ABSTRACT

Title of dissertation: HIGHER ORDER ASYMPTOTICS FOR THE CENTRAL LIMIT THEOREM AND LARGE DEVIATION PRINCIPLES

Buddhima Kasun Fernando Akurugodage
Doctor of Philosophy, 2018

Dissertation directed by: Professor Dmitry Dolgopyat
Department of Mathematics

First, we present results that extend the classical theory of Edgeworth expansions to independent identically distributed non-lattice discrete random variables. We consider sums of independent identically distributed random variables whose distributions have $d+1$ atoms and show that such distributions never admit an Edgeworth expansion of order d but for almost all parameters the Edgeworth expansion of order $d-1$ is valid and the error of the order $d-1$ Edgeworth expansion is typically $\mathcal{O}(n^{-d/2})$ but the $\mathcal{O}(n^{-d/2})$ terms have wild oscillations.

Next, going a step further, we introduce a general theory of Edgeworth expansions for weakly dependent random variables. This gives us higher order asymptotics for the Central Limit Theorem for strongly ergodic Markov chains and for piece–wise expanding maps. In addition, alternative versions of asymptotic expansions are introduced in order to estimate errors when the classical expansions fail to hold. As applications, we obtain Local Limit Theorems and a Moderate Deviation Principle.

Finally, we introduce asymptotic expansions for large deviations. For sufficiently regular weakly dependent random variables, we obtain higher order asymptotics (similar to Edgeworth Expansions) for Large Deviation Principles. In particular, we obtain asymptotic expansions for Cramér's classical Large Deviation Principle for independent identically distributed random variables, and for the Large Deviation Principle for strongly ergodic Markov chains.

HIGHER ORDER ASYMPTOTICS FOR THE CENTRAL LIMIT THEOREM AND LARGE DEVIATION PRINCIPLES

by

Buddhima Kasun Fernando Akurugodage

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Prof. Dmitry Dolgopyat, Chair/Adviser
Prof. Leonid Koralov
Prof. Mark Freidlin
Prof. Rodrigo Treviño
Prof. Edward Ott, Dean's Representative

Dedication

To the memory of my uncle, Ivan Fernando, who urged me to seek truth.

Acknowledgments

I am forever grateful to Dima, for taking me under his wing and showing me how to find my way in the mathematical landscape, for the insightful conversations and for his patience and thoughtfulness. His mathematical brilliance and humility are things that I will always look up to.

I am also thankful to my collaborators, Carlangelo and Pratima for keeping me on the right track and, to Lashi, Sam, Alex and Peter for being positive role models. I am grateful to all my friends at UMD: Hamid, Danul, Micah, Steven, Pratima, Jerry, Jenny, Hsin-Yi, Shujie, David, Corry, Jing, Phil, Minsung, Patrick, JP, Ryan, Jacky and Nick, for being great companions.

I extend my gratitude to Larry, without whom I would not even be in the program, to Vadim for initiating me into Dynamics and to Leonid for initiating me into Probability. I also appreciate the work of the awesome administrative staff at the UMD Mathematics Department, including Haydee, Celeste, Liliana, Sharon, Bill and Cristina. Thank you for turning paperwork into a minor problem and for all the reminders about deadlines.

Finally, I am grateful to my parents and Kasunka, for everything they have done. I would not have been able to get through my PhD if not for their encouragement and support.

Table of Contents

List of Abbreviations

CLT	Central Limit Theorem
i.i.d.	independent and identically distributed
LCLT	Local Central Limit Theorem
LDP	Large Deviation Principle
LDCT	Lebesgue Dominated Convergence Theorem
LHS	Left hand side
LLT	Local Limit Theorem
PDE	Partial differential equation
RHS	Right hand side
WLOG	Without loss of generality

Chapter 1: Introduction

The Central Limit Theorem (CLT) is one of the most fundamental concepts in probability which was introduced by the work of Laplace and Bernoulli. It describes the long term behaviour of random trials repeated under uniform conditions.

Let $S_N = \sum_{n=1}^{N} X_n$ be a sum of random variables. We say that S_N satisfies the CLT if there are real constants A and $\sigma > 0$ such that

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{S_N - NA}{\sqrt{N}} \leq z\right) = \mathfrak{N}(z) \tag{1.1}$$

where $\mathfrak{N}(z) = \int_{-\infty}^{z} \mathfrak{n}(y)dy$ and $\mathfrak{n}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}}$.

The usefulness of the CLT and related limit theorems depends on rapid convergence of distributions of normalized partial sums to the limiting distribution. This is because limit theorems are primarily used for approximating distributions of sums of large but finite number of random variables. Therefore, an important problem is to estimate the rate of convergence of (1.1).

In this regard, an asymptotic expansion as a series of increasing powers of order $n^{-1/2}$ (now commonly referred to as the Edgeworth expansion) was formally derived by Chebyshev in [8]. Kolmogorov and Gnedenko emphasize the importance of these expansion in their monograph [23] by stating that the Edgeworth Expansion is “the most powerful and general method of finding such corrections.”

Definition 1. *S_N admits Edgeworth expansion of order r if there are polynomials $P_1(z), \dots, P_r(z)$ such that*

$$\mathbb{P}\left(\frac{S_N - NA}{\sqrt{N}} \le z\right) = \underbrace{\mathfrak{N}(x) + \sum_{p=1}^{r} \frac{P_p(z)}{N^{p/2}} \mathfrak{n}(z)}_{\mathcal{E}_{r,N}(z)} + o\left(N^{-r/2}\right) \tag{1.2}$$

uniformly for $z \in \mathbb{R}$.

Remark 1.1. *It is an easy observation that Edgeworth expansion of S_N, if it exists, is unique. Suppose $\{P_p(z)\}_p$ and $\{\tilde{P}_p(z)\}_p$, $1 \le p \le r$ are polynomials corresponding to two Edgeworth expansions. Then,*

$$\sum_{p=1}^{r} \frac{P_p(z)}{N^{p/2}} \mathfrak{n}(z) = \sum_{p=1}^{r} \frac{\tilde{P}_p(z)}{N^{p/2}} \mathfrak{n}(z) + o\left(N^{-r/2}\right)$$

Multiplying by $\sqrt{N}$ taking the limit $N \to \infty$ we have $P_1(z) = \tilde{P}_1(z)$. Therefore,

$$\sum_{p=2}^{r} \frac{P_p(z)}{N^{p/2}} \mathfrak{n}(z) = \sum_{p=2}^{r} \frac{\tilde{P}_p(z)}{N^{p/2}} \mathfrak{n}(z) + o\left(N^{-r/2}\right)$$

Then, multiplying by N and taking $N \to \infty$, $P_2(z) = \tilde{P}_2(z)$. Continuing this r times we can conclude $P_p(z) = \tilde{P}_p(z)$ for $1 \le p \le r$.

Here and in what follows, A is the asymptotic mean i.e. $A = \lim_{N\to\infty} \mathbb{E}\left(\frac{S_N}{N}\right)$.

Work of Lyapunov, Edgeworth and Cramér focus on the problem of finding higher order asymptotics in the CLT. Their main focus was on independent and identically distributed (i.i.d.) sequences of random variables. In 1928, Cramér introduced a theory of Edgeworth expansions for a broad class of random variables. For the first rigorous derivation of this expansion see [10]. The monograph [11] by Cramér also gives a detailed account of his theory of Edgeworth expansions.

Theorem 1.1 (Cramér)**.** *Let X be a centred random variable with $\mathbb{E}(X^2) = \sigma^2 > 0$ and $r+2$ absolute moments. Let $X_1, \dots, X_N, \dots$ be sequence of i.i.d. copies of X.*

Assume further that

$$\limsup_{|t|\to\infty} |\mathbb{E}(e^{itX})| < 1. \tag{1.3}$$

Then, S_N *satisfies* (1.2).

Many refinements of this result appear in later literature. A good introduction to this theory and later developments can be found in [3, 11, 20, 23].

In the i.i.d. case, P_p's are polynomials such that the characteristic function $\phi(t) = \mathbb{E}(e^{itX})$ and the Fourier transform $\hat{\mathcal{E}}_{r,N}$ of $\mathcal{E}_{r,N}$ satisfy

$$\phi\left(\frac{t}{\sigma\sqrt{N}}\right)^N - \hat{\mathcal{E}}_{r,N}(t) = o\left(N^{-r/2}\right).$$

For example, $\mathcal{E}_{1,n}(z) = \mathfrak{N}(z) + \mathfrak{n}(z)\dfrac{\mathbb{E}(X^3)}{6\sigma^3\sqrt{n}}(1-z^2)$ and

$$\begin{aligned}\mathcal{E}_{2,n}(z) = \mathfrak{N}(z) + \mathfrak{n}(z)\Bigg[&\frac{\mathbb{E}(X^3)}{6\sqrt{n}\sigma^3}(1-z^2) + \frac{\mathbb{E}(X^4)-3\sigma^4}{24n\sigma^4}(3z-z^3)\\ &-\frac{\mathbb{E}(X^3)^2}{72n\sigma^6}(15z-10z^3+z^5)\Bigg].\end{aligned}$$

Since all distributions with an absolutely continuous component satisfy (1.3), this theorem covers a large class of random variables. However, (1.3) indicates that the common distribution of X_n's is far from being discrete. In fact, (1.3) fails when random variables are purely discrete. Surprisingly, not much had been explored in the case of discrete random variables, except in the lattice case. The purpose of my first project [16], joint with Dmitry Dolgopyat, was to address this issue. A detailed discussion about this can found in Chapter 2.

When X_n's are i.i.d., it is known that the order 1 Edgeworth expansion exists if and only if the distribution is non-lattice (see [19]). Therefore, the following asymptotic expansion for the Local Central Limit Theorem (LCLT) for lattice random variables is also useful.

Definition 2. *Suppose that X_n's are integer valued. We say that S_N admits a lattice Edgeworth expansion of order r, if there are polynomials $P_{0,d}, \ldots, P_{r,d}$ and a number A such that*

$$\sqrt{N}\mathbb{P}(S_N = k) = \mathfrak{n}\left(\frac{k - NA}{\sqrt{N}}\right) \sum_{p=0}^{r} \frac{P_{p,d}((k - NA)/\sqrt{N})}{N^{p/2}} + o\left(N^{-r/2}\right)$$

uniformly for $k \in \mathbb{Z}$.

Remark 1.2. *Here, the subscript d in $P_{p,d}$ refers to the fact that the expansion is for discrete lattice-valued random variables. A priori, there is no reason for the polynomials P_p in Definition 1 to be related to $P_{p,d}$. In Section 3.3, we show how these two polynomials are related.*

As in remark 1.1, we can prove the uniqueness of this expansion. Because $P_{p,d}$'s have finite degree, say at most q, choose N large enough so that S_N has more than q values. Then the argument in remark 1.1 applies.

During the 20th century, the work of Lyapunov, Edgeworth, Cramér, Kolmogorov, Esséen, Petrov, Bhattacharya and many others led to the development of the theory of these two asymptotic expansions. See [26, 31] and references therein, for more details.

It is immediate that S_N admits an order r Edgeworth expansion if

$$\lim_{N\to\infty} N^{r/2} \left[\mathbb{P}\left(\frac{S_N - NA}{\sqrt{N}} \leq z\right) - \mathcal{E}_{r,N}(z)\right] = 0. \tag{1.4}$$

uniformly in z. [3, 4] discuss weak Edgeworth expansions where the LHS of (1.4) is convolved with smooth compactly supported functions. These expansions yield the asymptotics of $\mathbb{E}(f(S_N))$.

To introduce these expansions, suppose $(\mathcal{F}, \|\cdot\|)$ is a function space.

Definition 3. *S_N admits weak global Edgeworth expansion of order r for $f \in \mathcal{F}$ if there are polynomials $P_{0,g}(z), \dots P_{r,g}(z)$ and A (which are independent of f) such that*

$$\mathbb{E}(f(S_N - NA)) = \sum_{p=0}^{r} \frac{1}{N^{\frac{p}{2}}} \int P_{p,g}(z)\mathfrak{n}(z) f\big(z\sqrt{N}\big) dz + \|f\| \cdot o\left(N^{-(r+1)/2}\right).$$

Definition 4. *S_N admits weak local Edgeworth expansion of order r for $f \in \mathcal{F}$ if there are polynomials $P_{0,l}(z), \dots P_{r,l}(z)$ and A (which are independent of f) such that*

$$\sqrt{N}\mathbb{E}(f(S_N - NA)) = \frac{1}{2\pi} \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^p} \int P_{p,l}(z) f(z) dz + \|f\| \cdot o\left(N^{-r/2}\right).$$

We also introduce the following asymptotic expansion which yields an averaged form of the error of approximation.

Definition 5. *S_N admits averaged Edgeworth expansion of order r if there are polynomials $P_{1,a}(z), \dots P_{r,a}(z)$ and numbers k, m such that for $f \in \mathcal{F}$ we have*

$$\begin{aligned} &\int \left[\mathbb{P}\left(\frac{S_N - NA}{\sqrt{N}} \leq z + \frac{y}{\sqrt{N}}\right) - \mathfrak{N}\left(z + \frac{y}{\sqrt{N}}\right)\right] f(y) dy \\ &\qquad = \sum_{p=1}^{r} \frac{1}{N^{p/2}} \int P_{p,a}\left(z + \frac{y}{\sqrt{N}}\right) \mathfrak{n}\left(z + \frac{y}{\sqrt{N}}\right) f(y)\, dy + \|f\| \cdot o\left(N^{-r/2}\right). \end{aligned}$$

Remark 1.3. *Here, the subscripts g, l, a refer to global, local and averaged respectively and used to distinguish the polynomials appearing each definition. In Section 3.3, we show how these two polynomials are related.*

All of these weak forms of expansions are unique provided that $\mathcal{F}$ is dense in C_c^∞. If there are two different weak global expansions with polynomials $\{P_{p,g}\}$ and

$\{\tilde{P}_{p,g}\}$, *the argument in remark 1.1 yields,*

$$\int P_{p,g}(z)\mathfrak{n}(z)f\big(z\sqrt{N}\big)\,dz = \int \tilde{P}_{p,g}(z)\mathfrak{n}(z)f\big(z\sqrt{N}\big)\,dz$$

for all $f \in C_c^\infty$ *which gives us the equality,* $P_{p,g}(z) = \tilde{P}_{p,g}(z)$. *The same idea works for the other two expansions.*

We have seen that these asymptotic expansions are unique. They also form a hierarchy. We discuss this in Appendix A.2. Due to this hierarchy, in the absence of one, others can be useful in extracting information about the rate of convergence in (1.1).

Previous results on existence of Edgeworth expansions, for example in [11, 20, 23], assume independence of random variables X_n. For many applications the independence assumption of random variables is too restrictive. Because of this reason, there have been attempts to develop a theory of Edgeworth expansions for weakly dependent random variables where weak dependence often refers to asymptotic decorrelation. See [9, 22, 29, 40, 41] for such examples. Their focus is on the *classical* expansions introduced in Definition 1 and Definition 2.

Except in [9], the sequences of random variables considered are uniformly ergodic Markov processes with strong recurrent properties or processes approximated by such Markov processes. In [9], the authors consider aperiodic subshifts of finite type endowed with a stationary equilibrium state and give explicit construction of the order 1 Edgeworth expansion. They also prove the existence of higher order classical Edgeworth expansions under a rapid decay assumption on the tail of the characteristic function.

The goal of [21], a joint work with Carlangelo Liverani, is to generalize these results and to provide sufficient conditions that guarantee the existence of Edgeworth expansions for weakly dependent random variables including observations arising from sufficiently chaotic dynamical systems, and strongly ergodic Markov chains. In fact, we introduce a widely applicable theory for both classical and weak forms of Edgeworth expansions and significantly improve existing results. This work is discussed in detail in Chapter 3.

The CLT and related asymptotic expansions provide accurate descriptions *only* of typical events. For example, if X_n's are centered i.i.d. random variables then for all $a > 0$, $\lim_{N\to\infty} \mathbb{P}(S_N \geq aN) = 0$, due to the Law of Large Numbers i.e. $\dfrac{S_N}{N} \to 0$ in probability. Large Deviation Principles (LDPs) give better descriptions of these non–typical events by specifying the exponential rate at which their probabilities decay.

Before we present results related to LDPs, we recall the following definitions, and facts whose proofs can be found in [17, 30]. Given a function $f : \mathbb{R} \to (-\infty, \infty]$ with $f \neq \infty$, define its effective domain to be $D_f = \{x \in \mathbb{R} | f(x) < \infty\}$ and its Legendre transform by $f^*(x) = \sup_{t\in\mathbb{R}} [tx - f(t)]$. Then, f^* is convex and lower semi-continuous. Therefore, D_{f^*} is an interval and f^* is continuous on $\overline{D}_{f^*}$.

In addition, suppose f is convex, lower semi-continuous with $\mathring{D}_f = (a, b)$ and $f \in C^2(a, b)$ with $f'' > 0$ on (a, b) (possibly $a = -\infty$ or $b = +\infty$). Then, $\mathring{D}_{f^*} = (A, B)$ where $A = \lim_{t\to a+} f'(t)$ and $B = \lim_{t\to b-} f'(t)$, f^* is continuously differentiable on (A, B) and $(f^*)' = (f')^{-1}$. For any f satisfying the above properties, for any $x \in \mathring{D}_{f^*}$ the supremum in the definition of $f^*(x)$ is achieved at the unique point

$t \in \mathring{D}_f$ which solves $f'(t) = x$ and hence, $f^*(x) = \sup_{t \in \mathring{D}_f} [tx - f(t)]$. Also, f is called steep if $\lim_{t \to a} |f'(t)| = \lim_{t \to b} |f'(t)| = \infty$.

The following classical result, due to Cramér, is one of the fundamental results in the theory of Large Deviations.

Theorem 1.2 (Cramér). *Let X be a real valued random variable with mean A and variance $\sigma^2 > 0$. Suppose that the logarithmic moment generating function of X, $\log \mathbb{E}(e^{tX})$, is finite in a neighbourhood of 0. Let X_n be a sequence of i.i.d. copies of X. Then,*

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{P}(S_N \geq Nz) = -I(z), \text{ if } z > A$$

and

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{P}(S_N \leq Nz) = -I(z), \text{ if } z < A$$

where I is given by $I(z) = \sup_{\lambda \in \mathbb{R}} \left[\lambda z - \log \mathbb{E}(e^{\lambda X})\right]$ (the Legendre transform of the logarithmic moment generating function of X).

From the hypothesis it is immediate that I is convex and lower semi-continuous. Also, $I'' > 0$ on $\mathring{D}_I = (\inf(\text{supp } X), \sup(\text{supp } X))$, therefore I is strictly convex on $\mathring{D}_I$, $I(z) = 0 \iff z = \mu$ and there is a unique λ^* such that $I(z) = \lambda^* z - \log \mathbb{E}(e^{\lambda^* X})$.

Cramér's LDP has an extension to the non–i.i.d. case. We refer the reader to [6][Chapter V.6] for a proof of the following result.

Theorem 1.3 (Local Gärtner–Ellis). *Let X_n be a sequence of random variables not necessarily i.i.d. Suppose there exists $\delta > 0$ such that for $\lambda \in (-\delta, \delta)$,*

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}(e^{\lambda S_N}) = \Omega(\lambda) \tag{1.5}$$

where Ω is strictly convex continuously differentiable function with $\Omega'(0) = 0$. Then, for all $z \in \left(0, \dfrac{\Omega(\delta)}{\delta}\right)$,

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}(S_N \geq Nz) = -I(z) \tag{1.6}$$

where $I(z) = \sup_{\lambda\in(-\delta,\delta)} [z\lambda - \Omega(\lambda)]$.

Remark 1.4.

1. *If the limit (1.5) exists for all $\lambda \in \mathbb{R}$. Then, $B = \lim_{\delta\to\infty} \dfrac{\Omega(\delta)}{\delta}$ exists and (1.6) holds for all $z \in (0, B)$.*

2. *The function I appearing in the theorem is called the rate function because it gives us the exponential rate of decay of tail probabilities.*

In an on-going joint work with Pratima Hebbar, we develop a theory of higher order asymptotics for LDPs, using the weak forms of Edgeworth expansions and extensions of results in [27, Chapter VIII]. As in the CLT case, higher order asymptotics are given as expansions.

Definition 6. *Suppose S_N satisfies an LDP with rate function I. Then, S_N admits strong asymptotic expansion of order r for large deviations in the range $(0, L)$ if there are functions $C_p : (0, L) \to \mathbb{R}$ for $0 \leq p < \dfrac{r}{2}$ and $A > 0$ such that for each $a \in (0, L)$,*

$$\mathbb{P}(S_N - AN \geq aN)e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{C_p(a)}{N^{p+1/2}} + C_{r,a} \cdot o\left(\frac{1}{N^{\frac{r+1}{2}}}\right).$$

These expansions are in the spirit of the higher order expansions found [1] for i.i.d. sequences of random variables. In [7], authors refer to these expansions as strong large deviation results. [7, 32] establish the order 1 expansions under certain assumptions on the behaviour of the moment generating functions. These strengthen the results of [1] but only in the order 1 case. Here, we give an alternative way to establish the so-called strong large deviation results of all orders. We also manage to recover the results in [1] in the non-lattice setting. For applications of these results to statisitcs, see examples listed in [1, 7, 32] and references therein.

We also introduce the following *weak* form of the expansion for LDPs. As in the CLT case, we define these expansions over a function space $(\mathcal{F}, \|\cdot\|)$.

Definition 7. *Suppose S_N satisfies an LDP with rate function I. Then, S_N admits weak asymptotic expansion of order r for large deviations in the range $(0, L)$ for $f \in \mathcal{F}$, if there are functions $D_p : (0, L) \to \mathbb{R}$ for $0 \le p < \frac{r}{2}$ and $A > 0$ such that for each $a \in (0, L)$,*

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{D_p(a)}{N^{p+1/2}} + C_{r,a} \cdot o\left(\frac{1}{N^{\frac{r+1}{2}}}\right).$$

In fact, our results show that for a sequence X_n of i.i.d. $l-$Diophantine random variables with all exponential moments, for every r, S_N admits weak asymptotic expansions of order r for large deviations on $(0, \infty)$ for sufficiently regular f. This is a refinement of the LDP by Cramér for a broad class of random variables.

We also obtain similar results for certain classes of non–i.i.d. random variables. As an application, we obtain asymptotic expansions for the LDP in the case of Markov chains with smooth densities. In particular, let x_n be a time homogeneous Markov chain on a compact connected manifold $\mathcal{M}$ with a smooth transition density and $h : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ be smooth with non-degenerate critical points. Then $X_n = h(x_n, x_{n-1})$ admits asymptotic expansions for large deviations of all orders. These results are presented in Chapter 4.

Chapter 2: Central Limit Theorem: Discrete Random Variables.

2.1 Overview and main results.

Let X be a random variable with zero mean and positive variance σ^2. Let $S_n = \sum_{n=1}^{n} X_j$ where X_j are independent identically distributed and have the same distribution as X. Then, it is well-known that S_n satisfies the CLT with $A = 0$ and σ as in (1.1).

In this chapter, we consider a case which is opposite to X having a density, namely we suppose that X has a discrete distribution with $d+1$ atoms where $d \geq 2$. $d = 2$ is the simplest non-trivial case since distributions with two atoms are lattice and as a result they do not admit even the first order Edgeworth expansion.

Thus, we suppose that X takes values $a_1, \dots, a_{d+1}$ with probabilities $p_1, \dots, p_{d+1}$ respectively. Since X should have zero mean we suppose that our $2(d+1)-$tuple $(\mathbf{a}, \mathbf{p})$ belongs to the set

$$\Omega = \{p_i > 0, \quad p_1 + \cdots + p_{d+1} = 1, \quad p_1 a_1 + \cdots + p_{d+1} a_{d+1} = 0\}.$$

It is easy to see that S_n never admits the order d Edgeworth expansion. Indeed

$$\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \leq z) = \sum_{\substack{m_i \geq 0,\ \sum m_i = n \\ \sum m_i a_i \leq z}} \frac{n!}{m_1! \dots m_{d+1}!} p_1^{m_1} \dots p_{d+1}^{m_{d+1}}. \tag{2.1}$$

Applying the Local Central Limit Theorem to the time homogeneous $\mathbb{Z}^d$-random walk which jumps to $\mathbf{e}_i$ from the origin $\mathbf{0}$ with probability p_i for $i = 1, \dots, d$ and stays at $\mathbf{0}$ with probability p_{d+1} we conclude that if

$$\sum m_i a_i = n \sum a_i p_i + \mathcal{O}(\sqrt{n})$$

then

$$n^{d/2} \frac{n!}{m_1! \dots m_{d+1}!} p_1^{m_1} \dots p_{d+1}^{m_{d+1}}$$

is uniformly bounded from below. Accordingly $\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \le z)$ has jumps of order $n^{-d/2}$. On the other hand $\mathcal{E}_d(z)$ is a smooth function of z. So, it cannot approximate both $\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \le z - 0)$ and $\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \le z + 0)$ at the points of jumps.

Here we show that for typical $(\mathbf{a}, \mathbf{p})$ the order d Edgeworth expansion just barely fails. We present two results in this direction. For the first result let

$$b_j = a_j - a_1, \text{ for } j = 2 \dots d + 1.$$

Set

$$d(s) = \max_{j \in \{2, \dots d+1\}} \operatorname{dist}(b_j s, 2\pi\mathbb{Z}).$$

We say that $\mathbf{a}$ is β-Diophantine if there is a constant K such that for $|s| > 1$,

$$d(s) \ge \frac{K}{|s|^\beta}.$$

It is easy to see that almost all $\mathbf{a}$ is β-Diophantine provided that $\beta > (d-1)^{-1}$ (see [36, 47]).

Theorem 2.1.1. *If $\mathbf{a}$ is β-Diophantine and*

$$2\left(R - \frac{1}{2}\right)\beta < 1 \tag{2.2}$$

then

$$\lim_{n\to\infty} n^R \left[\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) - \mathcal{E}_{d-1}(z)\right] = 0.$$

Thus for almost every $\mathbf{a}$ the order $(d-1)$ Edgeworth expansion approximates the distribution of $\frac{S_n}{\sigma\sqrt{n}}$ with error $\mathcal{O}(n^{\varepsilon-d/2})$ for any ε.

Note that Theorem 2.1.1 applies for all βs, in particular for βs which are much larger than $(d-1)^{-1}$. However if β is large, then the statement of the theorem can be simplified. Namely, let r be the integer such that $r < 2R \leq r+1$. (Note that (2.2) can be rewritten as $2R < \frac{1}{\beta} + 1$ so provided that $2R$ is suffciently close to $\frac{1}{\beta} + 1$ we can take $r = \lfloor \beta^{-1} \rfloor + 1$. Then,

$$\begin{aligned}\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) &= \mathcal{E}_{d-1}(z) + o\left(\frac{1}{n^R}\right)\\ &= \mathcal{E}_r(z) + o\left(\frac{1}{n^R}\right) + \mathcal{O}\left(\mathcal{E}_{d-1}(z) - \mathcal{E}_r(z)\right).\end{aligned}$$

Since $\frac{r+1}{2} > R$ the first error term dominates the second and we obtain the following result.

Corollary 2.1.1.

$$\lim_{n\to\infty} n^R \left[\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) - \mathcal{E}_r(z)\right] = 0$$

provided that $\mathbf{a}$ *is* β*-Diophantine,* $r = 1 + \lfloor \beta^{-1} \rfloor$*, and* $r < 2R < \frac{1}{\beta} + 1$.

Theorem 2.1.1 shows that for almost every $\mathbf{a}$ and for $r \in \{1, \ldots, d-1\}$, the order r Edgeworth expansion is valid. Results that follow show that,

$$\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) - \mathcal{E}_d(z) \tag{2.3}$$

is typically of order $\mathcal{O}(n^{-d/2})$ but the $\mathcal{O}(n^{-d/2})$ term has wild oscillations. To formulate this result precisely we suppose that our $2(d+1)$-tuple is chosen at random according to an absolutely continuous distribution $\mathbf{P}$ on Ω. Thus (2.3) becomes a random variable.

Theorem 2.1.2. *There exists a smooth function $\Lambda(\mathbf{a},\mathbf{p})$ such that for each z the random variable*

$$e^{z^2/2}\frac{n^{d/2}}{\Lambda(\mathbf{a},\mathbf{p})}\left[\mathcal{E}_d(z)-\mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}}\le z\right)\right]$$

converges in law to a non-trivial random variable $\mathcal{X}$.

More precisely we have,

$$\Lambda(\mathbf{a},\mathbf{p})=\frac{|a_{d+1}-a_1|}{2^d\pi^{d+\frac{1}{2}}\sqrt{\det(D_{\mathbf{a},\mathbf{p}})}\;\;\sigma(\mathbf{a},\mathbf{p})} \tag{2.4}$$

where $D_{\mathbf{a},\mathbf{p}}$ is a $(d-1)\times(d-1)$ matrix defined by equations (2.37)–(2.38) of Section 2.5, $\sigma(\mathbf{a},\mathbf{p})$ denotes the standard deviation of the distribution of the random variable taking value a_j with probability p_j and $\mathcal{X}$ is defined as follows.

Let $\mathbb{M}$ be the space of pairs $(\mathcal{L},\chi)$ where $\mathcal{L}$ is a unimodular lattice in $\mathbb{R}^d$ and χ is a homeomorphism $\chi:\mathcal{L}\to\mathbb{T}$. In the formulas below, we identify $\mathbb{T}$ with segment $[0,1)$ equipped with addition modulo one. Given a vector $\mathbf{w}\in\mathbb{R}^d$ we denote by $y(\mathbf{w})$ its first coordinate and by $\mathbf{x}(\mathbf{w})$ its last $d-1$ coordinates.

Lemma 2.1.2. *For almost every pair $(\mathcal{L},\chi)\in\mathbb{M}$ with respect to the Haar measure the following limit exists*

$$\mathcal{X}(\mathcal{L},\chi)=\lim_{R\to\infty}\sum_{\mathbf{w}\in\mathcal{L}\backslash\{\mathbf{0}\},\ ||\mathbf{w}||\le R}\frac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})}e^{-||\mathbf{x}(\mathbf{w})||^2}. \tag{2.5}$$

In order to simplify the notation we will abbreviate expressions such as (2.5) by

$$\mathcal{X}(\mathcal{L},\chi)=\sum_{\mathbf{w}\in\mathcal{L}\backslash\{\mathbf{0}\}}\frac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})}e^{-||\mathbf{x}(\mathbf{w})||^2}. \tag{2.6}$$

The Haar measure on $\mathbb{M}$ can be defined in two equivalent ways. First, note that χ is of the form $\chi(\mathbf{w}) = e^{i\tilde{\chi}(\mathbf{w})}$ for some linear functional $\tilde{\chi} \in (\mathbb{R}^d)^*$. $SL_d(\mathbb{R})$ acts on $\mathbb{R}^d \oplus (\mathbb{R}^d)^*$ by the formula,

$$A(\mathbf{w}, \tilde{\chi}) = (A\mathbf{w}, \tilde{\chi}A^{-1}).$$

Observe that if $A(\mathbf{w}, \tilde{\chi}) = (\hat{\mathbf{w}}, \hat{\chi})$ then,

$$\tilde{\chi}(\mathbf{w}) = \hat{\mathbf{w}}(\hat{\chi}). \tag{2.7}$$

The above action of $SL_d(\mathbb{R})$ induces the following action of $SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^*$ on $\mathbb{M}$ given by,

$$(A, \tilde{\chi})(\mathcal{L}, \chi) = (A\mathcal{L}, e^{2\pi i t\tilde{\chi}} \cdot (\chi \circ A^{-1})).$$

This action is transitive because $SL_d(\mathbb{R})$ acts transitively on unimodular lattices and $(\mathbb{R}^d)^*$ acts transitively on characters. This allows us to identify $\mathbb{M}$ with $(SL_d(\mathbb{R}) \ltimes \mathbb{R}^d)/(SL_d(\mathbb{Z}) \ltimes \mathbb{Z}^d)$ and so $\mathbb{M}$ inherits the Haar measure from $SL_d(\mathbb{R}) \ltimes \mathbb{R}^d$.

The second way to define the Haar measure is to note that the space $\mathcal{M}$ of unimodular lattices is naturally identified with $SL_d(\mathbb{R})/SL_d(\mathbb{Z})$ and so it inherits the Haar measure from $SL_d(\mathbb{R})$. Next for a fixed $\mathcal{L}$ the set of homeomorphisms $\chi : \mathcal{L} \to \mathbb{T}$ is a d dimensional torus so it comes with its own Haar measure.

Now, if we want to compute the average of a function $\Phi(\mathcal{L}, \chi)$ with respect to the Haar measure then we can first compute its average $\bar{\Phi}(\mathcal{L})$ in each fiber and then integrate the result with respect to the Haar measure on the space of lattices. In the proof of Lemma 2.1.2 given in Section A.1 the averaging inside a fiber will be denoted by $\mathbf{E}_\chi$ and the averaging with respect to the Haar measure on the space of lattices will be denoted by $\mathbf{E}_\mathcal{L}$.

If we assume that the pair $(\mathcal{L}, \chi)$ is distributed according to the Haar measure on $\mathbb{M}$ then $\mathcal{X}$, defined in Lemma 2.1.2, becomes a random variable. This is the variable mentioned in Theorem 2.1.2. Note that the distribution of $\mathcal{X}$ depends neither on $\mathbf{P}$ nor on z.

Using the second representation of the Haar measure we can also describe $\mathcal{X}$ as follows. Let $\mathbf{w}_1, \ldots, \mathbf{w}_d$ be the shortest spanning set of $\mathcal{L}$. That is $\mathbf{w}_1$ is the shortest non zero vector in $\mathcal{L}$ and, for $j > 1$, $\mathbf{w}_j$ is the shortest vector which is linearly independent of $\mathbf{w}_1, \ldots, \mathbf{w}_{j-1}$. Given $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{Z}^d$ let $(y, \mathbf{x})(\mathbf{m})$, $y \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^{d-1}$, denote the point

$$m_1\mathbf{w}_1 + \cdots + m_d\mathbf{w}_d \in \mathcal{L}. \tag{2.8}$$

Let $\theta_j = \chi(\mathbf{w}_j)$. Then θ_j are uniformly distributed on $\mathbb{T}$ and independent of each other. Set $\theta(\mathbf{m}) = m_1\theta_1 + \cdots + m_d\theta_d$. Now $\mathcal{X}$ (see definition in Lemma 2.1.2) can be rewritten as

$$\mathcal{X} = \sum_{\mathbf{m}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} \frac{\sin(2\pi\theta(\mathbf{m}))}{y(\mathbf{m})} e^{-||\mathbf{x}(\mathbf{m})||^2} \tag{2.9}$$

where $\mathcal{L}$ is uniformly distributed on the space of lattices, $(y, \mathbf{x})(\mathbf{m})$ is defined by (2.8), and $(\theta_1, \ldots \theta_d)$ is uniformly distributed on $\mathbb{T}^d$ and independent of $\mathcal{L}$.

