Thus, consider the simple 2-dimensional random walk \(S_n = \sum_{i=1}^{n} X_i\), where \(X_i = (\pm 1, 0)\) or \((0, \pm 1)\) with probability 1/4 for each direction. The cumulants of \(X\) are

\[
\kappa^{(4)}((\mathbb{R} X)^{\otimes 4}) = \kappa^{(4)}((\mathbb{R} X)^{\otimes 2}, (\mathbb{R} X)^{\otimes 2}) = -\frac{1}{4},
\]

so one has mod-Gaussian convergence of \(n^{-1/4}S_n\) with parameters \(t_n = n^{1/2}\) and limiting function

\[
\psi(z) = \exp \left( -\frac{(z^{(1)})^4 + (z^{(2)})^4 + 6(z^{(1)}z^{(2)})^2}{96} \right).
\]
Therefore, for every cone $C(\theta_1, \theta_2, r n^{3/4}) = \{ \text{Re}^{i\theta} \in \mathbb{R}^2 \mid R \geq r, \theta \in (\theta_1, \theta_2) \}$, Theorem 7.5 gives the estimate:

$$
P[S_n \in C(\theta_1, \theta_2, r n^{3/4})] = e^{-\frac{n^{1/2} r^2}{2\pi}} \left( \int_{\theta_1}^{\theta_2} \psi(\text{Re}^{i\theta}) d\theta \right) (1 + o(1)).$$

This leads to the following limiting result: if $S_n = R_n e^{i\theta_n}$ with $\theta_n \in (0, 2\pi)$, then

$$
\lim_{n \to \infty} P[\theta_n \in (\theta_1, \theta_2) \mid R_n \geq r n^{3/4}] = \int_{\theta_1}^{\theta_2} \frac{F(r, \theta)}{\int_0^{2\pi} \psi(\text{Re}^{i\theta}) d\theta} d\theta
$$

with $F(r, \theta) = \frac{\exp \left( -\frac{r^4 (\sin \theta)^2}{2 \eta} \right)}{\int_0^{2\pi} \exp \left( -\frac{r^4 (\sin \theta)^2}{2 \eta} \right) d\theta}$ drawn hereafter.

This function gets concentrated around the two axes of $\mathbb{R}^2$ when $r \to \infty$, whence a loss of symmetry in comparison to the behavior of the 2-dimensional Brownian motion (the scaling limit of the random walk). In dimension $d \geq 3$, one obtains the similar result

$$
\lim_{n \to \infty} P\left[ \frac{S_n}{\|S_n\|} \in A \mid \|S_n\| \geq r n^{3/4} \right] = K(r) \int_A \exp \left( -\frac{r^4}{12 d} \sum_{1 \leq i < j \leq d} (x^{(i)} x^{(j)})^2 \right) \mu(dx)
$$

for any Jordan measurable set $A \subset S^{d-1}$. The conditional probability is therefore concentrated around the axes of $\mathbb{R}^d$. All these estimates hold up to $r = o(n^{1/12})$.

Another similar setting in which our Theorems apply readily is the Poisson approximation of a sum of Bernoulli variables with “small” parameters $p_k$. Set $X_n = \sum_{k=1}^n B(p_k)$, where $B(p)$ is the random variable equal to 1 with probability $p$ and to 0 with probability $1 - p$; and where these Bernoulli variables are supposed independent. Under the assumptions $\sum_{k=1}^\infty p_k = +\infty$ and $\sum_{k=1}^\infty (p_k)^2 < +\infty$, the Laplace transform of
X_n writes as
\[ \mathbb{E}[e^{zX_n}] = \prod_{k=1}^{n}(1 + p_k(e^z - 1)) = e^{t_n(e^z - 1)} \left( \prod_{k=1}^{\infty}(1 + p_k(e^z - 1)) e^{-p_k(e^z - 1)} \right) (1 + o(1)), \]
where \( t_n = \sum_{k=1}^{n} p_k \to \infty \). The infinite product \( \psi(z) \) is indeed convergent, because
\[
(1 + p_k(e^z - 1)) e^{-p_k(e^z - 1)} = 1 - \frac{(p_k)^2}{2} (e^z - 1) + o((p_k)^2)
\]
and \( \sum_{k=1}^{\infty} (p_k)^2 \) is finite. So, one has mod-Poisson convergence, and by Theorem 3.3 the Poisson approximation holds with precise large deviations
\[
\mathbb{P}[X_n \geq (1 + \varepsilon) t_n] \simeq \frac{e^{-t_n((1+\varepsilon) \log(1+\varepsilon)-\varepsilon)}}{\sqrt{2\pi t_n}} \frac{1+\varepsilon}{\varepsilon} \psi(\varepsilon).
\]

8.2. Logarithmic combinatorial structures. The previous example is a toy-model for the so-called logarithmic combinatorial structures, see [ABT03, FS90]. Non-trivial examples falling in this framework are the number of cycles of a random permutation (Example 3.8), and the number of distinct prime divisors of a random integer or of a random polynomial over a finite field. Denote \( \mathbb{P} \) the set of prime numbers, \( \omega(k) \) the number of distinct prime divisors of an integer \( k \), and \( \omega_n \) the random variable \( \omega(k) \) with \( k \) random integer uniformly chosen in \( |n| = \{1, 2, \ldots, n\} \). The random variable \( \omega_n \) satisfies the Erdős-Kac central limit theorem (cf. [EK40]):
\[
\frac{\omega_n - \log \log n}{\sqrt{\log \log n}} \to \mathcal{N}(0,1).
\]
Indeed, the Selberg-Delange method (see [Ten95, §2.5] and the next paragraph) yields
\[
\mathbb{E}[e^{z\omega_n}] = e^{(\log \log n + \gamma)(e^z - 1)} \Pi_{p}(e^z - 1) \Pi_{N^*}(e^z - 1) (1 + o(1)),
\]
where \( \gamma = 0.577... \) is the Euler-Mascheroni constant, and \( \Pi_{A}(x) = \prod_{p \in A}(1 + \frac{z}{p}) e^{-\frac{z}{p}} \) for \( A \) part of \( N^* = \left[ 1, +\infty \right) \). This mod-Poisson convergence yields the following large deviation result:
\[
\mathbb{P}[\omega_n \geq (1 + \varepsilon)(\log \log n + \gamma)] \simeq \frac{e^{-(\log \log n + \gamma)((1+\varepsilon) \log(1+\varepsilon)-\varepsilon)}}{\sqrt{2\pi \log \log n}} \frac{1+\varepsilon}{\varepsilon} \Pi_{\mathbb{P}}(\varepsilon) \Pi_{N^*}(\varepsilon).
\]
Similarly, denote \( \omega_{n,q} \) the number of distinct irreducible divisors of a random monic polynomial of degree \( n \) over the finite field \( \mathbb{F}_q \) with \( q \) elements. It is shown in [KN10, Theorem 6.1] that the characteristic function of \( \omega_{n,q} \) has for asymptotics:
\[
\mathbb{E}[e^{z\omega_{n,q}}] \simeq e^{(\log n)(e^z - 1)} \frac{1}{\Gamma(e^z)} \prod_{\pi} \left( 1 - \frac{1}{q_{\deg \pi}} \right) e^{z \left( 1 + \frac{e^z}{q_{\deg \pi} - 1} \right)},
\]
where the product runs over monic irreducible polynomials over \( \mathbb{F}_q \). Hence, if \( I_{q,d} = \frac{1}{d} \sum_{d \mid \mu} q^{d/\varepsilon} = \frac{q^d}{\pi} + O(q^{d/2}) \) is the number of irreducible polynomials of degree \( d \), then
\[
\mathbb{E}[e^{z\omega_{n,q}}] \simeq e^{(\log n)(e^z - 1)} \frac{1}{\Gamma(e^z)} \prod_{d=1}^{\infty} \left( 1 - \frac{1}{q^d} \right) I_{q,d} e^{z \left( 1 + \frac{e^z}{q^d - 1} \right) I_{q,d}},
\]
from which one deduces that $\mathbb{P}[\omega_{n,q} \geq (1 + \varepsilon) \log n]$ is equivalent to
\[
e^{-\log n((1+\varepsilon)\log(1+\varepsilon)-(1-\varepsilon))}\frac{\sqrt{1+\varepsilon}}{\varepsilon\Gamma(1+\varepsilon)} \prod_{d=1}^{\infty} \left(\left(1 + \frac{\varepsilon}{q^d}\right)\left(1 - \frac{1}{q^d}\right)^{\varepsilon}\right)^{-d/2}\Gamma(1+\varepsilon)\prod_{d=1}^{\infty} \left(1 + \frac{\varepsilon}{q^d}\right)^{-d/2}\Gamma(1+\varepsilon)\prod_{d=1}^{\infty} \left(1 - \frac{1}{q^d}\right)^{\varepsilon}.\]
In particular, the cardinality $q$ only plays a role in the large deviations of $\omega_{n,q}$, but not in the central limit theorem.

8.3. Additive arithmetic functions of random integers. The previous example concerning the number of prime divisors of a random integer can be generalized as follows. Let $f : \mathbb{N} \to \mathbb{R}$ be an arithmetic function with the following properties:

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

(ii) $f(p^k) = g(p,k)$ is bounded as a function of prime numbers $p$ and integers $k$ (we shall denote $C$ a bound);

(iii) $f(p) = g(p,1) = 1 + O\left(\frac{\log p}{p^\alpha}\right)$ for some $\alpha > 0$.

For the last condition, by dividing $f$ by some constant, one can assume more generally $f$ “almost constant on primes”. Denote then
\[
F(s,w) = \sum_{n=1}^{\infty} \frac{w f(n)}{n^s},
\]
the Dirichlet series of $(w f(n))_{n \in \mathbb{N}}$, which is well-defined for $w \in \mathbb{C} \setminus \mathbb{R}$ and $\sigma = \Re(s) > 1$. One forms the ratio
\[
G(s,w) = \frac{F(s,w)}{\zeta(s)^w},
\]
where $\zeta(s)$ is the Riemann zeta function $\sum_{n=1}^{\infty} \frac{1}{n^s}$. The choice of complex logarithms implied by the writing $\zeta(s)^w$ is the one of [Ten95, II.5.1]. On the other hand, it is well-known that there exists a constant $\frac{1}{2} > c > 0$ such that $\zeta(s)$ does not vanish on the domain
\[
D = \left\{ s = \sigma + i\tau \in \mathbb{C} \left| \sigma \geq 1 - \frac{c}{1 + \log^+ |\tau|} \right. \right\} \subset \left\{ s \left| \sigma > \frac{1}{2} \right. \right\},
\]
see e.g. [Ten95, Theorem 3.15]. We shall assume $c < \alpha$ in the following.

**Proposition 8.1.** Suppose $w = e^z$ with $|\Re(z)| < \frac{\log^2 2}{2c}$ (hence, $w$ stays in a circular band). The map $s \mapsto G(s,w)$ is holomorphic on the domain $D$.

**Proof.** Since $f$ is additive, one can write for $\sigma > 1$ the Euler product
\[
F(s,w) = \prod_{p \in \mathbb{P}} \left(1 + \frac{w f(p,1)}{p^s} + \frac{w f(p,2)}{p^{2s}} + \cdots \right).
\]
Notice that by hypothesis on $w$ and $s$, $1 + \frac{w f(p)}{p^s}$ is non zero for all $p \in \mathbb{P}$, since
\[
\left|\frac{w f(p)}{p^s}\right| \leq \frac{e^{C |\Re(z)| \log^2 2 c}}{2^{\Re(s)}} \leq \frac{1}{2^{\Re(s)}} = 1.
\]
Therefore, the Euler product can be rewritten as
\[ F(s, w) = \tilde{F}(s, w) \prod_{p \in \mathbb{P}} \left( 1 + \frac{w_f(p)}{p^s} \right), \]
where \( \tilde{F}(s, w) \) is uniformly convergent and holomorphic on the domain \( \sigma > \frac{1}{2} \). Similarly,
\[ \zeta(s)^w = \prod_{p \in \mathbb{P}} \left( 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots \right)^w = Z(s, w) \prod_{p \in \mathbb{P}} \left( 1 + \frac{w}{p^s} \right), \]
with \( Z(s, w) \) holomorphic on the domain \( \sigma > \frac{1}{2} \), and non-vanishing on \( D \). It remains therefore to prove that
\[ \tilde{G}(s, w) = \prod_{p \in \mathbb{P}} \left( 1 + \frac{w_f(p)}{p^s} \right)^w \]
is itself holomorphic on the domain \( D \). This is clear, since the logarithm of the general term
\[ \log \left( 1 + \frac{w_f(p)}{p^s} \right) - \log \left( 1 + \frac{w}{p^s} \right) \]
behaves as
\[ \frac{w}{p^s} (e^{z(f(p)-1)} - 1) \sim \frac{w}{p^s} O \left( \frac{z}{p^a} \right) = O \left( \frac{1}{p^\Re(s)+a} \right), \]
and \( \Re(s) + a > \Re(s) + c \geq 1 \) on the domain \( D \). □

The previous Proposition allows one to apply [Ten95, II.5, Theorem 3], which is a Tauberian theorem for \( \sum_{k \leq n} w_f(k) \). Set
\[ Y(s, w) = \frac{1}{s} \{ (s-1) \zeta(s) \}^w = \sum_{j=0}^{\infty} \frac{\gamma_j(w)}{j!} (s-1)^j, \]
which is an analytic function in \( s \) for \( |s-1| < 1 \).

**Proposition 8.2.** Denote \( X_n \) a random integer uniformly chosen in \([1, n]\). One has the mod-Poisson convergence
\[ \mathbb{E}[e^{z f(X_n)}] = \frac{1}{n} \sum_{k=1}^{n} w_f(k) = e^{(e^z-1) \log \log n} \left( \sum_{k=0}^{\nu} \frac{\lambda_k(e^z)}{(\log n)^k} + o \left( \frac{1}{(\log n)^k} \right) \right), \]
where the coefficients \( \lambda_k(w) \) are defined by
\[ \lambda_k(w) = \frac{1}{\Gamma(w-k)} \sum_{h+j=k} \frac{1}{h! j!} \frac{\partial^h G(1, w)}{\partial s^h} \gamma_j(w). \]
In particular, \( \lambda_0(w) = \lim_{s \to 1} \frac{F(s, w)}{\Gamma(w) (\zeta(s))^w} \).

**Corollary 8.3.** For \(|\varepsilon| < \frac{\log 2}{2C}\), one has the asymptotics
\[ \mathbb{P}[f(X_n) \geq (1 + \varepsilon) \log \log n] \simeq \frac{e^{-(\log \log n)((1+\varepsilon) \log (1+\varepsilon) - \varepsilon)}}{\sqrt{2\pi (\log \log n)}} \sqrt{1 + \frac{\varepsilon}{\varepsilon}} \lambda_0(\varepsilon). \]
The same estimates have essentially been obtained by Radziwill in [Rad09]. On the other hand, our reasoning makes clear the link between the notion of mod-Poisson convergence, which is a statement on the ratio
\[
\frac{\mathbb{E}[e^{zX_n}]}{(\mathbb{E}[e^{zP}])^{\lambda n}}
\]
and the Selberg-Delange method, which is a statement on the ratio
\[
\frac{F(s, e^z)}{(\zeta(s))e^z}
\]
with \(F\) the Dirichlet series of an additive arithmetic function.

8.4. Statistics of random combinatorial objects and singularity analysis. We now present the method of singularity analysis of generating series which is in spirit comparable to the Selberg-Delange method introduced above but which is used to cover a variety of situations which do not fall in the domain of applications of the Selberg-Delange method. The idea is to use contour integration techniques in order to extract the asymptotic behavior of the coefficients of some generating series of interest. As in the Selberg-Delange method described in the previous section, we shall have two complex variables and the coefficients of our generating series can be themselves the moment generating functions of some random variables. Rather than stating general and abstract results, we would rather illustrate the power of this method (and at the same time the fact that mod-Poisson convergence is genuinely a higher order central limit theorem) by considering the example of the total number of cycles of random permutations under the so called weighted probability measure. For more details the reader can look at the monograph [FS09] or the paper [FO90].

From now on we shall follow the presentation in [NZ13]. Denote \(X_n(\sigma)\) the number of disjoint cycles (including fixed points) of a permutation \(\sigma\) in the symmetric group \(S_n\). We write \(C_j(\sigma)\) for the number of cycles of length \(j\) in the decomposition of \(\sigma\) as a product of disjoint cycles. Let \(\Theta = (\theta_m)_{m \geq 1}\) be given with \(\theta_m \geq 0\). The generalized weighted measure is defined as the probability measure \(\mathbb{P}_\Theta\) on the symmetric group \(S_n\):
\[
\mathbb{P}_\Theta[\sigma] = \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}
\]
with \(h_n\) a normalization 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 generalizes 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 (as in the Ewens measure case) 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 \(\theta_m\)'s). The difficulty is arising from the fact the there is nothing such as the Feller coupling anymore and that the probabilistic methods fail to hold here. We now show how the method of singularity analysis allows us to prove mod-Poisson convergence, and hence the central limit theorem, but also distributional approximations and precise...
large deviations. We first consider the generating series

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

It is well known that

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

Note that in general $h_n$ is not known. Our goal is to obtain an asymptotic for $h_n$ and for the moment generating function of $X_n$. We note the radius of convergence of $g_\Theta(t)$. The idea of singularity analysis is to introduce a properly chosen holomorphic function $S(t, w)$ 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, and then to consider the function $F(t, w) = \exp(wg(t))S(t, w)$ with the goal of extracting precise asymptotic information (with an error term) for the coefficient of $t^n$ in the series expansion as powers of $t$ for $F(t, w)$. This can be carried out if one makes 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 8.4.** 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 12. 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

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

![Figure 12. Domain $\Delta_0(r, R, \phi)$.](image)

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

(i) $g(t)$ is of class $F(r, \theta)$,

(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

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

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 8.5, we obtain an asymptotic for $h_n$.

Lemma 8.6. Let $\theta > 0$ and assume that $g_\Theta(t)$ is of class $F(r, \theta)$. We then have

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

Proof. We have already noted that $\sum_{n=1}^\infty h_n t^n = \exp(g_\Theta(t))$. We can apply Theorem 8.5 with $g(t) = g_\Theta(t), w = 1$ and $S(t, w) = 1$. □

Now using elementary combinatorial arguments (which are detailed in [NZ13]), one can prove for each $w \in \mathbb{C}$ the following identity as formal power series:

$$\sum_{n=0}^\infty h_n \mathbb{E}_\Theta[\exp(wX_n)] t^n = \exp(e^w g_\Theta(t)).  \quad (25)$$

Using this identity and Theorem 8.5, we can show:

Theorem 8.7 (Nikeghbali-Zeindler, [NZ13]). If $g_\Theta(t)$ is of class $F(r, \theta)$, then

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

Consequently the sequence $(X_n)$ converges in the mod-Poisson sense with parameters $K + \theta \log n$ and limiting function $\frac{\Gamma(\theta)}{\Gamma(\theta e^w)}$.

Proof. An application of Theorem 8.5 yields

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

with $O(\cdot)$ uniform for bounded $w \in \mathbb{C}$. Now a combination of identity (25) and Lemma 8.6 gives the desired result. □

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 3.3 and refer to [NZ13] for the distributional approximations results.
**Proposition 8.8** (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

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

In fact an application of Theorem 3.3 would immediately yield an arbitrary long expansion for \( P[Y_n = x t_n] \) and also for \( P[Y_n \geq x t_n] \) since the speed of convergence is fast enough.

---

### 8.5. Characteristic polynomials of random matrices in a compact Lie group.

Introduce the classical compact Lie groups of type A, C, D:

- \( U(n) = \{ g \in GL(n, \mathbb{C}) \mid g g^\dagger = g^\dagger g = I_n \} \) \hspace{1cm} (unitary group)
- \( USp(n) = \{ g \in GL(n, \mathbb{H}) \mid g g^* = g^* g = I_n \} \) \hspace{1cm} (compact symplectic group)
- \( SO(2n) = \{ g \in GL(2n, \mathbb{R}) \mid g g^t = g^t g = I_{2n} ; \det g = 1 \} \) \hspace{1cm} (special orthogonal group)

where for compact symplectic groups \( g^* \) denotes the transpose conjugate of a quaternionic matrix, the conjugate of a quaternionic number \( a + ib + jc + kd \) being \( a - ib - jc - kd \). 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 \( S^1 \), and for \( g \in G \) random taken under the Haar measure

\[
\log \det(e^{i \theta} - g) =_{(law)} \log \det e^{i \theta} + \log \det(1 - g)
\]

\[
= \begin{cases} 
  n i \theta + \log \det(1 - g) & \text{if } G = U(n), \\
  2n i \theta + \log \det(1 - g) & \text{if } G = SO(2n) \text{ or } USp(n), 
\end{cases}
\]

so the study of the random characteristic polynomials \( \det(z - g) \) reduces to the study of

\[
Y_{n}^{A,C,D} = \log \det(1 - g),
\]

which is complex-valued in type A and real-valued in type C, D. In the unitary case, we shall identify \( \mathbb{C} \) and \( \mathbb{R}^2 \) in order to be able to use the multi-dimensional framework introduced in §7. The mean of \( Y_n \) is

\[
\mathbb{E}[Y_n] = \begin{cases} 
  0 & \text{for } U(n); \\
  \frac{1}{2} \log \frac{\pi n}{2} & \text{for } USp(n); \\
  \frac{1}{2} \log \frac{8 \pi n}{n} & \text{for } SO(2n),
\end{cases}
\]

and it is proved in [KN12, DKN11] (after [KS00a, KS00b, HKO01]) that \( X_n = Y_n - \mathbb{E}[Y_n] \) converges in the mod-Gaussian sense with parameters

\[
t_n = \begin{cases} 
  \frac{\log n}{2} & \text{for } U(n) \\
  \log n & \text{for } USp(n) \text{ and } SO(2n)
\end{cases}
\]
\[
\psi = \begin{cases}
\frac{G(1+(x+iy)/2)G(1+(x-iy)/2)}{G(1+x)} & \text{for } U(n) \\
\frac{G(3/2)}{G(1/2)} & \text{for } USp(n) \\
\frac{G(3/2+x)}{G(1/2+x)} & \text{for } SO(2n).
\end{cases}
\]

Here \(G\) denotes Barnes’ \(G\)-function, which is the entire solution of the functional equation \(G(z+1) = \Gamma(z) G(z)\) with \(G(1) = 1\). Thus, our large deviations theorems apply and one obtains:

**Theorem 8.9.** Fix \(x > 0\). Over \(USp(n)\),

\[
\mathbb{P}_n \left[ \det(1-g) \geq \sqrt{\frac{\pi}{2} n^{1+x}} \right] \approx \frac{G(3/2)}{x G(3/2+x) \sqrt{2\pi \log n}} \exp \left( -\frac{x^2 \log 2n}{2} \right);
\]

\[
\mathbb{P}_n \left[ \det(1-g) \leq \sqrt{\frac{\pi}{2} n^{1-x}} \right] \approx \frac{G(3/2)}{x G(3/2-x) \sqrt{2\pi \log n}} \exp \left( -\frac{x^2 \log 2n}{2} \right).
\]

Over \(SO(2n)\),

\[
\mathbb{P}_n \left[ \det(1-g) \geq \sqrt{8\pi n^{-1+x}} \right] \approx \frac{G(1/2)}{x G(1/2+x) \sqrt{2\pi \log n}} \exp \left( -\frac{x^2 \log 2n}{2} \right);
\]

\[
\mathbb{P}_n \left[ \det(1-g) \leq \sqrt{8\pi n^{-1-x}} \right] \approx \frac{G(1/2)}{x G(1/2-x) \sqrt{2\pi \log n}} \exp \left( -\frac{x^2 \log 2n}{2} \right).
\]

Finally, over \(U(n)\),

\[
\mathbb{P}_n [ |\log \det(1-g)| \geq x \log n] \approx \left( \int_0^{2\pi} \frac{G(1+x e^{i\theta}) G(1+x e^{-i\theta})}{G(1+2x \cos \theta)} \frac{d\theta}{2\pi} \right) \exp(-x^2 \log n).
\]

**Proof.** These are immediate computations by using Theorem 4.3 in type C and D, and Theorem 7.5 in type A. In this last case, we also get

\[
\mathbb{P}_n [ |\det(1-g)| \geq n^x] = \mathbb{P}_n [\Re(\log \det(1-g)) \geq x \log n]
\]

\[
\approx \frac{G(1+x)^2}{G(1+2x)} \frac{1}{\sqrt{4\pi \log n} x} \exp(-x^2 \log n)
\]

by restriction to the first variable. \(\Box\)

The analogue of Theorem 8.9 in the setting of random matrices in the \(\beta\)-ensembles, or of general Wigner matrices, has been studied in the recent paper [DE13]. 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.
9. DEPENDENCY GRAPHS AND MOD-GAUSSIAN CONVERGENCE

Dependency graphs are a classical tool in the literature to prove convergence in distribution towards a Gaussian law of the sum of partly dependent random variables. They are used in various domains, such as random graphs [JLR00, pages 147-152], random polytopes [IV07], patterns in random permutations [Bón10]. As dependency graphs give a natural framework dealing in a uniform way with different kinds of objects, a natural question is the following: when we have a dependency graph with good properties, can we obtain more precise or other results than convergence in distribution? Here is a brief presentation of the literature around this question.

- In [BR89], P. Baldi and Y. Rinott give precise estimates for the total variation distance between the relevant sequence of random variables and the Gaussian distribution.
- In [Jan04], S. Janson has established some large deviation result involving the fractional chromatic number of the dependency graph.
- More recently, H. Döring and P. Eichelsbacher have shown how dependency graphs can be used to obtain some moderate deviation principles [DE12, Section 2].

Here, we shall see a link between dependency graphs and mod-Gaussian convergence. This gives us a large collection of examples, for which the material of this article gives automatically some precise moderate deviation results. Our deviation result has a larger domain of validity than the one of Döring and Eichelsbacher — see below.

In this section, we establish a general result involving dependency graphs (Theorem 9.3). In the next two sections, we focus on examples and derive the mod-Gaussian convergence of the following renormalized statistics:

- subgraph count statistics in Erdös-Rényi random graphs (Section 10);
- random character values from central measures on partitions (Section 11).

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 $Y_{\alpha}$ indexed by a set $V$.

**Definition 9.1.** 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 $\sigma$-algebras generated by these sets are independent).

Dependency graphs are used (among other things) to bound the cumulants of $X$, as we shall see below. 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. However, this is not interesting: indeed, the sparser is the dependency
Theorem 9.2. 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 \( Y_\alpha \) have all finite moments and that there exists a constant \( A \) such that, for all \( \alpha \in V \),

\[
\|Y_\alpha\|_r = (\mathbb{E}[|Y_\alpha|^r])^{1/r} \leq A.
\]

Then, if \( X = \sum_\alpha Y_\alpha \), one has:

\[
|\kappa^{(r)}(X)| \leq C_r N (D + 1)^{r-1} A^r.
\]

This theorem is often used to prove some central limit theorem. In [DE12], Döring and Eichelsbacher have analysed Janson’s original proof and have established that the theorem holds with \( C_r = (2e)^r (r!)^3 \). Then they have 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 9.3. Theorem 9.2 holds with \( C_r = 2^{r-1} r^{-2} \).

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. The end of this section is devoted to the proof of Theorem 9.3.

9.2. Joint cumulants. There exists a multivariate version of cumulants, called joint cumulants, that we shall use to prove Theorem 9.3. 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 [JLR00, Proposition 6.16]).

9.2.1. Preliminaries: set-partitions. We denote by \([n]\) the set \( \{1, \ldots, n\} \). 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 \( \Omega(n) \) the set of set partitions of \([n]\). Then \( \Omega(n) \) may be endowed with a natural partial order: the refinement order. We say that \( \pi \) is finer than \( \pi' \) or \( \pi' \) coarser than \( \pi \) (and denote \( \pi \leq \pi' \)) if every part of \( \pi \) is included in a part of \( \pi' \).

Lastly, denote \( \mu \) the Möbius function of the poset \( \Omega(n) \). In this paper, we only use evaluations of \( \mu \) at pairs \((\pi, [n])\) (the second argument is the partition of \([n]\) in only one part, which is the maximum element of \( \Omega(n) \)), 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:

\[
\mu(\pi) = (\mu(\pi, [n]) = (-1)^{\#(\pi)-1}(\#(\pi) - 1)!. \tag{26}
\]
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

$$\kappa(X_1, \ldots, X_r) = [t_1 \cdots t_r] \log \left( \mathbb{E} \left[ e^{t_1 X_1 + \cdots + t_r X_r} \right] \right).$$

(27)

As usual, $[t_1 \cdots t_r]F$ 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, see Example 2.3. Using set-partitions, joint cumulants can be expressed in terms of joint moments, and vice-versa:

$$\mathbb{E}[X_1 \cdots X_r] = \sum_{\pi \in \Omega(r)} \prod_{C \in \pi} \kappa(X_i; i \in C);$$

(28)

$$\kappa(X_1, \ldots, X_r) = \sum_{\pi \in \Omega(r)} \mu(\pi) \prod_{C \in \pi} \mathbb{E} \left[ \prod_{i \in C} X_i \right].$$

(29)

In these equations, $C \in \pi$ shall be understood as “$C$ is a part of the set partition $\pi$”. Recall that $\mu(\pi)$ has an explicit expression given by Equation (26). 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:

$$\kappa(X_1) = \mathbb{E}[X_1] \quad ; \quad \kappa(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2].$$

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 9.4. 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 [JLR00, Proposition 6.16 (v)]. We will not need this property here. In fact, we will prove a more precise version of it, see equation (35).