Theorems 2.1.1 and 2.1.2 have analogues when we consider probabilities that S_n belongs to finite intervals. In particular, our results have applications to the Local Limit Theorem.

Theorem 2.1.3. *Let $z_1(n)$ and $z_2(n)$ be two uniformly bounded sequences such that $|z_1(n) - z_2(n)|n^{d/2} \to \infty$. Then, the random vector,*

$$\frac{n^{d/2}}{\Lambda(\mathbf{a},\mathbf{p})}\left(e^{z_1^2/2}\left[\mathcal{E}_d(z_1) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z_1\right)\right], e^{z_2^2/2}\left[\mathcal{E}_d(z_2) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z_2\right)\right]\right) \tag{2.10}$$

converges in law to a random vector $(\mathcal{X}(\mathcal{L},\chi_1),\mathcal{X}(\mathcal{L},\chi_2))$ *where* $\mathcal{X}(\mathcal{L},\chi)$ *is defined by* (2.6) *and the triple* $(\mathcal{L},\chi_1,\chi_2)$ *is uniformly distributed on* $(SL_d(\mathbb{R})/SL_d(\mathbb{Z}))\times\mathbb{T}^d\times\mathbb{T}^d$.

Here and below the uniform distribution of $(\mathcal{L},\chi_1,\chi_2)$ means that $\mathcal{L}$ is uniformly distributed on the space of lattices and for a given lattice, χ_1 and χ_2 are chosen independently and uniformly from the space of characters.

Theorem 2.1.4. *Let* $z_1(n) < z_2(n)$ *be two uniformly bounded sequences such that* $l_n = z_2(n) - z_1(n) \to 0$.

(*a*) *If* $l_n \geq Cn^{\varepsilon - d/2}$ *for some* $\varepsilon > 0$ *then*

$$\frac{\mathbb{P}_{\mathbf{a},\mathbf{p}}(z_1 < \frac{S_n}{\sigma\sqrt{n}} < z_2)}{l_n\mathfrak{n}(z_1)} \to 1 \textit{ almost surely.}$$

(*b*) *If* $l_n n^{d/2} \to \infty$ *then*

$$\frac{\mathbb{P}_{\mathbf{a},\mathbf{p}}(z_1 < \frac{S_n}{\sigma\sqrt{n}} < z_2)}{l_n\mathfrak{n}(z_1)} \Rightarrow 1$$

(here and below "⇒" denotes the convergence in law).

(*c*) *If* $l_n = \dfrac{c|a_{d+1}-a_1|}{\sigma(\mathbf{a},\mathbf{p})n^{d/2}}$ *then*

$$2^{d-\frac{3}{2}}\pi^d\sqrt{\det(D_{\mathbf{a},\mathbf{p}})}\left[\frac{\mathbb{P}_{\mathbf{a},\mathbf{p}}(z_1 < \frac{S_n}{\sigma\sqrt{n}} < z_2)}{l_n\mathfrak{n}(z_1)} - 1\right] \Rightarrow \mathcal{Y}$$

where

$$\mathcal{Y}(\mathcal{L},\chi,c) = \sum_{\mathbf{w}\in\mathcal{L}\backslash\{\mathbf{0}\}} \frac{\sin(2\pi[\chi(\mathbf{w}) - cy(\mathbf{w})]) - \sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})} e^{-||\mathbf{x}(\mathbf{w})||^2}$$

and $\mathcal{L},\chi$ *are as in Theorem 2.1.2 and* $D_{\mathbf{a},\mathbf{p}}$ *given by equations* (2.37)–(2.38).

The intuition behind this result is the following. Call y_n δ-plausible if $\mathbb{P}(S_n = y_n) \geq \delta n^{-d/2}$. The discussion following (2.1) shows that for each δ there are about

$C(\delta)n^{d/2}$ δ-plausible values. Therefore, if $l_n \ll n^{-d/2}$ then the interval $[z_1(n), z_2(n)]$ would typically contain no plausible values. Hence, we should not expect the LLT to hold on that scale. Theorem 2.1.4 shows that as soon as interval $[z_1(n), z_2(n)]$ contains many plausible values then the LLT typically holds for this interval.

Recall that,

$$\mathbb{P}_{\mathbf{a},\mathbf{p}}(S_n \in [z_1, z_2]) = \sum_{\substack{m_i \geq 0,\, \sum m_i = n \\ z_1 \leq \sum m_i a_i \leq z_2}} \frac{n!}{m_1! \dots m_{d+1}!} p_1^{m_1} \dots p_{d+1}^{m_{d+1}}.$$

Thus, in Theorem 2.1.4 we just count the number of visits of a random linear form $\sum m_i a_i$ to a finite interval with weights given by multinomial coefficients. It is also interesting to consider counting with equal weight. In this case the analogue of Theorem 2.1.4(c) is obtained in [38] while for longer intervals only partial results are available, for example see [15, 34].

The chapter is organized as follows. Theorem 2.1.1 is proven in Section 2.2. The proof is a minor modification of the arguments of [20, Chapter XVI]. The bulk of the chapter is devoted to the proof of Theorem 2.1.2. In Section 2.3 we provide an equivalent formula for $\mathcal{X}$. This formula looks more complicated than (2.6) but it is easier to identify with the limit of the error term. Section 2.4 contains preliminary reductions. We show that the density ρ on Ω could be assumed smooth and the integration in the Fourier inversion formula could be restricted to a finite domain. In Section 2.5, we show that main contribution to the error term comes from resonances where characteristic function of S_n is close to 1 in absolute value. The proof relies on several technical estimates which are established in Section 2.6. In Section 2.7, we use dynamics on homogenuous spaces in order to show that the contribution of resonances converges to (2.6) completing the proof of Theorem 2.1.2. The proofs of Theorems 2.1.3 and 2.1.4 are similar to the proof of Theorem 2.1.2. The necessary modifications are explained in Section 2.8. We postpone the proof of Lemma 2.1.2 till Appendix A.1.

2.2 Edgeworth Expansion under Diophantine conditions.

Theorem 2.1.1 is a consequence of Theorem 2.2.1 below and the fact that in our case there is a positive constant c such that

$$|\phi(s)| \leq 1 - cd(s)^2. \tag{2.11}$$

(2.11) follows from inequality (2.35) proven in Section 2.5.

Theorem 2.2.1. *If the distribution of X has $d+2$ moments and its characteristic function satisfies*

$$|\phi(s)| \leq 1 - \frac{K}{|s|^\gamma} \tag{2.12}$$

and $R < \dfrac{d}{2}$ is such that

$$\left(R - \frac{1}{2}\right)\gamma < 1 \tag{2.13}$$

then

$$\lim_{n\to\infty} n^R \left[\mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) - \mathcal{E}_{d-1}(z)\right] = 0.$$

Theorem 2.2.1 follows easily from the estimates in [20, ChapterXVI] but we provide the proof here for completeness.

Proof. Denoting

$$\bar{\Delta}_n(\mathbf{a}, \mathbf{p}) = \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right) - \mathcal{E}_{d-1}(z)$$

we get by [20, Chapter XVI] that for each T

$$|\bar{\Delta}_n(\mathbf{a},\mathbf{p})| \leq \frac{1}{\pi}\int_{-\frac{T}{\sigma\sqrt{n}}}^{\frac{T}{\sigma\sqrt{n}}} \left|\frac{\phi^n(s)-\hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s}\right| ds + \frac{C}{T}. \tag{2.14}$$

Choose $T = Bn^R$ with $B = \dfrac{C}{\varepsilon}$. Then, $\dfrac{C}{T} = \dfrac{\varepsilon}{n^R}$. Take a small δ and split the integral in the RHS of (2.14) into two parts.

$$\frac{1}{\pi}\int_{-\delta}^{\delta}\left|\frac{\phi^n(s)-\hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s}\right| ds \;+\; \frac{1}{\pi}\int_{\delta<|s|<Bn^{R-1/2}/\sigma}\left|\frac{\phi^n(s)-\hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s}\right| ds. \tag{2.15}$$

Again, by [20, Chapter XVI], we have that the first integral of (2.15) is $\mathcal{O}\left(n^{-d/2}\right)$. Also, $\displaystyle\int_{|s|>\delta}\left|\frac{\hat{\mathcal{E}}_{d-1}(s\sigma\sqrt{n})}{s}\right| ds$ has exponential decay as $n\to\infty$. Put $J=\{s:\quad \delta < |s| < Bn^{R-1/2}/\sigma\}$. Thus, we only need to approximate,

$$\int_J\left|\frac{\phi^n(s)}{s}\right| ds \leq \frac{1}{\delta}\int_J |\phi^n(s)|\, ds \leq \frac{C}{\delta}\int_J \exp\left(-\bar{c}\; n^{1-\left(R-\frac{1}{2}\right)\gamma}\right) ds \tag{2.16}$$

where the last inequality is due to (2.12). By (2.13) the integral decay faster than any power of n. Because $R<\dfrac{d}{2}$ the contribution of $|s|\leq\delta$ is also under control. $\square$

2.3 Change of variables.

Here we deduce Theorem 2.1.2 from:

Theorem 2.1.2*. For each z the random variable

$$n^{d/2}\left[\mathcal{E}_d(z) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}}\leq z\right)\right]$$

converges in law to $\hat{\mathcal{X}}$ where

$$\hat{\mathcal{X}}(\mathfrak{a},\mathfrak{p},\mathcal{L},\chi) = e^{-z^2/2}\frac{|\mathfrak{a}_{d+1}-\mathfrak{a}_1|}{2\sigma(\mathfrak{a},\mathfrak{p})\sqrt{\pi^3}}\sum_{\mathbf{w}\in\mathcal{L}\setminus\{\mathbf{0}\}}\frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})}e^{-4\pi^2\mathbf{x}(\mathbf{w})^T D_{\mathfrak{a},\mathfrak{p}}\mathbf{x}(\mathbf{w})} \tag{2.17}$$

$\mathfrak{a} = (\mathfrak{a}_1, \ldots, \mathfrak{a}_{d+1})$, $\mathfrak{p} = (\mathfrak{p}_1, \ldots, \mathfrak{p}_{d+1})$ and $(\mathfrak{a}, \mathfrak{p}) \in \Omega$ are distributed according to $\mathbf{P}$ and $D_{\mathfrak{a},\mathfrak{p}}$ and $\sigma(\mathfrak{a}, \mathfrak{p})$ are defined immediately after (2.4).

In order to deduce Theorem 2.1.2 from Theorem 2.1.2* we need to show that $e^{z^2/2} \dfrac{\hat{\mathcal{X}}}{\Lambda(\mathfrak{a}, \mathfrak{p})}$ has the same distribution as $\mathcal{X}$. To this end we rewrite the sum in (2.17) as

$$\frac{1}{(2\pi)^{d-1} \det(\sqrt{D_{\mathfrak{a},\mathfrak{p}}})} \sum_{\mathbf{w} \in \mathcal{L} \setminus \{0\}} \frac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})/((2\pi)^{d-1} \det(\sqrt{D_{\mathfrak{a},\mathfrak{p}}}))} e^{-4\pi^2 ||(\sqrt{D_{\mathfrak{a},\mathfrak{p}}} \mathbf{x}(\mathbf{w}))||^2}. \quad (2.18)$$

Let A be the linear map such that

$$A(y, \mathbf{x}) = \left(\frac{y}{(2\pi)^{d-1} \sqrt{\det(D_{\mathfrak{a},\mathfrak{p}})}},\ 2\pi \sqrt{D_{\mathfrak{a},\mathfrak{p}}}\ \ \mathbf{x} \right).$$

Put $(\bar{\mathcal{L}}, \bar{\chi}) = A(\mathcal{L}, \chi)$. Then, using (2.7), (2.18) can be rewritten as:

$$\frac{1}{(2\pi)^{d-1} \det(\sqrt{D_{\mathfrak{a},\mathfrak{p}}})} \sum_{\bar{\mathbf{w}} \in \mathcal{L} \setminus \{0\}} \frac{\sin(2\pi\bar{\chi}(\bar{\mathbf{w}}))}{y(\bar{\mathbf{w}})} e^{-||\mathbf{x}(\bar{\mathbf{w}}))||^2}.$$

Since $\det(A) = 1$, the pair $(\bar{\mathcal{L}}, \bar{\chi})$ is distributed according to the Haar measure on $\mathbb{M}$ proving our formula for $\mathcal{X}$.

Sections 2.4–2.7 are devoted to the proof of Theorem 2.1.2*. Note that similarly to (2.9) we have

$$\hat{\mathcal{X}} = e^{-z^2/2} \frac{|\mathfrak{a}_{d+1} - \mathfrak{a}_1|}{2\sigma(\mathfrak{a}, \mathfrak{p})\sqrt{\pi^3}} \sum_{\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \frac{\sin 2\pi\theta(\mathbf{m})}{y(\mathbf{m})} e^{-4\pi^2 \mathbf{x}(\mathbf{m})^T D_{\mathfrak{a},\mathfrak{p}} \mathbf{x}(\mathbf{m})}.$$

The statements of Theorems 2.1.2 and 2.1.2* look similar, however, there is an important distinction. Namely the proof of Theorem 2.1.2* is constructive. In the course of the proof given n, $\mathbf{a}$ and z we construct a lattice $\mathcal{L}(\mathbf{a}, n)$ and a character $\chi(\mathbf{a}, \mathbf{p}, n, z)$ such that the expression $n^{-d/2} \hat{\mathcal{X}}(\mathbf{a}, \mathbf{p}, \mathcal{L}(\mathbf{a}, n), \chi(\mathbf{a}, \mathbf{p}, n, z))$ well-approximates the error in the Edgeworth expansion. We believe that such a construction could be made for more general distributions where the Edgeworth expansion fails, and this will be a subject of a future investigation. So the difference between Theorems 2.1.2 and 2.1.2* is that in the first case we have only an approximation in law while in the second case we are able to obtain an approximation in probability.

2.4 Cut off.

2.4.1 Density.

Here we show that it is enough to prove Theorem 2.1.2* under the assumption that $\mathbf{P}$ has smooth density supported on a subset

$$\Omega_\kappa = \{(\mathbf{a}, \mathbf{p}) \in \Omega : \forall i \, p_i \geq \kappa \quad \text{and} \quad \forall i \neq j \; |a_i - a_j| \geq \kappa\}$$

for some $\kappa > 0$. Indeed suppose that the theorem is true for such densities. Let $p(\mathbf{a}, \mathbf{p})$ the original density of $\mathbf{P}$. Let ϕ be a bounded continuous test function. Given ε we can find a smooth density $\tilde{p}(\mathbf{a}, \mathbf{p})$ supported on some Ω_κ such that $||p - \tilde{p}||_{L^1} \leq \varepsilon$. In Section 2.7 we prove that

$$\int \phi(n^{d/2}\Delta_n)\tilde{p}\, d\mathbf{a}\, d\mathbf{p} \to \iint \phi(\hat{\mathcal{X}}(\mathbf{a}, \mathbf{p}, \mathcal{L}, \boldsymbol{\theta}))\tilde{p}\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L}, \boldsymbol{\theta}) \tag{2.19}$$

where $\Delta_n = \mathcal{E}_d(z) - \mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \leq z\right)$ and μ is the Haar measure on $(SL_d(\mathbb{R})/SL_d(\mathbb{Z})) \times \mathbb{T}^d$. Let $p_m(\mathbf{a}, \mathbf{p})$ be the smooth density supported on Ω_κ corresponding to $\varepsilon = m^{-1}$. Passing to subsequence, $p_m \to p$ almost surely. Because $|p_m\phi| \leq \|\phi\||p_m| \in L^1$ and $|p\phi| \leq \|\phi\||p| \in L^1$ and $\|\phi\||p_m| \to \|\phi\||p|$ almost surely, Lebesgue Dominated Convergence Theorem gives

$$\iint \phi(\hat{\mathcal{X}}(\mathbf{a},\mathbf{p},\mathcal{L},\boldsymbol{\theta}))p_m\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L},\boldsymbol{\theta}) \\ \to \iint \phi(\hat{\mathcal{X}}(\mathbf{a},\mathbf{p},\mathcal{L},\boldsymbol{\theta}))p\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L},\boldsymbol{\theta}). \quad (2.20)$$

Combining (2.19) and (2.20) we have that,

$$\begin{aligned} \int \phi(n^{d/2}\Delta_n)p\, d\mathbf{a}\, d\mathbf{p} &= \int \phi(n^{d/2}\Delta_n)p_m\, d\mathbf{a}\, d\mathbf{p} + \mathcal{O}(m^{-1}\|\phi\|) \qquad (2.21) \\ &\xrightarrow{n\to\infty} \iint \phi(\hat{\mathcal{X}}(\mathbf{a},\mathbf{p},\mathcal{L},\boldsymbol{\theta}))p_m\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L},\boldsymbol{\theta}) + \mathcal{O}(m^{-1}\|\phi\|) \\ &\xrightarrow{m\to\infty} \iint \phi(\hat{\mathcal{X}}(\mathbf{a},\mathbf{p},\mathcal{L},\boldsymbol{\theta}))p\, d\mathbf{a}\, d\mathbf{p}\, d\mu(\mathcal{L},\boldsymbol{\theta}). \end{aligned}$$

2.4.2 Fourier transform.

As in the previous section let

$$\Delta_n = \mathcal{E}_d(z) - F_n(z) \quad \text{where} \quad F_n(z) = \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z\right).$$

Denote by $v_T(x) = \frac{1}{\pi}\cdot\frac{1-\cos Tx}{Tx^2}$ and let $\mathcal{V}(s,T) = \left(1 - \frac{|s|}{T}\right)\mathbb{1}_{|s|\le T}$ be its Fourier transform. Using the approach of [20, Section XVI.3] we let $T_2 = n^{2d+6}$ and decompose

$$\Delta_n = [\mathcal{E}_d - F_n] \star v_{T_2}(z) - [F_n - F_n \star v_{T_2}](z) + [\mathcal{E}_d - \mathcal{E}_d \star v_{T_2}](z). \qquad (2.22)$$

To estimate the last term we split

$$\begin{aligned} [\mathcal{E}_d - \mathcal{E}_d \star v_{T_2}](z) &= \int_{|x|\le 1} [\mathcal{E}_d(z) - \mathcal{E}_d(z-x)]\, v_{T_2}(x)dx \qquad (2.23) \\ &+ \int_{|x|\ge 1} [\mathcal{E}_d(z) - \mathcal{E}_d(z-x)]\, v_{T_2}(x)dx. \end{aligned}$$

Since v_T is even the first integral in (2.23) equals to

$$\int_{|x|\le 1} \mathcal{E}_d'(z) x v_{T_2}(x)dx - \int_{|x|\le 1} \frac{\mathcal{E}_d''(y(z,x))}{2} x^2 v_{T_2}(x)dx$$

$$= \int_{|x|\le 1} \frac{\mathcal{E}_d''(y(z,x))}{2} \frac{1-\cos T_2 x}{\pi T_2} dx = \mathcal{O}\left(\frac{1}{T_2}\right).$$

Since both $\mathcal{E}_d$ and cosine are bounded the second integral in (2.23) is bounded by

$$C \int_{|x|\ge 1} \frac{dx}{T_2 x^2} = \frac{C}{T_2}.$$

Thus the last term in (2.22) is $\mathcal{O}\left(T_2^{-1}\right)$. To estimate the second term in (2.22) we split the integral in $F_n \star v_{T_2}$ into regions $\{|x| \ge 1/\sqrt{T_2}\}$ and $\{|x| \le 1/\sqrt{T_2}\}$. The contribution of $\{|x| \ge 1/\sqrt{T_2}\}$ is bounded by

$$C \int_{1/\sqrt{T_2}}^{\infty} \frac{dx}{T_2 x^2} = \frac{C}{\sqrt{T_2}}.$$

On the other hand

$$\int_{|x|\le 1/\sqrt{T_2}} \left[F_n(z) - F_n(z-x)\right] V_{T_2}(x) dx = 0$$

unless there is a point of increase of F_n inside $\left[z - 1/\sqrt{T_2}, z + 1/\sqrt{T_2}\right]$. The probability that such a point exists is bounded by

$$\sum_{m_1+\cdots+m_{d+1}=n} \mathbf{P}\left(m_1 a_1 + \cdots + m_{d+1} a_{d+1} \in \left[z - 1/\sqrt{T_2}, z + 1/\sqrt{T_2}\right]\right). \qquad (2.24)$$

Note that for each fixed $(m_1, \ldots, m_{d+1})$ the random variable

$$m_1 a_1 + \cdots + m_{d+1} a_{d+1}$$

has a bounded density with respect to the uniform distribution on the segment of length $\mathcal{O}\left(\sqrt{m_1^2 + \cdots + m_{d+1}^2}\right)$ and so

$$\mathbb{P}(\mathbf{m}.\mathbf{a} \in J) = \mathcal{O}\left(\frac{|J|}{\|\mathbf{m}\|}\right)$$

for any interval J. Hence each term in (2.24) is $\mathcal{O}\left(\frac{1}{n\sqrt{T_2}}\right)$ and so the sum is $\mathcal{O}\left(\frac{n^d}{n\sqrt{T_2}}\right)$. Thus with probability $1-\mathcal{O}\left(\frac{1}{n^4}\right)$ we have that $\Delta_n = \Delta_{n,2}+\mathcal{O}\left(T_2^{-1/2}\right)$ where

$$\begin{aligned}\Delta_{n,2} &= \frac{1}{2\pi}\int_{-T_2}^{T_2} \frac{\left[\phi^n\left(\frac{t}{\sqrt{n}}\right) - \hat{\mathcal{E}}_d(t)\right]}{it}\mathcal{V}(t,T_2)e^{-itz}dt \\ &= \frac{1}{2\pi}\int_{-\frac{T_2}{\sigma\sqrt{n}}}^{\frac{T_2}{\sigma\sqrt{n}}} e^{-isz\sigma\sqrt{n}}\ \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{is}\mathcal{V}(s,n,T_2)ds\,,\end{aligned}$$

$\mathcal{V}(s,n,T) = 1 - \left|\frac{s\sigma\sqrt{n}}{T}\right|$ and $\phi(s)$ is the characteristic function of X given by

$$\phi(s) = p_1e^{isa_1} + \cdots + p_{d+1}e^{isa_{d+1}}.$$

Let $T_1 = K_1n^{d/2}$ and define

$$\Delta_{n,1} = \frac{1}{2\pi}\int_{-\frac{T_1}{\sigma\sqrt{n}}}^{\frac{T_1}{\sigma\sqrt{n}}} e^{-isz\sigma\sqrt{n}}\ \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{is}\ \mathcal{V}(s,n,T_2)\ ds.$$

Let $\Gamma_n = \Delta_{n,2} - \Delta_{n,1}$. Put

$$\tilde{\Gamma}_n = \frac{1}{2\pi}\int_{|s|\in\left[T_1/(\sigma\sqrt{n}),T_2/(\sigma\sqrt{n})\right]} e^{-isz\sigma\sqrt{n}}\ \frac{\phi^n(s)}{is}\ \mathcal{V}(s,n,T_2)\ ds.$$

Then, we have $\Gamma_n = \tilde{\Gamma}_n + \mathcal{O}\left(e^{-\varepsilon T_1^2}\right)$ due to the exponential decay of $\hat{\mathcal{E}}_d$.

The main result of Subsection 2.4.2 is the following.

Proposition 2.4.1.

$$\left\|\tilde{\Gamma}_n\right\|_{L^2} \leq \frac{C}{\sqrt{T_1n^d}}. \tag{2.25}$$

Proof.

$$\mathbf{E}(\tilde{\Gamma}_n^2) = \iint \mathbf{E}\left(e^{-i(s_1+s_2)z\sigma\sqrt{n}}\phi^n(s_1)\phi^n(s_2)\mathcal{V}(s_1,n,T_2)\mathcal{V}(s_2,n,T_2)\right)\frac{ds_1}{s_1}\frac{ds_2}{s_2}.$$

We split this integral into two parts.

(1) In the region where $|s_1+s_2| \leq 1$ we use Corollary 2.5.2 proven in Section 2.5 to estimate the integral by

$$\mathcal{O}\left(\int_{|s|\in\left[T_1/(\sigma\sqrt{n}),T_2/(\sigma\sqrt{n})\right]} \frac{1}{\sqrt{n}s_1^2}\mathbf{E}\left(|\phi^n(s_1)|\right)\, ds_1\right). \tag{2.26}$$

The next result will be proven in Section 2.6.

Lemma 2.4.2.

$$\mathbf{E}\left(|\phi^n(s_1)|\right) \leq \frac{C}{n^{d/2}}.$$

Plugging the estimate of Lemma 2.4.2 into (2.26) and integrating we see that the contribution of the first region to $\mathbf{E}(\tilde{\Gamma}_n^2)$ is $\mathcal{O}\left(\frac{1}{T_1 n^{d/2}}\right)$.

(2) Consider now the region where $|s_1+s_2| \geq 1$. Denote

$$b_{d+1} = a_{d+1} - a_1,\ \ldots,\ b_2 = a_2 - a_1.$$

Then

$$\phi(s) = e^{isa_1}\psi(s) \quad \text{where} \quad \psi(s) = p_1 + p_2 e^{isb_2} + \cdots + p_{d+1}e^{isb_{d+1}}.$$

Denote $\boldsymbol{\nu} = (p_1, \ldots, p_d, b_2, \ldots, b_d)$. Then there exists a compactly supported density $\rho = \rho(a_1, \boldsymbol{\nu})$ such that the contribution of the second region is

$$\iint_{|s_1+s_2|\geq 1}\left(\int e^{-i(s_1+s_2)z\sigma\sqrt{n}}e^{in(s_1+s_2)a_1}\psi^n(s_1)\psi^n(s_2)\mathcal{V}(s_1)\mathcal{V}(s_2)\rho\, da_1\, d\boldsymbol{\nu}\right)\frac{ds_1}{s_1}\frac{ds_2}{s_2}.$$

We are able to use a $2d$-dimensional coordinate system because on Ω

$$p_1 + \cdots + p_{d+1} = 1, \quad \text{and} \quad p_1a_1 + \cdots + p_{d+1}a_{d+1} = 0. \tag{2.27}$$

To estimate this integral we integrate by parts with respect to a_1. We use that

$$e^{isna_1}da_1 = \left[\frac{1}{isn}\frac{d}{da_1}\right]^k de^{isna_1}$$

for some large k (for example we can take $k = 2d + 1$). The integration by parts amounts to applying $\left(\frac{d}{da_1}\right)^k$ to $\left(e^{isz\sigma\sqrt{n}}\rho[\psi(s_1)\psi(s_2)]^n\right)$ which leads to the terms

$$\left\{\left(\frac{d}{da_1}\right)^{k_1}\left[e^{i(s_1+s_2)z\sigma\sqrt{n}}\right]\right\}\left\{\left(\frac{d}{da_1}\right)^{k_2}[\rho]\right\}\left\{\left(\frac{d}{da_1}\right)^{k_3}[[\psi(s_1)\psi(s_2)]^n]\right\}$$

where $k_1 + k_2 + k_3 = k$. (Note that both σ and ψ depend on a_1 implicitly due to the second equation in (2.27)). Thus, the contribution of the above term to the integral is bounded by

$$C\iint_{\substack{|s_1|,|s_2|\in[T_1/\sigma\sqrt{n},T_2/\sigma\sqrt{n}]\\ |s_1+s_2|\geq 1}} \frac{(s_1+s_2)^{k_1}\, n^{(k_1/2)+k_3}}{(s_1+s_2)^k\, n^k}\mathbf{E}\left(|\phi^n(s_1)|\right)\frac{ds_1}{s_1}\frac{ds_2}{s_2}.$$

Using Lemma 2.4.2 again we can estimate the above integral by

$$\begin{cases} \dfrac{C}{n^{k/2}} & \text{if} \quad k_1 \geq k-2 \\[2ex] \dfrac{C}{T_1 n^{k+d/2-k_1/2-k_3}} & \text{otherwise.} \end{cases}$$

Thus the main contribution comes from $k_1 = k_2 = 0$, $k_3 = k$ proving Proposition 2.4.1. $\square$

Proposition 2.4.1 shows that the contribution from $\tilde{\Gamma}_n$ to the L^2-limit of $n^{d/2}\Delta_n$ can be made arbitrarily small by choosing K_1 large. Also, on $|s| \leq T_1/\sigma\sqrt{n}$ we have

$$\mathcal{V}(s,n,T_2) = \left(1 - \frac{s\sigma\sqrt{n}}{T_2}\right)\mathbb{1}_{|s|<T_2/\sigma\sqrt{n}} = 1 - \frac{s\sigma}{n^{2d+\frac{11}{2}}}.$$

Hence $\Delta_{n,1} = \hat{\mathbf{\Delta}}_n + o(n^{-2d})$ where

$$\hat{\mathbf{\Delta}}_n := \frac{1}{2\pi}\int_{|s|\leq T_1/\sigma\sqrt{n}} \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{is}e^{-isz\sigma\sqrt{n}}ds$$

approximates well $\Delta_{n,1}$ and hence, Δ_n too. Also, the error from this approximation of $n^{d/2}\Delta_n$ converges to 0 in L^2. Hence, we only need to analyze $n^{d/2}\hat{\mathbf{\Delta}}_n$ for large n.

2.5 Simplifying the error.

Denote

$$s_k = \frac{2\pi k}{|b_{d+1}|}$$

and let I_k be the segment of length $\frac{2\pi}{|b_{d+1}|}$ centered at s_k. Put $K_2 \gg K_1$. Due to the results of the previous section it is sufficient to study

$$\hat{\mathbf{\Delta}}_n = \sum_{|k| \leq K_2\sqrt{n}} \tilde{\mathcal{I}}_k$$

where

$$\tilde{\mathcal{I}}_k = \frac{1}{2\pi i} \int_{I_k} e^{-isz\sigma\sqrt{n}} \, \frac{\phi^n(s) - \hat{\mathcal{E}}_d(s\sigma\sqrt{n})}{s} ds.$$

$\tilde{\mathcal{I}}_0 = \mathcal{O}(n^{-(d+1)/2})$ due to [20, Section XVI.2]. Next, $\hat{\mathcal{E}}_d(s\sigma\sqrt{n})$ decays exponentially with respect to n outside of I_0. So, its contribution to $\tilde{\mathcal{I}}_k$ is negligible for $k \neq 0$. Accordingly,

$$\hat{\mathbf{\Delta}}_n = \sum_{0<|k| \leq K\sqrt{n}} \mathcal{I}_k + \mathcal{O}\left(\frac{1}{n^{(d+1)/2}}\right)$$

where

$$\mathcal{I}_k = \frac{1}{2\pi i} \int_{I_k} e^{-isz\sigma\sqrt{n}} \, \frac{\phi^n(s)}{s} \mathbb{1}_{|s| \leq T_1/\sigma\sqrt{n}} \, ds.$$

Introduce the following notation

$$\bar{s}_k = \arg\max_{s \in I_k} |\phi(s)|, \quad \phi(\bar{s}_k) = r_k e^{i\phi_k}.$$

The following lemma is similar to the results of [12, Section 5.2].

Lemma 2.5.1. *Suppose that*

$$r_k^n \geq n^{-100} \tag{2.28}$$

and

$$\pm \frac{T_1}{\sigma\sqrt{n}} \notin I_k. \tag{2.29}$$

Then

$$\mathcal{I}_k = \frac{1}{i\sqrt{\pi n}\sigma}\frac{r_k^n}{\bar{s}_k} e^{-z^2/2}\ e^{in\phi_k - i\bar{s}_k z\sigma\sqrt{n}}(1+o_{n\to\infty}(1)).$$

Proof. Let $e^{i\bar{s}_k a_j} = e^{i(\phi_k+\beta_j(k))}$. Then

$$r_k = \sum_{j=1}^{d+1} p_j \cos(\beta_j(k)) \tag{2.30}$$

and

$$\sum_{j=1}^{d+1} p_j \sin(\beta_j(k)) = 0. \tag{2.31}$$

Since (2.28) implies that $r_k \geq 1 - \frac{C\ln n}{n}$, (2.30) shows that $|\beta_j(k)| \leq C\sqrt{\frac{\ln n}{n}}$ and so (2.31) gives

$$\sum_{j=1}^{d+1} p_j\beta_j(k) = \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}}\right). \tag{2.32}$$

Now we use Taylor expansion

$$e^{i(\bar{s}_k+\delta)a_j} = e^{i\phi_k}\left(1 + i\beta_j(k) - \frac{\beta_j(k)^2}{2}\right)\left(1 + ia_j\delta - \frac{a_j^2\delta^2}{2}\right) + \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}} + \delta^3\right). \tag{2.33}$$

Thus,

$$\begin{aligned} \phi(\bar{s}_k+\delta) &= e^{i\phi_k}\sum_{j=1}^{d+1} p_j\left(\cos(\beta_j(k)) - \frac{a_j^2\delta^2}{2}\right) + \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}} + \delta^3\right) \\ &= r_k e^{i\phi_k}\left(1 - \frac{\sigma^2\delta^2}{2}\right) + \mathcal{O}\left(\frac{\ln^{3/2} n}{n^{3/2}} + \delta^3\right) \end{aligned} \tag{2.34}$$

where we have used (2.32) as well as

$$p_1a_1 + \cdots + p_{d+1}a_{d+1} = 0, \quad p_1a_1^2 + \cdots + p_{d+1}a_{d+1}^2 = \sigma^2.$$

Hence for large n, the main contribution to $\mathcal{I}_k$ equals to

$$\frac{r_k^n}{2\pi i \bar{s}_k} e^{i(n\phi_k - \sqrt{n}\sigma z \bar{s}_k)} \int \left(1 - \frac{\sigma^2\delta^2}{2}\right)^n e^{-i\sigma z \delta\sqrt{n}} d\delta$$
$$\approx \frac{r_k^n}{2\pi i \bar{s}_k} e^{i(n\phi_k - \sqrt{n}\sigma z \bar{s}_k)} \int e^{-\sigma^2\delta^2 n/2 - i\sigma\delta\sqrt{n} z} d\delta.$$

Making the change of variables $\sigma\delta\sqrt{n/2} = t$ we evaluate the last integral as $\dfrac{2\sqrt{\pi}e^{-z^2/2}}{\sigma\sqrt{n}}$.