9.2.3. Statement with joint cumulants. Theorems 9.2 and 9.3 have some analog with joint cumulants. Let $\{Y_\alpha\}_{\alpha \in V}$ be a family of random variables with dependency graph $G$. As in Theorem 9.2, we assume that the variables $Y_\alpha$ all have finite moments and that there exists a constant $A$ such that, for all $\alpha \in V$,

$$\|Y_\alpha\|_r = (\mathbb{E}[|Y_\alpha|^r])^{1/r} \leq A.$$

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 9.5.** 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 9.3. However, to simplify notations, we only prove the latter here.
9.3. Useful combinatorial lemmas. We start our proof of Theorem 9.3 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 notations $V(H)$ and $E(H)$ for their vertex and edge sets. For a graph $H$ and a set partition $\pi$ 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 $\pi$ (in this case we use the notation $i \sim_{\pi} j$). We introduce the following functional on graphs $H$:

$$ F_H = (-1)^{|V(H)|-1} \sum_{\pi \perp H} \mu(\pi). $$

**Lemma 9.6.** For any graph $H$, one has

$$ F_H = \sum_{E \subseteq E(H)} (-1)^{|E|-|V(H)|+1}. $$

*Proof.* To simplify notations, suppose $V(H) = \{r\}$. We denote $1_{(P)}$ the characteristic function of the property $(P)$. By inclusion-exclusion,

$$ (-1)^{|V(H)|-1} F_H = \sum_{\pi \in \Omega(r)} \left( \prod_{(i,j) \in E(H)} 1_{i \sim_{\pi} j} \right) \mu(\pi) = \sum_{\pi \in \Omega(r)} \left( \prod_{(i,j) \in E(H)} (1 - 1_{i \sim_{\pi} j}) \right) \mu(\pi) $$

$$ = \sum_{E \subseteq E(H)} \sum_{\pi \in \Omega(r)} (-1)^{|E|} \left( \prod_{(i,j) \in E} 1_{i \sim_{\pi} j} \right) \mu(\pi) $$

$$ = \sum_{E \subseteq E(H)} (-1)^{|E|} \left[ \sum_{\pi \text{ such that } \forall (i,j) \in E, i \sim_{\pi} j} \mu(\pi) \right]. $$

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 $H$. \hfill \Box

**Corollary 9.7.** 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 $e$. \hfill \Box

This induction (over-)determines $F_H$ together with the initial conditions:

$$ \begin{cases} 
F_\bullet = 1, \\
F_\circ = F_\bullet = \cdots = 0, \\
F_H = 0 \quad \text{if } H \text{ is disconnected.}
\end{cases} $$
Corollary 9.8. 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:

$$
\begin{cases}
ST\cdot = 1, \\
ST\varnothing = ST\circ \cdot = \cdots = 1, \\
ST_H = 0 \text{ if } H \text{ is disconnected}.
\end{cases}
$$

\[\square\]

Remark 9.9. If $H$ is connected, both $F_H$ and $ST_H$ are actually specializations of the bi-variate Tutte polynomial $T_H(x,y)$ of $H$ (cf. [Bol98, Chapter X]):

$$
F_H = T_H(1,0) ; \quad ST_H = T_H(1,1).
$$

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 9.10. Fix a Cayley tree $T$ of size $r$ and a graph $G$ with $N$ vertices and maximal degree $D$. 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

$$
N(D + 1)^{r-1}.
$$

Proof. Lists $(v_1, \ldots, v_r)$ as in the lemma are constructed as follows. First choose any vertex $v_1$ among the $N$ vertices. Then consider a neighbor $j$ of $1$ in $T$. As we require $G[v_1, \ldots, v_r]$ to contain $T$, they must also be neighbor in $G[v_1, \ldots, v_r]$, that is to say that $v_j = v_1$ 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 neighbors of $1$ and then for all neighbors of neighbors of $1$ and so on. \[\square\]

We have the following immediate consequence.

Corollary 9.11. Let $G$ be a graph with $n$ vertices and maximal degree $D$ and $r \geq 1$. 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

$$
r^{r-2} N(D + 1)^{r-1}.
$$
9.3.3. Spanning trees and set partitions of vertices. Recall that \( \text{ST}_H \) 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 \( \text{ST}^{\pi_i}(H) = \text{ST}_{H[\pi_i]} \) the number of spanning trees of the graph induced by \( H \) on the vertex set \( \pi_i \). We also use the multiplicative notation 
\[
\text{ST}^\pi(H) = \prod_{j=1}^t \text{ST}^{\pi_j}(H).
\]
We can also consider the contraction \( H/\pi \) of \( H \) with respect to \( \pi \). 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 \( \pi \) and, for \( i \neq j \), there are as many edges between \( i \) and \( j \) as edges between a vertex of \( \pi_i \) and a vertex of \( \pi_j \) in \( H \). Denote \( \text{ST}_\pi(H) = \text{ST}_{H/\pi} \) the number of spanning trees of this contracted graph (multiple edges are here important). This should not be confused with \( \text{ST}^\pi(H) \): in the latter, \( \pi \) is placed as an exponent because the quantity is multiplicative with respect to the part of \( \pi \).

Note that the union of a spanning tree \( \overline{T} \) of \( H/\pi \) 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 bicoloration of its edges. Edges of color 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 color 1 correspond to a collection of spanning trees \( T_i \) of \( H[\pi_i] \) (for \( 1 \leq i \leq t \)). Besides, edges of color 2 define a spanning tree \( T \) on \( H/\pi \).

Therefore, we have described a bijection between spanning trees \( T_0 \) of \( H \) with a bicoloration of their edges and triples \((\pi, \overline{T}, (T_i)_{1 \leq i \leq t})\) where:

- \( \pi \) 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 \( H/\pi \);
- 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:
\[
2^{r-1} \text{ST}_H = \sum_{\pi} \text{ST}_\pi(H) \text{ST}^{\pi}(H),
\]
where the sum runs over all set partitions \( \pi \) of \([r]\).

Our bijection is illustrated on Figure 13, with the following conventions.

- On the left, blue plain edges are edges of color 1 in the tree \( T_0 \); on the right, these blue plain edges are the edges of the spanning trees \( T_i \).
- On the left, green dashed edges are edges of color 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 \( H, H/\pi, H[\pi_i] \) but not to their spanning tree \( T_0, \overline{T}, T_i \).
Note that in this example the graph $H/\pi$ 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 $T$ is the edge $\{3,5\}$. If we had chosen another edge, the tree on the left-hand side would have been different. Hence, in the Equality (30), the multiple edges of $H/\pi$ must be taken into account: in our example, $ST_\pi(H) = 4$.

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:

$$\kappa^{(r)}(X) = \sum_{\alpha_1, \ldots, \alpha_r} \kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r}).$$

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 (29) 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 $M_{\pi}$ 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 $\pi$ 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

$$\mathbb{E} \left[ \prod_{i \in B} Y_{\alpha_i} \right] = \mathbb{E} \left[ \prod_{i \in B_1} Y_{\alpha_i} \right] \times \mathbb{E} \left[ \prod_{i \in B_2} Y_{\alpha_i} \right].$$

Therefore, $M_{\pi} = M_{\phi_{H}(\pi)}$ where $H = G[\alpha_1, \ldots, \alpha_r]$ and $\phi_{H}(\pi)$ is the refinement of $\pi$ obtained as follows: for each part $\pi_i$ of $\pi$, consider the induced graph $H[\pi_i]$ and replace $\pi_i$ by the collection of vertex sets of the connected components of $H[\pi_i]$. This construction is illustrated on Figure 14.
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 14.** Illustration of the definition of $\phi_H$.

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 $H / \pi'$ of the graph $H$ with respect to the partition $\pi'$.

**Lemma 9.12.** Let $\pi'$ 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 $\phi_H^{-1}(\pi') = \emptyset$ in this case.

If all induced graphs are connected, let us try to describe $\phi_H^{-1}(\pi')$. All set partitions $\pi$ of this set are coarser than $\pi'$, so can be seen as set partitions of the index set $[r]$ of the parts of $\pi'$. This identification does not change their Möbius functions, which depends only on the number of parts. Then, it is easy to see that $\pi$ lies in $\phi_H^{-1}(\pi')$ if and only if $\pi$ is coarser than $\pi'$ and two elements in the same part of $\pi$ never share an edge in $H / \pi'$ (here, $\pi$ is seen as a set partition of $[r]$). In other words, $\pi$ lies in $\phi_H^{-1}(\pi')$ if and only if $\pi \perp (H / \pi')$. This implies the Lemma. $\square$

Consequently,

$$\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r}) = \sum_{\pi'} (-1)^{\ell(\pi') - 1} M_{\pi'} F_{H/\pi'} \left( \prod_{i=1}^t \mathbb{1}_{H[\pi'_i]} \text{ connected} \right), \quad (33)$$

where the sum runs over all set partitions $\pi'$ of $[r]$.

9.4.2. **Bounding all the relevant quantities.** The following bound for $M_{\pi}$ follows directly from Hölder inequality and the assumption $\|Y_\alpha\|_r \leq A$ (for every $\alpha \in V$):

$$|M_{\pi}| \leq A^r. \quad (34)$$

Finally, to bound each summand $\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r})$, we shall use the following bounds:

$$|F_{H/\pi'}| \leq ST_{H/\pi'}, \quad \text{by Corollary 9.8};$$

$$\mathbb{1}_{H[\pi'_i]} \text{ connected} \leq ST_{H[\pi'_i]}.$$  

Thus, Equation (33) gives

$$|\kappa(Y_{\alpha_1}, \ldots, Y_{\alpha_r})| \leq A^r \sum_{\pi'} ST_{H/\pi'} \left( \prod_{i=1}^t ST_{H[\pi'_i]} \right) = A^r 2^{r-1} ST_H, \quad (35)$$
the last equality corresponding to Equation (30). Using Equation (31), we get that the cumulant \( \kappa^{(r)}(X) \) is smaller than \( 2^{r-1}A^r \times \text{number of couples} \((\alpha_1, \ldots, \alpha_r), T\) \) where \((\alpha_1, \ldots, \alpha_r)\) is a list of vertices of the dependency graph \( G \) and \( T \) a spanning tree of the induced subgraph \( G[\alpha_1, \ldots, \alpha_r] \).

Corollary 9.11 now ends the proof of Theorem 9.3.

9.5. **Sums of random variables with a sparse dependency graph.** An immediate application of Theorem 9.3 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 degree \( D_n - 1 \), and satisfy \( \|Y_i\|_r \leq A \) for some \( A \geq 0 \) and every \( r \geq 1 \).

Assume that \( X_n \) is not deterministic so that its variance is non-zero.

**Theorem 9.13.** We assume that the dependency graph is sparse, in the sense that \( \lim \frac{D_n}{N_n} = 0 \).

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
   \[
   \sigma^2_n = \frac{D_n}{N_n} \kappa^{(2)} \left( \frac{X_n}{D_n} \right), \quad L_n = \sigma^2_n \frac{D_n}{N_n} \kappa^{(3)} \left( \frac{X_n}{D_n} \right).
   \]
   If they have limits \( \sigma^2 > 0 \) and \( L \), then
   \[
   \frac{X_n - \mathbb{E}[X_n]}{\sqrt{\frac{D_n}{N_n}}} \left( \frac{D_n}{N_n} \right)^{1/3}
   \]
   converges in the mod-Gaussian sense with parameters \( \left( \frac{N_n}{D_n} \right)^{1/3} \) and limiting function \( \psi(z) = \exp(\frac{Lz^3}{6}) \).

3. If furthermore,
   \[
   \sigma_n^2 = \sigma^2 \left( 1 + o \left( \frac{D_n}{N_n} \right)^{1/3} \right),
   \]
   then the variable
   \[
   \frac{X_n - \mathbb{E}[X_n]}{\sqrt{\frac{D_n}{N_n}}} \left( \frac{D_n}{N_n} \right)^{1/3}
   \]
   converges in the mod-Gaussian sense with parameters \( \left( \frac{N_n}{D_n} \right)^{1/3} \) and limiting function \( \psi(z) = \exp(\frac{Lz^3}{6}) \) (the difference between (2) and (3) is that \( \sigma_n \) has been replaced by \( \sigma \)).

Of course, if (2) holds, one has therefore, for \( x > 0 \),
\[
\mathbb{P}[X_n - \mathbb{E}[X_n] \geq \left( \frac{N_n}{D_n} \right)^{2/3} \left( \frac{D_n}{N_n} \right)^{1/3} \sigma_n x] = e^{-(\frac{N_n}{D_n})^{1/3} \frac{x^2}{2}} \cdot \exp \left( \frac{Lx^3}{6} \right) \left( 1 + o(1) \right).
\]
The same holds replacing \( \sigma_n \) by \( \sigma \) if one has the assumption on the speed of convergence of \( \sigma_n \) towards \( \sigma \), given in (3) above.
These are the fluctuations at order \( O((N_n)^{2/3} (D_n)^{1/3}) \), and the fluctuations of lower order are described by the central limit theorem, see the summary diagram on Figure 15.

**Figure 15.** 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 one recovers the theory treated at the beginning of Section 8.

In the next Sections, we shall study some particular cases of Theorem 9.13. We will see that most computations are related to the asymptotic analysis of the first cumulants of \( X_n \).

10. **Subgraph count statistics in Erdős-Rényi random graphs**

In this section, we consider Erdős-Rényi model \( \Gamma(n, p_n) \) of random graphs. A random graph \( \Gamma \) 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 \( \gamma \) is a fixed graph of size \( k \), then \( X_{\gamma}^{(n)} \) is the number of copies of \( \gamma \) 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 [JLR00].

The first result on this parameter was obtained by P. Erdős and A. Rényi, cf. [ER60]. They proved that, if \( \gamma \) belongs to some particular family of graphs (called balanced), 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 generalized 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 \( \gamma \). 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. [JLR00, 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 Section 10.3.3. Denote \( \alpha_n = n^2 \) and \( \beta_n = n^{k-2} \), where \( k \) is the number of vertices of \( \gamma \). It is easy to check that

\[
\mathbb{E}[X_{\gamma}^{(n)}] = c \alpha_n \beta_n \quad \text{and} \quad \text{Var}(X_{\gamma}^{(n)}) = \sigma^2 \alpha_n (\beta_n)^2
\]

for some positive constants \( c \) and \( \sigma \) — see, e.g., [JLR00, 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 9.3, 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 10.1.** Let \( 0 < p < 1 \) and \( \gamma \) be a graph with \( k \) vertices. We consider \( X_{\gamma}^{(n)} \) the number of copies of \( \gamma \) contained in Erdős-Rényi random graph \( \Gamma(n, p) \). Let \( \alpha_n \) and \( \beta_n \) be defined as above.