□

Corollary 2.5.2. *If I is a finite interval of order 1. Then*

$$\int_I |\phi^n(s)| \mathbb{1}_{|s| \le T_1/\sigma\sqrt{n}}\, ds = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

Proof. We can cover I by a finite number of intervals I_k. The intervals where $r_k^n \le \dfrac{1}{n^{100}}$ contribute $\mathcal{O}\left(\dfrac{|I|}{n^{100}}\right)$ while the contribution of the intervals where $r_k^n \ge \dfrac{1}{n^{100}}$ is $\mathcal{O}\left(\dfrac{1}{\sqrt{n}}\right)$ due to Lemma 2.5.1. □

Because $r_k \approx 1$, $r_k = |\psi(\bar{s}_k)| = \left|p_1 + \sum_{j=2}^{d+1} p_j e^{ib_j \bar{s}_k}\right| \approx \sum p_j$. Therefore, $a_j \bar{s}_k \approx a_1 \bar{s}_k \mod 2\pi$ for all $j \ge 2$. Thus, $\dfrac{2\pi k b_j}{b_{d+1}} \approx 0 \pmod{2\pi}$ for all $2 \le j \le d$ and hence, $\phi(s_k) \approx 1$ which means s_k and $\bar{s}_k$ are close. Define, $\xi_k = \bar{s}_k - s_k$, $\eta_{j,k} = \dfrac{2\pi k b_j}{b_{d+1}} + 2\pi l_{j,k}$, $j = 1, \ldots, d$ where $l_{j,k}$ is the unique integer such that $\dfrac{2\pi k b_j}{b_{d+1}} + 2\pi l_{j,k} \approx 0$. Then,

$$r_k^2 = \sum_{j=1}^{d+1} p_j^2 + 2 \sum_{l>j, j\neq 1} p_l p_j \cos[(b_l - b_j)\xi_k + \eta_{l,k} - \eta_{j,k}] + 2p_{d+1}p_1 \cos b_{d+1}\xi_k$$
$$+ 2\sum_{j=2}^{d} p_j p_1 \cos(b_j \xi_k + \eta_{j,k}). \quad (2.35)$$

Therefore

$$r_k^2 = 1 - \sum_{l>j, j\neq 1} p_l p_j [(b_l - b_j)\xi_k + \eta_{l,k} - \eta_{j,k}]^2 - p_{d+1}p_1 b_{d+1}^2 \xi_k^2$$

$$- \sum_{j=2}^{d} p_j p_1 (b_j \xi_k + \eta_{j,k})^2 + \mathcal{O}\left(\xi_k^3 + \sum_{l=1}^{d} \eta_{l,k}^3 \right).$$

Taking $\eta_{1,k} = b_1 = 0$ we can write the above as,

$$r_k^2 = -\xi_k^2 \sum_{l>j} p_l p_j (b_l - b_j)^2 - 2\xi_k \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})$$

$$+ 1 - \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})^2 + \mathcal{O}\left(\xi_k^3 + \sum_{l=1}^{d} \eta_{l,k}^3 \right).$$

Since we have r_k^2 approximated by a quadratic polynomial of ξ_k (the unknown) we can approximate ξ_k by determining the maximizer of $r_k^2(\xi_k)$, obtaining

$$\xi_k = -\frac{\sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})}{\sum_{l>j} p_l p_j (b_l - b_j)^2} + \mathcal{O}\left(\|\boldsymbol{\eta}_k|^2 \right). \tag{2.36}$$

Substituting back we find r_k in terms of $\eta_{j,k}$ only. Ignoring higher order terms we compute the maximum to be:

$$r_k^2 = 1 - \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k})^2$$

$$+ \frac{\left[\sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j)(\eta_{l,k} - \eta_{j,k}) \right]^2}{\sum_{l>j} p_l p_j (b_l - b_j)^2} + \mathcal{O}\left(\sum_{l=1}^{d} \eta_{l,k}^3 \right)$$

Put $R = \left[\sum_{l>j} p_l p_j (b_l - b_j)^2 \right]^{-1}$. Then,

$$r_k^2 = 1 + \sum_{\substack{l>j \\ (l,j)\neq(d,1)}} p_l p_j (b_l - b_j) \left[p_l p_j (b_l - b_j) R - 1 \right] (\eta_{l,k} - \eta_{j,k})^2$$

$$+ \sum_{\substack{l>j, m>n \\ l\neq m, j\neq n \\ (l,j),(m,n)\neq(d,1)}} p_l p_j p_m p_n (b_l - b_j)(b_m - b_n) R (\eta_{l,k} - \eta_{j,k})(\eta_{m,k} - \eta_{n,k}) + \mathcal{O}\left(\sum_{l>j} \eta_{l,j}^3 \right)$$

$$:= 1 - 2\sum_{l,j=2}^{d} D_{l,j}(\mathbf{a},\mathbf{p})\eta_{l,k}\eta_{j,k} + \mathcal{O}\left(\sum_{l>j}\eta_{l,j}^3\right). \tag{2.37}$$

Thus,

$$r_k = 1 - \sum_{l,j=2}^{d} D_{l,j}(\mathbf{a},\mathbf{p})\eta_{l,k}\eta_{j,k} + \mathcal{O}\left(\sum_{l>j}\eta_{l,j}^3\right) = 1 - \boldsymbol{\eta}_k^T D_{\mathbf{a},\mathbf{p}}\boldsymbol{\eta}_k + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3)$$

where $D_{\mathbf{a},\mathbf{p}}$ is a $(d-1)\times(d-1)$ matrix with

$$[D_{\mathbf{a},\mathbf{p}}]_{i,j} = D_{i,j}(\mathbf{a},\mathbf{p}) \tag{2.38}$$

and $\boldsymbol{\eta}_k^T = (\eta_{2,k},\dots,\eta_{d,k})$. From this we have,

$$\mathcal{I}_k = \frac{e^{-z^2/2}}{i\sqrt{\pi n}\sigma}\frac{(1-\boldsymbol{\eta}_k^T D_{\mathbf{a},\mathbf{p}}\boldsymbol{\eta}_k + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3))^n}{s_k}\; e^{in\phi_k - i\bar{s}_k z\sigma\sqrt{n}}\;(1+o(1)).$$

Let $\mathcal{B}(\mathbf{a},\mathbf{p})$ be the contribution of the boundary terms $\pm\dfrac{T_1}{\sigma\sqrt{n}} \in I_k$.

Lemma 2.5.3.

$$\mathbf{E}(|\mathcal{B}|) \le \frac{C}{n^{(2d-1)/2}}.$$

Lemma 2.5.4. *Let*

$$\mathcal{I}_{k,l} = \mathcal{I}_k \mathbb{1}_{|k|^\alpha n^{1/4}\|\boldsymbol{\eta}_k\|\in[2^l,2^{l+1}]}.$$

with $\alpha = [2(d-1)]^{-1}$. *Then there is a constant* $\tilde{c}$ *such that*

$$\mathbf{E}\left(\sum_{0<|k|<Kn^{(d-1)/2}}\;\sum_{l>K}\;|\mathcal{I}_{k,l}|\right) = \mathcal{O}\left(\frac{1}{n^{d/2}}2^K\exp(-\tilde{c}2^{2K})\right).$$

Lemmas 2.5.3 and 2.5.4 will be proven in Section 2.6.

Next we prove a lemma that would allow us to further simplify $\hat{\boldsymbol{\Delta}}_n$.

Lemma 2.5.5. *(a)* $\bar{s}_k = s_k + \boldsymbol{\omega}^T\boldsymbol{\eta}_k + \mathcal{O}(\|\eta\|_k^2)$ *where* $\boldsymbol{\omega} = \boldsymbol{\omega}(\mathbf{a},\mathbf{p})$ *is a* $1\times(d-1)$ *vector.*

(b) If $\|\boldsymbol{\eta}\| = \mathcal{O}\left(\dfrac{\ln n}{\sqrt{n}}\right)$ *then* $n\phi_k = ns_ka_1 + np_2\eta_{2,k} + \cdots + np_d\eta_{d,k} + o(1)$.

Proof. Since $\overline{s}_k - s_k = \zeta_k$ part (a) follows by (2.36).

Next, by (2.34)

$$\phi_k = \arg\, \phi(s_k) + \mathcal{O}\left(\delta^3 + \frac{\ln^{3/2} n}{n^{3/2}}\right)$$

Note that,

$$\phi(s_k) = e^{is_k a_1}(p_1 + p_2 e^{i\eta_{2,k}} + \cdots + p_d e^{i\eta_{d,k}} + p_{d+1}).$$

Thus,

$$\begin{aligned}\arg(\phi(s_k)) &= s_k a_1 + \tan^{-1}\left(\frac{p_2 \sin\eta_{2,k} + \cdots + p_d \sin\eta_{d,k}}{p_1 + p_2\cos\eta_{2,k} + \cdots + p_d\cos\eta_{d,k} + p_{d+1}}\right)\\ &= s_k a_1 + \sum_{l=2}^{d} p_l \eta_{l,k} + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3)\end{aligned}$$

since the denominator in the first line is $1+\mathcal{O}(||\boldsymbol{\eta}||^2)$. Now part (b) follows easily. $\square$

Now, we continue the analysis of the leading term in $\hat{\boldsymbol{\Delta}}_n$. Pick a small δ and define

$$A_1 = \{(\mathbf{a}, \mathbf{p})|\ \mathcal{I}_{k,l} = 0\ \forall k, l \text{ s.t. } |k| < \delta n^{(d-1)/2} \text{ and } l < K\}.$$

Then

$$A_1^c = \{(\mathbf{a}, \mathbf{p})|\ \exists |k| < \delta n^{(d-1)/2},\ |k|^{\alpha} n^{1/4}\|\eta_k\| \le 2^K\}.$$

Thus,

$$\mathbb{P}(A_1^c) = \sum_{|k|<\delta n^{(d-1)/2}} \frac{C2^K}{|k|^{(d-1)\alpha} n^{(d-1)/4}} = \mathcal{O}(\sqrt{\delta}2^K)$$

if $\alpha = \dfrac{1}{2(d-1)}$. Hence, for a very large K and δ such that $\sqrt{\delta}2^K$ is very small, we can approximate Δ_n by the sum of $\mathcal{I}_k$'s with $\delta \le \dfrac{|k|}{n^{(d-1)/2}} \le K$ and $|k|^{\alpha} n^{1/4}\|\boldsymbol{\eta}_k\| \le 2^K$.

We define the random vector $X_k = \sqrt{n}\boldsymbol{\eta}_k$ and $Y_k = \dfrac{k}{n^{(d-1)/2}}$. Then, combining terms corresponding to k and $-k$, we obtain the following approximation to the distribution of Δ_n for large n

$$\frac{|b_{d+1}|e^{-z^2/2}}{n^{d/2}\sigma\sqrt{\pi^3}} \sum_{k\in S(n,\delta,K)} \frac{\sin(n\phi_k - \overline{s}_k z\sigma\sqrt{n})}{Y_k} e^{-X_k^T D_{\mathbf{a},\mathbf{p}} X_k}$$

where $S(n,\delta,K) = \{k > 0 | \delta < Y_k < K, \quad |Y_k|^\alpha \|X_k\| < 2^K\}$.

Define $\mathbf{q} = (p_2, \dots, p_d)$. Then, Lemma 2.5.5 shows that

$$\begin{aligned} n\phi_k - \overline{s}_k z\sigma\sqrt{n} &= s_k(na_1 - z\sigma\sqrt{n}) + n\mathbf{q}^T\boldsymbol{\eta}_k - z\sigma\sqrt{n}\boldsymbol{\omega}^T\boldsymbol{\eta}_k + o(1) \\ &= \frac{2\pi n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z\sigma)Y_k + (\sqrt{n}\mathbf{q} - z\sigma\boldsymbol{\omega})^T X_k + o(1). \end{aligned}$$

Therefore, for large n and K and δ such that $\sqrt{\delta}2^K$ is very small, the distribution of Δ_n is well approximated by

$$\tilde{\Delta}_n(\delta,K) = \frac{|b_{d+1}|e^{-z^2/2}}{n^{d/2}\sigma\sqrt{\pi^3}} \sum_{k\in S(n,\delta,K)} \frac{\sin\left(\frac{2\pi n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z\sigma)Y_k + (\sqrt{n}\mathbf{q} - z\sigma\boldsymbol{\omega})^T X_k\right)}{Y_k} e^{-X_k^T D_{\mathbf{a},\mathbf{p}} X_k}.$$

2.6 Expectation of characteristic function.

Proof of Lemma 2.4.2. Recall that $d(s) = \max\limits_{2\le j\le d+1} d(b_j s, 0)$ where the distance is computed on the torus $\mathbb{R}/(2\pi\mathbb{Z})$. Formula (2.35) shows that there are positive constants C, c such that

$$\frac{1}{C} \le \frac{|\phi^n(s)|}{e^{-cnd(s)^2}} < C. \tag{2.39}$$

To prove the lemma we decompose $\mathbf{E}\big(e^{-cnd(s)^2}\big)$ into the pieces where $d(s)\sqrt{n}$ is of order 2^l for some $l \leq (\log_2 n)/2$. and use the fact that ∂ has a bounded density.

$$\begin{aligned}\mathbf{E}\left(\phi^n(s)\right) &\leq C\mathbf{P}\left(d(s) < \frac{1}{\sqrt{n}}\right) + C\sum_{l=0}^{(\log_2 n)/2} \mathbf{P}\left(d(s)\sqrt{n} \in [2^l, 2^{l+1}]\right) e^{-c4^l} \\ &\leq \frac{C}{n^{d/2}} + C\sum_{l=0}^{(\log_2 n)/2} \frac{4^l}{n^{d/2}} e^{-c4^l} \leq \frac{C}{n^{d/2}}\end{aligned}$$

completing the proof. $\square$

Proof of Lemma 2.5.3. Let k be such that $\dfrac{T_1}{\sigma\sqrt{n}} \in I_k$. Then

$$\mathcal{I}_k = \int_{\pi(2k-1)/|b_{d+1}|}^{T_1/\sigma\sqrt{n}} e^{-isz\sigma\sqrt{n}} \frac{\phi^n(s)}{s}\, ds.$$

Because $T_1 = K_1 n^{d/2}$ and $s \in \left[\dfrac{\pi(2k-1)}{|b_{d+1}|}, \dfrac{T_1}{\sigma\sqrt{n}}\right]$ we have $s \approx n^{(d-1)/2}$. Thus

$$\mathbf{E}(|\mathcal{I}_k|) \leq \frac{C}{n^{(d-1)/2}} \mathbf{E}\left(\int_{\pi(2k-1)/|b_{d+1}|}^{T_1/\sigma\sqrt{n}} |\phi^n(s)|\, ds\right).$$

We claim that for all fixed b_d,

$$\iint e^{-cnd(s)^2}\, ds\, db_2 \ldots\, db_{d-1} \leq \frac{C}{n^{d/2}}. \tag{2.40}$$

If this is true then using that ρ is a smooth compactly supported density of b_d we have that,

$$\begin{aligned}\mathbf{E}\bigg(\int_{\pi(2k-1)/|b_{d+1}|}^{T_1/\sigma\sqrt{n}} |\phi^n(s)|\, ds\bigg) &= \iint \int_{\pi(2k-1)/|b_{d+1}|}^{T_1/\sigma\sqrt{n}} |\phi^n(s)|\, ds\, db_d\, db_{d-1} \ldots\, db_2 \\ &\leq C \iint \int_{\pi(2k-1)/|x|}^{T_1/\sigma\sqrt{n}} e^{-cnd(s)^2} \rho(x)\, ds\, dx\, db_{d-1} \ldots\, db_2 \\ &\leq C \int \iint e^{-cnd(s)^2}\, ds\, db_{d-1} \ldots\, db_2\, \rho(x)\, dx \\ &\leq \frac{C}{n^{d/2}} \int \rho(x)\, dx = \mathcal{O}\left(\frac{1}{n^{d/2}}\right).\end{aligned}$$

Thus

$$\mathbf{E}(|\mathcal{I}_k|) \leq \frac{C}{n^{(2d-1)/2}}. \tag{2.41}$$

Similarly, if $-\dfrac{T}{\sigma\sqrt{n}} \in \mathcal{I}_k$, then(2.41) holds. Hence, $\mathbf{E}(|\mathcal{B}|) \leq \dfrac{C}{n^{(2d-1)/2}}$ as required.

To prove (2.40) we decompose it into pieces where $d(s)\sqrt{n}$ is of order 2^l. Taking μ to be the product measure $ds\, db_{d-1} \ldots db_2$ from (2.39) we have

$$\begin{aligned}\iint e^{-cnd(s)^2}\, ds\, db_{d-1} \ldots db_2 &\leq C\mu\{(s, b_2, \ldots, b_{d-1}) | d(s) < 1/\sqrt{n}\} \\ &\quad + C \sum_{l=0}^{(\log_2 n)/2} \mu\{(s, b_2, \ldots, b_{d-1}) | d(s)\sqrt{n} \in [2^l, 2^{l+1}]\} e^{-c4^l} \\ &\leq \frac{C}{n^{d/2}} + C \sum_{l=0}^{(\log_2 n)/2} \frac{4^l}{n^{d/2}} e^{-c4^l} \leq \frac{C}{n^{d/2}}\end{aligned}$$

as required. □

Proof of Lemma 2.5.4. Because

$$r_k = 1 - \boldsymbol{\eta}_k^T D_{\mathbf{a},\mathbf{p}} \boldsymbol{\eta}_k + \mathcal{O}(\|\boldsymbol{\eta}_k\|^3) \text{ and } |k|^\alpha n^{1/4} \|\boldsymbol{\eta}_k\| \in [2^l, 2^{l+1}]$$

we can write

$$r_k = 1 - c\frac{4^l}{|k|^{2\alpha}\sqrt{n}} + \mathcal{O}(n^{-3/4}).$$

Accordingly

$$r_k^n \leq Ce^{-\frac{c2^{2l}\sqrt{n}}{|k|^{2\alpha}}}.$$

Also

$$\mathbf{P}(|k|^\alpha n^{1/4} \|\boldsymbol{\eta}\| \in [2^l, 2^{l+1}]) \leq \frac{C2^l}{\sqrt{|k|} n^{(d-1)/4}}.$$

Hence,

$$\mathbf{E}(\mathcal{I}_{k,l}) \leq \frac{Ce^{-\frac{c2^{2l}\sqrt{n}}{|k|^{2\alpha}}}}{\sqrt{n}|k|} \frac{2^l}{\sqrt{|k|} n^{(d-1)/4}} = \frac{C2^l e^{-\frac{c2^{2l}\sqrt{n}}{|k|^{2\alpha}}}}{|k|^{3/2} n^{(d+1)/4}}.$$

Thus

$$\sum_{l>K} \mathbf{E}(\mathcal{I}_{k,l}) \leq \frac{C2^K e^{-\frac{c2^{2K}\sqrt{n}}{|k|^{2\alpha}}}}{|k|^{3/2} n^{(d+1)/4}}.$$

Therefore we need to estimate

$$\sum_{0<|k|<Kn^{(d-1)/2}} \frac{C2^K e^{-\frac{c2^{2K}\sqrt{n}}{|k|^{2\alpha}}}}{|k|^{3/2} n^{(d+1)/4}} =$$

$$\frac{C}{n^{d/2}} \sum_{0<|k|<Kn^{(d-1)/2}} \frac{1}{|k|} \sqrt{\frac{2^{2K} n^{(d-1)/2}}{|k|}} e^{-\frac{c2^{2K}\sqrt{n}}{|k|^{2\alpha}}}. \tag{2.42}$$

Split the sum over

$$|k| \in \left[\frac{Kn^{(d-1)/2}}{2^{s+1}}, \frac{Kn^{(d-1)/2}}{2^s} \right] \tag{2.43}$$

for $s \in \mathbb{N}$. Then, for a fixed s we have

$$|k|^{2\alpha} = \mathcal{O}\left(\frac{K^{\frac{1}{d-1}} \sqrt{n}}{2^{\frac{s}{d-1}}} \right),$$

so each term in the sum (2.42) is of order

$$\frac{2^{K+(3s/2)}}{K^{3/2} n^{(d-1)/2}} \ \exp\left(-\frac{c2^{2K+\frac{s}{d-1}}}{K^{\frac{1}{d-1}}} \right).$$

But the number of such terms is of order $\dfrac{n^{(d-1)/2}}{2^s}$. Hence, the sum over k in (2.43) is

$$\mathcal{O}\left(\frac{2^{K+s/2}}{K^{3/2}} \ \exp\left(-\frac{c2^{2K+\frac{s}{d-1}}}{K^{\frac{1}{d-1}}} \right) \right).$$

Summing over s we obtain the result. $\square$

2.7 Relation to homogeneous flows.

Given $\mathbf{u} \in \mathbb{R}^{d-1}, v \in \mathbb{R}$ consider the following function on space $\mathcal{M}$ of unimodular lattices in $\mathbb{R}^d$:

$$\mathcal{Z}(L) = \sum_{(y,\mathbf{x}) \in L \setminus \{\mathbf{0}\}} \frac{\sin 2\pi(\mathbf{u}^T \mathbf{x} + vy)}{y} \; e^{-4\pi^2 \mathbf{x}^T D_{\mathbf{a},\mathbf{p}} \mathbf{x}} \; \mathbb{1}_{\{\delta < y < K, y^\alpha \|\mathbf{x}\| < 2^K\}}. \quad (2.44)$$

Define $\boldsymbol{\gamma} = \frac{1}{k}\boldsymbol{\eta}$ and introduce the matrices, $H_{\boldsymbol{\gamma}} = \begin{pmatrix} 1 & \gamma \\ \mathbf{0}^T & I_{d-1} \end{pmatrix}$, $G_t = \begin{pmatrix} e^{-(d-1)t} & \mathbf{0} \\ \mathbf{0}^T & e^t I_{d-1} \end{pmatrix}$.

Then, we have

$$n^{d/2} \tilde{\Delta}_n = -\frac{|b_{d+1}| e^{-z^2/2}}{\sigma \sqrt{\pi^3}} \mathcal{Z}(\mathbb{Z}^d \; H_{\boldsymbol{\gamma}} \; G_{\frac{\ln n}{2}}),$$

where

$$\mathbf{u} = \sqrt{n}\mathbf{q} - z\sigma\boldsymbol{\omega} \text{ and } v = \frac{n^{d/2}}{|b_{d+1}|}(\sqrt{n} a_1 - z\sigma)$$

and $\mathbf{q}$ and $\boldsymbol{\omega}$ are defined at the end of Section 2.5. Let $\mathcal{L}(n, \mathbf{a})$ be the unimodular lattice $\mathbb{Z}^d \; H_{\boldsymbol{\gamma}} \; G_{\frac{\ln(n)}{2}}$. Let

$$\mathbf{w}_j(n, \mathbf{a}) = (y_j(n, \mathbf{a}), \mathbf{x}_j(n, \mathbf{a})), \; j = 1, \ldots, d$$

with $y_j \in \mathbb{R}$ and $\mathbf{x}_j \in \mathbb{R}^{d-1}$ be the shortest spanning set of $\mathcal{L}$. Put,

$$\theta_j(n, (\mathbf{a}, \mathbf{p})) = \mathbf{u}^T \mathbf{x}_j(n, \mathbf{a}) + v y_j(n, \mathbf{a}), \; j = 1, \ldots, d.$$

Proposition 2.7.1. *If* $(\mathbf{a}, \mathbf{p})$ *is distributed according to* $\mathbf{P}$ *then the distribution of the random vector*

$$((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}(n, (\mathbf{a}, \mathbf{p})))$$

converges to $\mathbf{P} \times \mu$ *as* $n \to \infty$, *where* μ *is the Haar measure on* $[SL_d(\mathbb{R})/SL_d(\mathbb{Z})] \times \mathbb{T}^d$.

If we restrict our attention only to $((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))$ then the result is standard (see [39, Theorem 5.8], as well as [18, 35, 45]). The proof in the general case follows the approach of the proof of Proposition 5.1 in [14].

Proof. We need to show that for each bounded smooth test function f,

$$\int_\Omega f((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}),\boldsymbol{\theta})\,d\mathbf{P} \to \int_{\Omega\times\mathcal{M}\times\mathbb{T}^d} f((\mathbf{a},\mathbf{p}),\mathcal{L},\boldsymbol{\theta})\,d\mathbf{P}\,d\mathcal{L}\,d\boldsymbol{\theta} \tag{2.45}$$

as $n\to\infty$. Write the Fourier series expansion of f:

$$f((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}),\boldsymbol{\theta}) = \sum_{\mathbf{k}=(k_1,\dots,k_d)\in\mathbb{Z}^d} f_{\mathbf{k}}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\; e^{2\pi i\mathbf{k}^T\boldsymbol{\theta}}. \tag{2.46}$$

Then, it is enough to prove (2.45) for individual terms in (2.46).

If $\mathbf{k}=\mathbf{0}$ then by [39, Theorem 5.8] we can conclude that

$$\int_\Omega f_{\mathbf{0}}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\,d\mathbf{P} \to \int_{\Omega\times\mathcal{M}\times\mathbb{T}^d} f_{\mathbf{0}}((\mathbf{a},\mathbf{p}),\mathcal{L})\,d\mathbf{P}\,d\mathcal{L}\,d\boldsymbol{\theta}.$$

Now assume that $\mathbf{k}\neq\mathbf{0}$. Since Ω is $2d$ dimensional, we can use $(p_1,\dots,p_d,a_1,b_2,\dots,b_d)$ as local coordinates. In these coordinates $\mathcal{L}$ is independent of a_1. Hence, y_j's and $\mathbf{x}_j$'s are independent of a_1. Put $\boldsymbol{\nu}=(p_1,\dots,p_d,b_2,\dots,b_d)$. Then there exists a compactly supported density ρ such that,

$$\begin{aligned} J_{n,\mathbf{k}} &= \int f_{\mathbf{k}}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\; e^{2\pi i\mathbf{k}^T\boldsymbol{\theta}}\,d\mathbf{P} \qquad (2.47)\\ &= \int f_{\mathbf{k}}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\exp 2\pi i\left(\sqrt{n}\sum k_j\mathbf{q}^T\mathbf{x}_j\right)\\ &\qquad\times\left[\int \rho(a_1,\boldsymbol{\nu})\exp 2\pi i\left(\frac{n^{d/2}}{|b_{d+1}|}\left(\sqrt{n}a_1 - z\sigma\right)\sum y_jk_j - z\sigma\sum k_j\boldsymbol{\omega}^T\mathbf{x}_j\right)da_1\right]d\boldsymbol{\nu}. \end{aligned}$$

Note that,

$$\int_{\mathbb{T}^d\times\Omega\times\mathcal{M}} f_{\mathbf{k}}((\mathbf{a},\mathbf{p}),\mathcal{L}(n,\mathbf{a}))\; e^{2\pi i\mathbf{k}^T\boldsymbol{\theta}}d\theta_1\dots d\theta_d\,d\mathbf{P}\,d\mathcal{L} = 0$$

because

$$\int_{\mathbb{T}^d} e^{2\pi i \mathbf{k}^T \boldsymbol{\theta}} d\theta_1 \dots d\theta_d = 0.$$

Therefore, it is enough to prove that $J_{n,\mathbf{k}}$ converges to 0 as $n \to \infty$. To prove this we use integration by parts as follows. Put,

$$g(a_1, \boldsymbol{\nu}) = \exp i \left(\frac{2\pi n^{(d+1)/2} \sum y_j k_j}{|b_{d+1}|} a_1 \right) = \exp i \left(n^{(d+1)/2} \phi(\boldsymbol{\nu}) a_1 \right)$$

where $\phi(\boldsymbol{\nu}) = \dfrac{2\pi \sum y_j k_j}{|b_{d+1}|}$ and,

$$h(a_1, \boldsymbol{\nu}) = \rho(a_1, \boldsymbol{\nu}) \exp \left[-i \left(\frac{2\pi n^{d/2} \sum y_j k_j}{|b_{d+1}|} + 4\pi \sum k_j \boldsymbol{\omega}^T \mathbf{x}_j \right) z\sigma(a_1, \boldsymbol{\nu}) \right]$$

Then, the inner integral in (2.47) is $\displaystyle\int g(a_1, \boldsymbol{\nu}) h(a_1, \boldsymbol{\nu}) \, da_1$. Let $\varepsilon > 0$. On the set $Q_{\mathbf{k}} = \{\phi(\boldsymbol{\nu}) > \varepsilon\}$ we can write

$$g(a_1, \boldsymbol{\nu}) \, da_1 = \frac{1}{i\phi(\boldsymbol{\nu}) n^{(d+1)/2}} \, d \exp \left(i a_1 n^{(d+1)/2} \phi(\boldsymbol{\nu}) \right).$$

Integrating by parts on $Q_{\mathbf{k}}$ (note that h has compact support) and using trivial bounds on $Q_{\mathbf{k}}^c$, we can conclude that

$$\begin{aligned} |J_{n,\mathbf{k}}| &\le \left| \int \frac{\exp \left(i a_1 n^{(d+1)/2} \phi(\boldsymbol{\nu}) \right)}{i\phi(\boldsymbol{\nu}) n^{(d+1)/2}} \, h'(a_1, \boldsymbol{\nu}) \, da_1 \right| + C\mathbf{P}(\{\phi(\boldsymbol{\nu}) \le \varepsilon\}) \\ &\le \frac{1}{\varepsilon n^{(d+1)/2}} \int |h'(a_1, \boldsymbol{\nu})| \, da_1 + C\mathbf{P}(\{\phi(\boldsymbol{\nu}) \le \varepsilon\}) \end{aligned}$$

for small enough ε. But $h'(a_1, \boldsymbol{\nu}) = \mathcal{O}(n^{d/2})$, hence the first term is $\mathcal{O}(1/\sqrt{n})$. Therefore, first taking $n \to \infty$ and then taking $\varepsilon \to 0$ we have the required result.

□

Proposition 2.7.1 implies that as $n \to \infty$ the distribution of $n^{d/2} \tilde{\boldsymbol{\Delta}}_n(\delta, K)$ converges to the distribution of

$$e^{-z^2/2} \frac{|\mathfrak{a}_{d+1} - \mathfrak{a}_1|}{2\sigma(\mathfrak{a}, \mathfrak{p}) \sqrt{\pi^3}} \sum_{\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \frac{\sin 2\pi\theta(\mathbf{m})}{y(\mathbf{m})} e^{-4\pi^2 \mathbf{x}^T D_{\mathfrak{a},\mathfrak{p}} \mathbf{x}} \mathbb{1}_{\{\delta < |y(\mathbf{m})| < K, \, |y(\mathbf{m})|^\alpha \|\mathbf{x}(\mathbf{m})\| < 2^K\}}. \quad (2.48)$$

Next we let $\delta \to 0$ and $K \to \infty$ in such a way that $\sqrt{\delta}2^K \to 0$. Then,

$$\mathbb{1}_{\{\delta<|y(\mathbf{m})|<K,\ |y(\mathbf{m})|^\alpha|\mathbf{x}(\mathbf{m})|<2^K\}} \to 1.$$

Thus, (2.48) converges to $\hat{\mathcal{X}}$ proving Theorem 2.1.2*.

2.8 Finite intervals.

The proofs of Theorems 2.1.3 and 2.1.4 are similar to the proofs of Theorems 2.1.1 and 2.1.2 so we just explain the necessary changes leaving the details to the readers.

Proof of Theorem 2.1.3. The random vector (2.10) can be approximated by $(\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)})$ where $\mathcal{Z}^{(i)}$ are defined as in (2.44) with $\mathbf{u}$ and v replaced by

$$\mathbf{u}^{(i)} = \sqrt{n}\mathbf{q} - z_i\sigma\boldsymbol{\omega} \quad \text{and} \quad v^{(i)} = \frac{n^{d/2}}{|b_{d+1}|}(\sqrt{n}a_1 - z_i\sigma)$$

respectively. Define $\boldsymbol{\theta}^{(i)}$ as in Proposition 2.7.1 but $\mathbf{u}$ and v replaced by $\mathbf{u}^{(i)}$ and $v^{(i)}$. To complete the proof we prove an analogue of Proposition 2.7.1. Namely that $((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}^{(1)}(n, (\mathbf{a}, \mathbf{p})), \boldsymbol{\theta}^{(2)}(n, (\mathbf{a}, \mathbf{p})))$ converges to $\mathbf{P} \times \mu'$ as $n \to \infty$ where μ' is the Haar measure on $[SL_d(\mathbb{R})/SL_d(\mathbb{Z})] \times \mathbb{T}^d \times \mathbb{T}^d$.

As in the proof of Proposition 2.7.1 we prove that for individual terms in the Fourier series of a smooth function f on $[SL_d(\mathbb{R})/SL_d(\mathbb{Z})] \times \mathbb{T}^d \times \mathbb{T}^d$

$$\sum_{(\mathbf{k}_1,\mathbf{k}_2)\in\mathbb{Z}^d\times\mathbb{Z}^d} f_{\mathbf{k}_1,\mathbf{k}_2}((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}))\ e^{2\pi i[\mathbf{k}_1^T\boldsymbol{\theta}^{(1)}+\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)}-\boldsymbol{\theta}^{(2)})]}$$

we have

$$J_{n,\mathbf{k}_1,\mathbf{k}_2} := \int_\Omega f_{\mathbf{k}_1,\mathbf{k}_2}((\mathbf{a}, \mathbf{p}), \mathcal{L}(n, \mathbf{a}))e^{2\pi i[\mathbf{k}_1^T\boldsymbol{\theta}^{(1)}+\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)}-\boldsymbol{\theta}^{(2)})]}\, d\mathbf{P}$$

$$\xrightarrow{n\to\infty} \int_{\Omega\times\mathcal{M}\times\mathbb{T}^d\times\mathbb{T}^d} f_{\mathbf{k}_1,\mathbf{k}_2}((\mathbf{a},\mathbf{p}),\mathcal{L}) e^{2\pi i[\mathbf{k}_1^T\boldsymbol{\theta}_1+\mathbf{k}_2^T(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2)]}\, d\mathbf{P}\, d\mathcal{L}\, d\boldsymbol{\theta}_1 d\boldsymbol{\theta}_2.$$

The case $\mathbf{k}_1 = \mathbf{k}_2 = 0$ follows from [39, Theorem 5.8]. Note that

$$\mathbf{k}_2^T(\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)})] = (z_2(n) - z_1(n)) \left(\frac{2\pi n^{d/2}}{|b_{d+1}|} \sum y_j k_{2,j} + \sum k_{2,j} \boldsymbol{\omega}^T \mathbf{x}_j \right) \sigma.$$

If $\mathbf{k}_1 = 0$ choose appropriate local-coordinates in which σ is a coordinate. Integrating by parts with respect to $\sigma = \sigma(\mathbf{a},\mathbf{p})$ and using $|z_1(n) - z_2(n)|n^{d/2} \to \infty$ we see that $J_{n,\mathbf{0},\mathbf{k}_2} \to 0$ as $n \to \infty$.

If $\mathbf{k}_1 \neq 0$ then using the same local coordinates $(a_1, \boldsymbol{\nu})$ as in the proof of Proposition 2.7.1 we can integrate by parts to conclude that $J_{n,\mathbf{k}_1,\mathbf{k}_2} \to 0$ as $n \to \infty$. The proof follows through because the leading term of $\mathbf{k}_1^T\boldsymbol{\theta}^{(1)} + \mathbf{k}_2^T(\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)})$ is still $n^{(d+1)/2}\phi(\boldsymbol{\nu})a_1$. □

Proof of Theorem 2.1.4. To prove part (a) pick $\bar{\varepsilon} < \varepsilon$. Applying Theorem 2.1.1 we obtain that for almost every $(\mathbf{a},\mathbf{p})$

$$\mathbb{P}_{(\mathbf{a},\mathbf{p})}\left(z_1 \le \frac{S_n}{\sigma\sqrt{n}} \le z_2 \right) = \mathcal{E}_{d-1}(z_2) - \mathcal{E}_{d-1}(z_1) + \mathcal{O}\left(n^{-(d-\bar{\varepsilon})/2}\right)$$

$$= \mathfrak{n}(z_1)l_n + \mathcal{O}(l_n^2) + \mathcal{O}(l_n/\sqrt{n}) + \mathcal{O}\left(n^{-(d-\bar{\varepsilon})/2}\right).$$

According to the assumptions of part (a) the first term is much larger than the remaining terms proving the result.

The proof of part (b) is similar except that we apply Theorem 2.1.3 instead of Theorem 2.1.1 so we only get convergence in probability.

To prove part (c) we first prove the following analogue of Theorem 2.1.3 in case $z_2 = z_1 + \dfrac{c|a_{d+1} - a_1|}{n^{d/2}\sigma}$

$$\frac{n^{d/2}}{\Lambda(\mathbf{a},\mathbf{p})} \left(e^{z_1^2/2} \left[\mathcal{E}_d(z_1) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z_1\right) \right], e^{z_2^2/2} \left[\mathcal{E}_d(z_2) - \mathbb{P}_{\mathbf{a},\mathbf{p}}\left(\frac{S_n}{\sigma\sqrt{n}} \le z_2\right) \right] \right)$$

converges in law to a random vector $(\tilde{\mathcal{X}}_1, \tilde{\mathcal{X}}_2)(\mathcal{L}, \theta, c)$ where

$$(\tilde{\mathcal{X}}_1, \tilde{\mathcal{X}}_2)(\mathcal{L}, \theta, c) = \sum_{\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \frac{e^{-4\pi^2 ||\mathbf{x}(\mathbf{m})||^2}}{y(\mathbf{m})} \Big(\sin \theta(\mathbf{m}), \sin(\theta(\mathbf{m}) - cy(\mathbf{m})) \Big).$$

Once this convergence is established the proof of part (c) is the same as the proof of part (b). The proof of convergence is similar to the proof of Theorem 2.1.3 except that $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$ are now not independent. Namely using the same notation as in the proof of Theorem 2.1.3 we have that $\mathbf{u}^{(2)} = \mathbf{u}^{(1)} + o(1)$, while $v^{(2)} = v^{(1)} - c + o(1)$. Following the same argument as in the proof of Proposition 2.7.1 we obtain that $(\mathcal{L}(n, \mathbf{a}), \boldsymbol{\theta}^{(1)}(n, \mathbf{a}), [\boldsymbol{\theta}^{(2)} - \boldsymbol{\theta}^{(1)}](n, \mathbf{a}))$ converges as $n \to \infty$ to $(\mathcal{L}^*, \boldsymbol{\theta}^*, \hat{\boldsymbol{\theta}}^*)$ where $(\mathcal{L}^*, \boldsymbol{\theta}^*)$ is distributed according to the Haar measure on $SL_d(\mathbb{R})/SL_d(\mathbb{Z}) \times \mathbb{T}^d$ and $\hat{\boldsymbol{\theta}}^*_j = \boldsymbol{\theta}^*_j - cy_j$. This justifies the formula for $(\tilde{\mathcal{X}}_1, \tilde{\mathcal{X}}_2)$. $\square$

Chapter 3: Central Limit Theorem: Weakly Dependent Random Variables.

3.1 Overview and main results.

Let $S_N = \sum_{n=1}^{N} X_n$ be a sum of random variables. We assume that there is a Banach space $\mathbb{B}$ and a family of bounded linear operators $\mathcal{L}_t : \mathbb{B} \to \mathbb{B}$ and vectors $v \in \mathbb{B}, \ell \in \mathbb{B}'$ such that

$$\mathbb{E}\left(e^{itS_N}\right) = \ell(\mathcal{L}_t^N v),\ t \in \mathbb{R}. \tag{3.1}$$

We will make the following assumptions on the family $\mathcal{L}_t$.

(A1) $t \mapsto \mathcal{L}_t$ is continuous and there exists $s \in \mathbb{N}$ and $\delta > 0$ such that $t \mapsto \mathcal{L}_t$ is s times continuously differentiable for $|t| \le \delta$.

(A2) 1 is an isolated and simple eigenvalue of $\mathcal{L}_0$, all other eigenvalues of $\mathcal{L}_0$ have absolute value less than 1 and its essential spectrum is contained strictly inside the disk of radius 1 (spectral gap).

(A3) For all $t \neq 0$, $\text{sp}(\mathcal{L}_t) \subset \{|z| < 1\}$.

(A4) There are positive real numbers K, r_1, r_2 and N_0 such that $\left\|\mathcal{L}_t^N\right\| \le \dfrac{1}{N^{r_2}}$ for all t satisfying $K \le |t| \le N^{r_1}$ and $N > N_0$.

Remark 3.1.1.

1. *In practice we would check* $(A3)$ *by showing that when* $t \neq 0$, *the spectral radius of* $\mathcal{L}_t$ *is at most* 1 *and no eigenvalue of* $\mathcal{L}_t$ *is on the unit circle. Because the spectrum of a linear operator is a closed set this would imply that* $\mathrm{sp}(\mathcal{L}_t)$ *is contained in a closed disk strictly inside the unit disk.*

2. *Suppose* $(A4)$ *holds. Let* $N_1 > N_0$ *be such that* $N_1^{(r_1-\epsilon)/r_1} > N_0$. *Then, for all* $N > N_1$,

$$\begin{aligned}\|\mathcal{L}_t^N\| &\leq \|(\mathcal{L}_t^{\lceil N^{(r_1-\epsilon)/r_1}\rceil})^{N_1^{\epsilon/r_1}}\| \leq \|(\mathcal{L}_t^{\lceil N^{(r_1-\epsilon)/r_1}\rceil})\|^{N_1^{\epsilon/r_1}} \\ &\leq \frac{1}{\lceil N^{(r_1-\epsilon)/r_1}\rceil^{r_2 N_1^{\epsilon/r_1}}} \text{ for } K \leq |t| \leq N^{r_1-\epsilon} \\ &\leq \frac{1}{N^{r_2 K_{N_1}}}\end{aligned}$$

where $K_{N_1} = \dfrac{r_1-\epsilon}{r_1} N^{\epsilon/r_1}$. *Therefore fixing* N_1 *large enough we can make* $r_2 K_{N_1}$ *as large as we want. Hence, given* $(A4)$, *by slightly reducing* r_1, *we may assume* r_2 *is sufficiently large.*

3. *Suppose* $(A1),(A2)$ *and* $(A3)$ *are satisfied with* $s \geq 3$. *Then, [24, Theorem 2.4] implies that there exists* $A \in \mathbb{R}$ *and* $\sigma^2 \geq 0$ *such that*

$$\frac{S_N - NA}{\sqrt{N}} \xrightarrow{d} \mathcal{N}(0, \sigma^2). \tag{3.2}$$

Our interest is in S_N *that satisfies the CLT i.e. the case* $\sigma^2 > 0$. *Since in applications we specify conditions which guarantee this, in the following theorems we always assume that* $\sigma^2 > 0$.

This is essentially an extension of Nagaev-Guivarc'h method. Some of the spectral assumptions in the theorem can be found in the proofs of decay of correlations and the CLT using transfer operators. For example, see [24, 29, 37]. The key novelty here is the condition (A4) which guarantees a sufficient control over the characteristic function for intermediate values of t. This is analogous to the condition (1.3) in Theorem 1.1. In addition, parallels can be drawn between the moment condition in Theorem 1.1 with the condition $s = r + 2$. The proof of the result is based on classical perturbation theory in [33], applicable due to (A1), (A2) and (A3), which provides the actual expansion and control of the error near 0, the Berry-Esseen inequality (see (3.4) below) which reduces that error to a Fourier inversion integral over an interval of size $\mathcal{O}(n^{r/2})$ and the condition (A4).

Now we are in a position to state our first result on the existence of the classical Edgeworth expansion for random variables satisfying (A1) through (A4) which we refer to as weakly dependent random variables.

Theorem 3.1.1. *Let $r \in \mathbb{N}$ with $r \geq 2$. Suppose (A1) through (A4) hold with $s = r + 2$ and $r_1 > \dfrac{r-1}{2}$. Then S_N admits Edgeworth expansion of order r.*

Next, we examine the error of the order 1 Edgeworth expansion in more detail. We first show that the order 1 expansion exists if (A1) through (A3) hold with $s = 3$. Then, we show that the error of approximation can be improved if (A4) holds.

Theorem 3.1.2. *Suppose (A1) through (A3) hold with $s \geq 3$. Then, the order 1 Edgeworth expansion exists.*

Theorem 3.1.3. *Suppose (A1) through (A4) hold with $s \geq 4$. Then,*

$$\mathbb{P}\left(\frac{S_N - NA}{\sqrt{N}} \leq z\right) = \mathfrak{N}(z) + \frac{P_1(z)}{N^{1/2}}\mathfrak{n}(z) + \mathcal{O}\left(\frac{1}{N^q}\right)$$

where $q = \min\left\{1, \dfrac{1}{2} + r_1\right\}$.

As one would expect, more precise asymptotics than the usual $o(N^{-\frac{1}{2}})$ are available when the characteristic function has better decay. The proof shows that the error depends mostly on the expansion of the characteristic function at 0. This is an indication that the error in Theorem 3.1.2 cannot be improved more than by a factor of $\dfrac{1}{\sqrt{N}}$ even when r_1 is large.

In [9], analogous results are obtained for subshifts of finite type in the stationary case and an explicit description of the first order Edgeworth expansion is given. Here, we consider a wider class of (not necessarily stationary) sequences and give explicit descriptions of higher order Edgeworth polynomials by relating the coefficients to asymptotic moments. Also, we improve the condition

$$H_r: \ |\mathbb{E}(e^{itS_N})| \leq K\left(1 - \frac{c}{|t|^\alpha}\right)^n, \ \frac{\alpha(r-1)}{2} < 1, \ |t| > K$$

found in [9] by replacing it with (A4). In addition, this allows us to obtain better asymptotics for the first order expansion.

We also extend the results in [4] on the existence of weak Edgeworth expansions for i.i.d. random variables. In section 3.5.1, we compare their results with the ours.

Before we mention our results, we define the space F_k^m of functions. Put

$$C^m(f) = \max_{0 \leq j \leq m} \|f^{(j)}\|_{\mathrm{L}^1} \ \text{ and } \ C_k(f) = \max_{0 \leq j \leq k} \|x^j f\|_{\mathrm{L}^1}.$$

Define

$$C_k^m(f) = C^m(f) + C_k(f).$$

We say $f \in F_k^m$ if f is m times continuously differentiable and $C_k^m(f) < \infty$.

Theorem 3.1.4. *Suppose* $(A1)$ *through* $(A4)$ *hold with* $s = r+2$. *Choose* $q \in \mathbb{N}$ *such that* $q > \dfrac{r+1}{2r_1}$. *Then, for* $f \in F_{r+1}^{q+2}$, S_N *admits weak local Edgeworth expansion of order* r.

Theorem 3.1.5. *Suppose* $(A1)$ *through* $(A4)$ *hold with* $s = r+2$. *Choose* $q \in \mathbb{N}$ *such that* $q > \dfrac{r+1}{2r_1}$. *Then, for* $f \in F_0^{q+2}$, S_N *admits weak global Edgeworth expansion of order* r.

In Theorem 3.1.4 and Theorem 3.1.5, f is required to have at least three derivatives in order to guarantee the integrability of Fourier transforms of f and its derivatives. In addition to (A1) through (A4), if we have,

(A5) There exists $C, \alpha > 0$ and N_1 such that $\|\mathcal{L}_t^N\| \leq \dfrac{C}{t^\alpha}$ for $|t| > N^{r_1}$ for $N > N_1$.

then we can improve this assumption to f having only one continuous derivative.

Theorem 3.1.4*. Suppose $(A1)$ through $(A5)$ hold with $s = r+2$ and $\alpha > \dfrac{r+1}{2r_1}$ for sufficiently large N. Then, for $f \in F_{r+1}^1$, S_N admits weak local Edgeworth expansion of order r.

Theorem 3.1.5*. Suppose $(A1)$ through $(A5)$ hold with $s = r+2$ and $\alpha > \dfrac{r+1}{2r_1}$ for sufficiently large N. Then, for $f \in F_0^1$, S_N admits weak global Edgeworth expansion of order r.

The proofs of these theorems are minor modifications of the proofs of the previous two theorems. This is described in remark 3.2.2 appearing after the proofs.

The next theorem gives sufficient conditions for the existence of the averaged Edgeworth expansion.

Theorem 3.1.6. *Suppose* $(A1)$ *through* $(A4)$ *hold with* $s = r+2$. *Choose* $q \in \mathbb{N}$

such that $q > \dfrac{r}{2r_1}$. *Then,* S_N *admits averaged Edgeworth expansion of order* r *for* $f \in F_0^q$.

We note that for integer valued random variable assumptions (A3) and (A4) cannot hold since the characteristic function of S_N is 2π-periodic. Therefore we replace (A3) by,

$\widetilde{\text{(A3)}}$ When $t \notin 2\pi\mathbb{Z}$, $\text{sp}(\mathcal{L}_t) \subset \{|z| < 1\}$ and when $t \in 2\pi\mathbb{Z}$, $\text{sp}(\mathcal{L}_t) \subset \{|z| < 1\} \cup \{1\}$. Also, because of periodicity of the characteristic function, an assumption similar to (A4) is not required.

The following theorem provides conditions for the existence of asymptotic expansions for the LCLT for weakly dependent integer valued random variables. A similar result for X_n's that are $\mathbb{Z}^d$-valued, is obtained in [42]. Compare with Proposition 4.2 and 4.4 therein.

Theorem 3.1.7. *Suppose* X_n *are integer valued,* $(A1), (A2)$ *and* $\widetilde{(A3)}$ *are satisfied with* $s = r + 2$. *Then* S_N *admits order* r *lattice Edgeworth expansion.*

The layout of the rest of the chapter is as follows. In section 3.2 we prove the results mentioned earlier by constructing the Edgeworth polynomials using characteristic functions and concluding that they satisfy the required asymptotics. In section 3.3 we relate the coefficients of these polynomials to moments of S_N and provide an algorithm to compute coefficients. A few applications of the Edgeworth expansions such as the Local Central Limit Theorem and Moderate Deviations, are discussed in section 3.4. In the last section we give examples of sequences of random variables for which our theory can be applied. First, we revisit the i.i.d. case and recover previous results. Then, we focus on non-trivial examples like observations arising from piece-wise expanding maps of an interval, Markov chains with finitely many states and markov processes which are strongly ergodic.

3.2 Proofs of the main results.

Here we prove the results mentioned earlier. From now on we work in the setting described in section 3.1.

Proof of Theorem 3.1.1. We seek polynomials $P_p(x)$ with real coefficients such that

$$\mathbb{P}\left(\frac{S_n - nA}{\sqrt{n}} \leq x\right) - \mathfrak{N}(x) = \sum_{p=1}^{r} \frac{P_p(x)}{n^{p/2}}\mathfrak{n}(x) + o\left(n^{-r/2}\right). \tag{3.3}$$

Once we have found suitable candidates for $P_p(x)$ we can apply the Berry-Esseen inequality,

$$|F_n(x) - \mathcal{E}_{r,N}(x)| \leq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\widehat{F}_n(t) - \widehat{\mathcal{E}}_{r,n}(t)}{t}\right| dt + \frac{C_0}{T}, \tag{3.4}$$

where

$$F_n(x) = \mathbb{P}\left(\frac{S_n - nA}{\sqrt{n}} \leq x\right), \quad \mathcal{E}_{r,n}(x) = \mathfrak{N}(x) + \sum_{p=1}^{r} \frac{P_p(x)}{n^{p/2}}\mathfrak{n}(x),$$

and C_0 is independent of T. We refer the reader to [20, Chapter XVI.3] for a proof of (3.4). What follows is a formal derivation of $P_p(x)$. Later, we will use (3.4) along with other estimates to prove (3.3).

It follows from (A1), (A2) and classical perturbation theory (see [33, IV.3.6 and VII.1.8]) that there exist $\delta > 0$ such that for $|t| \leq \delta$, $\mathcal{L}_t$ has a top eigenvalue $\mu(t)$ which is simple and the remainder of the spectrum is contained in a strictly smaller disk. One can express $\mathcal{L}_t$ as

$$\mathcal{L}_t = \mu(t)\Pi_t + \Lambda_t \tag{3.5}$$

where Π_t is the eigenprojection to the top eigenspace of $\mathcal{L}_t$ and $\Lambda_t = (I - \Pi_t)\mathcal{L}_t$. Because $\Lambda_t\Pi_t = \Pi_t\Lambda_t = 0$, iterating (3.5), we obtain

$$\mathcal{L}_t^n = \mu^n(t)\Pi_t + \Lambda_t^n.$$

Using (A3) and compactness, there exist C (which does not depend on n and t) and $0 < r < 1$ such that $\|\Lambda_t^n\| \leq Cr^n$ for all $|t| \leq \delta$. By (3.1),

$$\mathbb{E}(e^{itS_n/\sqrt{n}}) = \mu\Big(\frac{t}{\sqrt{n}}\Big)^n \ell\big(\Pi_{t/\sqrt{n}}v\big) + \ell\big(\Lambda_{t/\sqrt{n}}^n v\big). \tag{3.6}$$

Now, we focus on the first term of (3.6). Put

$$Z(t) = \ell(\Pi_t v). \tag{3.7}$$

Then, substituting $t = 0$ in (3.6) yields $1 = Z(0) + \ell(\Lambda_0^n v)$. Also, we know that $\lim\limits_{n\to\infty} \|\Lambda_0^n v\| = 0$. This gives $\lim\limits_{n\to\infty} \ell(\Lambda_0^n v) = 0$. Therefore, $Z(0) = 1$ and $Z(t) \neq 0$ when $|t| < \delta$. Also, this shows that $\ell(\Lambda_0^n v) = 0$ for all n. Next, note that $t \mapsto \mu(t)$ and $t \mapsto \Pi_t$ are $r+2$ times continuously differentiable on $|t| < \delta$ (see [33, IV.3.6 and VII.1.8]). Therefore, $Z(t)$ is $r+2$ times continuously differentiable on $|t| < \delta$.

Now we are in a position to compute $P_p(x)$. To this end we make use of ideas in [20, Chapter XVI] (where the Edgeworth expansions for i.i.d. random variables are constructed) and [24] (where the CLT is proved using Nagaev-Guivarc'h method).

Consider the function ψ such that,

$$\log\mu\Big(\frac{t}{\sqrt{n}}\Big) = \frac{iAt}{\sqrt{n}} - \frac{\sigma^2 t^2}{2n} + \psi\Big(\frac{t}{\sqrt{n}}\Big) \iff \mu^n\Big(\frac{t}{\sqrt{n}}\Big) = e^{\frac{inAt}{\sqrt{n}} - \frac{\sigma^2 t^2}{2}} \exp\Big(n\psi\Big(\frac{t}{\sqrt{n}}\Big)\Big).$$

where $A = \lim_{n\to\infty} \mathbb{E}\Big(\frac{S_n}{n}\Big)$ is the asymptotic mean and $\sigma^2 = \lim_{n\to\infty} \mathbb{E}\Big(\Big[\frac{S_n - nA}{\sqrt{n}}\Big]^2\Big)$ is the asymptotic variance. (For details see section 3.3.)

By (3.6) we have,

$$\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) = e^{-\frac{\sigma^2 t^2}{2}} \exp\Big(n\psi\Big(\frac{t}{\sqrt{n}}\Big)\Big) Z\Big(\frac{t}{\sqrt{n}}\Big) + e^{-\frac{inAt}{\sqrt{n}}} \ell\Big(\Lambda^n_{\frac{t}{\sqrt{n}}} v\Big) \tag{3.8}$$

Notice that $\psi(0) = \psi'(0) = 0$ and $\psi(t)$ is $r+2$ times continuously differentiable. Now, denote by $t^2\psi_r(t)$ the order $(r+2)$ Taylor approximation of ψ. Then, ψ_r is the unique polynomial such that $\psi(t) = t^2\psi_r(t) + o(|t|^{r+2})$. Also, $\psi_r(0) = 0$ and ψ_r is a polynomial of degree r. In fact, we can write $\psi(t) = t^2\psi_r(t) + t^{r+2}\tilde{\psi}_r(t)$ where $\tilde{\psi}_r$ is continuous and $\tilde{\psi}_r(0) = 0$. Thus,

$$\exp\left(n\psi\left(\frac{t}{\sqrt{n}}\right)\right) = \exp\Big(t^2\psi_r\Big(\frac{t}{\sqrt{n}}\Big) + \frac{1}{n^{r/2}} t^{r+2}\tilde{\psi}_r\Big(\frac{t}{\sqrt{n}}\Big)\Big).$$

Denote by $Z_r(t)$ the order$-r$ Taylor expansion of $Z(t) - 1$. Then, $Z_r(0) = 0$ and $Z(t) = 1 + Z_r(t) + t^r\tilde{Z}_r(t)$ with twice continuously differentiable $\tilde{Z}_r(t)$ such that $\tilde{Z}_r(0) = 0$. Then, to make the order $n^{-j/2}$ terms explicit, we compute:

$$\begin{aligned}
&e^{\frac{\sigma^2 t^2}{2}} \mu^n\Big(\frac{t}{\sqrt{n}}\Big) Z\Big(\frac{t}{\sqrt{n}}\Big) \\
&\quad = e^{\frac{\sigma^2 t^2}{2}} \mu^n\Big(\frac{t}{\sqrt{n}}\Big) \exp\log Z\Big(\frac{t}{\sqrt{n}}\Big) \\
&\quad = \exp\Big(t^2\psi_r\Big(\frac{t}{\sqrt{n}}\Big) + \frac{1}{n^{r/2}} t^{r+2}\tilde{\psi}_r\Big(\frac{t}{\sqrt{n}}\Big) \\
&\qquad\qquad\qquad - \sum_{k=1}^{r} \frac{(-1)^{k+1}}{k}\Big[Z_r\Big(\frac{t}{\sqrt{n}}\Big)\Big]^k - \frac{1}{n^{r/2}} t^r \overline{Z}_r\Big(\frac{t}{\sqrt{n}}\Big)\Big)
\end{aligned}$$

$$
\begin{aligned}
&= 1 + \sum_{m=1}^{r} \frac{1}{m!}\Big[t^2\psi_r\Big(\frac{t}{\sqrt{n}}\Big) - \sum_{k=1}^{r} \frac{(-1)^{k+1}}{k}\Big[Z_r\Big(\frac{t}{\sqrt{n}}\Big)\Big]^k\Big]^m \\
&\qquad\qquad + \frac{1}{n^{r/2}} t^{r+2}\tilde{\psi}_r\Big(\frac{t}{\sqrt{n}}\Big) - \frac{1}{n^{r/2}} t^r \overline{Z}_r\Big(\frac{t}{\sqrt{n}}\Big) + t^{r+1}\mathcal{O}\big(n^{-\frac{r+1}{2}}\big) \\
&= \sum_{k=0}^{r} \frac{A_k(t)}{n^{k/2}} + \frac{t^r}{n^{r/2}}\varphi\Big(\frac{t}{\sqrt{n}}\Big) + t^{r+1}\mathcal{O}\big(n^{-\frac{r+1}{2}}\big) \qquad (3.9)
\end{aligned}
$$

where $A_0 \equiv 1$, $\varphi(t) = t^2\tilde{\psi}_r(t) - \overline{Z}_r(t)$ is continuous and $\varphi(0) = 0$. Here $\overline{Z}_r$ is the remainder of $\log Z(t)$ when approximated by powers of Z_r. Next write,

$$
Q_n(t) = \sum_{k=1}^{r} \frac{A_k(t)}{n^{k/2}}. \qquad (3.10)
$$

Notice that

$$
A_k \text{ and } k \text{ have the same parity.} \qquad (3.11)
$$

This can be seen directly from the construction, because we collect terms with the same power of $n^{-1/2}$, ψ_r and Z_r are a polynomial in $\dfrac{t}{\sqrt{n}}$ with no constant term and we take powers of $t^2\psi_r(t)$ and $Z_r(t)$, the resulting A_k will contain terms of the form $c_s t^{2s+k}$.

We claim that,

$$
\begin{aligned}
&\int_{|t|<\delta\sqrt{n}} \left| \frac{\mu^n\big(\frac{t}{\sqrt{n}}\big)Z\big(\frac{t}{\sqrt{n}}\big) - e^{-\frac{t^2\sigma^2}{2}} - e^{-\frac{t^2\sigma^2}{2}}Q_n(t)}{t} \right| dt \qquad (3.12) \\
&\qquad = \int_{|t|<\delta\sqrt{n}} e^{-\frac{t^2\sigma^2}{2}} \left| \frac{\exp\big[n\psi\big(\frac{t}{\sqrt{n}}\big) + \log Z\big(\frac{t}{\sqrt{n}}\big)\big] - 1 - Q_n(t)}{t} \right| dt \\
&\qquad = o\left(n^{-r/2}\right).
\end{aligned}
$$

We note that from the choice of Q_n,

$$
\frac{\exp\Big[n\psi\big(\frac{t}{\sqrt{n}}\big) + \log Z\big(\frac{t}{\sqrt{n}}\big)\Big] - 1 - Q_n(t)}{t} = \frac{1}{n^{r/2}}\Big(t^{r-1}\varphi\Big(\frac{t}{\sqrt{n}}\Big) + t^r\mathcal{O}\big(n^{-\frac{r+1}{2}}\big)\Big)
$$

where $\varphi(t) = o(1)$ as $t \to 0$. As a result, for all $\varepsilon > 0$ the integrand of (3.12) can be made smaller than $\frac{\varepsilon}{n^{r/2}}(t^{r-1} + t^r)e^{-\frac{t^2\sigma^2}{2}}$ by choosing δ small enough. This proves the claim.

Even though the following derivation is only valid for $|t| < \delta\sqrt{n}$, once the polynomial function $Q_n(t)$ is obtained as above, we can consider it to be defined for all $t \in \mathbb{R}$.

Suppose $|t| \leq \delta$. From classical perturbation theory (see [33, Chapter IV] and [29, Section 7]) we have

$$\Lambda_t^n = \frac{1}{2\pi i}\int_\Gamma z^n(z - \mathcal{L}_t)^{-1}\,dz \tag{3.13}$$

where Γ is the positively oriented circle centered at $z = 0$ with radius ε_0. Here ε_0 is uniform in t and $0 < \varepsilon_0 < 1$. Now,

$$\begin{aligned}\Lambda_t^n - \Lambda_0^n &= \frac{1}{2\pi i}\int_\Gamma z^n[(z - \mathcal{L}_t)^{-1} - (z - \mathcal{L}_t)^{-1}]\,dz\\ &= \frac{1}{2\pi i}\int_\Gamma z^n[(z - \mathcal{L}_0)^{-1}(\mathcal{L}_t - \mathcal{L}_0)(z - \mathcal{L}_t)^{-1}]\,dz.\end{aligned}$$

Because $\mathcal{L}_t - \mathcal{L}_0 = \mathcal{O}(|t|)$ we have that $\frac{\Lambda_t^n - \Lambda_0^n}{|t|} = \mathcal{O}(\varepsilon_0^n)$. $\ell \in \mathbb{B}'$ and $\ell(\Lambda_0^n v) = 0$ implies that

$$\begin{aligned}\int_{|t|<\delta\sqrt{n}}\left|\frac{e^{-\frac{inAt}{\sqrt{n}}}\ell(\Lambda_{t/\sqrt{n}}^n v)}{t}\right| dt &= \int_{|t|<\delta\sqrt{n}}\left|\frac{e^{-\frac{inAt}{\sqrt{n}}}\ell(\Lambda_{t/\sqrt{n}}^n v - \Lambda_0^n v)}{t}\right| dt\\ &\leq C\int_{|t|<\delta}\left|\frac{\Lambda_t^n - \Lambda_0^n}{t}\right| dt = \mathcal{O}(\varepsilon_0^n).\end{aligned}$$

This decays exponentially fast to 0 as $n \to \infty$. This allows us to control the second term in the RHS of (3.6). Combining this with (3.12) we can conclude that,

$$\int_{|t|<\delta\sqrt{n}}\left|\frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}} - e^{-\frac{t^2\sigma^2}{2}}Q_n(t)}{t}\right| dt = o(n^{-r/2}). \tag{3.14}$$

Observe that,

$$(it)^k e^{-\frac{\sigma^2 t^2}{2}} = \frac{1}{\sqrt{2\pi\sigma^2}} \widehat{\frac{d^k}{dt^k} e^{-\frac{t^2}{2\sigma^2}}} = \widehat{\frac{d^k}{dt^k}\mathfrak{n}(t)}$$

where $\widehat{f}(x) = \displaystyle\int e^{-itx} f(t)\, dt$ is the Fourier transform of f. Therefore,

$$R_j(t)\mathfrak{n}(t) = \frac{1}{\sqrt{2\pi\sigma^2}} A_j \left(-i\frac{d}{dt}\right) \left[e^{-\frac{t^2}{2\sigma^2}}\right]. \tag{3.15}$$

Then, the required $P_p(x)$ for $p \geq 1$, can be found using the relation,

$$\mathfrak{n}(x)R_p(x) = \frac{d}{dx}\Big[\mathfrak{n}(x)P_p(x)\Big]. \tag{3.16}$$

For more details, we refer the reader to [20, Chapter XVI.3,4].

Given $\varepsilon > 0$, choose $B > \dfrac{C_0}{\varepsilon}$ where C_0 is as in (3.4). Let $r \in \mathbb{N}$. Then we choose polynomials $P_p(x)$ as described above. Then, from (3.4) it follows that,

$$\begin{aligned}
|F_n(x) - \mathcal{E}_{r,n}(x)| &\leq \frac{1}{\pi}\int_{-Bn^{r/2}}^{Bn^{r/2}} \left| \frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}}(1+Q_n(t))}{t} \right| dt + \frac{C_0}{Bn^{r/2}} \\
&\leq I_1 + I_2 + I_3 + \frac{\varepsilon}{n^{r/2}}
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{\pi}\int_{|t|<\delta\sqrt{n}} \left| \frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}}(1+Q_n(t))}{t} \right| dt \\
I_2 &= \frac{1}{\pi}\int_{\delta\sqrt{n}<|t|<Bn^{r/2}} \left| \frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t} \right| dt \\
I_3 &= \frac{1}{\pi}\int_{|t|>\delta\sqrt{n}} e^{-\frac{t^2\sigma^2}{2}} \left| \frac{1+Q_n(t)}{t} \right| dt.
\end{aligned}$$

From (3.12) we have that I_1 is $o(n^{-r/2})$. Because our choice of $\varepsilon > 0$ is arbitrary the proof is complete, if I_2 and I_3 are also $o(n^{-r/2})$. These follow from (3.18), (3.19) and (3.17) below.

It is easy to see that,

$$\int_{|t|>\delta\sqrt{n}} e^{-\frac{t^2\sigma^2}{2}} \left|\frac{1+Q_n(t)}{t}\right| dt = \mathcal{O}(e^{-cn}) \tag{3.17}$$

for some $c > 0$. Thus, we only need to control,

$$\begin{aligned} I_2 &= \int_{\delta\sqrt{n}<|t|<Bn^{r/2}} \left|\frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t}\right| dt \\ &= \int_{\delta\sqrt{n}<|t|<\bar{\delta}\sqrt{n}} \left|\frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t}\right| dt + \int_{\bar{\delta}\sqrt{n}<|t|<Bn^{r/2}} \left|\frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t}\right| dt \end{aligned}$$

where $\bar{\delta} > \max\{\delta, K\}$ with K as in (A4).

By (A3) the spectral radius of $\mathcal{L}_t$ has modulus strictly less than 1. Because $t \mapsto \mathcal{L}_t$ is continuous, for all $p < q$, there exists $\gamma < 1$ and $C > 0$, such that $\|\mathcal{L}_t^m\| \leq C\gamma^m$ for all $p \leq |t| \leq q$ for sufficiently large m. Then using (3.1) for sufficiently large n we have,

$$\int_{\delta\sqrt{n}<|t|<\bar{\delta}\sqrt{n}} \left|\frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t}\right| dt \leq \frac{1}{\delta\sqrt{n}} \int_{\delta\sqrt{n}<|t|<\bar{\delta}\sqrt{n}} \|\mathcal{L}^n_{t/\sqrt{n}}\|\, dt \leq \frac{C\gamma^n}{\sqrt{n}}. \tag{3.18}$$

This shows that the integral converges to 0 faster than any inverse power of $\sqrt{n}$. Next for sufficiently large n,

$$\begin{aligned} \int_{\bar{\delta}\sqrt{n}<|t|<Bn^{r/2}} \left|\frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t}\right| dt &\leq \frac{1}{\bar{\delta}\sqrt{n}} \int_{\bar{\delta}\sqrt{n}<|t|<Bn^{r/2}} |\ell(\mathcal{L}^n_{t/\sqrt{n}}v)|\, dt \\ &\leq \frac{2Bn^{r/2}}{\bar{\delta}n^{r_2+1/2}} \|\ell\|\|v\| \\ &= Cn^{\frac{r-1}{2}-r_2} = o(n^{-r/2}). \end{aligned} \tag{3.19}$$

The second inequality is due to assumption (A4) i.e. $\|\mathcal{L}^n_{t/\sqrt{n}}\| \leq \dfrac{1}{n^{r_2}}$ where $r_2 > \dfrac{r-1}{2}$ (we can assume $r_2 > \dfrac{r-1}{2}$ for large n due to Remark 3.1.1) and $K \leq \bar{\delta} < \dfrac{|t|}{\sqrt{n}} < Bn^{\frac{r-1}{2}} \leq n^{r_1}$ for $n \in \mathbb{N}$ with $n^{r_1 - \frac{r-1}{2}} \geq B$. $\square$

The proof of Theorem 3.1.2 follows the same idea. We include its proof for completion.