1. The renormalized 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 \( \sigma \) and \( L \) are computed in §10.3.2.

2. Besides, if \( (\alpha_n)^{1/2} \ll T \ll (\alpha_n)^{3/4} \), the following precise moderate deviations principle holds:

\[
\mathbb{P} \left[ \frac{X_{\gamma}^{(n)} - \mathbb{E}[X_{\gamma}^{(n)}]}{\beta_n} \geq T \right] \sim \frac{e^{-\frac{L T^3}{6 \sigma^6 (\alpha_n)^2}}}{\sqrt{2\pi \frac{T^2}{\sigma^2 \alpha_n}}} \left( 1 + o(1) \right).
\]

A similar result has been obtained by H. Döring and P. Eichelsbacher in [DE12, 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 their proof works for \( T \ll n^{6/5} \) while ours is valid for \( T \ll n^{3/2} \).

Unfortunately, we cannot get deviation results when \( T \sim x n^2 \) for some real number \( x \); this would correspond to evaluate \( \mathbb{P}[X_{\gamma}^{(n)} > (1 + \epsilon) \mathbb{E}[X_{\gamma}^{(n)}]] \). For large deviations equivalents of

\[
\log \mathbb{P}[X_{\gamma}^{(n)} > (1 + \epsilon) \mathbb{E}[X_{\gamma}^{(n)}]],
\]
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 $n^{|k|} = n(n-1) \cdots (n-k+1)$. Let $A = (a_1, \ldots, a_k)$ be an arrangement in $[n]$ of length $k$, and $\gamma$ be a fixed graph with vertex set $[k]$. Recall that $\Gamma = \Gamma(n,p_n)$ is a random Erdős-Rényi graph on $[n]$. We denote $\delta_{\gamma}(A)$ the following random variable:

$$\delta_{\gamma}(A) = \begin{cases} 1 & \text{if } \gamma \subseteq \Gamma[a_1, \ldots, a_k]; \\ 0 & \text{else.} \end{cases}$$

(36)

Here $\Gamma[a_1, \ldots, a_k]$ denotes the graph induced by $\Gamma$ 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 $\gamma$. Then the inclusion should be understood as inclusion of edge sets.

For any graph $\gamma$ 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 10.2. It would also be natural to replace in Definition (36) 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 [JLR00, Theorem 6.52]). Notice that if $\gamma$ is a complete graph, then both definitions coincide.

10.1.2. A dependency graph of the subgraph count statistics. Fix some graph $\gamma$ 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 Bernoulli 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 $\gamma$. In particular, if, for two arrangements $A$ and $A'$, one has $|E(A) \cap E(A')| = 0$ (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 $\mathfrak{B}$ is a dependency graph for the family of variables $\{\delta_1(A)\}_{A \in \mathfrak{A}(n,k)}$:

- 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., [JLR00, Example 1.6].

All variables in this graph are Bernoulli variables and, hence, bounded by 1. Besides the graph $\mathfrak{B}$ has $N = n^k$ vertices, and is regular of degree $D$ smaller than

$$\binom{k}{2} 2 (n-2)(n-3) \cdots (n-k-1) < k^4 n^{k-2}.$$

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 9.3 and we get:

**Proposition 10.3.** Fix a graph $\gamma$ of vertex set $[k]$. For any $r \leq 1$, one has

$$|k^{(r)}(X^{\gamma}_{\kappa})| \leq 2^{r-1} r^{r-2} n^k (k^4 n^{k-2})^{r-1}.$$

10.2. **Polynomiality of cumulants.**

10.2.1. **Dealing with several arrangements.** Consider a list $(A^1, \ldots, A^m)$ 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 $G_\mathfrak{A}$ has vertex set
  $$V_k = V(G_\mathfrak{A}) = \{(t,i) \mid 1 \leq i \leq r, 1 \leq t \leq k_i\}$$
  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 $A^r$’s are arrangements, then the graph $G_\mathfrak{A}$ is endowed with a natural proper $r$-coloring, $(t,i)$ being of color $i$.
- the graph $H_\mathfrak{A}^m$ has vertex set $[r]$ and an edge between $i$ and $j$ if $|A^i \cap A^j| \geq m$.

Notice that $H_\mathfrak{A}^1$ is the contraction of the graph $G_\mathfrak{A}$ by the map $\varphi : (t,i) \mapsto i$ from the vertex set of $G_\mathfrak{A}$ to the vertex of $H_\mathfrak{A}^1$. Indeed,

$$(i,j) \in E(H_\mathfrak{A}^1) \Leftrightarrow \exists v_i \in \varphi^{-1}(i), v_j \in \varphi^{-1}(j) \text{ such that } (v_i, v_j) \in E(G_\mathfrak{A}) .$$

An example of a graph $G_\mathfrak{A}$ and of its 1- and 2-contractions $H_\mathfrak{A}^1$ and $H_\mathfrak{A}^2$ is drawn on Figure 16. For $m \geq 2$, the definition of $H_\mathfrak{A}^m$ is less common. We call it the $m$-contraction of $G_\mathfrak{A}$. The 2-contraction is interesting for us. It corresponds exactly to the graph induced by the dependency graph $\mathfrak{B}$ on the list of arrangement $\mathfrak{A}$, considered in the proof of Theorem 9.3.
Remark 10.4. 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 $G_A$ and their contractions. The dependency graph will always be called $\mathcal{B}$ 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 $G_A$. Here, we state a few consequences of this invariance property that will be useful in the next Section.

Lemma 10.5. Fix a graph $\gamma$ of size $k$, the quantity

$$\mathbb{E} \left[ \delta_\gamma(A^1) \cdots \delta_\gamma(A^r) \right]$$

depends only on the graph $G_A$ 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 (29).

□

Corollary 10.6. Fix some graph $\gamma$. Then the joint cumulant $\kappa(X^{(n)}_\gamma, \ldots, X^{(n)}_\gamma)$ is a polynomial in $n$.

Proof. Using Lemma 10.5, we can rewrite the expansion (31) as

$$\kappa(X^{(n)}_\gamma, \ldots, X^{(n)}_\gamma) = \sum_G \kappa(G) N_G,$$

(37)
where:

- the sum runs over graphs $G$ of vertex set $V_k$ 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 $\kappa(\delta_\gamma(A^1), \ldots, \delta_\gamma(A^r))$, 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 10.1. We would like to apply Proposition 5.4 to the sequence $S_n = X^{(n)}_\gamma - \mathbb{E}[X^{(n)}_\gamma]$ 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 10.3; we may even choose $C = 2k^4$.

(2) We also have to check the speed of convergence:

$$\kappa^{(2)}(S_n) = \sigma^2 \alpha_n (\beta_n)^2 (1 + o(\alpha_n^{-5/12})) ; \quad \kappa^{(3)}(S_n) = L \alpha_n (\beta_n)^3 (1 + o(\alpha_n^{-1/6})).$$ (38)

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 10.6.

Finally, the mod-Gaussian convergence follows from the observations in Section 5, and we can apply Proposition 5.4 to get the moderate deviation statement. This ends the proof of Theorem 10.1. □

Remark 10.7. Using Theorem 9.5, 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 and vectors of subgraph count statistics, in the spirit of Section 7. However, as we do not have a specific motivation for that and as 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 $\sigma^2$ and $L$. The proof above does not give an explicit value for $\sigma^2$ 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 10.8. Let $\gamma$ be a graph with $k$ vertices and $h$ edges. Then the positive number $\sigma$ appearing in Theorem 10.1 is given by

$$\sigma^2 = 2h^2 p^{2h-1} (1 - p).$$
Proof. By definition, $\sigma^2$ is the coefficient of $n^{2k-2}$ in $\kappa^{(2)}(X_\gamma^{(n)})$. As seen in Equation (37), 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 (35) — 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 17.

![Figure 17. Graphs involved in the computation of the main term in $\kappa^{(2)}(X_\gamma^{(n)})$.](image)

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 graph, $N_G = n(n-1)\ldots(n-2k+3) = n^{2k-2}(1 + o(1))$. Therefore, the coefficient of $n^{2k-2}$ in $\kappa^{(2)}(X_\gamma^{(n)})$ is $2h^2 p^{2h-1}(1-p)$, as claimed. \qed

The number $L$ can be computed by the same method.

**Lemma 10.9.** Let $\gamma$ be a graph with $k$ vertices and $h$ edges. Then the number $L$ appearing in Theorem 10.1 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 18.

![Figure 18.](image)

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 12$h^3(h-1)$ 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 4$h^3$ 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.

**Example 10.10.** Denote $T_n$ the number of triangles in a random Erdős-Rényi graph $\Gamma(n, p)$ (each triangle being counted 6 times). According to the previous Lemmas, the parameters $\sigma^2$ and $L$ are respectively

$$\sigma^2 = 18 p^5 (1-p) \quad \text{and} \quad L = 108 p^7 (1-p)(7-8p).$$

Moreover, $\mathbb{E}[T_n] = n^{13} p^3 = n^3 p^3 - 3 n^2 p^3 + O(n)$. So,

$$\mathbb{P}[T_n \geq n^3 p^3 + n^2 (v - 3p^3)] \approx \sqrt{\frac{9p^5(1-p)}{\pi v^2}} \exp\left( -\frac{v^2}{36 p^5(1-p)} + \frac{(7-8p) v^3}{324 n p^6(1-p)^2} \right)$$

for $1 \ll v \ll n^{1/2}$.

### 10.3.3. Case of a non-constant sequence $p_n$. Proposition 10.3 still holds when $p_n$ is a non-constant sequence (the particularly interesting case is $p_n \to 0$). One can even sharpen a little this bound by replacing Inequality (34) by the trivial bound

$$\left| \kappa^{(r)}(X^{(n)}_\gamma) \right| \leq C^r r^{-2} n^{r(k-2)+2} (p_n)^h.$$

But, unlike in the case $p_n = p$ constant, this bound is not always optimal (up to a multiplicative constant) for a fixed $r$. For example, if $p_n = n^{-\epsilon}$ with $\epsilon > 0$ sufficiently small, then — see [JLR00, Lemma 3.5] —

$$\text{Var}(X^{(n)}_\gamma) \sim \text{const.} \times n^{2k-2}(p_n)^{2h-1} \ll n^{2k-2}(p_n)^h.$$

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 some deviation result. Let $\alpha_n$ and $\beta_n$ 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

$$
\left| \kappa^{(r)}(X_{\gamma}^{(n)}) \right| \leq (C^4 r')^r \alpha_n(\beta_n)^r.
$$

Unfortunately the convergence speed hypotheses (38) are not always satisfied, so one cannot apply Proposition 5.4 in general. However, one can see 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 $p_n$ is of order $n^{-\epsilon}$ with $0 < \epsilon < 1/6h$, then Conditions (38) are satisfied and Theorem 10.1 still holds in that case.

11. RANDOM CHARACTER VALUES FROM CENTRAL MEASURES ON PARTITIONS

Our Theorem 9.3 can also be used to study certain models of random integer partitions. Recall that if $G$ is a finite group and if $\tau$ 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 $\tau$ can be expanded uniquely as a linear combination of normalized irreducible characters:

$$
\tau = \sum_{\lambda \in \hat{G}} P_{\tau}[\lambda] \hat{\chi}^\lambda,
$$

where $\hat{G}$ is the (finite) set of isomorphism classes of irreducible representations of $G$ and $\hat{\chi}^\lambda$ the normalized (i.e. divided by the dimension of the space) character of the irreducible representation associated to $\lambda$.

**Definition 11.1.** The map $\lambda \mapsto P_{\tau}[\lambda]$ is called the spectral measure of the trace $\tau$. 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 $\hat{G}$.

When $G = \mathcal{S}(n)$ is the symmetric group of order $n$, the irreducible representations are indexed by integer partitions of size $n$, that is non-increasing sequences of positive integers $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ with $\sum_{i=1}^{\ell} \lambda_i = n$. In the following we denote $\mathcal{P}(n) = \hat{\mathcal{G}}(n)$ the set of integer partitions of size $n$, and $\ell(\lambda) = \ell$ the length of a partition. We shall also consider the infinite symmetric group $\mathcal{G}(\infty) = \bigcup_{n \geq 1} \mathcal{S}(n)$, which is the group of permutations of the set of natural numbers that have a finite support.

**Definition 11.2.** A central measure on partitions is a family $(P_{\tau,n})_{n \in \mathbb{N}}$ of spectral measures on the sets $\mathcal{P}(n)$ that come from the same trace of the infinite symmetric group $\mathcal{G}(\infty)$. In other words $(P_{\tau,n})_{n \in \mathbb{N}}$ is a central measure if there exists a trace $\tau : \mathcal{G}(\infty) \to \mathbb{C}$ such that

$$
\tau|_{\mathcal{S}(n)} = \sum_{\lambda \in \mathcal{P}(n)} P_{\tau,n}[\lambda] \hat{\chi}^\lambda.
$$
Example 11.3. The regular trace $\tau(\sigma) = 1_{\sigma = \text{id}}$ corresponds to the Plancherel measures of the symmetric groups, given by the formula $P_n[\lambda] = (\dim V^\lambda)^2 / n!$, where $V^\lambda$ is the $\mathfrak{S}(n)$-irreducible module of label $\lambda$. They have been extensively studied in connection with Ulam’s problem of the longest increasing subsequence and with random matrix theory, see e.g. [BDJ99, BOO00, Oko00, IO02].

A central measure $(P_{\tau,n})_{n \in \mathbb{N}}$ 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]. Call extremal a non-negative trace on $\mathfrak{S}_\infty$ that is not a positive linear combination of 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 $\omega$ is given by

$$\tau_\omega(\sigma) = \prod_{c \in C(\sigma)} p_{|c|}(\omega) \quad \text{with} \quad p_1(\omega) = 1, \quad p_{k \geq 2}(\omega) = \sum_{i=1}^\infty (\alpha_i)^k + (-1)^{k-1}(\beta_i)^k, \quad (39)$$

$C(\sigma)$ denoting the set of cycles of $\sigma$. 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 $\lambda$ chosen according to the central measure $P_{\omega,n}$ 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, Mé12]. The aim of this section is to use the techniques of §9-10 in order to prove the following:

Theorem 11.4. Fix a parameter $\omega \in \Omega$, and denote $\rho$ a $k$-cycle in $\mathfrak{S}(\infty)$ (for $k \geq 2$). Assume $p_{2k-1}(\omega) - (p_k(\omega))^2 \neq 0$. Denote then $X^{(n)}_p$ the random character value $\widehat{\chi}^\lambda(\rho)$, where:

- $\lambda \in \mathfrak{P}(n)$ is picked randomly according to the central measure $P_{\omega,n}$,
- and $\widehat{\chi}^\lambda$ denotes the normalized irreducible character indexed by $\lambda$ of $\mathfrak{S}(n)$, that is

$$\widehat{\chi}^\lambda(\rho) = \frac{\text{tr} \Pi^\lambda(\rho)}{\dim V^\lambda} \quad \text{with} \quad (V^\lambda, \Pi^\lambda) \text{ irreducible representation of } \mathfrak{S}(n).$$

The rescaled random variable $n^{2/3} (X^{(n)}_p - \tau_\omega(\rho))$ 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

$$\sigma^2 = k^2 \left( p_{2k-1}(\omega) - (p_k(\omega))^2 \right);$$

$$L = k^3 \left( (3k - 2) p_{3k-2}(\omega) - (6k - 3) p_{2k-1}(\omega) p_k(\omega) + (3k - 1) p_k(\omega)^3 \right).$$

Remark 11.5. The condition $\sigma^2 > 0$ is satisfied as soon as the sequence $\alpha \sqcup \beta$ associated to the Thoma parameter $\omega$ contains two different non-zero coordinates. Indeed, for
any summable non-increasing non-negative sequence \(\gamma = (\gamma_1, \gamma_2, \ldots)\) 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 \(\varepsilon\) on the left-hand side is

\[
\sum_{i<j} (\gamma_i \gamma_j)^k \log \left(\frac{\gamma_i}{\gamma_j}\right) \left(\left(\frac{\gamma_i}{\gamma_j}\right)^\varepsilon - (\frac{\gamma_j}{\gamma_i})^\varepsilon\right).
\]

Applying the result to \(\gamma = \alpha \uplus \beta\) 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 \(0 < \sum_{i=1}^{\infty} \alpha_i + \beta_i < 1\).

**Corollary 11.6.** The random character value \(X_k^{(n)} = X_\rho^{(n)}\) with \(\rho\) a \(k\)-cycle satisfies the principle of moderate deviations

\[
P\left[|X_k^{(n)} - p_k(\omega)| \geq n^{-1/2} x\right] \leq \frac{2 e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi \frac{x^2}{\sigma^2}}} \cosh \left(\frac{L x^3}{6 \sigma^6 n^{1/2}}\right) \left(1 + o(1)\right),
\]

for \(1 \ll x \ll n^{1/4}\), with \(\sigma^2\) and \(L\) as in Theorem 11.4.

This Section is organized as follows. In §11.1, we present the necessary material. In §11.2, we then prove bounds on these cumulants similar to those of Proposition 10.3, and we compute the limits of the second and third cumulants. This will allow us to use in §11.3 the framework of §5.3 in order to prove the results stated above. We shall also detail some consequences of these results for the shapes of the random partitions \(\lambda \sim \mathcal{P}_{n,\omega}\) viewed as Young diagrams, in the spirit of [FM12, Mél12].

11.1. **Preliminaries.**

11.1.1. **Non-commutative probability theory.** The originality of this section 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 11.7.** A non-commutative probability space is a complex unital algebra \(\mathcal{A}\) with some linear functional \(\varphi : \mathcal{A} \to \mathbb{C}\) such that \(\varphi(1) = 1\).

This generalizes the notion of probability space: elements of \(\mathcal{A}\) should be thought as random variables, while \(\varphi\) should be thought as the expectation. The difference is that, unlike random variables in usual probability theory, elements of \(\mathcal{A}\) are not assumed to commute. Usually one also assumes that \(\varphi\) is tracial, i.e., \(\varphi(ab) = \varphi(ba)\) for any \(a, b \in \mathcal{A}\).
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 11.8.** Two subalgebras $A_1$ and $A_2$ of $A$ are independent if and only if, for any sequence $a_1, \ldots, a_r$ with, for each $i$, $a_i \in A_1$ or $a_i \in A_2$, one has

$$\varphi(a_1 \ldots a_r) = \prod_{1 \leq i \leq r, a_i \in A_1} \varphi(a_i) \prod_{1 \leq i \leq r, a_i \in A_2} \varphi(a_i).$$

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 $A$, we set

$$\kappa(a_1, \ldots, a_r) = \sum_{\pi} \mu(\pi) \prod_{B \in \pi} \varphi\left(\prod_{i \in B} a_i\right).$$

Note that, in the proof of Theorem 9.3, independence is only used in equation (32). By definition of tensor independence, equation (32) also holds in the non-commutative setting and hence, so does Theorem 9.3.

### 11.1.2. Two probability spaces

From now on, we fix an element $\omega$ in the Thomas simplex $\Omega$. Denote $\mathbb{C}S(n)$ the group algebra of $S(n)$. The function $\tau_\omega$, defined by Equation (39), can be linearly extended to $\mathbb{C}S(n)$. Then $(\mathbb{C}S(n), \tau_\omega)$ is a non-commutative probability space.

Note that we are now working with two probability spaces: the non-commutative probability space $(\mathbb{C}S(n), \tau_\omega)$ 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 $\mathbb{C}S(n)$ and define the random variable (for the usual probability space) $X_y$ by $X_y(\lambda) = \hat{\chi}(y)$. Then one has

$$\mathbb{E}_{\mathbb{P}_{\omega, n}}(X_y) = \sum_{\lambda \vdash n} \mathbb{P}_{\omega, n}(\lambda) \hat{\chi}(y) = \tau_\omega(y).$$

In other sense, 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 $\tau_\omega$, is considered as an expectation).

Besides, if we restrict to the center of $(\mathbb{C}S(n), \tau_\omega)$, 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 $(\mathbb{C}S(n), \tau_\omega)$, their joint moments (and joint cumulants) are the same as those of $X_{y_1}, \ldots, X_{y_r}$.
11.1.3. Renormalized 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

\[
\Sigma_{\mu,n} = \sum_A \rho_{\mu}(A) \text{ where } \rho_{\mu}(A) = (a_1, \ldots, a_{\mu_1})(a_{\mu_1+1}, \ldots, a_{\mu_1+\mu_2}) \ldots.
\]

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 \( n^{\left< k \right.} \) elements of cycle-type \( \mu \), hence, if \( \rho \) is a fixed permutation of type \( \mu \), one has:

\[
X_{\Sigma_{\mu,n}} = n^{\left< k \right.} X_{\rho}.
\]

Note that considering these elements \( \Sigma_{\mu,n} \) and their normalized characters is a classical trick in the study of central measures on Young diagrams; see [IO02, Šni06b, 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.

\[
n^{\left< k_1 \right.} \ldots n^{\left< k_r \right.} \kappa\left( X_{\rho_1}^{(n)}, \ldots, X_{\rho_r}^{(n)} \right) = \kappa(\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r}) = \sum_{A^1 \in \mathfrak{A}(n,k_1)} \kappa\left( \rho_{\mu^1}(A^1), \ldots, \rho_{\mu^r}(A^r) \right) \quad (40)
\]

As in the framework of subgraph count statistics, the invariance of \( \tau_\nu \) 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 \( G_\mathfrak{A} \) associated to the family of arrangement \( \mathfrak{A} = (A^1, \ldots, A^r) \). Copying the proof of Corollary 10.6, we get:

**Lemma 11.9.** Fix some integer partitions \( \mu^1, \ldots, \mu^r \). Then the rescaled joint cumulant

\[
\kappa(\Sigma_{\mu^1}, \ldots, \Sigma_{\mu^r}) = n^{\left< k_1 \right.} \ldots n^{\left< k_r \right.} \kappa\left( X_{\rho_1}^{(n)}, \ldots, X_{\rho_r}^{(n)} \right)
\]

is a polynomial in \( n \).

### 11.2. Bounds and limits of the cumulants.

11.2.1. The dependency graph of the 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 (39), 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 \( \mathcal{B} \) 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 $\mathcal{B}$ is obviously regular with degree strictly smaller than $k^2 n^{i_k-1}$. On the other hand, all joint moments of the family $(\rho_{\mu}(A))_{A \in \mathcal{A}(n,k)}$ are normalized characters of single permutations and hence bounded by 1 in absolute value. So one can once again apply Theorem 9.3 and we get:

**Proposition 11.10.** Fix a partition $\mu$ of size $k$. For any $r \leq 1$, one has

$$|\kappa^{(r)}(\Sigma_{\mu})| \leq 2^{r-1}r^{r-2}n^{i_k} (k^2 n^{i_k-1})^{r-1};$$

$$|\kappa^{(r)}(X^{(n)}_{\rho_{\mu}})| \leq r^{r-2} \left(\frac{2k^2}{n}\right)^{r-1}.$$  

**Remark 11.11.** Using Theorem 9.5, one can also obtain a bound for joint cumulants of $X^{(n)}_{\rho_{\mu}}$, namely,

$$|\kappa(X^{(n)}_{\rho_{\mu_1}}, \ldots, X^{(n)}_{\rho_{\mu_k}})| \leq k_1 \cdots k_r (r \cdot \max_{1 \leq i \leq r} k_i)^{r-2} \left(\frac{2}{n}\right)^{r-1}$$

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 $X^{(n)}_{\rho_{\mu}}$, it will be a bit clearer to manipulate joint cumulants of variables $\Sigma_{\mu_1}, \ldots, \Sigma_{\mu_k}$ with arbitrary integer partitions.

11.2.2. **Limits of the second and third cumulants.** Because of Lemma 11.9 and Proposition 11.10, for any fixed integer partitions,

$$\kappa(X^{(n)}_{\rho_{\mu_1}}, \ldots, X^{(n)}_{\rho_{\mu_k}}) n^{1-r} \simeq \kappa(\Sigma_{\mu_1}, \ldots, \Sigma_{\mu_k}) n^{k_1+\cdots+k_r-(r-1)}$$

converges to a constant. Let us compute this limit when $r = 2$ or 3; we use the same reasoning as in §10.3.2.

As $\kappa$ is invariant by simultaneous conjugacy of its arguments, the summand in Equation (40) depends only on the graph $G = G_{\mathcal{A}}$ associated to the collection $\mathcal{A} = (A_1, \ldots, A^\ell)$, 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 Section 10, $N_G$ denotes the number of list $\mathcal{A}$ of arrangements with associated graph $G$.

When $r = 2$, we have to look for graphs $G$ on vertex set $V_k = [k_1] \cup [k_2]$ 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

$$(\mu_1 \ast \mu_2)(i, j) = (\mu_1 \setminus \mu_i^1) \sqcup (\mu_2 \setminus \mu_j^2) \sqcup \{\mu_i^1 + \mu_j^2 - 1\}.$$  

This is the cycle type of a permutation $\rho_{\mu_1}(A^1) \rho_{\mu_2}(A^2)$, where $G_{\mathcal{A}}$ is the graph with one edge joining an element of $A^1$ in the cycle of length $\mu_i^1$ with an element of $A^2$ in the
cycle of length $\mu_j^2$. These graphs are the only ones involved in our computation, and they yield

$$\kappa(G) = p_{(\mu_1 \times \mu_2)(i,j)}(\omega) - p_{\mu_1 \cup \mu_2}(\omega),$$

where for a partition $\mu$ we denote $p_\mu(\omega)$ the product $\prod_{i=1}^{\ell(\mu)} p_{(i)}(\omega)$. So,

**Proposition 11.12.** For any partitions $\mu$ and $\nu$, the limit of $n \kappa(X_{\rho_\mu}^{(n)}, X_{\rho_\nu}^{(n)})$ is

$$\sum_{i=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \mu_i \nu_j \left( p_{(\mu \times \nu)(i,j)}(\omega) - p_{\mu \cup \nu}(\omega) \right).$$

In particular, for cycles $\mu = (k)$ and $\nu = (l)$,

$$\lim_{n \to \infty} n \kappa(X_k^{(n)}, X_l^{(n)}) = kl \left( p_{k+l-1}(\omega) - p_{k,l}(\omega) \right).$$

On the other hand, if $\mu = \nu$, then the limit of $n \kappa^{(2)}(X_{\rho_\mu}^{(n)})$ is

$$(p_\mu(\omega))^2 \sum_{1 \leq i,j \leq \ell(\mu)} \mu_i \mu_j \left( \frac{p_{(\mu \times \mu^{-1})(i,j)}(\omega)}{p_\mu(\omega) p_{\mu^{-1}}(\omega)} - 1 \right).$$

When $r = 3$, we look for graphs $G$ on vertex set $V_k = [k_1] \sqcup [k_2] \sqcup [k_3]$ 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 $\rho_{\mu^1}(A^1) \rho_{\mu^2}(A^2) \rho_{\mu^3}(A^3)$ a permutation of cycle type $(\mu_1 \times \mu_2 \times \mu_3)(i,j,k) = (\mu_1 \setminus \mu_1^1) \sqcup (\mu_2 \setminus \mu_2^2) \sqcup (\mu_3 \setminus \mu_3^3) \sqcup \{\mu_i^1 + \mu_j^2 + \mu_k^3 - 2\};$

    and the corresponding cumulant is

    $$\kappa(G) = p_{\mu_1 \cup \mu_2 \cup \mu_3}(\omega) + p_{(\mu_1 \times \mu_2 \times \mu_3)(i,j,k)}(\omega)$$

    $$- p_{((\mu_1 \times \mu_2)(i,j)) \cup \mu_3}(\omega) - p_{((\mu_2 \times \mu_3)(j,k)) \cup \mu_1}(\omega).$$

    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 $a_1^1 = a_2^2 = a_3^3$. This gives again for $\rho_{\mu^1}(A^1) \rho_{\mu^2}(A^2) \rho_{\mu^3}(A^3)$ a permutation of cycle type $(\mu_1 \times \mu_2 \times \mu_3)(i,j,k)$, but the corresponding cumulant takes now the form

    $$\kappa(G) = 2 p_{\mu_1 \cup \mu_2 \cup \mu_3}(\omega) + p_{(\mu_1 \times \mu_2 \times \mu_3)(i,j,k)}(\omega) - p_{((\mu_1 \times \mu_2)(i,j)) \cup \mu_3}(\omega)$$

    $$- p_{((\mu_2 \times \mu_3)(j,k)) \cup \mu_1}(\omega) - p_{((\mu_1 \times \mu_3)(i,k)) \cup \mu_2}(\omega).$$

    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

$$\left(\mu^1 \times \mu^2 \times \mu^3\right)(i; j; k, l)$$

$$= (\mu^1 \setminus \mu^1_j) \sqcup (\mu^2 \setminus \{\mu^2_j, \mu^2_k\}) \sqcup (\mu^3 \setminus \mu^3_j) \sqcup \{\mu^1_j + \mu^2_j - 1, \mu^2_k + \mu^3 - 1\}.$$ 

The cumulant corresponding to this last case is

$$\kappa(G) = p_{\mu^1\setminus\mu^2\setminus\mu^3}(\omega) + p_{(\mu^1\mu^2\mu^3)(i; j; k, l)}(\omega)$$

$$- p_{((\mu^1\mu^2)(i; j))\setminus\mu^3}(\omega) - p_{((\mu^2\mu^3)(i; j))\setminus\mu^1}(\omega),$$

and again one can permute cyclically the indices 1, 2, 3 to get 3 different graphs.