Proof of Theorem 3.1.2. Because (A1) through (A3) hold with $s \geq 3$, we have (3.9) where φ is continuous, $\varphi(0) = 0$ and $r = 1$. Given $\varepsilon > 0$, choose $B > \dfrac{C_0}{\varepsilon}$. Then,

$$\begin{aligned}
|F_n(x) - \mathcal{E}_{1,n}(x)| &\leq \frac{1}{\pi}\int_{-B\sqrt{n}}^{B\sqrt{n}} \left| \frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}}(1 + Q_n(t))}{t} \right| dt + \frac{C_0}{B\sqrt{n}} \\
&\leq I_1 + I_2 + I_3 + \frac{\varepsilon}{B\sqrt{n}}.
\end{aligned}$$

Because, $\varphi(t) = o(1)$ as $t \to 0$ and

$$\frac{\exp\left[n\psi\left(\frac{t}{\sqrt{n}}\right) + \log Z\left(\frac{t}{\sqrt{n}}\right)\right] - 1 - Q_1(t)}{t} = \frac{1}{\sqrt{n}}\varphi\Big(\frac{t}{\sqrt{n}}\Big) + t\,\mathcal{O}\Big(\frac{1}{n}\Big)$$

we have that,

$$I_1 = \int_{|t|<\delta\sqrt{n}} \left| \frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}} - e^{-\frac{t^2\sigma^2}{2}}Q_1(t)}{t} \right| dt = o(n^{-1/2}).$$

Also, $I_3 = \mathcal{O}(e^{-cn})$. Finally, because of (A3) there is $\gamma < 1$ such that,

$$\int_{\delta\sqrt{n}<|t|<B\sqrt{n}} \left| \frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t} \right| dt = \int_{\delta<|t|<B} \left| \frac{\mathbb{E}(e^{itS_n})}{t} \right| dt \leq C \sup_{\delta \leq |t| \leq B} \|\mathcal{L}_t^n\| \leq C\gamma^n$$

Combining these estimates we have the result. □

A slight modification of the previous proof gives us the proof of Theorem 3.1.3. Higher regularity assumption gives us better asymptotics near 0 and the assumption on the faster decay of the characteristic function gives us more control in the mid range.

Proof of Theorem 3.1.3. Because (A1) through (A4) hold with $s \geq 4$, we have (3.9) where φ is C^1, $\varphi(0) = 0$ and $r = 1$. Then,

$$|F_n(x) - \mathcal{E}_{1,n}(x)| \leq \frac{1}{\pi}\int_{-n^{1/2+r_1}}^{n^{1/2+r_1}} \left| \frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}}(1 + Q_n(t))}{t} \right| dt + \frac{C_0}{n^{1/2+r_1}}$$

$$\leq I_1 + I_2 + I_3 + \frac{C_0}{n^{1/2+r_1}}$$

Because, $\varphi(\frac{t}{\sqrt{n}}) \sim \frac{t}{\sqrt{n}}$ near 0 and

$$\frac{\exp\left[n\psi\left(\frac{t}{\sqrt{n}}\right) + \log Z\left(\frac{t}{\sqrt{n}}\right)\right] - 1 - Q_1(t)}{t} = \frac{1}{\sqrt{n}}\varphi\Big(\frac{t}{\sqrt{n}}\Big) + t\,\mathcal{O}\Big(\frac{1}{n}\Big)$$

we have that,

$$I_1 = \int_{|t|<\delta\sqrt{n}} \left| \frac{\mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) - e^{-\frac{t^2\sigma^2}{2}} - e^{-\frac{t^2\sigma^2}{2}} Q_1(t)}{t} \right| dt = \mathcal{O}\Big(\frac{1}{n}\Big).$$

Also, $I_3 = \mathcal{O}(e^{-cn})$. As before, (3.18) holds for $\overline{\delta} > \max\{\delta, K\}$.

$$\|\mathcal{L}_t^n\| \leq \frac{1}{n^{r_2}} \text{ where } K \leq \overline{\delta} < |t| < n^{r_1}.$$

$$\int_{\overline{\delta}\sqrt{n}<|t|<n^{1/2+r_1}} \left| \frac{\mathbb{E}(e^{itS_n/\sqrt{n}})}{t} \right| dt = \int_{\overline{\delta}<|t|<n^{r_1}} \left| \frac{\mathbb{E}(e^{itS_n})}{t} \right| dt \leq C n^{r_1 - r_2 + \frac{1}{2}}$$

Because r_2 can be made arbitrarily large by choosing n large enough, $I_2 = \mathcal{O}(\frac{1}{n})$.

Therefore,

$$|F_n(x) - \mathcal{E}_{1,n}(x)| = \mathcal{O}\Big(\frac{1}{n^s}\Big)$$

where $s = \min\big\{1, \frac{1}{2} + r_1\big\}$ and we have the required conclusion. □

Remark 3.2.1. *In the proof above, I_1 gives the contribution to the error from the expansion of the characteristic function near 0. This dominates when $r_1 \geq \frac{1}{2}$.*

Weak forms of Edgeworth expansions are discussed in detail in [4]. We adapt the ideas found in [4] to our proofs of Theorems 3.1.4 and 3.1.5. One key difference is the requirement on f to have two more derivatives than required in [4]. This compensates for the lack of control over the tail of the characteristic function of S_N. In fact, it is enough to assume $1 + \alpha$ more derivatives. But to avoid technicalities we stick to the stronger regularity assumption. In the i.i.d. case as shown in [4], a Diophantine assumption takes care of this. See section 3.5.1 for a detailed discussion.

Proof of Theorem 3.1.4. Recall that $\widehat{f}(t) = \displaystyle\int e^{-itx} f(x)\, dx$ and pick A as in (3.2). Then by Plancherel theorem,

$$\begin{aligned} \mathbb{E}(f(S_n - nA)) &= \frac{1}{2\pi}\int \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\, dt \\ \implies \sqrt{n}\mathbb{E}(f(S_n - nA)) &= \frac{1}{2\pi}\int \widehat{f}\left(\frac{t}{\sqrt{n}}\right)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\, dt. \end{aligned} \tag{3.20}$$

We first estimate RHS away from 0. Fix small $\delta > 0$. (A particular δ is chosen later). Notice that for all $\delta \le |t| \le K$ (where K as in (A4)), there exists $c_0 \in (0,1)$ such that $\|\mathcal{L}_t^n\| \le c_0^n$. Thus,

$$\left| \int_{\delta<|t|<K} \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\, dt \right| \le \int_{\delta<|t|<K} \left| \widehat{f}(t)\ell(\mathcal{L}_t^n v) \right| dt \le C\|f\|_1 c_0^n.$$

By Remark 3.1.1, for large n we can assume $r_2 > r_1 + (r+1)/2$. Therefore,

$$\begin{aligned} \left| \int_{K<|t|<n^{r_1}} \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\, dt \right| \le \|f\|_1\|\ell\|\|v\| \int_{K<|t|<n^{r_1}} \|\mathcal{L}_t^n\|\, dt &\le \frac{C\|f\|_1}{n^{r_2-r_1}} \\ &= \|f\|_1 o(n^{-(r+1)/2}). \end{aligned}$$

Because $f \in F_{r+1}^{q+2}$, we have that $t^q\widehat{f}(t) = (-i)^q\widehat{f^{(q)}}(t)$ and $\widehat{f^{(q)}}$ is integrable. In fact, $|\widehat{f^{(q)}}(t)| \le \dfrac{C}{(1+|t|)^2}$. Note that we are using only the fact that f is $q+2$ times continuously differentiable with integrable derivatives. Therefore for this to be true $f \in F_0^{q+2}$ is sufficient. Integrability of $\widehat{f^{(q)}}$ along with $q > \dfrac{r+1}{2r_1}$ implies,

$$\left| \int_{|t|>n^{r_1}} \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\, dt \right| \le \int_{|t|>n^{r_1}} |\widehat{f}(t)|\, dt \le \int_{|t|>n^{r_1}} \left| \frac{\widehat{f^{(q)}}(t)}{t^q} \right| dt \tag{3.21}$$

$$\leq \frac{\|\widehat{f^{(q)}}\|_1}{n^{r_1 q}} = \|\widehat{f^{(q)}}\|_1 o(n^{-(r+1)/2}).$$

Therefore,

$$\left| \int_{|t|>\delta} \widehat{f}(t)\mathbb{E}(e^{it(S_n - nA)})\, dt \right| = o(n^{-(r+1)/2}). \tag{3.22}$$

From (3.8), for $|t| \leq \delta\sqrt{n}$, we have,

$$\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) = e^{-\frac{\sigma^2 t^2}{2}} e^{t^2\mathcal{O}(\delta)}(1+\mathcal{O}(\delta)) + \mathcal{O}(\epsilon_0^n).$$

Thus, choosing small δ, for large n when $|t| < \delta\sqrt{n}$ there exist $c, C > 0$ such that

$$\left|\mathbb{E}\left(e^{it\frac{S_n-nA}{\sqrt{n}}}\right)\right| \leq Ce^{-ct^2}.$$

Then,

$$\sqrt{D\log n} < |t| < \delta\sqrt{n} \implies \left|\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\right| \leq Ce^{-cD\log n} = \frac{C}{n^{cD}}$$

and

$$\begin{aligned}\left| \int_{\sqrt{\frac{D\log n}{n}}<|t|<\delta} \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\, dt \right| &= \left| \int_{\sqrt{D\log n}<|t|<\delta\sqrt{n}} \widehat{f}\left(\frac{t}{\sqrt{n}}\right)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\frac{dt}{\sqrt{n}} \right| \\ &\leq \frac{C}{n^{cD}} \int_{\sqrt{\frac{D\log n}{n}}<|t|<\delta} |\widehat{f}(t)|\, dt = \frac{2\delta C\|f\|_1}{n^{cD}}.\end{aligned}$$

Combining this with (3.22) and choosing D such that, $cD > (r+1)/2$ we have that,

$$\left| \int_{|t|>\sqrt{\frac{D\log n}{n}}} \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\, dt \right| = o(n^{-(r+1)/2}). \tag{3.23}$$

Next, suppose $|t| < \sqrt{\dfrac{D\log n}{n}}$. Then,

$$\widehat{f}(t) = \sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!}t^j + \frac{t^{r+1}}{(r+1)!}\widehat{f}^{(r+1)}(\epsilon(t))$$

where $0 \leq |\epsilon(t)| \leq |t|$. Note that,

$$|\widehat{f}^{(r+1)}(\epsilon(t))| = \left| \int x^{r+1}e^{-i\epsilon(t)x}f(x)\, dx \right| \leq \int |x^{r+1}f(x)|\, dx \leq C_{r+1}(f).$$

Therefore,

$$\begin{aligned}
&\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt \\
&\qquad = \sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!n^{j/2}} \int_{|t|<\sqrt{D\log n}} t^j \mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt \\
&\qquad\qquad + \frac{1}{n^{(r+1)/2}}\frac{1}{(r+1)!}\int_{|t|<\sqrt{D\log n}} \mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})t^{r+1}\widehat{f}^{(r+1)}\Big(\epsilon\Big(\frac{t}{\sqrt{n}}\Big)\Big)\,dt
\end{aligned}$$

where

$$\Big|\int_{|t|<\sqrt{D\log n}} \mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})t^{r+1}\widehat{f}^{(r+1)}\Big(\epsilon\Big(\frac{t}{\sqrt{n}}\Big)\Big)\,dt\Big| \leq C_{r+1}(f)\int |t|^{r+1}e^{-ct^2}\,dt$$

for large n. Hence,

$$\begin{aligned}
&\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt \\
&\qquad = \sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!n^{j/2}} \int_{|t|<\sqrt{D\log n}} t^j \mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt + C_{r+1}(f)\mathcal{O}(n^{-(r+1)/2}). \qquad (3.24)
\end{aligned}$$

Because $s = r+2$, from (3.9),

$$\begin{aligned}
e^{\frac{\sigma^2t^2}{2}}\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) &= \exp\Big(n\psi\Big(\frac{t}{\sqrt{n}}\Big)\Big)Z\Big(\frac{t}{\sqrt{n}}\Big) + e^{-\frac{inAt}{\sqrt{n}}+\frac{\sigma^2t^2}{2}}\ell\big(\Lambda^n_{t/\sqrt{n}}v\big) \\
&= \sum_{k=0}^{r}\frac{A_k(t)}{n^{k/2}} + \frac{t^r}{n^{r/2}}\varphi\Big(\frac{t}{\sqrt{n}}\Big) + \mathcal{O}\Big(\frac{\log^{(r+1)/2}(n)}{n^{(r+1)/2}}\Big). \qquad (3.25)
\end{aligned}$$

Substituting this in (3.24),

$$\begin{aligned}
&\int_{|t|<\sqrt{D\log n}} \widehat{f}\left(\frac{t}{\sqrt{n}}\right)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt \qquad (3.26)\\
&= \sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!n^{j/2}} \int_{|t|<\sqrt{D\log n}} t^j e^{-\sigma^2t^2/2}\sum_{k=0}^{r}\frac{A_k(t)}{n^{k/2}}\,dt + \mathcal{O}\Big(\frac{\log^{(r+1)/2}(n)}{n^{(r+1)/2}}\Big) \\
&= \sum_{k=0}^{r}\sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!n^{(k+j)/2}} \int_{|t|<\sqrt{D\log n}} t^j A_k(t)e^{-\sigma^2t^2/2}\,dt + o(n^{-r/2}).
\end{aligned}$$

Recall from (3.11) that A_k and k have the same parity. Therefore, if $k+j$ is odd then

$$\int_{|t|<\sqrt{D\log n}} t^j A_k(t)e^{-\sigma^2t^2/2}\,dt = 0.$$

So only integral powers of n^{-1} will remain in the expansion. Also there is C that depends only on r such that,

$$\int_{|t|\geq\sqrt{D\log n}} t^j A_k(t)e^{-\sigma^2t^2/2}\,dt \leq C\int_{|t|\geq\sqrt{D\log n}} t^{4r}e^{-\sigma^2t^2/2}\,dt \leq \frac{C}{e^{\sigma^2 D\log(n)/4}} = \frac{C}{n^{\sigma^2D/4}}.$$

Choosing D such that $2\sigma^2 D > (r+1)/2$,

$$\int_{\mathbb{R}} t^j A_k(t)e^{-\sigma^2t^2/2}\,dt = \int_{|t|\leq\sqrt{D\log n}} t^j A_k(t)e^{-\sigma^2t^2/2}\,dt + o(n^{-r/2}).$$

Therefore, fixing D large, we can assume the integrals to be over the whole real line.

Now, define

$$a_{k,j} = \int_{\mathbb{R}} t^j A_k(t)e^{-\sigma^2t^2/2}\,dt$$

and substitute

$$\widehat{f}^{(j)}(0) = \int_{\mathbb{R}} (-it)^j f(t)\,dt$$

in (3.26) to obtain,

$$\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt = \sum_{k=0}^{r}\sum_{j=0}^{r} a_{k,j}\frac{1}{j!n^{(k+j)/2}}\int_{\mathbb{R}}(-it)^j f(t)\,dt + o(n^{-r/2}) \tag{3.27}$$

$$= \sum_{p=0}^{r}\frac{1}{n^p}\int_{\mathbb{R}} f(t)\sum_{k+j=2p}\frac{a_{k,j}}{j!}(-it)^j\,dt + o(n^{-r/2})$$

$$= \sum_{p=0}^{\lfloor r/2\rfloor}\frac{1}{n^p}\int_{\mathbb{R}} f(t)P_{p,l}(t)\,dt + o(n^{-r/2})$$

where

$$P_{p,l}(t) = \sum_{k+j=2p} \frac{a_{k,j}}{j!}(-it)^j. \tag{3.28}$$

The final simplification was done by absorbing the terms corresponding to higher powers of n^{-1} into the error term. Note that $P_{p,l}$ is a polynomial of degree at most $2p$ and that once we know $A_0, \ldots, A_{2p}$ we can compute $P_{p,l}$.

Finally combining (3.27) and (3.23) substituting in (3.20) we obtain the required result as shown below.

$$\begin{aligned}
\sqrt{n}\mathbb{E}(f(S_n - nA)) &= \frac{1}{2\pi}\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt \\
&\qquad + \frac{\sqrt{n}}{2\pi}\int_{|t|>\sqrt{\frac{D\log n}{n}}} \widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\,dt \\
&= \frac{1}{2\pi}\sum_{p=0}^{\lfloor r/2\rfloor}\frac{1}{n^p}\int_{\mathbb{R}} f(t)P_{p,l}(t)\,dt + o(n^{-r/2}) + \sqrt{n}\;o(n^{-(r+1)/2}) \\
&= \frac{1}{2\pi}\sum_{p=0}^{\lfloor r/2\rfloor}\frac{1}{n^p}\int_{\mathbb{R}} f(t)P_{p,l}(t)\,dt + o(n^{-r/2}).
\end{aligned}$$

□

The proof of Theorem 3.1.5 uses the relation (3.25) derived in the previous proof. But we do not use the Taylor expansion of $\widehat{f}$, so differentiability of $\widehat{f}$ is not required. So the assumption on the decay of f at infinity can be relaxed.

Proof of Theorem 3.1.5. Multiplying (3.25) by $\widehat{f}$ and integrating we obtain,

$$\begin{aligned}
&\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt \\
&\qquad = \sum_{k=0}^{r}\frac{1}{n^{k/2}}\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)A_k(t)e^{-\frac{\sigma^2t^2}{2}}\,dt + \|f\|_1 o(n^{-r/2}).
\end{aligned}$$

As in the proof of Theorem 3.1.4 the integrals above can be replaced by integrals over $\mathbb{R}$ without altering the order of the error because

$$\int_{|t|\geq\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)A_k(t)e^{-\frac{\sigma^2t^2}{2}}\,dt \leq \|f\|_1\, o(n^{-r/2})$$

for D such that $2\sigma^2D > (r+1)/2$. Therefore,

$$\int_{|t|<\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt = \sum_{k=0}^{r}\frac{1}{n^{k/2}}\int_{\mathbb{R}}\widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)A_k(t)e^{-\frac{\sigma^2t^2}{2}}\,dt + \|f\|_1 o(n^{-r/2}).$$

We pick R_p as in (3.15) and claim $P_{p,g} = R_p$.

Note that $\sqrt{n}f(t\sqrt{n}) \longleftrightarrow \widehat{f}(t/\sqrt{n})$. So by the Plancherel theorem,

$$\int_{\mathbb{R}}\sqrt{n}f\left(t\sqrt{n}\right)R_k(t)\mathfrak{n}(t)\,dt = \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)A_k(t)e^{-\frac{\sigma^2t^2}{2}}\,dt.$$

Thus,

$$\begin{aligned}\frac{1}{2\pi\sqrt{n}}&\int_{|t|<\sqrt{D\log n}}\widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\,dt\\ &= \frac{1}{\sqrt{n}}\Big(\sum_{p=0}^{r}\frac{1}{n^{p/2}}\int_{\mathbb{R}}\sqrt{n}f\left(t\sqrt{n}\right)R_p(t)\mathfrak{n}(t)\,dt + \|f\|_1 o(n^{-r/2})\Big)\\ &= \sum_{p=0}^{r}\frac{1}{n^{p/2}}\int_{\mathbb{R}}f\left(t\sqrt{n}\right)R_p(t)\mathfrak{n}(t)\,dt + \|f\|_1 o(n^{-(r+1)/2}).\end{aligned}\tag{3.29}$$

Note that (3.23) holds because $f \in F_0^{q+2}$. Now, combining (3.29) with the estimate (3.23) completes the proof. $\square$

Remark 3.2.2. *Proofs of both the Theorem 3.1.4* and Theorem 3.1.5* are almost identical except the estimate (3.21). In order to obtain the same asymptotics, the assumption on the integrability of $\widehat{f^{(q)}}$ can be replaced by (A5) and the fact that $|\widehat{f}(t)| \sim \dfrac{1}{t}$ for as $t \to \pm\infty$.*

$$\Big|\int_{|t|>n^{r_1}}\widehat{f}(t)\mathbb{E}(e^{it(S_n-nA)})\,dt\Big| \leq C\int_{|t|>n^{r_1}}|\widehat{f}(t)|\|\mathcal{L}_t^n\|\,dt$$

$$\leq C\|f\|_1 \int_{|t|>n^{r_1}} \frac{1}{t^{1+\alpha}}\, dt$$
$$\leq \frac{C\|f\|_1}{n^{r_1(\alpha-\epsilon)}} \int \frac{1}{t^{1+\epsilon}}\, dt$$

Since, $r_1\alpha > \dfrac{r+1}{2}$ choosing ϵ small enough we can make the expression $\|f\|_1\, o(n^{-(r+1)/2})$ as required.

Proof of Theorem 3.1.6. Select A as in (3.2). Define P_p by (3.15) and (3.16) and $\tilde{f}_n(x) = f(-\sqrt{n}x)$. Then the change of variables $-\dfrac{y}{\sqrt{n}} \to y$ yields,

$$\int \Big[\mathbb{P}\Big(\frac{S_n - nA}{\sqrt{n}} \leq x + \frac{y}{\sqrt{n}}\Big) - \mathfrak{N}\Big(x + \frac{y}{\sqrt{n}}\Big) - \mathcal{E}_{r,n}\Big(x + \frac{y}{\sqrt{n}}\Big)\Big] f(y)dy = \sqrt{n}\Delta_n * \tilde{f}_n(x).$$

where $\mathcal{E}_{r,n}(x) = \displaystyle\sum_{p=1}^{r} \frac{1}{n^{p/2}} P_p(x)\mathfrak{n}(x)$.

Notice that $\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\widehat{\tilde{f}}_n \in L^1$. Therefore,

$$(F_n * \tilde{f}_n)'(x) = \frac{1}{2\pi} \int e^{-itx} \mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}})\widehat{\tilde{f}}_n(t)\, dt.$$

Also,

$$\Big[\mathfrak{n} + \Big(\sum_{p=1}^{r} \frac{1}{n^{p/2}} R_p \mathfrak{n}\Big)\Big] * \tilde{f}_n(x) = \frac{1}{2\pi} \int e^{-itx} e^{-\frac{\sigma^2 t^2}{2}} \big(1 + Q_n(t)\big)\widehat{\tilde{f}}_n(t)\, dt$$

where R_p's are polynomials given by (3.15) and $Q_n(t)$ is given by (3.10). From these we conclude that,

$$(\Delta_n * \tilde{f}_n)'(x) = \frac{1}{2\pi} \int e^{-itx} \big(\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2 t^2}{2}} \big(1 + Q_n(t)\big)\widehat{\tilde{f}}_n(t)\, dt. \tag{3.30}$$

We claim that,

$$(\Delta_n * \tilde{f}_n)(x) = \frac{1}{2\pi} \int e^{-itx} \frac{\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2 t^2}{2}} \big(1 + Q_n(t)\big)}{-it} \widehat{\tilde{f}}_n(t)\, dt. \tag{3.31}$$

Indeed, if the right side of (3.31) converges absolutely, then Riemann-Lebesgue Lemma gives us that it converges 0 as $|x| \to \infty$. Differentiating (3.31) we obtain (3.30). Thus the two sides in (3.31) can differ only by a constant. Since both are 0 at $\pm\infty$, this constant is 0 and (3.31) holds.

Now, we are left with the task of showing that the right side of (3.31) converges absolutely. From the definition of $\tilde{f}_n$ it follows that, $\widehat{\tilde{f}}_n(t) = \frac{1}{\sqrt{n}}\widehat{f}\big(-\frac{t}{\sqrt{n}}\big)$. Combining this with (3.14), we have that,

$$\begin{aligned}
&\Big|\int_{|t|<\delta\sqrt{n}} e^{-itx}\frac{\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2t^2}{2}}\big(1+Q_n(t)\big)}{-it}\widehat{\tilde{f}}_n(t)\,dt\Big| \\
&\qquad\le \int_{|t|<\delta\sqrt{n}}\Big|\frac{\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2t^2}{2}}\big(1+Q_n(t)\big)}{t}\widehat{\tilde{f}}_n(t)\Big|\,dt \\
&\qquad\le \frac{\|f\|_1}{\sqrt{n}}\int_{|t|<\delta\sqrt{n}}\Big|\frac{\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2t^2}{2}}\big(1+Q_n(t)\big)}{t}\Big|\,dt \\
&\qquad= \|f\|_1 o(n^{-(r+1)/2}).
\end{aligned}$$

Note that,

$$\begin{aligned}
&\Big|\int_{|t|>\delta\sqrt{n}} e^{-itx}\frac{\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2t^2}{2}}\big(1+Q_n(t)\big)}{-it}\widehat{\tilde{f}}_n(t)\,dt\Big| \\
&\qquad\le \int_{|t|>\delta\sqrt{n}}\Big|\frac{\mathbb{E}(e^{it\frac{S_n-nA}{\sqrt{n}}}) - e^{-\frac{\sigma^2t^2}{2}}\big(1+Q_n(t)\big)}{t}\widehat{f}\Big(-\frac{t}{\sqrt{n}}\Big)\Big|\,dt \\
&\qquad\le \frac{1}{\sqrt{n}}\int_{|t|>\delta}\Big|\frac{\mathbb{E}(e^{-it(S_n-nA)}) - e^{-\frac{n^2\sigma^2t^2}{2}}\big(1+Q_n(-\sqrt{n}t)\big)}{t}\widehat{f}(t)\Big|\,dt \\
&\qquad\le \frac{1}{\sqrt{n}}\int_{|t|>\delta}\Big|\frac{\mathbb{E}(e^{-it(S_n-nA)})}{t}\widehat{f}(t)\Big|\,dt + \mathcal{O}(e^{-cn^2}).
\end{aligned}$$

Put,

$$J_n = \frac{1}{\sqrt{n}}\int_{|t|>\delta}\Big|\frac{\mathbb{E}(e^{-it(S_n-nA)})}{t}\widehat{f}(t)\Big|\,dt.$$

We claim $J_n = o(n^{-(r+1)/2})$. This proves that (3.31) converges absolutely as required.

To conclude the asymptotics of J_n, choose $\overline{\delta} > \max\{\delta, K\}$ where K as in (A4). From (A3) there exists $\gamma < 1$ such that $\|\mathcal{L}_t^n\| \le \gamma^n$ for all $\delta \le |t| \le \overline{\delta}$ for sufficiently large n. Then, using (3.1) for sufficiently large n we have,

$$\frac{1}{\sqrt{n}} \int_{\delta<|t|<\overline{\delta}} \left| \frac{\mathbb{E}(e^{-it(S_n - nA)})}{t} \widehat{f}(t) \right| dt \le \frac{C\|f\|_1}{\delta\sqrt{n}} \int_{\delta<|t|<\overline{\delta}} \|\mathcal{L}_t^n\| \, dt = \mathcal{O}(\gamma^n).$$

Next, for $K \le \overline{\delta} \le |t| \le n^{r_1}$, $\|\mathcal{L}_t^n\| \le \dfrac{1}{n^{r_2}}$. Hence, for n sufficiently large so that $r_2 > \dfrac{r}{2}$,

$$\begin{aligned}\frac{1}{\sqrt{n}} \int_{\overline{\delta}<|t|<n^{r_1}} \left| \frac{\mathbb{E}(e^{-it(S_n - nA)})}{t} \widehat{f}(t) \right| dt &\le \frac{C}{\delta\sqrt{n}} \int_{\overline{\delta}<|t|<n^{r_1}} \|\mathcal{L}_t^n\| |\widehat{f}(t)| \, dt \\ &\le \frac{C\|\widehat{f}\|_1}{n^{r_2+1/2}} = o(n^{-(r+1)/2}).\end{aligned}$$

Since $q > \dfrac{r}{2r_1}$, we have that,

$$\begin{aligned}\frac{1}{\sqrt{n}} \int_{|t|>n^{r_1}} \left| \frac{\mathbb{E}(e^{-it(S_n - nA)})}{t} \widehat{f}(t) \right| dt \le \frac{\|f^{(q)}\|_1}{\sqrt{n}} \int_{|t|>n^{r_1}} \frac{1}{|t|^{q+1}} \, dt &\le \frac{C\|f^{(q)}\|_1}{n^{qr_1+1/2}} \\ &= o(n^{-(r+1)/2}).\end{aligned}$$

Combining the above estimates, $J_n = C^q(f) o(n^{-(r+1)/2})$.

This completes the proof that $(\Delta_n * \tilde{f}_n)(x) = o(n^{-(r+1)/2})$. Hence,

$$\begin{aligned}&\int \Big[\mathbb{P}\Big(\frac{S_n - nA}{\sqrt{n}} \le x + \frac{y}{\sqrt{n}}\Big) - \mathfrak{N}\Big(x + \frac{y}{\sqrt{n}}\Big)\Big)\Big] f(y) dy \\ &\qquad = \int \mathcal{E}_{r,n}\Big(x + \frac{y}{\sqrt{n}}\Big) f(y) \, dy + \sqrt{n}\Delta_n * \tilde{f}_n(x) \\ &\qquad = \sum_{p=1}^{r} \frac{1}{n^{p/2}} \int P_p\Big(x + \frac{y}{\sqrt{n}}\Big) \mathfrak{n}(x) f(y) \, dy + C^q(f) o(n^{-r/2})\end{aligned}$$

as required. ∎

In the lattice case, periodicity allows us to simplify the proof significantly although the idea behind the proof is similar to the previous proofs.

Proof of Theorem 3.1.7. Under assumptions (A1) and (A2) we have the CLT for S_n. Put A as in (3.2). We observe that,

$$2\pi\mathbb{P}(S_n = k) = \int_{-\pi}^{\pi} e^{-itk}\mathbb{E}(e^{itS_n})\,dt = \int_{-\pi}^{\pi} e^{-itk}\ell(\mathcal{L}_t^n v)\,dt.$$

After changing variables and using (3.6), (3.7) we have,

$$2\pi\sqrt{n}\mathbb{P}\left(S_n = k\right) = \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} e^{-\frac{itk}{\sqrt{n}}}\mu\Big(\frac{t}{\sqrt{n}}\Big)^n Z\Big(\frac{t}{\sqrt{n}}\Big)\,dt + \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} e^{-\frac{itk}{\sqrt{n}}}\ell\big(\Lambda_{t/\sqrt{n}}^n v\big)\,dt. \tag{3.32}$$

By $\widetilde{\text{(A3)}}$ there exists $C > 0$ and $r \in (0,1)$ (both independent of t) such that $|\ell\left(\Lambda_t^n v\right)| \leq Cr^n$ for all $t \in [-\pi,\pi]$. Therefore the second term of (3.32) decays exponentially fast to 0 as $n \to \infty$.

Now, we focus on the first term. Using the same strategy as in the proof of Theorem 3.1.1 we have,

$$\mu\Big(\frac{t}{\sqrt{n}}\Big)^n Z\Big(\frac{t}{\sqrt{n}}\Big) = e^{\frac{inAt}{\sqrt{n}} - \frac{\sigma^2t^2}{2}}\left[1 + Q_n(t) + o(n^{-r/2})\right] \tag{3.33}$$

where $Q_n(t)$ is as in (3.10). Define R_j as in (3.15).

$$\begin{aligned}
&2\pi\sqrt{n}\mathbb{P}(S_n = k) - 2\pi\bigg\{\frac{1}{\sqrt{2\pi}}e^{-\frac{(k-nA)^2}{2\sigma^2 n}}\left(1 + \sum_{j=1}^{r}\frac{(R_p(k-nA)/\sqrt{n})}{n^{j/2}}\right)\bigg\}\\
&\quad= \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} e^{-\frac{itk}{\sqrt{n}}}\mu\Big(\frac{t}{\sqrt{n}}\Big)^n Z\Big(\frac{t}{\sqrt{n}}\Big)\,dt\\
&\qquad\quad - \int_{-\infty}^{\infty} e^{-\frac{it(k-nA)}{\sqrt{n}}}e^{-\sigma^2t^2/2}\,dt - \int_{-\infty}^{\infty} e^{-\frac{itk}{\sqrt{n}}}e^{-\frac{\sigma^2t^2}{2}}Q_n(t)\,dt + o(n^{-r/2}).
\end{aligned}$$

We estimate the RHS by estimating the three integrals given below,

$$I_1 = \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} e^{-\frac{itk}{\sqrt{n}}}\mu\Big(\frac{t}{\sqrt{n}}\Big)^n Z\Big(\frac{t}{\sqrt{n}}\Big) - e^{-\frac{it(k-nA)}{\sqrt{n}}}e^{-\frac{\sigma^2t^2}{2}}[1 + Q_n(t)]\,dt$$

$$I_2 = \int_{\delta\sqrt{n}<|t|<\pi\sqrt{n}} e^{-\frac{itk}{\sqrt{n}}}\mu\Big(\frac{t}{\sqrt{n}}\Big)^n Z\Big(\frac{t}{\sqrt{n}}\Big)\,dt$$

$$I_3 = \int_{|t|>\delta\sqrt{n}} e^{-\frac{it(k-nA)}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2}} [1 + Q_n(t)] \, dt.$$

Clearly, $|I_3|$ decays to 0 exponentially fast as $n \to \infty$. Also, $|\mu(2\pi)| = 1$ and $|\mu(t)| \in (0, 1)$ for $0 < |t| < 2\pi$. Therefore, there exists $\epsilon > 0$ such that $|\mu(t)| < \epsilon$ on $\delta \le |t| \le \pi$. Put $M = \max_{\delta \le |t| \le \pi} |Z(t)|$. Then,

$$|I_2| \le M\sqrt{n} \int_{\epsilon<|t|<\pi} |\mu(t)|^n \, dt \le 2M(\pi - \delta)\sqrt{n}\epsilon^n.$$

Hence, $|I_2|$ decays to 0 exponentially fast as $n \to \infty$. From (3.33), we have that

$$e^{-\frac{itk}{\sqrt{n}}} \Big[\mu\Big(\frac{t}{\sqrt{n}}\Big)^n Z\Big(\frac{t}{\sqrt{n}}\Big) - e^{\frac{inAt}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2}} [1 + Q_n(t)]\Big] = e^{-\frac{\sigma^2 t^2}{2}} o(n^{-r/2}).$$

This implies $|I_1| = o(n^{-r/2})$. Combining these estimates we have the required result.

□

3.3 Computing coefficients.

Since $\int_{|t|>\delta} \mathbb{E}(e^{itS_n}) \, dt$ decays sufficiently fast, the Edgeworth expansion, and hence its coefficients, depend only on the Taylor expansion of $\mathbb{E}(e^{itS_n})$ about 0. Here we relate the coefficients of Edgeworth polynomials to the asymptotics of moments of S_n by relating them to derivatives of $\mu(t)$ and $Z(t)$ at 0.

Suppose (A1) through (A4) are satisfied with $s = r + 2$. Recall (3.6):

$$\mathbb{E}(e^{itS_n}) = \mu(t)^n \, \ell(\Pi_t v) + \ell(\Lambda_t^n v). \tag{3.34}$$

Put $Z(t) = \ell(\Pi_t v)$ as before. Also write $U_n(t) = \ell(\Lambda_t^n v)$. We already know that $\mu(t), Z(t)$ and $U(t)$ are $r + 2$ times continuously differentiable. Using (3.13) one can show further that the derivatives of $U_n(t)$ satisfy:

$$\sup_{|t|\le\delta} \|U_n^{(k)}\| \le C\varepsilon_0^n$$

for all n and for all $1 \le k \le r+2$.