Consequently:

**Proposition 11.13.** For any partitions $\mu, \nu$ and $\delta$, the limit of $n^2 \kappa(X_{\mu}^{(n)}, X_{\nu}^{(n)}, X_{\mu_{\nu}}^{(n)})$ is

$$\sum_{Z/3Z} \left( \sum_{l=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \sum_{k=1}^{\ell(\delta)} \mu_i \nu_j \delta_k \left( -p_{\mu\nu\mu\nu\delta}(\omega) + p_{(\mu\mu\nu\nu\delta)(i; j; k, l)}(\omega) \right. \right.$$ 

$$+ \sum_{l=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \sum_{k=1}^{\ell(\delta)} \mu_i \nu_j \nu_k \delta_l \left( -p_{\mu\nu\mu\nu\delta}(\omega) + p_{(\mu\mu\nu\nu\delta)(i; j; k, l)}(\omega) \right)$$

$$+ \sum_{l=1}^{\ell(\mu)} \sum_{j=1}^{\ell(\nu)} \sum_{k=1}^{\ell(\delta)} \mu_i \nu_j \delta_l \left( -p_{\mu\nu\mu\nu\delta}(\omega) - p_{(\mu\mu\nu\nu\delta)(i; j; k, l)}(\omega) \right),$$

where $\sum_{Z/3Z}$ means that one permutes cyclically the partitions $\mu, \nu, \delta$.

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

$$k\ell m \left( (k + l + m - 1) p_{k; l; m}(\omega) + (k + l + m - 2) p_{k+l+m-2}(\omega) \right. \right.$$ 

$$- (k + l - 1) p_{k+l-1; m}(\omega) - (l + m - 1) p_{1; m-1; k}(\omega) - (l + m - 1) p_{k+m-1; l}(\omega).$$

One recovers for $k = l = m$ the values of $\sigma^2$ and $L$ announced in Theorem 11.4. On the other hand, if $\mu = \nu = \delta$, then the limit of $n^2 \kappa^{(3)}(X_{\mu}^{(n)})$ is

$$3 \left( p_{\mu}(\omega) \right)^3 \left( \sum_{1 \leq i, j, k \leq \ell(\mu)} \mu_i \mu_j \mu_k \left( 1 + \frac{p_{\mu_i+\mu_j+\mu_k-2}(\omega)}{p_{\mu_i}(\omega) p_{\mu_j}(\omega) p_{\mu_k}(\omega)} - \frac{p_{\mu_i+\mu_j-1}(\omega)}{p_{\mu_i}(\omega) p_{\mu_j}(\omega)} \right) \right.$$ 

$$+ \sum_{1 \leq i, j, k, l \leq \ell(\mu)} \mu_i \mu_j \mu_k \mu_l \left( 1 - \frac{p_{\mu_i+\mu_j-1}(\omega)}{p_{\mu_i}(\omega) p_{\mu_j}(\omega)} \right) \left( 1 - \frac{p_{\mu_k+\mu_l-1}(\omega)}{p_{\mu_k}(\omega) p_{\mu_l}(\omega)} \right) \right)$$

$$+ \left( p_{\mu}(\omega) \right)^3 \sum_{1 \leq i, j, k \leq \ell(\mu)} \mu_i \mu_j \mu_k \left( 3 \frac{p_{\mu_i+\mu_j-1}(\omega)}{p_{\mu_i}(\omega) p_{\mu_j}(\omega)} - 2 \frac{p_{\mu_i+\mu_j+\mu_k-2}(\omega)}{p_{\mu_i}(\omega) p_{\mu_j}(\omega) p_{\mu_k}(\omega)} - 1 \right).$$
11.3. Asymptotics of the random character values and partitions. Fix an integer \( k \geq 2 \), and consider the random variable

\[
V_{n,k} = n^{2/3} \left( X^{(n)}_{\rho} - p_k(\omega) \right),
\]

where \( \rho \) 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 \( \sigma \) and \( L \) are the quantities given in the statement of Theorem 11.4. This ends the proof of Theorem 11.4 and Corollary 11.6. 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 11.9; therefore, one can indeed apply Proposition 5.4 with \( \alpha_n = n \) and \( \beta_n = n^{-1} \). The theorem can be extended to other permutations \( \rho \) of \( \mathfrak{S}(\infty) \) than cycles: if \( \rho \) has cycle-type \( \mu \), then define

\[
V_{n,\mu} = n^{2/3} \left( X^{(n)}_{\rho} - p_{\mu}(\omega) \right).
\]

Then 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 \( \sigma^2(\mu) = \lim_{n \to \infty} n \kappa^{(2)}(X^{(n)}_{\rho}) \) and \( L(\mu) = \lim_{n \to \infty} n^2 \kappa^{(3)}(X^{(n)}_{\rho}) \) are the limiting quantities given in §11.2.2. 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 11.5. Under these conditions, one can also easily establish mod-Gaussian convergence for every vector of renormalized random character values \((V_{n,\mu_1}, \ldots, V_{n,\mu_\ell})\).

Remark 11.14. 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

\[
\frac{n^{k/2} \hat{\chi}^\lambda(c_k)}{\sqrt{k}}
\]

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 renormalizations 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 [´Sni06a].
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 $\lambda = (\lambda_1, \ldots, \lambda_i)$ of size $n$, the Frobenius coordinates of $\lambda$ 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 $\lambda'_1, \lambda'_2, \ldots$ are the sizes of the columns of the Young diagram of $\lambda$, 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
\[
X_\lambda = \sum_{i=1}^{d} \frac{a_i}{n} \delta_{(\frac{a_i}{n})} + \sum_{i=1}^{d} \frac{b_i}{n} \delta_{(-\frac{b_i}{n})}.
\]
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 $X_\lambda$ encodes the geometry of the Young diagram $\lambda$. 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 $X_\lambda$. Here, $\gamma = 1 - \sum_{i=1}^{\infty} \alpha_i - \sum_{i=1}^{\infty} \beta_i$. Notice that $E_{n,\omega}[X_k^{(n)}] = \tau_\omega(c_k) = X_\omega(x^{k-1})$.

It is shown in [IO02] that for any partition $\lambda$ of size $n$ and for any $k$,
\[
p_k(\lambda) = \Sigma_k(\lambda) + \text{remainder},
\]
where the remainder is a linear combination of symbols $\Sigma_\mu$ with $|\mu| < k$. It follows that the cumulants of the $p_k$’s satisfy the same estimates as the cumulants of the $\Sigma_k$’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 $\psi(z) = \exp(L \frac{z^3}{\sigma})$, where $\sigma^2$ and $L$ are given by the same formula as in the case of the random character value $X_k^{(n)}$, that is to say
\[
\sigma^2 = k^2 (p_{2k-1}(\omega) - p_k(\omega)^2);
\]
\[
L = k^3 ((3k - 2) p_{3k-2}(\omega) - (6k - 3) p_{2k-1}(\omega) p_k(\omega) + (3k - 1) p_k(\omega)^3).
\]
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 moderate deviations 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 $a_1 > a_2 > \cdots > a_i > \cdots$ and $b_1 > b_2 > \cdots > b_i > \cdots$ (see [Mél12], and also [Buf12]). Indeed, for any smooth test function $\psi_i$ equal to 1 around $\alpha_i$ and to 0 outside a neighborhood of this point, one has
\[
X_\lambda(\psi_i) - X_\omega(\psi_i) = \frac{a_i}{n} - \alpha_i
\]
(41)
with probability going to 1, and then the quantities in the left-hand side renomalized by \( \sqrt{n} \) converge jointly towards a Gaussian vector with covariance

\[
\kappa(i, j) = \delta_{ij} \alpha_i - \alpha_i \alpha_j.
\] (42)

So, the fluctuations

\[
\sqrt{n} \left( \frac{\lambda_i}{n} - \alpha_i \right)
\]

of the rows of the random partitions taken under central measures \( P_{n, \omega} \) converge jointly towards a gaussian vector with covariances given by Equation (42), 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 (41), 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 \textit{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 \( \alpha_i \) and \( \beta_j \) 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 generalized riffle shuffle.

12. Appendix: Berry-Esseen estimates for mod-convergence in \( \mathbb{R}^d \)

In this additional Section, we place ourselves in the setting of multi-dimensional mod-Gaussian convergence, and describe various Berry-Esseen estimates that are cruder than the one of Theorem 7.3, but hold in more general settings. Thus, we denote \( X_n \) a sequence of random variables in \( \mathbb{R}^d \) such that

\[
\phi_n(z) = \mathbb{E} \left[ e^{\sum_{i=1}^d X_n(z^{(i)})} \right] = e^{\frac{1}{2} \sum_{i=1}^d (z^{(i)})^2} \psi(z) \left( 1 + o(1) \right)
\]

locally uniformly on \( \mathbb{C}^d \). We then set \( Y_n = X_n / \sqrt{n} \), and denote \( \mu_n \) the law of \( Y_n \), and we want to estimate the difference \( \mu_n - G \), where \( G \) is the standard \( d \)-dimensional Gaussian distribution, of density \( g(x) \).
12.1. Properties of certain smoothing kernels. A preliminary step for our computations consists in smoothing the distribution of the $Y_n$’s. To this purpose, we shall use standard methods of convolution. Set

$$
\Delta_T^{(k)}(x) = \frac{1}{I(k)} \left( \frac{1 - \cos Tx}{Tx^2} \right) \left( \frac{\sin Tx}{Tx} \right)^{2k} = \frac{\left( 1 - \cos \frac{Tx}{T^2} \right) \left( \frac{\sin \frac{Tx}{T}}{\frac{Tx}{T}} \right)^{2k}}{\int_{-\infty}^{\infty} \left( \frac{1 - \cos \frac{x}{T^2}}{\frac{x}{T^2}} \right) \left( \frac{\sin \frac{x}{T}}{\frac{x}{T}} \right)^{2k} dx};
$$

and in dimension $d$, $\Delta_{T,R^d}^{(k)}(x) = \prod_{i=1}^{d} \Delta_T^{(k)}(x^{(i)})$. One defines the convolution of $\Delta_{T,R^d}^{(k)}$ with a test function $F$ by

$$
(\Delta_{T,R^d}^{(k)} * F)(x) = \int_{R^d} F(x - y) \Delta_{T,R^d}^{(k)}(y) dy.
$$

Since $\Delta_{T,R^d}^{(k)}$ is non-negative and of total mass $\int_{R^d} \Delta_{T,R^d}^{(k)}(y) dy = 1$, if $\mu$ is a probability measure, then one can define a positive normalized linear form on the space $C_c(R^d)$ of continuous compactly supported functions by

$$
(\Delta_{T,R^d}^{(k)} * \mu)(F) = \mu(\Delta_{T,R^d}^{(k)} * F),
$$

and it corresponds by Riesz’ representation theorem to a unique probability measure on $R^d$, the convolution $\mu_{T}^{(k)} = \Delta_{T,R^d}^{(k)} * \mu$ of $\mu$ with the kernel $\Delta_{T,R^d}^{(k)}$. This new measure has the following additional properties:

**Lemma 12.1 (Kernel with compactly supported and smooth Fourier transform).**

1. Suppose $F$ Lipschitz with constant $m$ (w.r.t. the euclidian norm on $R^d$) and bounded in absolute value by $C$. Then,

$$
|\mu_{T}^{(k)}(F) - \mu(F)| \leq \frac{27}{2} d^{\frac{2k + 3}{4k + 4}} \left( \frac{C m^{2k + 1}}{T^{2k + 1}} \right)^{\frac{1}{2k + 2}}.
$$

2. The Fourier transform of $\Delta_{T,R^d}^{(k)}$ takes its values in $[0, 1]$, is supported by the hypercube $[-(2k + 1)T, (2k + 1)T]^d$ and is of class $C^{2k}$ on $R^d$.

**Proof.** On the hypercube $[-\varepsilon, \varepsilon]^d$, $|F(x - y) - F(x)|$ is smaller than $m d^{1/2} \varepsilon$, whereas on the complementary, the mass of the integral kernel is smaller than

$$
\frac{2d}{I(k)} \int_{\varepsilon}^{\infty} \frac{1}{T^{2k+1} x^{2k+2}} dx = \frac{2d (2k + 1)}{I(k) (T \varepsilon)^{2k+1}}.
$$

It can be shown that $I(k)$ is always a rational multiple of $\pi$, decreasing with $k$ — for instance, $I(0) = \pi$ and $I(1) = \frac{5\pi}{12}$. On $[-\frac{\pi}{2\sqrt{2k+1}}, \frac{\pi}{2\sqrt{2k+1}}]$,

$$
\frac{1 - \cos(x)}{x^2} \geq \frac{1}{2} - \frac{\pi^2}{96(2k+1)} \geq \frac{1}{2} \left( 1 - \frac{\pi^2}{24(2k+1)} \right);
$$

$$
\frac{\sin x}{x} \geq 1 - \frac{\pi^2}{24(2k+1)}.
$$
So,

\[ I(k) \geq \frac{\pi}{2^{2k+1}} \left( 1 - \frac{\pi x}{x^2} \right) \left( \frac{\sin x}{x} \right)^{2k} dx \]

\[ \geq \frac{\pi}{2 \sqrt{2k+1}} \left( 1 - \frac{\pi^2}{24(2k+1)} \right) \geq \frac{\pi}{2 \sqrt{2k+1}} \frac{1 - \frac{\pi^2}{24}}{2 \sqrt{2k+1}}, \]

and therefore, \( \int_{[-\varepsilon,\varepsilon]^d} \Delta_{T,R}^{(k)}(x) dx \leq \frac{5d (2k+1)^{3/2}}{2(T \varepsilon)^{2k+1}}. \) Then,

\[ |\mu_T^{(k)}(F) - \mu(F)| \leq \|\Delta_{T,R}^{(k)} * F - F\|_\infty \]

\[ \leq \sup_{x \in R^d} \left( \int_{[-\varepsilon,\varepsilon]^d} + \int_{[-\varepsilon,\varepsilon]^d} \right) (|F(x - y) - F(x)| \Delta_{T,R}^{(k)}(y) dy) \]

\[ \leq \inf_{\varepsilon > 0} \left( md^{1/2} + \frac{5C d (2k+1)^{3/2}}{(T \varepsilon)^{2k+1}} \right) \leq \frac{27}{2} d^{2k+1} \left( \frac{C m^{2k+1}}{T^{2k+1}} \right)^{\frac{1}{2k+2}} \]

by choosing on the last line the minimizer \( \varepsilon = \left( (5C d^{1/2} (2k+1)^{3/2}) / (mT^{2k+1}) \right)^{\frac{1}{2k+2}} \) of the previous bound, and simplifying a bit the constants.