Taking the first derivative of (3.34) at $t=0$ we have:

$$i\mathbb{E}(S_n) = n\mu'(0) + Z'(0) + U_n'(0) \implies \lim_{n\to\infty} i\mathbb{E}\Big(\frac{S_n}{n}\Big) = \mu'(0).$$

In fact, using the Taylor expansion of $\log\mu(t)$ and above limit one can conclude that the number A we used in the statement of the CLT in (3.2), is given by

$$A = \lim_{n\to\infty} \mathbb{E}\Big(\frac{S_n}{n}\Big).$$

Therefore one can rewrite (3.6) as

$$\mathbb{E}(e^{it(S_n-nA)}) = e^{-nt\mu'(0)}\mu(t)^n Z(t) + \overline{U}_n(t) \tag{3.35}$$

where $\overline{U}_n(t) = e^{-nt\mu'(0)}U_n(t)$. Also note that its derivatives satisfy $\|\overline{U}_n^{(k)}\|_\infty = \mathcal{O}(\varepsilon_0^n)$ for all $1 \le k \le r+2$.

From (3.35), it follows that moments of $S_n - nA$ can be expanded in powers of n with coefficients depending on derivatives of μ and Z at 0. However, only powers of n upto order $k/2$ will appear. We prove this fact below.

Lemma 3.3.1. *Let $1 \le k \le r+2$. Then for large n,*

$$\mathbb{E}\big(\,[S_n - nA]^k\,\big) = \sum_{j=0}^{\lfloor k/2 \rfloor} a_{k,j} n^j + \mathcal{O}(\epsilon_0^n). \tag{3.36}$$

Proof. We first note that taking the kth derivative of (3.35) at $t=0$,

$$i^k\mathbb{E}\big(\,[S_n - nA]^k\,\big) = \frac{d^k}{dt^k}\bigg|_{t=0} \Big[e^{-nt\mu'(0)}\mu(t)^n Z(t)\Big] + \overline{U}^{(k)}(0)$$

$$= \frac{d^k}{dt^k}\bigg|_{t=0} \left[e^{-nt\mu'(0)}\mu(t)^n Z(t)\right] + \mathcal{O}(\epsilon_0^n).$$

Observe that all the derivatives of $e^{-nt\mu'(0)}\mu(t)^n Z(t)$ will only have positive integral powers of n (possibly) up to order k. Therefore, $\frac{d^k}{dt^k}\big|_{t=0} \left[e^{-nt\mu'(0)}\mu(t)^n Z(t)\right] = \sum_{j=0}^{k} a_{k,j} n^j$. We claim that for $j > k/2$, $a_{k,j} = 0$. This claim proves the result.

We notice that the first derivative of $e^{-t\mu'(0)}\mu(t)$ at $t = 0$ is 0. Thus we prove the more general claim that if $g(0) = 1$ and $g'(0) = 0$ then $\frac{d^k}{dt^k}\big|_{t=0}[g(t)^n Z(t)]$ has no terms with powers of n greater than $k/2$. From the Leibniz rule,

$$\frac{d^k}{dt^k}\bigg|_{t=0}[g(t)^n Z(t)] = \sum_{l=0}^{k} \binom{k}{l} Z^{(k-l)}(0) \frac{d^l}{dt^l}\bigg|_{t=0}[g(t)^n].$$

Therefore it is enough to prove that $\frac{d^l}{dt^l}\big|_{t=0}[g(t)^n]$ has no powers of n greater than $l/2$.

To this end we use the order l Taylor expansion of $g(t)$ about $t = 0$. Since $g'(0) = 0$ and g is $r + 2$ times continuously differentiable for $l \leq r + 2$ there exists $\phi(t)$ continuous such that,

$$\begin{aligned}
&g(t) = 1 + a_2 t^2 + \cdots + a_l t^l + t^{l+1}\phi(t) \\
\implies &g(t)^n = \sum_{k_0+k_2+\cdots+k_{l+1}=n} \frac{n!}{k_0! k_2! \ldots k_{l+1}!} (a_2 t^2)^{k_2} \ldots t^{(l+1)k_{l+1}} \phi(t)^{k_{l+1}} \\
&\qquad = \sum_{k_0+k_2+\cdots+k_{l+1}=n} \frac{C_{k_0 k_2 \ldots k_{l+1}} n!}{k_0! k_2! \ldots k_{l+1}!} t^{2k_2+\cdots+(l+1)k_{l+1}} \phi(t)^{k_{l+1}}.
\end{aligned}$$

After combining and rearranging terms according to powers of t, we can obtain the order l Taylor expansion of $g(t)^n$. Notice that if $k_{l+1} \geq 1$ then $2k_2 + \cdots + (l + 1)k_{l+1} \geq l + 1$. Terms with $k_{l+1} \geq 1$ are part of the error term of the order l Taylor expansion of $g(t)^n$. Since our focus is on the derivative at $t = 0$, the only terms that matter are terms with $k_{l+1} = 0$ and $2k_2 + \cdots + lk_l = l$. This implies that $k_2 + \cdots + k_l \leq \dfrac{l}{2}$. Because k_i's are non-negative integers, this means $k_2 + \cdots + k_l \leq \lfloor \dfrac{l}{2} \rfloor$. Hence, $k_0 \geq n - \lfloor \dfrac{l}{2} \rfloor$.

This analysis shows that the largest contribution to $\dfrac{d^l}{dt^l}\Big|_{t=0}[g(t)^n]$ comes from the term,

$$\frac{C_{(n-\lfloor \frac{l}{2} \rfloor),1,\ldots,1,0,\ldots,0}\ n!}{\left(n - \lfloor \frac{l}{2} \rfloor\right)!}\ t^l$$

whose kth derivative at 0 is,

$$\frac{C_{(n-\lfloor \frac{l}{2} \rfloor),1,\ldots,1,0,\ldots,0}\ l!\ n!}{\left(n - \lfloor \frac{l}{2} \rfloor\right)!} = C_{(n-\lfloor \frac{l}{2} \rfloor),1,\ldots,1,0,\ldots,0}\ l!\ n \ldots \left(n - \left\lfloor \frac{l}{2} \right\rfloor + 1\right) = \mathcal{O}(n^{\lfloor \frac{l}{2} \rfloor}).$$

Therefore,

$$\frac{d^l}{dt^l}\bigg|_{t=0}[g(t)^n] = \mathcal{O}(n^{\lfloor \frac{l}{2} \rfloor}).$$

$\square$

It is immediate from the proof that the coefficients $a_{k,j}$ are determined by the derivatives of $\mu(t)$ and $Z(t)$ near 0. For example, the constant term $a_{k,0} = (-i)^k Z^{(k)}(0)$. This follows from these three facts. The expansion (3.36) is the kth derivative of the product of the three functions $e^{-nt\mu'(0)}, \mu(t)^n$ and $Z(t)$ at $t = 0$. All derivatives of $\mu(t)^n$ and $e^{-nt\mu'(0)}$ at $t = 0$ contain powers of n and thus, $a_{k,0}$ corresponds to the term $Z(t)$ being differentiated k times in the Leibneiz rule. Both $e^{-nt\mu'(0)}$ and $\mu(t)^n$ are 1 at $t = 0$. We will see later that the other coefficients $a_{k,j}$ are combinations of $\mu'(0) = iA$, higher order derivatives of μ at 0 upto order k and derivatives of Z at 0 upto order $k - 1$.

As a corollary to Lemma 3.3.1, we conclude that asymptotic moments of orders upto $r + 2$ exist. These provide us an alternative way to describe $a_{k,j}$.

Corollary 3.3.2. *For all* $1 \leq m \leq r+2$ *and* $0 \leq j \leq \dfrac{m}{2}$,

$$a_{m,j} = \lim_{n\to\infty} \frac{\mathbb{E}\left([S_n - nA]^m\right) - n^{j+1}a_{m,j+1} - \cdots - n^{\lfloor \frac{m}{2} \rfloor}a_{m,\lfloor \frac{m}{2} \rfloor}}{n^j}.$$

Proof. When $m = 1$, $\mathbb{E}([S_n - nA]) = a_{1,0} + \mathcal{O}(\epsilon_0^n)$ and it is immediate that $a_{1,0} = \lim_{n\to\infty} \mathbb{E}([S_n - nA])$. For arbitrary k we have,

$$\mathbb{E}\Big([S_n - nA]^k \Big) = a_{k,\lfloor k/2 \rfloor} n^{\lfloor k/2 \rfloor} + a_{k,\lfloor k/2 \rfloor -1} n^{\lfloor k/2 \rfloor -1} + \cdots + a_{k,0} + \mathcal{O}(\epsilon_0^n)$$

and dividing by n we obtain,

$$\frac{\mathbb{E}\Big([S_n - nA]^k \Big)}{n^{\lfloor k/2 \rfloor}} = a_{k,\lfloor k/2 \rfloor} + \mathcal{O}\Big(\frac{1}{n}\Big).$$

Now, it is immediate that,

$$a_{k,\lfloor k/2 \rfloor} = \lim_{n\to\infty} \frac{\mathbb{E}\Big([S_n - nA]^k \Big)}{n^{\lfloor k/2 \rfloor}}.$$

Having computed $a_{k,j}$, for $r \leq j \leq \lfloor \dfrac{k}{2} \rfloor$, we can write,

$$\mathbb{E}\Big([S_n - nA]^k \Big) - a_{k,\lfloor k/2 \rfloor} n^{\lfloor k/2 \rfloor} - \cdots - a_{k,r} n^r = a_{k,r-1} n^{r-1} + \cdots + a_{k,0} + \mathcal{O}(\epsilon_0^n).$$

Dividing by n^{r-1}, we obtain,

$$\frac{\mathbb{E}\Big([S_n - nA]^k \Big) - n^r a_{k,r} - \cdots - n^{\lfloor k/2 \rfloor} a_{k,\lfloor k/2 \rfloor}}{n^{r-1}} = a_{k,r-1} + \mathcal{O}\Big(\frac{1}{n}\Big).$$

Now, we can compute $a_{m+1,r-1}$,

$$a_{k,r-1} = \lim_{n\to\infty} \frac{\mathbb{E}\Big([S_n - nA]^k \Big) - n^r a_{k,r} - \cdots - n^{\lfloor k/2 \rfloor} a_{k,\lfloor k/2 \rfloor}}{n^{r-1}}.$$

This proves the Corollary for arbitrary $k \in \{1, \ldots, r+2\}$. $\square$

Because the coefficients of polynomials $A_p(t)$ (see (3.10)) are combinations of derivatives of $\mu(t)$ and $Z(t)$ at $t = 0$, we can write them explicitly in terms of $a_{k,j}$, and hence, by applying Corollary 3.3.2, the coefficients of Edgeworth polynomials can be expressed in terms of moments of S_n. Next, we will introduce a recursive algorithm to do this and illustrate the process by computing the first and second Edgeworth polynomials.

Taking the first derivative of (3.35) at $t = 0$,

$$i\mathbb{E}([S_n - nA]) = Z'(0) + \overline{U}'_n(0).$$

Then,

$$a_{1,0} = \lim_{n\to\infty} \mathbb{E}([S_n - nA]) = -iZ'(0).$$

Next, taking the second derivative of (3.35) at $t = 0$ we have,

$$i^2\mathbb{E}([S_n - nA]^2) = n[\mu''(0) - \mu'(0)^2] + Z''(0) + \overline{U}''_n(0).$$

Therefore, dividing by n and taking the limit we have,

$$a_{2,1} = \sigma^2 = \lim_{n\to\infty} \mathbb{E}\left(\left[\frac{S_n - nA}{\sqrt{n}}\right]^2\right) = \mu'(0)^2 - \mu''(0). \tag{3.37}$$

Once we have found $a_{2,1}$ we can find

$$a_{2,0} = \lim_{n\to\infty} \left(\mathbb{E}([S_n - nA]^2) - n\sigma^2\right) = -Z''(0).$$

We can repeat this procedure iteratively. For example, after we compute the 3rd derivative of (3.35) at $t = 0$:

$$i^3\mathbb{E}([S_n - nA]^3) = Z^{(3)}(0) + n\mu'(0)[2\mu'(0)^2 - 3\mu''(0)] + n\mu^{(3)}(0)$$

$$+ 3nZ'(0)[\mu'(0)^2 - \mu''(0)] + \overline{U}_n^{(3)}(0)$$

we get that,

$$a_{3,1} = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left([S_n - nA]^3\right) = -A(3\sigma^2 + A^2) + i\mu^{(3)}(0) - 3i\sigma^2 Z'(0)$$

$$= -A(3\sigma^2 + A^2) + i\mu^{(3)}(0) + 3\sigma^2 a_{1,0}.$$

This gives us $\mu^{(3)}(0)$ and $Z^{(3)}(0)$ in terms of asymptotics of moments of S_n:

$$i\mu^{(3)}(0) = a_{3,1} + A(3\sigma^2 + A^2) - 3\sigma^2 a_{1,0}$$

$$iZ^{(3)}(0) = \lim_{n\to\infty} \left(\mathbb{E}([S_n - nA]^3) - na_{3,1}\right).$$

Given that we have all the coefficients $a_{k,j}$, $1 \leq k \leq m$ computed and $\mu^{(k)}(0), Z^{(k)}(0)$ for $1 \leq k \leq m$ expressed in terms of the former, we can compute $a_{m+1,j}$ and express $\mu^{(m+1)}(0), Z^{(m+1)}(0)$ in terms of $a_{k,j}$, $1 \leq k \leq m+1$.

To see this note that $\mu^{(m+1)}(0)$ appears only as a result of $\mu^n(t)$ being differentiated $m+1$ times. So, $\mu^{(m+1)}(0)$ only appears in derivatives of order $m+1$ and higher. It is also easy to see that it appears in the form $n\mu^{(m+1)}(0)$ in the $(m+1)$th derivative of (3.35). Thus, it is a part of $a_{m+1,1}$ and all the other terms in $a_{m+1,1}$ are products of $\mu^{(k)}(0), Z^{(k)}(0)$ for $1 \leq k \leq m$ whose orders add upto $m+1$ and hence they are products of $a_{k,j}$, $1 \leq k \leq m$.

Also, $Z^{m+1}(0)$ appears only in $a_{m+1,0}$. This is because $Z^{m+1}(0)$ appears only as a result of $Z(t)$ being differentiated $m+1$ times. Thus, it appears only in derivatives of (3.35) of order $m+1$ or higher. In the $(m+1)$th derivative of (3.35), there is only one term containing $Z^{(m+1)}(t)$ and it is $e^{-nt\mu'(0)}\mu(t)^n Z^{m+1}(t)$. So $a_{m+1,0} = (-i)^{m+1} Z^{m+1}(0)$.

Using Corollary 3.3.2, we have,

$$a_{m+1,\lfloor\frac{m+1}{2}\rfloor} = \lim_{n\to\infty} \frac{\mathbb{E}\left([S_n - nA]^{m+1}\right)}{n^{\lfloor\frac{m+1}{2}\rfloor}}.$$

Having computed $a_{m+1,j}$, for $r \le j \le \lfloor\dfrac{m+1}{2}\rfloor$, we compute $a_{m+1,r-1}$:

$$a_{m+1,r-1} = \lim_{n\to\infty} \frac{\mathbb{E}\left([S_n - nA]^{m+1}\right) - n^r a_{m+1,r} - \cdots - n^{\lfloor\frac{m+1}{2}\rfloor} a_{m+1,\lfloor\frac{m+1}{2}\rfloor}}{n^{r-1}}.$$

This gives us $Z^{(m+1)}(0) = i^{m+1} a_{m+1,0}$ and $\mu^{m+1}(0)$ in terms of $a_{m+1,1}$ and $a_{k,j}$, $1 \le k \le m$ i.e. explicitly in terms of moments of S_n. Proceeding inductively we can compute all the derivatives upto order r of $\mu(t)$ and $Z(t)$ at $t = 0$ in this manner by taking derivatives up to order r of (3.35) at $t = 0$. This is possible because our assumptions guarantee the existence of the first $r + 2$ derivatives of (3.35) near $t = 0$.

Remark 3.3.1. *This representation of $\mu^{(k)}(0)$ and $Z^{(k)}(0)$ in terms of $a_{k,j}$ is not unique. However, it is convenient to choose the $a_{k,j}$'s with the lowest possible indices. The inductive procedure explained above yields exactly this representation.*

We will illustrate how the first and the second order Edgeworth expansion can be computed explicitly once we have $\mu^{(4)}(0), \mu^{(3)}(0), Z''(0)$ and $Z'(0)$ in terms of asymptotic moments of S_n. Because $A_0(t) = 1$ we have $R_0(t) = 1$. From the derivation of (3.9) we have,

$$\begin{aligned} A_1(t) = (\log\mu)^{(3)}(0)\frac{t^3}{6} - Z'(0)t &= (\mu^{(3)}(0) - 3\mu''(0)\mu'(0) + 2\mu'(0)^3)\frac{t^3}{6} - Z'(0)t \\ &= \left(\mu^{(3)}(0) + iA(3\sigma^2 + A^2)\right)\frac{t^3}{6} - Z'(0)t \\ &= (a_{3,1} - 3\sigma^2 a_{1,0})\frac{(it)^3}{6} - a_{1,0}(it). \end{aligned}$$

After taking the inverse Fourier transform as shown in (3.15) we have,

$$R_1(x) = \frac{(a_{3,1} - 3\sigma^2 a_{1,0})}{6\sigma^6} x(3\sigma^2 - x^2) + \frac{a_{1,0}}{\sigma^2} x.$$

Using (3.16) we obtain the first Edgeworth polynomial,

$$P_1(x) = \frac{(a_{3,1} - 3\sigma^2 a_{1,0})}{6\sigma^4} (\sigma^2 - x^2) - \frac{a_{1,0}}{\sigma}.$$

Similar calculations give us,

$$\begin{aligned} A_2(t) = (a_{3,1} + 3\sigma^2 a_{1,0})^2 \frac{(it)^6}{72} + \Big[& A^2(6\sigma^2 + A^4) + 4a_{3,1}(A - 2a_{1,0}) \\ & - 3\sigma^2(2a_{2,0} - 4Aa_{1,0} + \sigma^2) + a_{4,1} \Big] \frac{(it)^4}{24} + (2a_{1,0}^2 - a_{2,0}) \frac{(it)^2}{2}. \end{aligned}$$

From (3.15) and (3.16) we have,

$$\begin{aligned} R_2(t) =& (a_{3,1} + 3\sigma^2 a_{1,0})^2 \frac{x^6 - 15\sigma^2 x^4 + 45\sigma^4 x^2 - 15\sigma^6}{72\sigma^{12}} \\ & + \Big[A^2(6\sigma^2 + A^4) + 4a_{3,1}(A - 2a_{1,0}) - 3\sigma^2(2a_{2,0} - 4Aa_{1,0} + \sigma^2) + a_{4,1} \Big] \\ & \times \frac{(x^4 - 6\sigma^2 x^2 + 3\sigma^2)}{24\sigma^8} + (2a_{1,0}^2 - a_{2,0}) \frac{(x^2 - \sigma^2)}{2\sigma^4}, \end{aligned}$$

$$\begin{aligned} P_2(t) =& (a_{3,1} + 3\sigma^2 a_{1,0})^2 \frac{x(15\sigma^2 - 10\sigma^2 x^2 + x^6)}{72\sigma^{10}} \\ & + \Big[A^2(6\sigma^2 + A^4) + 4a_{3,1}(A - 2a_{1,0}) - 3\sigma^2(2a_{2,0} - 4Aa_{1,0} + \sigma^2) + a_{4,1} \Big] \\ & \times \frac{x(3\sigma^2 - x^2)}{24\sigma^6} + (2a_{1,0}^2 - a_{2,0}) \frac{x}{2\sigma^2}. \end{aligned}$$

Remark 3.3.2. *Once we have R_p for $p \in \mathbb{N}_0$ and P_p for $p \in \mathbb{N}$, the polynomials $P_{p,g}, P_{p,d}$ and $P_{p,a}$ are given by $P_{p,g} = P_{p,d} = R_p$ and $P_{p,a} = P_p$. These relations were obtained in the proofs in section 3.2.*

Also, one can compute $P_{p,l}$ using (3.28):

$$P_{p,l}(x) = \sum_{l+j=2p} \frac{(-ix)^j}{j!} \int t^j A_l(t) e^{-\frac{\sigma^2 t^2}{2}} \, dt.$$

For example,

$$P_{0,l}(x) = \int A_0(t)e^{-\frac{\sigma^2t^2}{2}}\,dt = \sqrt{\frac{2\pi}{\sigma^2}}.$$

$$\begin{aligned}
P_{1,l}(x) &= \int A_2(t)e^{-\frac{\sigma^2t^2}{2}}\,dt\, - ix\int tA_1(t)e^{-\frac{\sigma^2t^2}{2}}\,dt - \frac{x^2}{2}\int t^2A_0(t)e^{-\frac{\sigma^2t^2}{2}}\,dt \\
\frac{P_{1,l}(x)}{\sqrt{2\pi}} =&(a_{3,1}+3\sigma^2a_{1,0})^2\frac{5}{24\sigma^7} \\
&+\Big[A^2(6\sigma^2+A^4)+4a_{3,1}(A-2a_{1,0})-3\sigma^2(2a_{2,0}-4Aa_{1,0}+\sigma^2)+a_{4,1}\Big]\frac{1}{8\sigma^5} \\
&-(2a_{1,0}^2-a_{2,0})\frac{1}{2\sigma^6}-\bigg((a_{3,1}-3\sigma^2a_{1,0})\frac{1}{\sigma^5}+\frac{2a_{1,0}}{\sigma^3}\bigg)\frac{x}{2}-\frac{x^2}{2\sigma^3}
\end{aligned}$$

Higher order Edgeworth polynomials can be computed similarly.

We can compare our results with the centered i.i.d. case. Then, we have that $A=0$, $a_{1,0}=0$ because the sequence is stationary. Also, $a_{3,1}=\lim_{n\to\infty}\frac{1}{n}\mathbb{E}([S_n-nA]^3)=\mathbb{E}((X_1-A)^3)$, $a_{2,0}=0$ and $a_{4,1}=\mathbb{E}(X_1^4)$. So, the above polynomials reduce to,

$$\begin{aligned}
A_1(t) &= \frac{\mathbb{E}(X_1^3)}{6}(it)^3,\ R_1(x)=\frac{\mathbb{E}(X_1^3)}{6\sigma^6}x(3\sigma^2-x^2),\ P_1(x)=\frac{\mathbb{E}(X_1^3)}{6\sigma^4}(\sigma^2-x^2) \\
A_2(t) &= \mathbb{E}(X_1^3)^2\frac{(it)^6}{72}+(\mathbb{E}(X_1^4)-3\sigma^4)\frac{(it)^4}{24} \\
\frac{P_{0,l}(x)}{\sqrt{2\pi}} &= \frac{1}{\sigma},\ \frac{P_{1,l}(x)}{\sqrt{2\pi}}=\frac{\mathbb{E}(X_1^3)^2}{\sigma^7}\frac{5}{24}+\Big(\frac{\mathbb{E}(X_1^4)}{\sigma^5}-\frac{3}{\sigma}\Big)\frac{1}{8}-\frac{\mathbb{E}(X_1^3)}{\sigma^5}\frac{x}{2}-\frac{1}{\sigma^3}\frac{x^2}{2}
\end{aligned}$$

These agree with the polynomials found in [20, Chapter XVI] (to see this one has to replace x by x/σ to make up for not normalizing by σ here) and [4]. The polynomials Q_k found in the latter are related to $P_{k,l}$ by $Q_k(x)=\frac{1}{2\pi}P_{k,l}(x)$.

It is also easy to see that these agree with previous work on non-i.i.d. examples. In both [9, 29] only the first order Edgeworth polynomial is given explicitly. In [9], because the sequence is stationary and centered, we can take $A=0$ and $a_{1,0}=0$.

Also, the pressure $P(t)$ given there, corresponds to $\log \mu(t)$ here. So we recover $A_1(t) = P'''(0)\dfrac{(it)^3}{6}$ in [9, Theorem 3]. In [29], sequence is centered but not assumed to be stationary. So $A = 0$ and $a_{1,0} \neq 0$ and the asymptotic bias appears in the expansion and $A_1(t) = i\mu^{(3)}(0)\dfrac{(it)^3}{6} - a_{1,0}(it)$ which agrees with [29, Theorem 8.1]. This dependence on initial distribution corresponds to presence of ℓ in (3.1).

3.4 Applications.

3.4.1 Local Limit Theorem.

Existence of the Edgeworth expansion allows us to derive Local Limit Theorems (LLTs). For example see [16, Theorem 4]. Also, as direct consequences of weak global Edgeworth expansions, an LCLT comparable to the one given in [27, Chapter II], holds. In fact, a stronger version of LCLT holds true in special cases.

To make the notation simpler, we assume that the asymptotic mean of S_N is 0. That is $A = \lim\limits_{N\to\infty} \mathbb{E}\big(\dfrac{S_N}{N}\big) = 0$.

Proposition 3.4.1. *Suppose that S_N satisfies the weak global Edgeworth expansion of order 0 for an integrable function $f \in (\mathcal{F}, \|\cdot\|)$ where $\|\cdot\|$ is translation invariant. Further, assume that $|xf(x)|$ is integrable. Then,*

$$\sqrt{N}\mathbb{E}(f(S_N - u)) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2N\sigma^2}} \int f(x)\, dx + o(1) \tag{3.38}$$

uniformly for $u \in \mathbb{R}$.

Proof. After the change of variables $z\sqrt{N} \to z$ in the RHS of the weak global

Edgeworth expansion,

$$\begin{aligned}
&\sqrt{N}\mathbb{E}(f(S_N-u)) \\
&\quad = \int \mathfrak{n}\Big(\frac{z}{\sqrt{N}}\Big) f(z-u)dz + \|f\|o(1) \\
&\quad = \int \Big[\mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big) + (z-u)\mathfrak{n}'\Big(\frac{z_u}{\sqrt{N}}\Big)\Big] f(z-u)dz + \|f\|o(1) \\
&\quad = \mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big)\int f(z-u)\,dz + \frac{C}{N}\int (z-u)\mathfrak{n}\Big(\frac{z_u}{\sqrt{N}}\Big) f(z-u)dz \, + \|f\|o(1)
\end{aligned}$$

Here z_u is between u and z and depends continuously on u.

Notice that,

$$\Big|\int (z-u)\mathfrak{n}\Big(\frac{z_u}{\sqrt{N}}\Big) f(z-u)dz\Big| \le \int |(z-u)f(z-u)|dz \le \|xf\|_1$$

Therefore, after a change of variables $z-u \to z$ in the RHS,

$$\sqrt{N}\mathbb{E}(f(S_N-u)) = \mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big)\int f(z)dz + \max\{\|xf\|_1, \|f\|\}\, o(1)$$

as required. $\square$

In particular, the result holds for $\mathcal{F} = F_0^1$. If the order 0 weak global Edgeworth expansion holds for all $f \in F_0^1$, then we have the following corollary. We note that this is indeed the case for faster decaying $|\mathbb{E}(e^{itS_N})|$ as in Markov chains and piecewise expanding maps described in sections 3.5.3.1, 3.5.3.2 and 3.5.4.

Corollary 3.4.2. *Suppose that S_N admits the weak global Edgeworth expansion of order 0 for all $f \in F_0^1$. Then, for all $a < b$,*

$$\frac{\sqrt{N}}{(b-a)}\mathbb{P}\Big(S_N \in (u+a, u+b)\Big) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2N\sigma^2}} + o(1)$$

uniformly in $u \in \mathbb{R}$.

Proof. Fix $a < b$. It is elementary to see that there exists a sequence $f_k \in F_0^1$ with compact support such that $f_k \to 1_{(u+a,u+b)}$ point-wise and f_k's are uniformly bounded in F_1^1. This bound can be chosen uniformly in u, call it C.

Therefore, from the proof of Proposition 3.4.1, we have,

$$\sqrt{N}\mathbb{E}(f_k(S_N - u)) = \mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big) \int f_k(z)dz + C_1^1(f_k)\, o(1)$$

Because $0 \le C_1^1(f_k) \le C$, taking the limit as $k \to \infty$ we conclude,

$$\sqrt{N}\mathbb{P}\Big(S_N \in (u+a, u+b)\Big) = \mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big) \int_{u+a}^{u+b} 1\, dz + C\, o(1)$$

and the result follows. □

In fact, u in the previous theorem need not be fixed. For example, for a sequence u_N with $\dfrac{u_N}{\sqrt{N}} \to u$, we have the following:

Corollary 3.4.3. *Suppose that S_N admits the weak global Edgeworth expansion of order 0 for all $f \in F_0^1$. Let u_N be a sequence such that $\lim_{N\to\infty} \dfrac{u_N}{\sqrt{N}} = u$. Then, for all $a < b$,*

$$\lim_{N\to\infty} \frac{\sqrt{N}}{(b-a)}\mathbb{P}\Big(S_N \in (u_N + a, u_N + b)\Big) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}}.$$

Now, we state the stronger version of LCLT in which we allow intervals to shrink.

Definition 8. *Given a sequence ϵ_N in $\mathbb{R}^+$ with $\epsilon_N \to 0$ as $N \to \infty$, we say that S_N admits an LCLT for ϵ_N if we have,*

$$\frac{\sqrt{N}}{2\epsilon_N}\mathbb{P}\Big(S_N \in (u - \epsilon_N, u + \epsilon_N)\Big) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2N\sigma^2}} + o(1)$$

uniformly in $u \in \mathbb{R}$.

The next proposition gives a existence of weak global Edgeworth expansions as a sufficient condition for S_N to admit a LCLT for a sequence ϵ_N. Notice that existence of higher order expansions allow ϵ_N to decay faster. In case expansions of all orders exist, ϵ_N can decay at any subexponential rate.

Proposition 3.4.4. *Suppose that S_N satisfies the weak global Edgeworth expansion of order r (≥ 1) for all $f \in F_0^1$. Let ϵ_N be a sequence of positive real numbers such that $\epsilon_N \to 0$ and $\epsilon_N N^{r/2} \to \infty$ as $N \to \infty$. Then, S_N admits an LCLT for ϵ_N.*

Proof. WLOG assume $\epsilon_N < 1$ for all N. As in the previous proof, there exists a sequence $f_k \in F_0^1$ with compact support such that $f_k \to 1_{(u-\epsilon_N, u+\epsilon_N)}$ point-wise and f_k's are uniformly bounded in F_0^1. This bound can be chosen uniformly in N and u, call it C.

Let $N \in \mathbb{N}$. Note that for all k,

$$\mathbb{E}(f_k(S_N)) = \sum_{p=0}^{r} \frac{1}{N^{\frac{p}{2}}} \int P_{p,g}(z)\mathfrak{n}(z) f_k\big(z\sqrt{N}\big) dz + C_0^1(f_k)\, o\left(N^{-(r+1)/2}\right).$$

By taking the limit as $k \to \infty$ and using the fact $0 \leq C_0^1(f_k) \leq C$, we conclude,

$$\mathbb{P}\Big(S_N \in (u-\epsilon_N, u+\epsilon_N)\Big) = \sum_{p=0}^{r} \frac{1}{N^{\frac{p}{2}}} \int_{\frac{u-\epsilon_N}{\sqrt{N}}}^{\frac{u+\epsilon_N}{\sqrt{N}}} P_{p,g}(z)\mathfrak{n}(z)\, dz + C\, o\left(N^{-(r+1)/2}\right).$$

After a change of variables $z \to \dfrac{z}{\sqrt{N}}$ in the $p = 0$ term and divide the whole equation by $2\epsilon_N$ to get,

$$\begin{aligned}
&\frac{\sqrt{N}}{2\epsilon_N}\mathbb{P}\Big(S_N \in (u-\epsilon_N, u+\epsilon_N)\Big) \\
&= \frac{1}{2\epsilon_N} \int 1_{J_N}(z-u)\mathfrak{n}\big(\frac{z}{\sqrt{N}}\big)\, dz + \sum_{p=1}^{r} \frac{\sqrt{N}}{2\epsilon_N N^{\frac{p}{2}}} \int_{\frac{u-\epsilon_N}{\sqrt{N}}}^{\frac{u+\epsilon_N}{\sqrt{N}}} P_{p,g}(z)\mathfrak{n}(z)\, dz + C\, o\left(\frac{1}{\epsilon_N N^{r/2}}\right)
\end{aligned}$$

where $J_N = (-\epsilon_N, \epsilon_N)$.

Note that for $p \geq 1$, there exists C_p such that $|P_{p,g}(z)\mathfrak{n}(z)| < C_p$. Therefore,

$$\left|\frac{\sqrt{N}}{2\epsilon_N N^{\frac{p}{2}}} \int_{\frac{u-\epsilon_N}{\sqrt{N}}}^{\frac{u+\epsilon_N}{\sqrt{N}}} P_{p,g}(z)\mathfrak{n}(z)\, dz\right| \leq \frac{C_p\sqrt{N}}{2\epsilon_N N^{\frac{p}{2}}} \int_{\frac{u-\epsilon_N}{\sqrt{N}}}^{\frac{u+\epsilon_N}{\sqrt{N}}} 1\, dz \leq \frac{C_p}{N^{p/2}} = o(1)$$

Also, as in the proof of Proposition 3.4.1,

$$\frac{1}{2\epsilon_N} \int 1_{J_N}(z-u)\mathfrak{n}\big(\frac{z}{\sqrt{N}}\big)\, dz = \frac{1}{2\epsilon_N}\mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big) \int_{u-\epsilon_N}^{u+\epsilon_N} 1\, dz$$
$$+ \frac{C}{2\epsilon_N N} \int_{u-\epsilon_N}^{u+\epsilon_N} (z-u)\mathfrak{n}\Big(\frac{z_u}{\sqrt{N}}\Big)\, dz$$

Note that,

$$\left|\frac{C}{2\epsilon_N N} \int_{u-\epsilon_N}^{u+\epsilon_N} (z-u)\mathfrak{n}\Big(\frac{z_u}{\sqrt{N}}\Big)\, dz\right| \leq \frac{C}{2\epsilon_N N} \int_{u-\epsilon_N}^{u+\epsilon_N} |z-u|\, dz = \frac{C\epsilon_N}{2N}$$

Therefore,

$$\frac{1}{2\epsilon_N} \int 1_{J_N}(z-u)\mathfrak{n}\big(\frac{z}{\sqrt{N}}\big)\, dz = \mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big) + o(1).$$

Combining these estimates with $\epsilon_N N^{r/2} \to \infty$ we have that,

$$\frac{\sqrt{N}}{2\epsilon_N}\mathbb{P}\Big(S_N \in (u-\epsilon_N, u+\epsilon_N)\Big) = \mathfrak{n}\Big(\frac{u}{\sqrt{N}}\Big) + o(1)$$

and it is straightforward from the proof that this is uniform. □

Remark 3.4.1. *We note that this result implies [16, Theorem 4] because existence of classical Edgeworth expansions imply the existence of the weak global Edgeworth expansion and this result is uniform in u.*

3.4.2 Moderate Deviations.

While the CLT describes the typical behaviour or ordinary deviations from the mean provided by the law of large numbers, it is not sufficient to understand properties of distribution of X_n completely. Therefore, the study of excessive deviations is important.

For example, deviations of order n are called large deviations. An exponential moment condition is required for a large deviation principle to hold, even for the i.i.d. case. However, when deviations are of order $\sqrt{n \log n}$ (moderate deviations) this is not the case. We show here that a moderate deviation principle holds for S_N under a weaker assumption than the exponential moment assumption.

It is also worth noting that moderate deviations have numerous applications in areas like statistical physics and risk analysis. For example, moderate deviations are greatly involved in the computation of Bayes risk efficiency. See [44] for details.

Proposition 3.4.5. *Suppose S_N admits the order r Edgeworth expansion. Then for all $c \in (0, r)$, when $1 \leq x \leq \sqrt{c\sigma^2 \ln N}$,*

$$\lim_{N \to \infty} \frac{1 - \mathbb{P}\left(\frac{S_N - AN}{\sqrt{N}} \leq x\right)}{1 - \mathfrak{N}(x)} = 1. \tag{3.39}$$

Proof. Note that,

$$\begin{aligned} 1 - \mathfrak{N}(x) - \left[1 - \mathbb{P}\left(\frac{S_N - AN}{\sqrt{N}} \leq x\right)\right] &= \mathbb{P}\left(\frac{S_N - AN}{\sqrt{N}} \leq x\right) - \mathfrak{N}(x) \\ &= \sum_{p=1}^{r} \frac{P_p(x)}{N^{p/2}} \mathfrak{n}(x) + o\left(N^{-r/2}\right) \end{aligned}$$

uniformly in x. So it is enough to show that for $1 \leq x \leq \sqrt{c\sigma^2 \ln N}$,

$$\lim_{N \to \infty} \frac{P_p(x)\mathfrak{n}(x)}{N^{p/2}(1 - \mathfrak{N}(x))} = 0 \text{ and } \frac{N^{-r/2}}{1 - \mathfrak{N}(x)} = o(1)$$

Note that for $x \geq 1$,

$$1 - \mathfrak{N}(x) = \frac{\sigma^2 \mathfrak{n}(x)}{x} + \mathcal{O}\left(\frac{\mathfrak{n}(x)}{x^3}\right).$$

Thus,

$$\frac{N^{-r/2}}{1-\mathfrak{N}(x)} \leq \frac{N^{-r/2}}{1-\mathfrak{N}(\sqrt{c\sigma^2 \ln N})} = \mathcal{O}\Big(\sqrt{\ln N}\frac{N^{-r/2}}{e^{-\frac{c}{2}\ln N}}\Big)$$
$$= \mathcal{O}\Big(\frac{\ln N}{N^{(r-c)/2}}\Big)$$

Say $P_p(x)$ is of degree q. Then for some C and K,

$$\Big|\frac{P_p(x)\mathfrak{n}(x)}{N^{p/2}(1-\mathfrak{N}(x))}\Big| \leq C\frac{(x^q+K)\mathfrak{n}(x)}{N^{p/2}(1-\mathfrak{N}(x))} = C\frac{(x^q+K)}{N^{p/2}}x\Big(1+\mathcal{O}\Big(\frac{1}{x^2}\Big)\Big)$$
$$\leq C\frac{(\ln N)^{q+1}}{N^{p/2}} \to 0 \text{ as } N \to \infty.$$

This completes the proof of (3.39). □

Proposition 3.4.5 is a generalization of the results on moderate deviations found in [43] to the non-i.i.d. case along with improvements on the moment condition. It should be noted that [4] contains an improvement of the moment condition for the i.i.d. case. But the proof we present here is different from the proof presented in [4].

As an immediate corollary to the above theorem, we can state the following first order asymptotic for probability of moderate deviations.

Corollary 3.4.6. *Assume S_N admits the order r Edgeworth expansion. Then for all $c \in (0, r)$,*

$$\mathbb{P}(S_N \geq AN + \sqrt{c\sigma^2 N \ln N}) \sim \frac{1}{\sqrt{2\pi c}}\frac{1}{\sqrt{N^c \ln N}}.$$

3.5 Examples

Here we give several examples of systems satisfying assumptions (A1)–(A4).

3.5.1 Independent variables.

Let X_n be i.i.d. with $r+2$ moments. In this case we can take $\mathbb{B} = \mathbb{R}$, and define $\mathcal{L}_t v = \mathbb{E}(e^{itX_1}v) = \phi(t)v$ where ϕ is the characteristic function of X_1. Here we have taken $\ell = 1$. Put $v = 1$. Then, the independence of the random variables gives us, $\mathcal{L}_t^n 1 = \mathbb{E}(e^{itS_n}) = \phi(t)^n$. Also, the moment condition implies $t \to \phi(t)$ is C^{r+2}. This means (A1) is satisfied. (A2) is clear.

Suppose X_1 is $l-$Diophantine. That is there exists $C > 0$ and $t_0 > 0$ such that for all $|t| > t_0$, $|\phi(t)| < 1 - \dfrac{C}{|t|^l}$. Then $|\phi(t)| \leq e^{-\frac{C}{|t|^l}}$. So $|\phi(t)| < 1$ for all $t \neq 0$. So we have (A3). Also, this implies that X_1 is non-lattice. An easy computation shows that when $r_1 < \dfrac{1}{l}$, there exists r_2 such that $t_0 < |t| < n^{r_1} \implies |\phi(t)|^n \leq n^{-r_2}$. In fact, $|\phi(t)|^n \leq e^{-cn^\alpha}$ where $\alpha = 1 - r_1 l > 0$. So, (A4) is satisfied with $r_1 < \dfrac{1}{l}$.

When $l = 0$ we see that (A4) is satisfied with $r_1 > \dfrac{r-1}{2}$. Hence, by Theorem 3.1.1 order r Edgeworth expansion for S_n exists. This is exactly the classical result of Cramér because the condition: $\limsup\limits_{|t|\to\infty} |\phi(t)| < 1$ corresponds to $l = 0$.

Choose $q > \dfrac{r+1}{2r_1} > \dfrac{(r+1)l}{2}$. Then, by Theorem 3.1.4 and Theorem 3.1.5 we have that S_n admits weak global expansion for $f \in F_0^{q+2}$ and weak local expansion for $f \in F_{r+1}^{q+2}$. These are similar to the results appearing in [4] but slightly weaker because we require one more derivative: $q + 2 > 2 + \dfrac{(r+1)l}{2}$ as opposed to $1 + \dfrac{(r+1)l}{2}$. This is because we do not use the optimal conditions for the integrability of the Fourier transform. If we required $f \in F_r^{q+1}$ and $f^{(q+1)}$ to be $\alpha-$Hölder for small α, then the proof would still hold true and we could recover the results in [4].

3.5.2 Finite state Markov chains.

Here we present a non-trivial example for which the weak Edgeworth expansions exist but the strong expansion does not exist.

Consider the Markov chain x_n with states $S = \{1, \dots, d\}$ whose transition probability matrix $P = (p_{jk})_{d\times d}$ is positive. Then, by the Perron-Forbenius theorem, 1 is a simple eigenvalue of P and all other eigenvalues are strictly contained inside the unit disk. Suppose $\mathbf{h} = (h_{jk})_{d\times d} \in \mathrm{M}(d, \mathbb{R})$ and that there does not exist constants c, r and a $d-$vector H such that

$$rh_{jk} = c + H(k) - H(j) \mod 2\pi$$

for all j, k. Put $X_n = h_{x_n x_{n+1}}$.

For the family of operators $\mathcal{L}_t : \mathbb{C}^d \to \mathbb{C}^d$,

$$(\mathcal{L}_t f)_j = \sum_{k=1}^{d} e^{ith_{jk}} p_{jk} f_k,\ j = 1, \dots, d \tag{3.40}$$

$v = 1$ and $\ell = \mu_0$, the initial distribution, we have (3.1).

Define $b_{r,j,k} = h_{rj} + h_{jk}$ for all $j, r = 1, \dots, d$ and $k = 2, \dots, d$. Put $d(s) = \max \{(b_{r,j,k} - b_{r,1,k})s\}$ where $\{\ .\ \}$ denotes the fractional part. We further assume that $\mathbf{h}$ is $\beta-$Diophantine, that is, there exists $K \in \mathbb{R}$ such that for all $|s| > 1$,

$$d(s) \geq \frac{K}{|s|^\beta}. \tag{3.41}$$

If $\beta > \dfrac{1}{d^2(d-1)-1}$ then almost all $\mathbf{h}$ are $\beta-$Diophantine.

Because S_n can take at most $\mathcal{O}(n^{d^2-1})$ distinct values, S_n has a maximal jump of order at least $n^{-(d^2-1)}$. Therefore, the process $X^{\mathbf{h}}_n = h_{x_n x_{n-1}}$ does not admit the order $2(d^2-1)$ Edgeworth expansion.

The Perron-Forbenius theorem implies that the operator $\mathcal{L}_0$ satisfies (A2). Because (3.40) is a finite sum, it is clear that $t \mapsto \mathcal{L}_t$ is analytic on $\mathbb{R}$. So we also have (A1). Also the spectral radius of $\mathcal{L}_t$ is at most 1. Assume $\mathcal{L}_t$ has an eigenvalue on the unit circle, say $e^{i\lambda}$, with eigenvector f, then,

$$e^{i\lambda} f_j = (\mathcal{L}_t f)_j = \sum_{k=1}^{d} e^{ith_{jk}} p_{jk} f_k$$

Assuming $\max_j |f_j| = |f_r|$,

$$|f_r| = |e^{i\lambda} f_r| = \left| \sum_{k=1}^{d} e^{ith_{jk}} p_{jk} f_k \right| \leq \sum_{k=1}^{d} p_{jk} |f_k| \implies \sum_{k=1}^{d} p_{jk}(|f_k| - |f_r|) \geq 0$$

Because $|f_k| - |f_r| \leq 0$ for all k and $p_{jk} \geq 0$ for all j and k we have $|f_k| = |f_r|$ for all k. Therefore, there exist a $d-$vector H such that $f_k = Re^{iH(k)}$ for all k. Then,

$$\begin{aligned} e^{i\lambda} Re^{iH(j)} &= \sum_{k=1}^{d} e^{ith_{jk}} p_{jk} Re^{iH(k)} \\ 0 &= \sum_{k=1}^{d} p_{jk} (e^{i(th_{jk} + H(k) - H(j) - \lambda)} - 1) \\ \implies th_{jk} &= \lambda + H(j) - H(k) \mod 2\pi \end{aligned}$$

But this is a contradiction. Therefore, (A3) holds. Next we notice that,

$$\begin{aligned} |(\mathcal{L}_t^2 f)_r| = \left| \sum_{j=1}^{d} \sum_{k=1}^{d} e^{it(h_{rj} + h_{jk})} p_{rj} p_{jk} f_k \right| &= \left| \sum_{k=1}^{d} \left(\sum_{j=1}^{d} e^{it(h_{rj} + h_{jk})} p_{rj} p_{jk} \right) f_k \right| \\ &\leq \|f\| \left(\sum_{k=1}^{d} \left| \sum_{j=1}^{d} e^{itb_{r,j,k}} p_{rj} p_{jk} \right| \right) \end{aligned} \tag{3.42}$$

Now we estimate $|b_{r,k}(t)|$ where

$$b_{r,k}(t) = \sum_{j=1}^{d} e^{itb_{r,j,k}} p_{rj} p_{jk} = e^{itb_{r,1,k}} \sum_{j=1}^{d} e^{it(b_{r,j,k} - b_{r,1,k})} p_{rj} p_{jk}$$

Then we have,

$$
\begin{aligned}
|b_{r,k}(t)|^2 &= \sum_{j=1}^{d} p_{rj}^2 p_{jk}^2 + 2\sum_{j>l}^{d} p_{rj}p_{jk}p_{rl}p_{lk}\cos((b_{r,j,k}-b_{r,l,k})t) \\
&= \left(\sum_{j=1}^{d} p_{rj}p_{jk}\right)^2 - 2\sum_{j>l}^{d} p_{rj}p_{jk}p_{rl}p_{lk}[1-\cos((b_{r,j,k}-b_{r,l,k})t)] \\
&= \left(\sum_{j=1}^{d} p_{rj}p_{jk}\right)^2 - 2Cd(t)^2 + \mathcal{O}(d(t)^3),\ C>0 \\
|b_{r,k}(t)| &= \sum_{j=1}^{d} p_{rj}p_{jk} - \tilde{C}d(t)^2 + \mathcal{O}(d(t)^3),\ \tilde{C}>0
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\sum_{k=1}^{d}\left|\sum_{j=1}^{d} e^{itb_{r,j,k}} p_{rj}p_{jk}\right| &= \sum_{k=1}^{d}\left(\sum_{j=1}^{d} p_{rj}p_{jk}\right) - \overline{C}d(t)^2 + \mathcal{O}(d(t)^3) \\
&= 1 - \overline{C}d(t)^2 + \mathcal{O}(d(t)^3),\ \overline{C}>0
\end{aligned}
$$

From the Diophantine condition (3.41), we can conclude that there exists $\theta > 0$ such that for all $|t| > 1$,

$$
\|\mathcal{L}_t^2\| \le 1-\theta d(t)^2 \implies \|\mathcal{L}_t^N\| \le \left(1-\theta d(t)^2\right)^{\lceil N/2 \rceil} \le e^{-\theta d(t)^2 N/2} \le e^{-\theta t^{-2\beta} N/2}
$$

When $1 < |t| < N^{\frac{1-\epsilon}{2\beta}}$, we have, $\|\mathcal{L}_t^N\| \le e^{-\theta N^{\epsilon}/2}$ which gives us (A4) with $r_1 = \dfrac{1-\epsilon}{2\beta}$ where $\epsilon > 0$ can be made as small as required. Because for small ϵ, $\lceil \dfrac{r+1}{2(1-\epsilon)} \rceil = \lceil \dfrac{r+1}{2} \rceil$, choosing $q > \dfrac{r+1}{2}\beta$, we conclude that for $f \in F_0^{q+2}$ weak global and for $f \in F_{r+1}^{q+2}$ weak local Edgeworth expansions of order r for the process $X_n^{\mathbf{h}}$ exist. Also, S_N admits averaged Edgeworth expansions of order r for $f \in F_0^2$. In the special case of $\beta > \dfrac{1}{d^2(d-1)-1}$, these hold for a full measure set of $\mathbf{h}$ even though the order r strong expansion does not exist for $r+1 \ge d^2$.

3.5.3 More general Markov chains.

3.5.3.1 Chains with smooth transition density.

First we consider the case where x_n is a time homogeneous Markov process on a compact connected manifold $\mathcal{M}$ with smooth transition density $p(x, y)$ which is bounded away from 0, and $X_n = h(x_{n-1}, x_n)$ for a piece-wise smooth function $h : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$. We assume that $h(x, y)$ can not be written in the form

$$h(x, y) = H(y) - H(x) + c(x, y) \tag{3.43}$$

where $c(x, y)$ is piece-wise constant.

In particular, there is no constant c and a function H such that $h(x, y) = H(y) - H(x) + c$. Also, the transition probability $P(x, dy)$ of X_n has a non-degenrate absolute continuous component. Then, by [25], the CLT holds with $\sigma^2 > 0$.

To check the assumption 3.43 we need the following:

Lemma 3.5.1. *(3.43) holds iff there exists $o \in \mathcal{M}$ such that the function $x \mapsto h(o, x) + h(x, y)$ is piece-wise constant for each y.*

Proof. If (3.43) holds then for each $o \in \mathcal{M}$

$$h(o, x) + h(x, y) = c(o, x) + c(x, y) + H(y) - H(o)$$

where $c(o, x) + c(x, y)$ is piece-wise constant in x for each y.

Conversely, suppose for some $o \in \mathcal{M}$, $x \mapsto h(o, x) + h(x, y)$ is piece-wise constant for each y. Fix y. Let $c = h(o, o)$ and $H(x) = h(o, x) - h(o, o)$. Then, $h(o, o) + h(o, y)$ and $h(o, x) + h(x, y)$ differ by a piece-wise constant function. Then

(3.43) holds because $h(o,x)+h(x,y)-(h(o,o)+h(o,y)) = h(x,y)+H(x)-H(y)-c$ is piecewise constant. $\square$

Let $\mathbb{B} = L^\infty(\mathcal{M})$ and consider the family of integral operators,

$$(\mathcal{L}_t u)(x) = \int p(x,y)e^{ith(x,y)}u(y)\,dy.$$

Let μ be the initial distribution of the Markov chain and $\{\mathcal{F}_n\}$ be the filtration adapted to the processes. Then, using the Markov property,

$$\mathbb{E}_\mu[e^{itS_n}] = \mathbb{E}_\mu[e^{itS_{n-1}}\mathcal{L}_t 1].$$

By induction we can conclude

$$\mathbb{E}_\mu(e^{itS_n}) = \int \mathcal{L}_t^n 1\,d\mu$$

Because h is bounded, expanding $e^{ith(x,y)}$ as a power series in t, we see that $t \mapsto \mathcal{L}_t$ is analytic for all t. This shows that (A1) is statisfied.

From the Weierstrass theorem there exist functions q_k, r_k on $\mathcal{M}$ such that $p(x,y)$ is a uniform limit of functions of the form $\sum_{k=1}^{n} q_k(x)r_k(y)$. Therefore, $\mathcal{L}_t$ is a uniform limit of finite rank operators and is compact. Compact operators have a point spectrum hence the essential spectral radius of $\mathcal{L}_t$ vanishes. It is also immediate that $\|\mathcal{L}_t\| \le 1$ for all t. Hence the spectrum is contained in the closed unit disk.

In addition, $\mathcal{L}_0 : L^\infty(\mathcal{M}) \to L^\infty(\mathcal{M})$ given by

$$(\mathcal{L}_0 u)(x) = \int p(x,y)u(y)\,dy$$

is a positive operator. Note that $(\mathcal{L}_0 1)(x) = 1$ for all x. Thus, 1 is an eigenvalue of $\mathcal{L}_0$ with eigenfunction 1. Also, eigenvalue 1 is simple and all other eigenvalues

β are such that $|\beta| < 1$. This follows from a direct application of Birkhoff Theory (see [2]). Thus, we have (A2).

Next we show that if $\beta \in \mathrm{sp}(\mathcal{L}_t)$, $t \neq 0$ then $|\beta| < 1$. If not, then there exists λ and $u \in L^\infty(\mathcal{M})$ such that

$$\int p(x,y)e^{ith(x,y)}u(y)\,dy = e^{i\lambda}u(x)$$

Suppose $\sup_x |u(x)| = R$ then for each $\epsilon > 0$ there exists x_ϵ such that

$$R - \epsilon \leq |u(x_\epsilon)| = |e^{i\lambda}u(x_\epsilon)| = \left|\int p(x,y)e^{ith(x,y)}u(y)\,dy\right| \leq \int p(x,y)|u(y)|\,dy$$

Therefore,

$$\int p(x,y)[|u(y)| - R]\,dy \geq -\epsilon,$$

But $|u(y)| - R \leq 0$. Hence, $|u(y)| = R$ a.e. Therefore, $u(y) = Re^{i\theta(y)}$ a.e. for some function θ and we may assume $\theta \in [0, 2\pi)$.

$$\begin{aligned} &\int p(x,y)e^{ith(x,y)}Re^{i\theta(y)}\,dy = Re^{i\lambda}e^{i\theta(x)} \\ \Longrightarrow &\int p(x,y)[e^{i(th(x,y)-\lambda+\theta(y)-\theta(x))} - 1]\,dy = 0 \\ \Longrightarrow\ & th(x,y) - \lambda + \theta(y) - \theta(x) \equiv 0 \mod 2\pi \end{aligned} \tag{3.44}$$

Fix y and t. Then, for all z, $x \mapsto h(y,x) + h(x,z)$ does not depend on x modulo 2π i.e. it is piece-wise constant for all $t \neq 0$. By Lemma 3.5.1, $h(x,y)$ satisfies (3.43). This contradiction proves (A3).

Recall that if $\mathcal{K}$ is integral operator

$$(\mathcal{K}u)(x) = \int k(x,y)u(y)dy$$

then

$$\|\mathcal{K}\| = \sup_x \int |k(x,y)|dy.$$

In our case $\mathcal{L}_t^2$ has the kernel,

$$\mathfrak{l}_t(x,y) = \int e^{it[h(x,z)+h(z,y)]}p(x,z)p(z,y)dz.$$

By Lemma 3.5.1 for each x and y the function $z \mapsto (h(x,z)+h(z,y))$ is not piecewise constant. So its derivative (whenever it exists) is not identically 0. Thus there is an open set $V_{x,y}$ and a vector field e such that $\partial_e[h(x,z)+h(z,y)] \neq 0$ on $V_{x,y}$. Integrating by parts in the direction of e we conclude that

$$\lim_{t\to\infty} \int_{V_{x,y}} e^{it[h(x,z)+h(z,y)]}p(x,z)p(z,y)dz = 0.$$

By compactness there are constants r_0, ε_0 such that for $|t| \geq r_0$ and all x and y in $\mathcal{M}$, $|\mathfrak{l}_t(x,y)| \leq \mathfrak{l}_0(x,y) - \varepsilon_0$. It follows that

$$\|\mathcal{L}_t^2\| = \sup_x \int_{\mathcal{M}} |\mathfrak{l}_t(x,y)|dy \leq \int_{\mathcal{M}} \mathfrak{l}_0(x,y)dy \ - \varepsilon_0. \tag{3.45}$$

The first term here equals

$$\iint_{\mathcal{M}\times\mathcal{M}} p(x,z)p(z,y)dzdy = 1.$$

Hence for $|t| \geq r_0$, $\|\mathcal{L}_t^2\| \leq 1 - \varepsilon_0$ and so $||\mathcal{L}_t^N|| \leq (1-\varepsilon_0)^{\lceil N/2 \rceil}$. This proves (A4) with no restriction on r_1. Therefore, S_N admits Edgeworth expansions of all orders.

Next we look at the case when (3.43) fails but the constants are not lattice valued. Then, arguments for (A1), (A2) and (A3) hold. In particular, (3.44) cannot hold since it implies that

$$\left(h(x,y) + \frac{\theta(y)}{t} - \frac{\theta(x)}{t}\right) \in \frac{\lambda}{t} + \frac{2\pi}{t}\mathbb{Z}$$

However, we have to impose a Diophantine condition on the values that $h(x, y)$ can take in order to obtain a sufficient control over $\|\mathcal{L}_t^N\|$ and obtain (A4).

For fixed x, y let the range of $z \mapsto h(x, z) + h(z, y)$ be $S = \{c_1, \ldots, c_d\}$. Note that these c_i's may depend on x and y. However, there can be at most finitely many values that $h(x, z) + h(z, y)$ can take as x and y vary on $\mathcal{M}$ because h is piece-wise smooth. So we might as well assume that S is this complete set of values. Also, take U_k to be the open set on which $z \mapsto h(x, z) + h(z, y)$ takes value c_k. Take $b_k = c_k - c_1$ and define $d(s) = \max\,\{b_k s\}$. Assume further that there exists $K > 0$ such that for all $|s| > 1$,

$$d(s) \geq \frac{K}{|s|^\beta}$$

If $\beta > (d-1)^{-1}$ for almost all $d-$tuples $\mathbf{c} = (c_1, \ldots, c_d)$, the above holds.

Note that,

$$\begin{aligned}
|\mathcal{L}_t^2 u(x)| &= \int \left| \int e^{it[h(x,z)+h(z,y)]} p(x,z)p(z,y)\, dz \right| |u(y)|\, dy \\
&\leq \|u\| \int \left| \sum_{k=1}^{d} e^{itc_k} \int_{U_k} p(x,z)p(z,y)\, dz \right| dy = \|u\| \int \left| \sum_{k=1}^{d} p_k e^{itb_k} \right| dy
\end{aligned}$$

where and $p_k = \displaystyle\int_{U_k} p(x,z)p(z,y)\, dz$. Therefore, $p_1 + \cdots + p_d = p(x, y)$.

Now the situation is similar to that of (3.42) and a similar calculation yields,

$$\left| \sum_{k=1}^{d} p_k e^{itb_k} \right| = p(x,y) - Cd(t)^2 + \mathcal{O}(d(t)^3),\ C > 0$$

Therefore,

$$\|\mathcal{L}_t^2\| \leq \int \Big[p(x,y) - Cd(t)^2 + \mathcal{O}(d(t)^3)\Big] dy = 1 - \tilde{C}d(s)^2$$

From this we can repeat the analysis done in the finite state Markov chains example following (3.42). In particular, when $1 < |t| < N^{\frac{1-\epsilon}{2\beta}}$, there exists $\theta > 0$ such that $\|\mathcal{L}_t^N\| \leq e^{-\theta N^\epsilon}$ which gives us (A4).

Finally, when (3.43) fails and h takes integer values with span 1, X_n is a lattice random variable and we can discuss the existence of the lattice Edgeworth expansion. In this case S_N admits the lattice expansion of all orders. To this end, only the condition $\widetilde{\text{(A3)}}$ needs to be checked. First note that $\mathcal{L}_0 = \mathcal{L}_{2\pi k}$ for all $k \in \mathbb{Z}$. Also, assuming $\mathcal{L}_t$ has an eigenvalue on the unit circle, we conclude (3.44),

$$th(x,y) - \lambda + \theta(y) - \theta(x) \equiv 0 \mod 2\pi$$

This implies $t(h(x,y) + h(y,x)) \in 2\pi\mathbb{Z} + 2\lambda$. Note that LHS belongs a lattice with span t and RHS is a lattice with span 2π. Because t is not a multiple of 2π this equality cannot happen. Therefore, when $t \notin 2\pi\mathbb{Z}$, $\text{sp}(\mathcal{L}_t) \subset \{|z| < 1\}$ and we have the claim.

3.5.3.2 Chains without densities.

We consider a more general case where transition probabilities may not have a density. We claim we can recover (A1)–(A4) if the transition operator takes the form

$$\mathcal{L}_0 = a\mathcal{J}_0 + (1-a)\mathcal{K}_0$$

where $a \in (0,1)$ and $\mathcal{J}_0$ and $\mathcal{K}_0$ are Markov operators on $L^\infty(\mathcal{M})$ (i.e. $\mathcal{J}_0 f \geq 0$ if $f \geq 0$ and $\mathcal{J}_0 1 = 1$ and similarly for $\mathcal{K}_0$),

$$\mathcal{J}_0 f(x) = \int p(x,y) f(y)\, d\mu(y)$$

and

$$\mathcal{K}_0 f(x) = \int f(y) Q(x, dy)$$

where p is a smooth transition density and Q is a transition probability measure. Let $h(x, y)$ be piece-wise smooth and put,

$$\mathcal{J}_t(f) = \mathcal{J}_0(e^{ith} f) \quad \text{and} \quad \mathcal{K}_t(f) = \mathcal{K}_0(e^{ith} f).$$

Defining $\mathcal{L}_t = a\mathcal{J}_t + (1-a)\mathcal{K}_t$ we can conclude $t \mapsto \mathcal{L}_t$ is analytic and that

$$\mathbb{E}_\mu(e^{itS_n}) = \int \mathcal{L}_t^n 1 \, d\mu.$$

Now we show that conditions (A2), (A3) and (A4) are satisfied. Because $\|\mathcal{J}_t\| \leq 1$ and $\|\mathcal{K}_t\| \leq 1$ we have $\|\mathcal{L}_t\| \leq 1$. Thus the spectral radius of $\mathcal{L}_t$ is ≤ 1. Because $a\mathcal{J}_t$ is compact, $\mathcal{L}_t$ and $(1-a)\mathcal{K}_t$ have the same essential spectrum. See [33, Theorem IV.5.35]. However the spectral radius of the latter is at most $(1-a)$. Hence, the essential spectral radius of $\mathcal{L}_t$ is at most $(1-a)$.

Because both $\mathcal{J}_0$ and $\mathcal{K}_0$ are Markov operators we can conclude that 1 is an eigenvalue of $\mathcal{L}_0$ with constant function 1 as the corresponding eigenfunction. From the previous paragraph the essential spectral radius of $\mathcal{L}_0$ is at most $(1-a)$. Because $\mathcal{L}^n$ is norm bounded it cannot have Jordan blocks. So 1 is semisimple.

Suppose, $\mathcal{L}_t u = e^{i\theta} u$. Without loss of generality we may assume $\|u\|_\infty = 1$. Assuming there exists a positive measure set Ω with $|u(x)| < 1 - \delta$ we can conclude that, for all x,

$$\begin{aligned}
|u(x)| = |L_t u(x)| &= |a\mathcal{J}_t u(x) + (1-a)\mathcal{K}_t u(x)| \\
&\leq a \int_\Omega |u(y)| p(x, y) d\mu(y) + a \int_{\Omega^c} |u(y)| p(x, y) d\mu(y) + (1-a)
\end{aligned}$$

$$\leq 1 - a\delta\mu(\Omega).$$

This is a contradiction. Therefore, $|u(x)| = 1$. Put $u(x) = e^{i\gamma(x)}$. Then,

$$1 = a \int e^{i(th(x,y)+\gamma(y)-\gamma(x)-\theta)} p(x,y) d\mu(y) + (1-a)e^{-i(\theta+\gamma(x))}\mathcal{K}_t u$$

Hence, $\int e^{i(th(x,y)+\gamma(y)-\gamma(x)-\theta)} p(x,y) d\mu(y) = 1 \implies \mathcal{J}_t u = e^{i\theta} u$. From section 3.5.3.1, this can only be true when $t = 0$ and in this case $\theta = 0$ and $u \equiv 1$. This concludes that $\mathcal{L}_t$, $t \neq 0$ has no eigenvalues on the unit disk and the only eigenvalue of $\mathcal{L}_0$ on the unit disk is 1 and its geometric multiplicity is 1. As 1 is semisimple, it is simple as required. This concludes proof of (A2) and (A3).

From the previous case, there exists $r > 0$ and $\epsilon \in (0,1)$ such that such that for all $|t| > r$ we have $\|\mathcal{J}_t^2\| \leq 1 - \epsilon$. From this we have,

$$\|\mathcal{L}_t^2\| = \|a^2\mathcal{J}_t^2 + a(1-a)\mathcal{J}_t\mathcal{K}_t + (1-a)a\mathcal{K}_t\mathcal{J}_t + (1-a)^2\mathcal{K}_t^2\| \leq 1 - a^2\epsilon.$$

Hence, for all $|t| > r$, for all N, $\|\mathcal{L}_t^N\| \leq (1 - a^2\epsilon)^{\lfloor N/2 \rfloor}$ which gives us (A4) with no restrictions on r_1. Therefore, S_N admits Edgeworth expansions of all orders as before.

As in the previous section, an analysis can be carried out when (3.43) fails. The conclusions are exactly the same.

3.5.4 One dimensional piecewise expanding maps.

Here we check assumptions (3.1), (A1)–(A4) for piecewise expanding maps of the interval using the results of [5, 37].

Let $f : [0,1] \to [0,1]$ be such that there is a finite partition $\mathcal{A}_0$ of $[0,1]$ (except possibly a measure 0 set) into open intervals such that for all $I \in \mathcal{A}_0$, $f|_I$ extends to a C^2 map on an interval containing $\overline{I}$. In other words f is a piece-wise C^2 map. Further, assume that $f' \geq \lambda > 1$ i.e. f is uniformly expanding. Next, let $\mathcal{A}_n = \bigvee_{k=0}^{n} T^{-j}\mathcal{A}_0$ and suppose for each n there is N_n such that for all $I \in \mathcal{A}_n$, $f^{N_n} I = [0,1]$. Such maps are called *covering*.

Statistical properties of piece-wise C^2 covering expanding maps of an interval, are well-understood. For example, see [37]. In particular, such a function f has a unique absolutely continuous invariant measure with a strictly positive density $h \in \mathrm{BV}[0,1]$ and the associated transfer operator

$$\mathcal{L}_0\varphi(x) = \sum_{y \in f^{-1}(x)} \frac{\varphi(y)}{f'(y)}$$

has a spectral gap.

Let g be C^2 except possibly at finite number of points and admitting a C^2 extension on each interval of smoothness. Define $X_n = g \circ f^n$ and consider it as a random variable with x distributed according to some measure $\rho(x)dx$, $\rho \in \mathrm{BV}[0,1]$.

Define a family of operators $\mathcal{L}_t : \mathrm{BV}[0,1] \to \mathrm{BV}[0,1]$ by

$$\mathcal{L}_t\varphi(x) = \sum_{y \in f^{-1}(x)} \frac{e^{itg(y)}}{f'(y)}\varphi(y)$$

where $t = 0$ corresponds to the transfer operator. Because g is bounded, writing $e^{itg(y)}$ as a power series we can conclude $t \to \mathcal{L}_t$ is analytic for all t. This gives (A1).

(A2) follows from the fact that $\mathcal{L}_0$ has a spectral gap. We further assume that

$$g \text{ is not cohomologous to a piece-wise constant function.} \tag{3.46}$$

In particular, g is not a BV coboundary.

The assumption (3.46) is reasonable. Indeed, suppose that g is piece-wise constant taking values $c_1, c_2 \dots c_k$. Then S_n takes less than n^{k-1} distinct values so the maximal jump is of order at least $n^{-(k-1)}$ so S_n can not admit Edgeworth expansion of order $(2k-2)$ in contrast to the case where (3.46) holds as we shall see below.

A direct computation gives,

$$\mathbb{E}(e^{itS_n/\sqrt{n}}) = \int_0^1 \mathcal{L}^n_{t/\sqrt{n}} \rho(x)\, dx.$$

Therefore, there exists A such that,

$$\lim_{n\to\infty} \mathbb{E}(e^{it\frac{S_n - nA}{\sqrt{n}}}) = e^{-t^2\sigma^2/2} \tag{3.47}$$

where $\sigma^2 \geq 0$. It is well know that $\sigma^2 > 0 \iff g$ is a BV coboundary (see [24]). From (3.47) it is clear that S_n satisfies the CLT.

To show (A3) holds, we first normalize the family of operators,

$$\overline{\mathcal{L}}_t v(x) = \sum_{f(y)=x} \frac{e^{itg(y)} h(y)}{f'(y) h\circ f(y)} v(y)$$

Then, $\overline{\mathcal{L}}_t = H^{-1} \circ \mathcal{L}_t \circ H$ where H is multiplication by the function h. Therefore, $\mathcal{L}_t$ and $\overline{\mathcal{L}}_t$ have the same spectrum. However, the eigenfunction corresponding to the eigenvalue 1 of $\overline{\mathcal{L}}_0$ changes to the constant function 1.