The second part of Lemma 12.1 relies on the following arguments. Denote \( M_T \) the operator on \( C^0(\mathbb{R}) \) defined by

\[ M_T f(x) = \frac{1}{2T} \int_{x-T}^{x+T} f(y) dy. \]

The image of \( C^k(\mathbb{R}) \) by \( M_T \) is included into \( C^{k+1}(\mathbb{R}) \), and on the other hand, if \( f \) is supported by \([a,b]\), then \( M_T f \) is supported by \([a - T, b + T]\). At the level of Fourier transforms, one has:

\[ (M_T \hat{f})(x) = \frac{1}{2T} \int_{x-T}^{x+T} \hat{f}(y) dy = \frac{1}{2T} \int \hat{f}(z) \left( \int_{x-T}^{x+T} e^{izy} dy \right) dz \]

\[ = \int \hat{f}(z) \frac{\sin Tz}{Tz} e^{izx} dz = (f \hat{g}_T)(x), \]

with \( \hat{g}_T(x) = \frac{\sin \frac{T x}{T}}{T}. \) As a consequence, if \( f \) is a non-negative function with Fourier transform supported by \([-T,T]\), then \( (\hat{g}_T)^{2k} f \) is a non-negative function with Fourier transform of class \( C^{2k} \) and supported by \([-2k+1)T, (2k+1)T\]. However,

\[ \Delta_T^{(0)}(x) = \frac{1}{\pi} \frac{1 - \cos \frac{T x}{Tx^2}}{T} \quad \Rightarrow \quad \Delta_T^{(0)}(x) = \left( 1 - \frac{|x|}{T} \right)_+ \]

which is supported by \([-T,T]\) and with values in \([0,1]\). So, the Fourier transform of \( \Delta_T^{(k)} \) takes its values in \([0,1]\), is supported by \([-2k+1)T, (2k+1)T\] and is of class \( C^{2k} \).

The result for \( \Delta_{T,R}^{(k)} \) follows immediately. \( \square \)
12.2. Estimates for test functions. By combining the properties of the kernels $\Delta^{(k)}_{T,R^d}$ and the hypotheses of mod-convergence, we get:

**Theorem 12.2** (Berry-Esseen estimates for Lipschitz test functions). Let $F$ be a function which is at the same time integrable, bounded and Lipschitz; $C = \|F\|_\infty$ and $M$ a Lipschitz bound for the fluctuations of $F$. For any $k \geq \lceil \frac{d+1}{2} \rceil$, there exists a constant $K_k$ such that

$$|\mu_n(F) - G(F)| \leq \frac{1}{(tn)^{1/2}} \left( 27(2k + 1) \sqrt{d} \left( CM^{2k+1} \right)^{\frac{1}{2k+2}} (dt_n)^{\frac{1}{4(k+1)}} + CK_k \right)$$

$$= \mathcal{O}_{M,C,k} \left( \frac{1}{(t_n)^{1/2 - \frac{1}{4k+4}}} \right),$$

and $K_k$ depends only on $k$, on the speed of convergence of the renormalized characteristic functions on the compact set $\Sigma([-1,1]^d)$, and on the behavior of $\psi$ on this compact set.

**Proof.** Fix $k \geq \lceil \frac{d+1}{2} \rceil$. If $\mu_n$ is the law of $Y_n$, then the difference $|\mu_n(F) - G(F)|$ is smaller than

$$|\mu_n(F) - \mu_{n,T}^{(k)}(F)| + |\mu_{n,T}^{(k)}(F) - G_T^{(k)}(F)| + |G_T^{(k)}(F) - G(F)| = \alpha + \beta + \gamma.$$

For $\alpha$ and $\gamma$, a correct bound is given by the property (1) of the previous lemma, and we cannot hope for much better since $\mu_n$ might be a singular measure (e.g. without density). Hence,

$$\alpha + \gamma \leq 27 d^{\frac{2k+3}{4k+4}} \left( \frac{CM^{2k+1}}{T^{2k+1}} \right)^{\frac{1}{4k+4}}.$$

As for $\beta$, we use Parseval’s theorem. Suppose first that $F$ is in $L^2(\mathbb{R}^d, dx)$; then,

$$|\mu_{n,T}^{(k)}(F) - G_T^{(k)}(F)| = \left| \int_{\mathbb{R}^d} F(x) \left( \frac{d\mu_{n,T}^{(k)}(x)}{dx} - \frac{dG_T^{(k)}(x)}{dx} \right) dx \right|$$

$$= \frac{1}{(2\pi)^d} \left| \int_{\mathbb{R}^d} \hat{F}(y) \Delta^{(k)}_{T,R^d}(y) \left( \hat{\mu}_n(y) - \hat{G}(y) \right) dy \right|.$$

The identity still holds for $F \in L^1(\mathbb{R}^d, dx)$, because $L^1(\mathbb{R}^d, dx) \cap L^2(\mathbb{R}^d, dx)$ is dense in $L^1(\mathbb{R}^d, dx)$. Denote $\psi_n(z)$ the left-hand side of the Equation appearing in $\S 7.1(a)$. With $T = \Theta \sqrt{t_n}$ and $y \in [-2k+1)T, (2k+1)T]^d$, one has

$$\hat{\mu}_n(y) - \hat{G}(y) = (\psi_n(i\Sigma_n y) - 1) e^{-\frac{|y|^2}{2}} = \int_0^1 \langle \psi'_n(t i \Sigma_n y) | i \Sigma_n y \rangle e^{-\frac{|y|^2}{2}} dt$$

$$= \frac{1}{\sqrt{t_n}} \int_0^1 \langle \Sigma \psi'_n(t i \Sigma_n y) | i y \rangle e^{-\frac{|y|^2}{2}} dt,$$

and therefore,

$$|\mu_{n,T}^{(k)}(F) - G_T^{(k)}(F)| \leq \frac{1}{(2\pi)^d \sqrt{t_n}} \left| \int_{\mathbb{R}^d} \hat{F}(y) \Delta^{(k)}_{T,R^d}(y) \langle \Sigma \psi'_n(t i \Sigma_n y) | iy \rangle e^{-\frac{|y|^2}{2}} dy \right| dt$$

$$\leq \frac{1}{\sqrt{t_n}} \int_0^1 \left| \int_{\mathbb{R}^d} F(x) G_{n,t}(x) dx \right| dt \leq \frac{\|F\|_\infty}{\sqrt{t_n}} \int_0^1 \int_{\mathbb{R}^d} |G_{n,t}(x)| dx dt.$$
where \( G_{n,t} \) is the inverse Fourier transform of \( \hat{\Delta}_{T,R^d}^{(k)}(y) \langle \Sigma \psi_n' \eta(t \Sigma_n y) \mid iy \rangle \), that is to say that
\[
G_{n,t}(x) = \int_{\mathbb{R}^d} \hat{\Delta}_{T,R^d}^{(k)}(y) \langle \Sigma \psi_n' \eta(t \Sigma_n y) \mid iy \rangle e^{-\frac{|y|^2}{2}} dy.
\]
However,

\begin{enumerate}
\item \( \hat{\Delta}_{T,R^d}^{(k)} \) is of class \( \mathcal{C}^{d+1} \) on \( \mathbb{R}^d \) and vanishes outside \([- (2k + 1)T, (2k + 1)T]^d \);
\item by the hypothesis of mod-Gaussian convergence, \( \psi_n'(t \Sigma_n y) \) and all its derivatives up to order \( (d + 1) \) are uniformly bounded on \([- (2k + 1)T, (2k + 1)T]^d \) (actually one does not need here the convergence of these quantities);
\item and \( y e^{-\frac{|y|^2}{2}} \) is a Schwartz function on \( \mathbb{R}^d \).
\end{enumerate}

Therefore,
\[
\sup_{t \in [0,1], \, n \in \mathbb{N}, \, |x| \leq d+1, \, l \leq d+1} \left\| (1 + \|y\|)^l D^n H_{n,t}(y) \right\|_\infty < \infty,
\]
so by \( (d + 1) \) integration by parts, \( |G_{n,t}(x)| \) is bounded by \( L (1 + \|x\|)^{-d-1} \), with \( L \) constant that only depends on \( k \) and on the behavior of the \( \psi_n \)'s and their derivatives up to order \( (d + 2) \) on \( \Sigma([- (2k + 1)i \Theta, (2k + 1)i \Theta])^d \). Since \( (1 + \|x\|)^{-d-1} \) is integrable on \( \mathbb{R}^d \), there is another finite constant \( K \) with the same properties of dependence and such that
\[
\sup_{t \in [0,1], \, n \in \mathbb{N}} \left( \int_{\mathbb{R}^d} |G_{n,t}(x)| dx \right),
\]
so \( \beta \leq \frac{CK}{\sqrt{n}} \).

To fix the ideas, set \( \Theta = 1/(2k + 1) \), and denote \( K_k \) the corresponding constant, which only depends on \( k \) and on the behavior of \( \psi \) and the \( \psi_n \)'s on \( \Sigma([-i,i]^d) \) (more precisely it suffices to have bound on their derivatives up to order \( d + 2 \)). Then,
\[
|\mathbb{E}[F(Y_n)] - G(F)| \leq 27(2k + 1) d^{\frac{2k+3}{4}} \left( \frac{CM^{2k+1}}{(t_n)^{\frac{2k+1}{4}}} + \frac{C K_k}{(t_n)^{1/2}} \right).
\]

\[ \square \]

**Corollary 12.3.** Under the same hypotheses, for any \( \epsilon > 0 \),
\[
|\mu_n(F) - G(F)| = O \left( \frac{1}{(t_n)^{1/2 - \epsilon}} \right),
\]
with constants in the \( O(\cdot) \) that only depend on \( \epsilon \) and \( M, C. \)

One might then wonder if it is possible to remove the \( \epsilon > 0 \) from the previous estimate. It is indeed possible, but with much stronger hypotheses than in Theorem 12.2. Hence, if \( F \) is a Schwartz function (or at least Schwartz “up to order \( d' \)” in the same sense as in the proof of Theorem 12.2), then by integration by parts one can:

- write \( \mu_n(F) \) as the integral of the \( d \)-th derivative of \( F \) against the cumulative distribution function of \( Y_n \).
• replace this cumulative distribution function by its Gaussian approximation, see Theorem 7.3;
• and undo the $d$ integrations by parts to obtain $G(F)$ up to a $O((t_n)^{-1/2})$, instead of a $O((t_n)^{-1/2+\varepsilon})$.

However we need $F$ to be much smoother than before, and this only removes the $-\varepsilon$ in the exponent of the estimates.

12.3. **Gaussian regular domains and convex bodies.** Starting from the previous results, one can estimate the difference of probabilities $|\mu_n(B) - G(B)|$ for $B$ domain in $\mathbb{R}^d$, but one needs $B$ to be “sufficiently regular”. The right condition to impose is the following. Denote $B^\varepsilon = \{x \mid d(x, B) \leq \varepsilon\}$ for $\varepsilon > 0$, and $B^{-\varepsilon} = ((B^c)^c)^c$. When $B$ is a convex body, these are the Minkowski sum and the Minkowski difference of $B$ with the ball $B_{(0,\varepsilon)}$.

**Definition 12.4.** A domain $B \subset \mathbb{R}^d$ will be called Gaussian regular with constant $G$ if for all $\varepsilon > 0$,

$$\frac{1}{(2\pi)^{d/2}} \int_{B^\varepsilon \setminus B} e^{-\frac{|x|^2}{2}} \, dx \leq G\varepsilon \quad ; \quad \frac{1}{(2\pi)^{d/2}} \int_{B \setminus B^{-\varepsilon}} e^{-\frac{|x|^2}{2}} \, dx \leq G\varepsilon.$$

**Example 12.5.** Consider an hypercube, possibly with infinite bounds and possibly rotated by some $u \in SO(d, \mathbb{R})$:

$$B = u \left( \prod_{i=1}^d \left[a^{(i)}, b^{(i)}\right] \right), \quad u \in SO(d, \mathbb{R}), \quad a^{(i)}, b^{(i)} \in \mathbb{R}.$$  

As shown by Figure 19, the first integral in Definition 12.4 is bounded by

$$\frac{2d}{(2\pi)^{d/2}} \sup_{a \in \mathbb{R}} \left( \int_a^{a+\varepsilon} e^{-\frac{x^2}{2}} \, dx \int_{\mathbb{R}^{d-1}} e^{-\frac{(x(2))^2 + \cdots + (x(d))^2}{2}} \, dx \right) \leq \sqrt{\frac{2}{\pi}} \varepsilon.$$  

![Figure 19](image.png)

**Figure 19.** For hypercubes, the Gaussian mass of the $\varepsilon$-boundary is bounded by the mass of $2d$ bands $H \times [0,\varepsilon]$ with $H$ hyperplane, and therefore by a constant times $\varepsilon$. 
The same bound holds for the second integral, so, hypercubes are Gaussian regular with constant $\sqrt{2/\pi} \cdot d$. More generally, the same discussion shows that any convex polytope with $F$ faces on its boundary is Gaussian regular with constant $\frac{F}{\sqrt{2\pi}}$.

A more general set of domains for which our results will apply consists in convex bodies, the asymptotics of Laplace or Fourier transforms of the indicators of convex bodies being a well-known problem (see e.g. [Ste93]). The Gaussian regularity of convex sets follows for instance from [BR10, Chapter I, Theorem 3.1]; for completeness we give here a shorter proof of this result. Let $B$ be a convex body in $\mathbb{R}^d$ containing 0 in its interior, homeomorphic to a closed ball and with a boundary $\partial B$ which we assume to be Lipschitz in the following sense: there exists a constant $L > 0$ such that the boundary $\partial B$ is parametrized by an homeomorphism $\chi : S^{d-1} \to \partial B$ with

$$
\chi(\theta) \in \mathbb{R}_+^\theta \quad ; \quad \frac{\|\chi(\theta) - \chi(\theta')\|}{\|\chi(\theta)\|} \leq L \|\theta - \theta'\|.
$$

Notice that the constant $L$ does not change if one renormalizes $B$ and looks at $\lambda B$ instead of $B$, $\lambda > 0$. In other words, $L$ only depends on the form of the convex body viewed from the origin. We refer to [Sch93, Hör94] for references on convex bodies. We claim that under these assumptions, $B$ is Gaussian regular with constant $4(1 + 2L)^2$. As will be clear from the proof, the origin 0 only plays a role of reference point, so the same holds for any convex body with boundary $L$-Lipschitz viewed from a given point in the interior of $B$.

First, for any direction $\theta \in S^{d-1}$, we claim that the intersection of $\mathbb{R}_+^\theta$ with $B^\epsilon \setminus B$ is an interval. Indeed, if two points $a, b$ on $\mathbb{R}_+^\theta$ are in $B^\epsilon \setminus B$, then there exists corresponding points $c, d$ in $B$ with $\|a - c\| \leq \epsilon$ and $\|b - d\| \leq \epsilon$, and then by convexity of $B$ all the points between $a$ and $b$ are in $B^\epsilon$. So, there exists real numbers $0 < \alpha < \beta$ such that

$$(B^\epsilon \setminus B) \cap \mathbb{R}_+^\theta = (\alpha \theta, \beta \theta],$$

and in fact one has necessarily $\alpha \theta \in \partial B$, and $\alpha \theta = \chi(\theta)$. Then, $\beta \theta$ is $\epsilon$-close to some $\chi(\theta')$, so

$$
\beta - \alpha = \|\beta \theta - \chi(\theta)\| \leq \|\beta \theta - \chi(\theta')\| + \|\chi(\theta') - \chi(\theta)\| \leq \epsilon + L \|\chi(\theta)\| \|\theta - \theta'\|.
$$

However, for two points $a$ and $b$ on two lines $\mathbb{R}_+^\theta$ and $\mathbb{R}_+^{\theta'}$, one has always $\|a - b\| \geq \frac{\|a\|}{2} \|\theta - \theta'\|$, the worst case being when $\theta = -\theta'$ and $b$ is near zero. Consequently,

$$
\beta - \alpha \leq \epsilon + 2L \frac{\|\chi(\theta)\|}{\|\beta \theta\|} \|\chi(\theta') - \beta \theta\| \leq (2L + 1) \epsilon.
$$

Then, one writes:

$$
\frac{1}{(2\pi)^{d/2}} \int_{B^\epsilon \setminus B} e^{-\frac{|x|^2}{2}} \ dx = \frac{1}{(2\pi)^{d/2}} \int_{S^{d-1}} \int_{\alpha(\theta)}^{\beta(\theta)} \left( e^{-\frac{r^2}{2}} r^{d-1} \ dr \right) d\mu_{S^{d-1}}(\theta)
\leq \frac{\text{vol}(S^{d-1})}{(2\pi)^{d/2}} \max_{a > 0} \left( \int_{a}^{a + (2L + 1) \epsilon} \left( e^{-\frac{r^2}{2}} r^{d-1} \ dr \right) \right)
\leq (2L + 1) \epsilon \ \frac{2^{1-d/2}}{\Gamma(d/2)} \max_{r \geq 0} \left( e^{-\frac{r^2}{2}} r^{d-1} \right) \leq 2 (2L + 1) \epsilon
$$

by using Stirling estimates.
Similarly, we claim that for any direction $\theta \in S^{d-1}$, the intersection of $\mathbb{R}_+ \theta$ with $B \setminus B^{-\varepsilon}$ is an interval. Fix $\varepsilon, \eta > 0$, a direction $\theta$ and consider the set

$$I_{\varepsilon,\eta,\theta} = \{x \in B \cap \mathbb{R} \theta \mid B(x,\varepsilon+\eta) \subset B\}.$$ 

This is a convex compact subset of $\mathbb{R} \theta$, as shown by Figure 21. However, $B^{-\varepsilon} = ((B^c)^c)^c = \bigcup_{\eta > 0} \{x \in B \mid B(x,\varepsilon+\eta) \subset B\}$, so $\mathbb{R} \theta \cap B^{-\varepsilon}$ is the increasing union of the segments $I_{\varepsilon,\eta,\theta}$. Therefore, there exists real numbers $0 \leq \gamma < \alpha$ such that

$$(B \setminus B^{-\varepsilon}) \cap \mathbb{R}_+ \theta = [\gamma \theta, \alpha \theta],$$

possibly with $\gamma = 0$. As before, $\alpha \theta = \chi(\theta)$. Denote $\theta'$ a direction such that $\gamma \theta$ is $\varepsilon$-close to a point of $\mathbb{R}_+ \theta'$ which is outside $B$; one can assume without loss of generality that this point is on the boundary of $B$, that is to say that $\|\gamma \theta - \chi(\theta')\| \leq \varepsilon$. Again, we want to bound $\alpha - \gamma$ by a constant times $\varepsilon$. 

**Figure 20.** The intersection of an half-line $\mathbb{R}_+ \theta$ with $B^c \setminus B$, $B$ convex body is a segment whose length is bounded by a multiple of $\varepsilon$ if $0 \in B^o$ and $\partial B$ is Lipschitz.

**Figure 21.** The intersection of a line $\mathbb{R} \theta$ with $B^{-\varepsilon} = ((B^c)^c)^c$ is the union of all sets $I_{\varepsilon,\eta,\theta}$, $\eta > 0$, and these sets are nested segments.
• If $\gamma \geq 4L\varepsilon$, then one has

$$\alpha - \gamma = \|\chi'(\theta) - \gamma\theta\| \leq \|\chi'(\theta') - \gamma\theta\| + \|\chi'(\theta') - \chi(\theta')\| \leq \varepsilon + L\|\chi(\theta)\| \|\theta - \theta'\|.$$ 

The second term in the right-hand side is smaller than $L\frac{\alpha}{\gamma}\|\gamma\theta - \chi(\theta')\| \leq \frac{2L\varepsilon\alpha}{\gamma}$, so

$$\alpha - \gamma \leq \varepsilon + \frac{2L\varepsilon\alpha}{\gamma} \quad ; \quad R \leq 1 + \frac{(2L+1)\varepsilon}{\gamma - 2L\varepsilon}$$

with $R = \frac{\varepsilon}{\gamma}$. Thus, $\alpha \leq \gamma + (2L+1)\varepsilon \frac{\gamma}{\gamma - 2L\varepsilon}$, and by the hypothesis $\gamma \geq 4L\varepsilon$, this leads to $\alpha - \gamma \leq 2(1+2L)\varepsilon$.

• Conversely, if $\gamma \leq 4L\varepsilon$, then $\|\chi'(\theta')\| \leq (4L+1)\varepsilon$, so

$$\alpha - \gamma = \|\chi'(\theta) - \gamma\theta\| \leq \|\chi'(\theta') - \gamma\theta\| + \|\chi'(\theta') - \chi(\theta')\| \leq \varepsilon + L\|\chi'(\theta)\| \|\theta - \theta'\|$$

$$\leq (1+2L(4L+1))\varepsilon \leq (8L^2 + 4L + 2)\varepsilon.$$ 

So in any case, $\alpha - \gamma \leq 2(1+2L)^2\varepsilon$.

Using then spherical coordinates to compute the integrals, we obtain the bound

$$\frac{1}{(2\pi)^{d/2}} \int_{B \setminus B^{-\varepsilon}} e^{-\frac{|x|^2}{2}} \, dx \leq 4(1+2L)^2\varepsilon.$$ 

Hence, we have indeed proved that convex bodies with boundary Lipschitz with constant $L$ viewed from a certain point of their interior are Gaussian-regular with constant $4(1+2L)^2$. Now, it should be noticed that any convex body $B$ with non-empty interior has a Lipschitz boundary for a certain constant $L$ and with respect to a certain reference point $z$ in $B^\circ$. However, this constant $L$ may be extremely big if one does not choose $z$ correctly (e.g. near the boundary of $B$). On the other hand, it follows from [BR10] that one can in fact choose a constant $G$ of Gaussian regularity that works for every convex body of $\mathbb{R}^d$, but the proof of this relies then on more complex arguments than before, and is much longer.

### 12.4. Estimates for Gaussian regular domains.

Working with the notion of Gaussian regularity, we can replace characteristic functions of domains by smooth approximations of them, and deduce from Theorem 12.2 the following:

**Theorem 12.6** (Berry-ESseen estimates for probabilities of Gaussian regular domains). Fix $\varepsilon > 0$. For $n \geq n_\varepsilon$ big enough and for all Gaussian regular domain $B$ with constant $G$,

$$|\mu_n(B) - G(B)| \leq \frac{\sqrt{G}}{(t_n)^{\frac{d+1}{2} - \varepsilon}}.$$ 

Here $n_\varepsilon$ depends on $d, \varepsilon, \Sigma$ and on the behavior of the $\psi_i$'s on $\Sigma([-i,i]^d)$, but is independent from $B$ and $G$.

**Proof.** Fix $k \geq \lceil \frac{d+1}{2} \rceil$. For $M > 0$, set

$$\psi_{B,M}(\cdot) = (1 - M d(\cdot, B))_+ \quad ; \quad \phi_{B,M} = 1 - (1 - M d(\cdot, B^c))_+.$$
These are Lipschitz functions with constant $M$ and bound $C = 1$. One has

$$\mu_n(B) - G(B) \leq \mu_n(\psi_{B,M}) - G(B) \leq |\mu_n(\psi_{B,M}) - G(\psi_{B,M})| + \int_{B^{1/M}\setminus B} \frac{e^{-\frac{|x|^2}{2}}}{(2\pi)^{d/2}} \, dx$$

$$\mu_n(B) - G(B) \geq \mu_n(\phi_{B,M}) - G(B) \geq -|\mu_n(\phi_{B,M}) - G(\phi_{B,M})| - \int_{B\setminus B^{1/M}(2\pi)^{d/2}} \, dx.$$ 

By hypothesis, the Gaussian integrals are bounded by $\frac{G}{M}$, and we can use Lemma 12.2 in order to bound the remaining terms:

$$|\mu_n(B) - G(B)| \leq \frac{27(2k+1)d^{2k+3}}{(t_n)^{2k+2}} \mu_n(B) \leq \frac{G}{M} + \frac{K_k}{(t_n)^{1/2}},$$

for any $k \geq \lceil \frac{d+1}{2} \rceil$. It suffices now to choose the best $M$, which is

$$M = \left( \frac{(2k+2)G}{27(2k+1)^2} \frac{t_n^{2k+1}}{d^{2k+2}} \right)^{2k+2},$$

and yields the inequality

$$|\mu_n(B) - G(B)| \leq 14.08 \sqrt{2k+1} \frac{G^{2k+3}}{(t_n)^{2k+2}} \sqrt{\frac{d^{2k+2}}{(2k+1)(2k+2)}} + \frac{K_k}{(t_n)^{1/2}},$$

the constant 14.08 corresponding to the largest value of $\frac{1}{\sqrt{2k+1}} \frac{4k+3}{2k+1} \left( \frac{27(2k+1)^2}{2k+2} \right)^{2k+2}$ (obtained when $k = 1$). Therefore, for $n \geq n_k$ with $n_k$ depending only on $k$ and on the behavior of the $\psi_n$’s on $\Sigma([-i, i]^d)$,

$$|\mu_n(B) - G(B)| \leq \frac{29}{2} \sqrt{\frac{d}{(2k+1)(2k+2)}} \frac{G^{2k+3}}{(t_n)^{2k+2}} \sqrt{\frac{4k+3}{d}},$$

making again some simplifications on the constants involved in this upper bound. Since $d \leq 2k - 1$, setting $\varepsilon = \frac{1}{16k}$, this gives

$$|\mu_n(B) - G(B)| \leq \frac{29}{16} \frac{\sqrt{G}}{(t_n)^{1/4 - \varepsilon} \frac{1}{4k+3}},$$

For $n$ big enough, $\frac{16}{29} \left( t_n^{\frac{1}{16}} \frac{1}{4k+3} \right)$ is bigger than 1, so, possibly raising the value of $n_k = n_\varepsilon$, we end up with

$$|\mu_n(B) - G(B)| \leq \frac{\sqrt{G}}{(t_n)^{1/4 - \varepsilon}},$$

and this holds for any Gaussian regular domain. \qed

For domains with smoother boundaries, it is reasonable to conjecture that a correct estimate is once again a $O(1/\sqrt{t_n})$, but we don’t know how to prove this but for hypercubes or polytopes (Theorem 7.3).
REFERENCES

[ABT03] R. Arratia, A. D. Barbour, and S. Tavaré. Logarithmic Combinatorial Structures: a Probabilistic Approach. EMS Monographs in Mathematics. European Mathematical Society, 2003.

[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.

[Ben13] C. G. Beneš. A local central limit theorem and loss of rotational symmetry of planar simple random walk. arXiv:1302.2971v1 [math.PR], 2013.

[Bin86] N. H. Bingham. Variants on the law of the iterated logarithm. Bull. London Math. Soc., 18(5):433–467, 1986.

[BKN13] A. D. Barbour, E. Kowalski, and A. Nikeghbali. Mod-discrete expansions, 2013.

[Bol98] B. Bollobás. Modern graph theory, volume 73 of Cambridge studies in advanced mathematics. Cambridge University Press, 2001. Second edition.

[Bon10] 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 Eliminationsresultante. 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.

[BR10] R. N. Bhattacharya and R. R. Rao. Normal Approximation and Asymptotic Expansions, volume 64 of Classics in Applied Mathematics. SIAM, 2010.

[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.

[Cha12] S. Chatterjee. The missing log in large deviations for triangle counts. Random Structures & Algorithms, 40(4):437–451, 2012.

[CV11] S. Chatterjee and S.R.S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. Eur. J. Comb., 32(7):1000–1017, 2011.

[DE12] H. Döring and P. Eichelsbacher. Moderate deviations via cumulants. To appear in Journal of Theor. Probability, 2012.

[DE13] H. Döring and P. Eichelsbacher. Moderate deviations for the determinant of wigner matrices. arXiv:1301:2915v2 [math.PR], 2013.

[DKN14] V. Féray, P.-L. Méliot, and A. Nikeghbali. Estimates of the speed of convergence and local limit theorems in the mod-stable framework. In preparation, 2014.

[FO90] P. Flajolet and A.M. Odlyzko. Singularity analysis of generating functions. SIAM J. Discrete Math., 3:216–240, 1990.
[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.

[Mur03] N. Muraki. The five independences as natural products. Infinite Dimensional Analysis, Quantum Probability and Related Topics, 6(3):337–371, 2003.

[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. Annales de l’Institut Henri Poincaré, to appear, 2013.

[Oko00] A. Okounkov. Random matrices and random permutations. Internat. Math. Res. Notices, 20:1043–1095, 2000.

[Pen02] M. Penrose. Random Geometric graphs, volume 5 of Oxford studies in probability. Oxford University Press, 2002.

[Pet75] V. V. Petrov. Sums of independent random variables, volume 82 of Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, 1975.

[Rad09] M. Radziwill. On large deviations of additive functions. B.Sc thesis, arXiv:09095274v4 [math.NT], 2009.

[RSS78] R. Rudzkiś, 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, 79(1):1–10, 1988.

[Sch93] R. Schneider. Convex Bodies: The Brunn-Minkowski Theory, volume 44 of Encyclopaedia of Mathematics and Its Applications. Cambridge University Press, 1993.

[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.

[´Sni06a] P. ´Sniady. Asymptotics of characters of symmetric groups, genus expansion and free probability. Discrete Math., 306(7):624–665, 2006.

[´Sni06b] P. ´Sniady. Gaussian fluctuations of characters of symmetric groups and of Young diagrams. Probab. Th. Rel. Fields, 136(2):263–297, 2006.

[SS91] L. Saulis and V. A. Statulevičius. Limit Theorems for Large Deviations. Kluwer Academic Publications, 1991.

[Ste93] E. M. Stein. Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, volume 43 of Princeton Mathematical Series. Princeton University Press, 1993.

[Sze75] G. Szegő. 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.

[Wol27] E. T. Whittaker and G. N. Watson. A Course of Modern Analysis. Cambridge University Press, 4th edition, 1927.