Assume $e^{i\theta}$ is an eigenvalue of $\overline{\mathcal{L}}_t$. Then, there exists $u \in \mathrm{BV}[0,1]$ with $\overline{\mathcal{L}}_t u(x) = e^{i\theta}u(x)$. Observe that,

$$\begin{aligned}\overline{\mathcal{L}}_0|u|(x) &= \sum_{f(y)=x} \frac{|u(y)|h(y)}{f'(y)h\circ f(y)} \\ &\geq \left| \sum_{f(y)=x} \frac{e^{itg(y)}u(y)h(y)}{f'(y)h\circ f(y)} \right| = |\overline{\mathcal{L}}_t u(x)| = |e^{i\theta}u(x)| = |u(x)|\end{aligned}$$

Also note that, $\overline{\mathcal{L}}_0$ is a positive operator. Hence, $\overline{\mathcal{L}}_0^n|u|(x) \geq |u(x)|$ for all n. However,

$$\lim_{n\to\infty}(\overline{\mathcal{L}}_0^n|u|)(x) = \int |u(y)|\cdot 1\, dy$$

because 1 is the eigenfunction corresponding to the top eigenvalue. So for all x,

$$\int |u(y)|\, dy \geq |u(x)|$$

This implies that $|u(x)|$ is constant. WLOG $|u(x)| \equiv 1$. So we can write $u(x) = e^{i\gamma(x)}$. Then,

$$\begin{aligned}\overline{\mathcal{L}}_t u(x) &= \sum_{f(y)=x} \frac{h(y)}{f'(y)h\circ f(y)} e^{i(tg(y)+\gamma(y))} = e^{i(\theta+\gamma(x))} \\ \implies &\sum_{f(y)=x} \frac{h(y)}{f'(y)h\circ f(y)} e^{i(tg(y)+\gamma(y)-\gamma(f(y))-\theta)} = 1\end{aligned}$$

for all x. Since,

$$\overline{\mathcal{L}}_0 1 = \sum_{f(y)=x} \frac{h(y)}{f'(y)h\circ f(y)} = 1$$

and $e^{i(tg(y)+\gamma(y)-\gamma(x)-\theta)}$ are unit vectors, it follows that

$$tg(y) + \gamma(y) - \gamma(f(y)) - \theta = 0 \mod 2\pi \tag{3.48}$$

for all y. Because g is not cohomologous to a piecewise constant function we have a contradiction. Therefore, $\overline{\mathcal{L}}_t$ and hence $\mathcal{L}_t$ does not have an eigenvalue on the unit circle when $t \neq 0$.

To complete the proof of (A3) one has to show that the spectral radius of $\mathcal{L}_t$ is at most 1 and that the essential spectral radius of $\mathcal{L}_t$ is strictly less than 1. This is clear from Lasota-Yorke type inequality in [5, Lemma 1]. In fact, there is a uniform $\kappa \in (0,1)$ such that $r_{ess}(\mathcal{L}_t) \leq \kappa$ for all t.

Next, we describe in detail how the estimate in [5, Proposition 1] gives us (A4). To make the notation easier we assume $t > 0$ and we replace $|t|$ by t. [5, Proposition 1] implies that there exist c and C such that if K_1 large enough (we fix one such K_1) then for all $t > K_1$,

$$\|\mathcal{L}_t^{\lceil c\ln t\rceil} u\|_t \leq e^{-C\lceil c\ln t\rceil}\|u\|_t \tag{3.49}$$

where $\|h\|_t = (1+t)^{-1}\|h\|_{\mathrm{BV}} + \|h\|_{\mathrm{L}^1}$. Therefore,

$$\|\mathcal{L}_t^{k\lceil c\ln t\rceil} u\|_t \leq e^{-C\lceil c\ln t\rceil}\|\mathcal{L}^{(k-1)\lceil c\ln t\rceil} u\|_t \leq \cdots \leq e^{-Ck\lceil c\ln t\rceil}\|u\|_t$$

Also, $\|\mathcal{L}_t\|_t \leq 1$. So, if $n = k\lceil c\ln t\rceil + r$ where $0 \leq r < \lceil c\ln t\rceil$ then

$$\|\mathcal{L}_t^n u\|_t \leq e^{-Ck\lceil c\ln t\rceil}\|\mathcal{L}_t^r u\|_t \leq e^{-Cn\frac{k\lceil c\ln t\rceil}{k\lceil c\ln t\rceil + r}}\|u\|_t \leq e^{-Cn\frac{k}{k+1}}\|u\|_t$$

However,

$$(1+t)^{-1}\|h\|_{\mathrm{BV}} \leq \|h\|_t \leq [1+(1+t)^{-1}]\|h\|_{\mathrm{BV}}$$

Therefore,

$$(1+t)^{-1}\|\mathcal{L}_t^n u\|_{\mathrm{BV}} \leq [1+(1+t)^{-1}]e^{-Cn\frac{k}{k+1}}\|u\|_{\mathrm{BV}}$$

which gives us

$$\|\mathcal{L}_t^n\|_{\mathrm{BV}} \leq (t+2)e^{-Cn\frac{k}{k+1}}$$

and here $k = k(n,t) = \lfloor \frac{n}{\lceil c \ln t \rceil} \rfloor$. When $K_1 \leq |t| \leq n^{r_1}$, $k_{\min} = \lfloor \frac{n}{\lceil c \ln n^{r_1} \rceil} \rfloor$ and $\frac{k_{\min}}{k_{\min}+1} \to 1$ as $n \to \infty$. Also, $1 \geq \frac{k}{k+1} \geq \frac{k_{\min}}{k_{\min}+1}$ and,

$$\|\mathcal{L}_t^n\|_{\mathrm{BV}} \leq (t+2)e^{-Cn\frac{k(n,t)}{k(n,t)+1}} \leq 2n^{r_1}e^{-Cn\frac{k_{\min}}{k_{\min}+1}}$$

Choosing n_0 such that for all $n > n_0$, $\frac{k_{\min}}{k_{\min}+1} > \frac{1}{2}$ (so this choice of n_0 works for all t) we can conclude that,

$$\|\mathcal{L}_t^n\|_{\mathrm{BV}} \leq 2n^{r_1}e^{-Cn/2}$$

This proves (A4) for all choices of r_1. In particular given r, we can choose $r_1 > \frac{r-1}{2}$ in the above proof. This implies that Edgeworth expansions of all orders exist.

3.5.5 Multidimensional expanding maps.

Let $\mathcal{M}$ be a compact Riemannian manifold and $f : \mathcal{M} \to \mathcal{M}$ be a C^2 expanding map. Let $g : \mathcal{M} \to \mathbb{R}$ be a C^2 function which is non homologous to constant. The proof of Lemma 3.13 in [13] shows that this condition is equivalent to g not being *infinitesimally integrable* in the following sense. The natural extension of f acts on the space of pairs $(\{y_n\}_{n\in\mathbb{N}}, x)$ where $f(y_{n+1}) = y_n$ for $n > 0$ and $fy_1 = x$. Given such pair let

$$\Gamma(\{y_n\}, x) = \lim_{n\to\infty} \frac{\partial}{\partial x}\left[\sum_{k=0}^{n-1} g(f^k y_n)\right] = \lim_{n\to\infty} \frac{\partial}{\partial x}\left[\sum_{k=1}^{n} g(y_k)\right] = \sum_{k=1}^{\infty} \frac{\partial}{\partial x} g(y_k).$$

g is called infinitesimally integrable if $\Gamma(\{y_n\}, x)$ actually depends only on x but not on $\{y_n\}$.

Let $X_n = g \circ f^n$. We want to verify (A1)–(A4) when x is distributed according to a smooth density ρ. Note that assumption (3.1) holds with $v = \rho$, ℓ being the

Lebesgue measure and

$$(\mathcal{L}_t\phi)(x) = \sum_{y\in f^{-1}(x)} \frac{e^{itg(y)}}{\left|\det\left(\frac{\partial f}{\partial x}\right)\right|}\phi(y).$$

We will check (A1)–(A4) for $\mathcal{L}_t$ acting on $C^1(\mathcal{M})$. The proof of (A1)–(A3) is the same as in section 3.5.4. In particular, for (A3) we need Lasota–Yorke inequality (see (3.52) below) which is proven in [13, equation (19)].

The proof of (A4) is also similar to section 3.5.4, so we just explain the differences. As before we assume that $t > 0$. Given a small constant κ let

$$\|\phi\|_t = \max\left(\|\phi\|_{C^0}, \frac{\kappa\|D\phi\|_{C^0}}{1+t}\right).$$

Then by [13, Proposition 3.16]

$$\|\mathcal{L}_t^n\phi\|_t \le \|\phi\|_t \tag{3.50}$$

provided that $n \ge C_1 \ln t$.

By [13, Lemma 3.18] if g is not infinitesimally integrable then there exists a constant $\eta < 1$ such that

$$\|\mathcal{L}_t^n\phi\|_{L^1} \le \eta^n\|\phi\|_t. \tag{3.51}$$

The Lasota–Yorke inequality says that there is a constant $\theta < 1$, such that

$$\|D\left(\mathcal{L}_t^n\phi\right)\|_{C^0} \le C_3\left(t\|\phi\|_{C^0} + \theta^n\|D\phi\|_{C^0}\right) \tag{3.52}$$

Also,

$$\|\mathcal{L}_t^n\phi\|_{C^0} \le \|\mathcal{L}_0^n(|\phi|)\|_{C^0} \le C_4\left(\|\,|\phi|\,\|_{L^1} + \theta^n\|\,|\phi|\,\|_{\mathrm{Lip}}\right) \tag{3.53}$$

where the last step relies on $\mathcal{L}_0$ having a spectral gap on the space of Lipshitz functions. Combing (3.50) through (3.53), we conclude that $\mathcal{L}_t$ satisfies (3.49). The rest of the argument is the same as in section 3.5.4.

Chapter 4: Large Deviation Principles.

4.1 Asymptotics for Cramér's Theorem.

In this section, we focus on sequences of i.i.d. random variables. First, we prove the existence of weak asymptic expansions for Cramér's LDP – Theorem 1.2. Next, we deduce existence of the strong expansion in special cases. As expected, a stronger assumption on the regularity of the law of the random variables is required for the second step.

4.1.1 Weak asymptotic expansions.

We recall that a random variable X is called $l-$Diophantine if there exist positive constants t_0 and C such that $|\mathbb{E}(e^{itX})| < 1 - \dfrac{C}{|t|^l}$ for $|t| > t_0$. It is known that when X is $l-$Diophantine and $r+2$ moments exist weak Edgeworth expansions exist. For example, see [4] and Section 3.5.1.

Given a random variable X with distribution function F, we define $Y_{X,\gamma}$ to be a random variable with distribution function G^γ given by,

$$dG^\gamma(y) = \frac{e^{y\gamma} dF(y)}{\mu(\gamma)} \tag{4.1}$$

where $\mu(\gamma) = \displaystyle\int e^{y\gamma} dF(y)$. Therefore,

$$\mathbb{E}[Y_{X,\gamma}] = \frac{\int y e^{y\gamma} dF(y)}{\int e^{y\gamma} dF(y)}. \tag{4.2}$$

In Section 3.1 we defined the function spaces F_k^m: $f \in F_k^m$ if f is m times continuously differentiable and $C_k^m(f) = \left(\max_{0\le j\le m} \|f^{(j)}\|_{\mathrm{L}^1} + \max_{0\le j\le k} \|x^j f\|_{\mathrm{L}^1}\right) < \infty$. We call a function f, (left) exponential of order α, if $\lim_{x\to-\infty} |e^{-\alpha x} f(x)| = 0$. Denote by $F_{m,\alpha}^k$ the collection of all $f \in F_m^k$ with $f^{(k)}$ is exponential of order α.

We note that due to assumption $f \in F_m^k$, $f^{(k)}$ being exponential of order α is enough to guarantee that $f^{(l)}$ is exponential of order α for all $0 \le l \le k$. To see this suppose $f, f' \in L^1$. Then, $\lim_{|x|\to\infty} f(x) = 0$. Suppose f' is exponential of order α. Then, given $\epsilon > 0$ there is $M > 0$ such that for $x < -M$, $-\epsilon e^{\alpha x} < f'(x) < \epsilon e^{\alpha x}$. So, $-\epsilon \displaystyle\int_{-\infty}^{x} e^{\alpha y}\, dy \le \int_{-\infty}^{x} f'(y)\, dy \le \epsilon \int_{-\infty}^{x} e^{\alpha y}\, dy \implies -\frac{\epsilon}{\alpha} e^{\alpha x} \le f(x) \le \frac{\epsilon}{\alpha} e^{\alpha x}$ for $x < -M$. So f is also of exponential order α. Since $f^{(l)} \in L^1$ for all $0 \le l \le k$, we can repeat the same argument starting from k and conclude that all lower order derivatives are of exponential order α.

It is clear that $F_{m,\alpha}^k \subset F_{m,\beta}^k$ if $\alpha > \beta$. Finally, define, $F_{m,\infty}^k = \bigcap_{\alpha>0} F_{m,\alpha}^k$. This intersection is non-empty. For example, the family of Gaussian functions and $C_c^\infty(\mathbb{R})$ are in $F_{m,\alpha}^k$ for all $\alpha > 0$.

Recall from Chapter 1 that for a function $f : \mathbb{R} \to (-\infty, \infty]$ with $f \ne \infty$, $D_f = \{x \in \mathbb{R} | f(x) < \infty\}$ and $f^*(x) = \sup_{t\in\mathbb{R}} [tx - f(t)]$. If f is convex, lower semi-continuous with $\mathring{D}_f = (a, b)$ and $f \in C^2(a, b)$ with $f'' > 0$ on (a, b) then, $\mathring{D}_{f^*} = (A, B)$ where $A = \lim_{t\to a+} f'(t)$ and $B = \lim_{t\to b-} f'(t)$, f^* is continuously differentiable on (A, B). For any f satisfying the above properties, for any $x \in \mathring{D}_{f^*}$ the supremum in the definition of $f^*(x)$ is achieved at a unique point. f is called steep if $\lim_{t\to a+} |f'(t)| = \lim_{t\to b-} |f'(t)| = \infty$.

Theorem 4.1.1. *Let X be a non-constant, real-valued, and centred random variable. Assume that the logarithmic moment generating function $h(\theta) = \log \mathbb{E}(e^{\theta X})$ is finite on a neighbourhood of 0. Further assume that there is $l \in \mathbb{N}$ such that for all $\theta \in \mathring{D}_h$, $Y_{X,\theta}$ is $l-$Diophantine. Let X_n be a sequence of i.i.d. copies of X. Let $r \in \mathbb{N}$ and $a \in (0, \sup(\text{supp } X))$. Let θ_a be the unique θ such that*

$$I(a) = \sup_{\theta \in \mathring{D}_h} \left(a\theta - \log \int e^{y\theta} dF(y) \right) = a\theta_a - \log \int e^{y\theta_a} dF(y).$$

Take $q > \dfrac{l(r+2)}{2} + 1$ and $\alpha > \theta_a$. Then, for every $f \in F^q_{r+1,\alpha}$ we have,

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+\frac{1}{2}}} \int P_p(z) f_{\theta_a}(z)dz + C^q_{r+1}(f_{\theta_a}) \cdot o_{r,\theta_a}\left(\frac{1}{N^{\frac{r+1}{2}}}\right) \quad (4.3)$$

where $f_\theta(x) = e^{-\theta x} f(x)$ and $P_p(z)$ polynomials depending on a.

Proof. Assuming F to be the distribution function of X we can define $Y_{X,\gamma}$ by (4.1). Let Y_i's be i.i.d. copies of $Y_{X,\gamma}$ and take $\tilde{S}_N = Y_1 + \cdots + Y_N$. A simple computation gives us,

$$dG^\gamma_N(y) = \frac{e^{y\gamma} dF_N(y)}{\mu(\gamma)^N}$$

where F_N is the distribution function of S_N and G^γ_N is the distribution function of $\tilde{S}_N$. Now, we formally compute,

$$\begin{aligned}
\mathbb{E}(f(S_N - aN))e^{a\gamma N} &= \mathbb{E}(e^{a\gamma N} f(S_N - aN)) \\
&= \mathbb{E}(e^{\gamma S_N} f_\gamma(S_N - aN)) \\
&= \int e^{\gamma y} 2\pi f_\gamma(y - aN) dF_N(y)
\end{aligned}$$

$$= \mu(\gamma)^N \int 2\pi f_\gamma(y - aN) dG_N^\gamma(y)$$

$$= \mu(\gamma)^N \mathbb{E}_\gamma(2\pi f_\gamma(\tilde{S}_N - aN))$$

where $f_\gamma(s) = \frac{1}{2\pi} e^{-s\gamma} f(s)$. Hence,

$$\mathbb{E}(f(S_N - aN))e^{(a\gamma - \log\mu(\gamma))N} = \mathbb{E}_\gamma(2\pi f_\gamma(\tilde{S}_N - aN)). \tag{4.4}$$

Put $\gamma = \theta_a$. Then, $Y_{X,\gamma}$ has mean a (see [17, Chapter 2]).

Since $f \in F^q_{r+1,\alpha}$ with $\theta_a < \alpha$ we have $f_{\theta_a} \in F^q_{r+1}$. We prove this when $r = 0$ and $q = 1$. The argument for general q and r is similar. Suppose, $f(x), f'(x), xf(x) \in L^1$ and $f'(x)$ is continuous. It is immediate that $(e^{-\theta_a x} f(x))' = -\theta_a e^{-\theta_a x} f(x) + e^{-\theta_a x} f'(x)$ is continuous. We need to show, $e^{-\theta_a x} f(x), (e^{-\theta_a x} f(x))', xe^{-\theta_a x} f(x) \in L^1$. Since f and f' are of exponential order, it is enough to show, $e^{-\theta_a x} g(x), xe^{-\theta_a x} g(x) \in L^1$ if g is exponential of order $\alpha(> \theta_a)$. This is true because there is $M > 0$ such that for $x < -M$, $|e^{-\theta_a x} f(x)| < e^{(\alpha - \theta_a)x}$ and $|xe^{-\theta_a x} f(x)| < -xe^{(\alpha-\theta_a)x}$.

Therefore, from [4], RHS of (4.4) admits the weak Edgeworth expansion whose coefficients are determined by moments of Y_{X,θ_a}. Therefore, we have that for all functions $f \in F^q_{r+1,\alpha}$

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+\frac{1}{2}}} \int P_{p,l}(z) f_{\theta_a}(z)\, dz + C^q_{r+1}(f_{\theta_a}) \cdot o\left(\frac{1}{N^{\frac{r+1}{2}}}\right).$$

□

Remark 4.1.1.

1. *The assumption of X being centred is just to simplify the notation. One can easily reformulate the results for non-centred X using the corresponding results*

for $X - \mathbb{E}(X)$. Therefore, from now on we discuss results for centred random variables only.

2. *A similar result holds for $a \in (\inf(\text{supp } X), 0)$. In fact, one can deduce the corresponding results for $a < 0$ by considering $-X$ and $(-a) > 0$. But, for simplicity we focus only on $a > 0$ hereafter.*
3. *Note that the requirement to expand $\mathbb{E}_\gamma(f_{\theta_a}(\tilde{S}_N - aN))$ is $f_{\theta_a} \in F^q_{r+1}$ which is indeed the case when $f \in F^q_{\theta_a,\alpha}$ for some $\alpha > \theta_a$. In particular, this result holds for $f \in C^q_c(\mathbb{R})$.*
4. *In addition, if $h(\theta)$ is steep then* $\sup(\text{supp } X) = \infty$ *(see [30, Chapter 1]) and the expansion holds for all $a > 0$.*

We note that for a large class of random variables X, $Y_{X,\theta}$ is $l-$Diophantine. For example, if X is $0-$Diophantine then so is $Y_{X,\theta}$ because X is absolutely continuous with respect to $Y_{X,\theta}$ (see [1, Lemma 4]). Also, we claim that if X is compactly supported and $l-$Diophantine for $l > 0$ then so is $Y_{X,\theta}$.

We recall from [4], that a random variable X with distribution function F is $l-$Diophantine if and only if there exists $C_1, C_2 > 0$ such that for all $|x| > C_1$,

$$\inf_{y \in \mathbb{R}} \int_{\mathbb{R}} \{ax + y\}^2 dF(a) \geq \frac{C_2}{|x|^l}$$

where $\{z\} = \text{dist}(z, \mathbb{Z})$. If X is compactly supported (say on $[c, d]$) then,

$$\begin{aligned}\int_{\mathbb{R}} \{ax + y\}^2 dG^\theta(a) &= \frac{1}{\int_c^d e^{\theta a} dF(a)} \int_c^d \{ax + y\}^2 e^{\theta a} dF(a) \\ &\geq \frac{e^{\theta c}}{\int_{\mathbb{R}} e^{\theta a} dF(a)} \int_c^d \{ax + y\}^2 \, dF(a).\end{aligned}$$

Thus, for all $|x| > C_1$,

$$\inf_{y\in\mathbb{R}} \int_{\mathbb{R}} \{ax+y\}^2 dG^\theta(a) \geq \frac{e^{\theta c}}{\int_c^d e^{\theta a} dF(a)} \frac{C_2}{|x|^l}.$$

So the random variable $Y_{X,\theta}$ with distribution function G^θ is $l-$Diophantine as claimed earlier. From this we obtain the following corollary.

Corollary 4.1.2. *Let X be a non-constant, real-valued, compactly supported and l-Diophantine centred random variable. Let X_n be a sequence of i.i.d. copies of X. Let $r \in \mathbb{N}$ and $a \in (0, \sup(\text{supp } X))$. Let θ_a be the unique θ such that*

$$I(a) = \sup_{\theta \in \mathring{D}_h} \left(a\theta - \log \int e^{y\theta} dF(y) \right) = a\theta_a - \log \int e^{y\theta_a} dF(y).$$

Then, for every $f \in F^q_{r+1,\alpha}$ with $q > \dfrac{l(r+2)}{2} + 1$ and $\alpha > \theta_a$ we have,

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+\frac{1}{2}}} \int P_p(z) f_{\theta_a}(z) dz + C^q_{r+1}(f_{\theta_a}) \cdot o_{r,\theta_a}\left(\frac{1}{N^{\frac{r+1}{2}}} \right)$$

for some polynomials $P_p(z)$ depending on a.

4.1.2 Strong asymptotic expansions.

We prove a lemma that gives conditions for the point-wise limit of a sequence of functions uniformly bounded in F^q_{r+1} to satisfy the asymptotic expansions.

Lemma 4.1.3. *Let $q \geq 0$. Suppose $\{f_k\}$ is a sequence in F^q_{r+1}, S_N admits the weak local Edgeworth expansion for f_k, $C^q_{r+1}(f_k) \leq C$ for all k, f_k are uniformly bounded in $L^\infty(\mathbb{R})$, $f_k \to f$ point-wise and for all p,*

$$\lim_{k\to\infty} \int P_p(z) f_k(z) dz = \int P_p(z) f(z) dz. \tag{4.5}$$

Then,

$$\sqrt{N}\mathbb{E}(f(S_N)) = \frac{1}{2\pi}\sum_{p=0}^{\lfloor r/2 \rfloor}\frac{1}{N^p}\int P_p(z)f(z)dz + C \cdot o_{r,\beta}(N^{-r/2}).$$

Proof. For large N,

$$\begin{aligned}\left|\sqrt{N}\mathbb{E}(f_k(S_N)) - \frac{1}{2\pi}\sum_{p=0}^{\lfloor r/2 \rfloor}\frac{1}{N^p}\int P_p(z)f_k(z)dz\right| &\le C^q_{r+1}(f_k)\cdot o_{r,\beta}(N^{-r/2}) \\ &\le C \cdot o_{r,\beta}(N^{-r/2}). \qquad (4.6)\end{aligned}$$

LDCT gives us that,

$$\lim_{k\to\infty}\mathbb{E}(f_k(S_N)) = \mathbb{E}(f(S_N))$$

This along with assumption (4.5) allows us to take the limit $k \to \infty$ in the RHS of (4.6) and to conclude,

$$\left|\sqrt{N}\mathbb{E}(f(S_N)) - \frac{1}{2\pi}\sum_{p=0}^{\lfloor r/2 \rfloor}\frac{1}{N^p}\int P_p(z)f(z)dz\right| \le C \cdot o_{r,\beta}(N^{-r/2})$$

which implies the result. □

Remark 4.1.2. *The same would hold if we replace weak local by weak global. However, our focus here is on weak local expansions.*

The next theorem specifies when the existence of weak expansions imply the existence of strong expansions.

Theorem 4.1.4. *Let X_n be a sequence of random variables not necessarily i.i.d. Suppose $S_N = X_1 + \cdots + X_N$ admits the weak asymptotic expansion of order r for large deviations in the range $(0, L)$ for $f \in F^1_{r+1,L_+}$ where $L_+ > L$ when $L < \infty$ and $L_+ = \infty$ if $L = \infty$. That is,*

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor}\frac{1}{N^{p+1/2}}\int P_p(z)f_{\theta_a}(z)dz + C^1_{r+1}(f_{\theta_a})\cdot o_{r,\theta_a}\left(\frac{1}{N^{\frac{r+1}{2}}}\right)$$

for all $a \in (0, L)$ where $I(a)$ and θ_a as in (4.11). Then, S_N admits the strong asymptotic expansion of order r for large deviation in $(0, L)$.

Proof. If $f \in C_c^\infty$ then $f_\theta \in F_{r+1}^1$ for all θ. Therefore, we approximate $1_{[0,\infty)}$ by a sequence f_k of C_c^∞ functions such that $(f_k)_{\theta_a}$ are uniformly bounded in F_{r+1}^1 (see Appendix A.3 for such a sequence) and invoke Lemma 4.1.3 to establish,

$$\mathbb{P}(S_N \geq aN)e^{I(a)N} = \frac{1}{2\pi} \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+1/2}} \int_0^\infty P_p(z) e^{-\theta_a z} dz + C \cdot o_{r,\theta_a}\left(\frac{1}{N^{\frac{r+1}{2}}}\right).$$

$\square$

Remark 4.1.3. *Note that the coefficients of the strong expansion are $C_p(a) = \frac{1}{2\pi}\int_0^\infty P_p(z)e^{-\theta_a z}\, dz$ obtained by replacing f with $1_{[0,\infty)}$ in coefficients of the weak expansions. Since f_k's are bounded in F_{r+1}^1, we can do this without altering the order of the error. However, for any $q > 1$, $1_{[0,\infty)}$ is not a pointwise limit of a sequence of functions f_k in F_r^q with $C_{r+1}^q(f_k)$ bounded. To see this, assume that $\|f_k\|_1, \|f_k'\|_1, \|f_k''\|_1$ are uniformly bounded and $f_k \to 1_{[0,\infty)}$ point-wise. Then, for all $\phi \in C_c^\infty(\mathbb{R})$,*

$$\int \delta' \ \phi = -\int \delta \ \phi' = \int 1_{[0,\infty)} \ \phi'' = \lim_{k\to\infty} \int f_k \ \phi'' = \lim_{k\to\infty} -\int f_k' \ \phi' = \lim_{k\to\infty} \int f_k'' \ \phi$$

This implies that $\dfrac{|\phi'(0)|}{\|\phi\|_\infty} \leq \sup_k \|f_k''\|_1$ for all $\phi \in C_c^\infty(\mathbb{R})$. Clearly, this is a contradiction. Therefore, Theorem 4.1.1 does not automatically give us strong expansions.

Now we are in a position to state and prove the main result of this section, which extends Cramér's LDP for i.i.d. random variables when the random variables have a sufficiently regular density.

Theorem 4.1.5. *Let X be a non-constant real valued centred random variable. Assume that the logarithmic moment generating function $h(\theta) = \log \mathbb{E}(e^{\theta X})$ is finite on a neighbourhood of 0. Further assume that, X is 0−Diophantine. Let $r \in \mathbb{N}$. Then for all $a \in (0, \sup(\mathrm{supp}X))$, there are constants $C_p(a)$ such that*

$$\mathbb{P}(S_N \geq aN)e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{C_p(a)}{N^{p+\frac{1}{2}}} + o\left(\frac{1}{N^{\frac{r+1}{2}}}\right)$$

where

$$C_p(a) = \frac{1}{2\pi} \int_0^\infty e^{-\theta_a z} P_p(z) dz$$

for some polynomials $P_p(z)$ depending on a,

$$I(a) = \sup_{\theta \in \mathbb{R}} \left(a\theta - \log \int e^{y\theta} dF(y) \right)$$

and θ_a is this unique point the supremum is achieved.

Proof. If X is 0−Diophantine then so is $Y_{X,\theta}$ as X is absolutely continuous with respect to $Y_{X,\theta}$ (see [1, Lemma 4]). Since, $Y_{X,\theta}$ has moments of all orders, $Y_{X,\theta}$ admits the strong Edgeworth expansion of all orders. Therefore, for each $r \in \mathbb{N}$, $Y_{X,\theta}$ admits the weak local Edgeworth expansion of order r for $f \in F_r^1$ (see Appendix A.2).

From (4.4) we know that,

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \mathbb{E}_\gamma(2\pi f_{\theta_a}(\tilde{S}_N - aN))$$

where summands of $\tilde{S}_N$ have mean a. The assumptions allow us to expand RHS using the weak local Edgeworth expansion and obtain,

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+\frac{1}{2}}} \int P_p(z) f_{\theta_a}(z) dz + C^1_{r+1}(f_{\theta_a}) \cdot o_{r,\beta}\left(N^{-r/2}\right).$$

for $f \in C_c^\infty(\mathbb{R})$.

Now, we approximate $1_{[0,\infty)}$, by a sequence $f_k \in C_c^\infty(\mathbb{R})$ such that $(f_k)_{\theta_a}$ are bounded in F^1_{r+1} (see Appendix A.3 for such a sequence) and use Theorem 4.1.4 to obtain the required expansion. □

Remark 4.1.4. *This gives us an alternative proof of [1, Theorem 2] for X satisfying the Cramér's condition (which corresponds to Case 1 there).*

There are two ways the coefficients $C_p(a)$ depend on a. First note that θ_a depends on the choice of a. Also, from Section 3.3, we know exactly how the coefficients of P_p depend on the first $p+2$ asymptotic moments of $\tilde{S}_N$ and thus, on the first $p+2$ moments of Y_{X,θ_a}. So the dependence of $C(a)$ on a is explicit and one can compute these coefficients. In addition, $C_p(a)$ does not depend on r because $P_p(z)$'s do not.

4.2 Higher order asymptotics in the non–i.i.d. case.

Let X_n be a sequence of random variables that are not necessarily i.i.d. with asymptotic mean 0. Suppose that there exist a Banach space $\mathbb{B}$, a family of bounded linear operators $\mathcal{L}_z : \mathbb{B} \to \mathbb{B}$ and vectors $v \in \mathbb{B}, \ell \in \mathbb{B}'$ such that

$$\mathbb{E}\left(e^{zS_N}\right) = \ell(\mathcal{L}_z^N v), \;\; z \in \mathbb{C} \tag{4.7}$$

and satisfying the following,

(B1) There exists $\delta > 0$ such that $z \mapsto \mathcal{L}_z$ is continuous on the strip $|\mathrm{Re}(z)| < \delta$ and holomorphic on the disc $|z| < \delta$.

(B2) 1 is an isolated and simple eigenvalue of $\mathcal{L}_0$, all other eigenvalues of $\mathcal{L}_0$ have absolute value less than 1 and its essential spectrum is contained strictly inside the disk of radius 1 (spectral gap).

(B1) and (B2) along with perturbation theory of operators (see [33]) imply that there is $\delta_0 \in (0, \delta)$ such that

$$\mathcal{L}_z = \mu(z)\Pi_z + \Lambda_z,\ |z| < \delta_0 \tag{4.8}$$

where $\mu(z)$ is the top eigenvalue of $\mathcal{L}_z$, Π_z is the corresponding eigen–projection, $\Pi_z\Lambda_z = \Lambda_z\Pi_z = 0$ and $z \mapsto \mu(z),\ z \mapsto \Pi_z$ and $z \mapsto \Lambda_z$ are holomorphic. In addition, $\left\|\frac{d^k}{dz^k}\Lambda_z^N\right\| < \beta_k^N$ with $0 < \beta_k < 1$. Therefore,

$$\mathcal{L}_z^N = \mu(z)^N\Pi_z + \Lambda_z^N$$

Combining this with (4.7) we have,

$$\mathbb{E}(e^{zS_N}) = \mu(z)^N\ell(\Pi_z v) + \ell(\Lambda_z v). \tag{4.9}$$

Then, plugging in $z = 0$ and taking $N \to \infty$, we conclude that $\ell(\Pi_0 v) = 1$. Also, taking the derivative at $z = 0$, dividing by N and taking the limit as $N \to \infty$, we obtain,

$$\frac{d}{dz}\mu(z)\Big|_{z=0} = \lim_{N\to\infty}\frac{\mathbb{E}(S_N)}{N} = 0.$$

Taking the second derivative at $z = 0$, dividing by N^2 and taking the limit as $N \to \infty$, we obtain,

$$\frac{d^2}{dz^2}\mu(z)\Big|_{z=0} = \lim_{N\to\infty}\frac{\mathbb{E}(S_N^2)}{N^2}$$

In addition, it follows from [24][Theorem 2.4] that there exists $\sigma^2 \geq 0$ such that $\dfrac{S_N}{\sqrt{N}} \xrightarrow{d} \mathcal{N}(0, \sigma^2)$. Since our interest is in S_N that satisfies the CLT we would asumme from now on that $\sigma^2 > 0$.

We also assume the following:

(B3) $\mu(\theta) > 0$ for all $\theta \in (-\delta_0, \delta_0)$ (Here δ_0 as in (4.8)).

Define $\Omega(\theta) = \log \mu(\theta)$ for $\theta \in (-\delta_0, \delta_0)$. Then, $\Omega(0) = \log \mu(0) = 0$ and $\Omega'(0) = \dfrac{\mu'(0)}{\mu(0)} = 0$. Also, $\Omega''(0) = \dfrac{\mu''(0)\mu(0) - \mu'(0)^2}{\mu(0)^2} = \mu''(0) = \sigma^2 > 0$. Since Ω'' is continuous, there exists $\delta_1 \in (0, \delta_0)$ such that Ω is strictly convex on $(-\delta_1, \delta_1)$. Note that due to convexity, $\Omega'(-\delta_1) < 0 < \Omega'(\delta_1)$. In addition, when $\theta \neq 0$, $\mu(\theta) > \mu(0) = 1$ by convexity.

Next, we consider the Legendre transform of Ω, I given by,

$$I(a) = \sup_{\theta \in (-\delta_1, \delta_1)} [a\theta - \Omega(\theta)], \text{ for } a \in [0, \Omega'(\delta_1))$$

which itself is a strictly convex function.

Because Ω' is strictly increasing and continuous on $[0, \Omega'(\delta_1)]$, $a - \Omega'(\theta) = 0$ has a unique solution θ_a which depends continuously on a. Note that $I(a) \geq 0$ for all a and $I(a) = 0 \iff a = 0$. Also, $I(a)$ is continuous because I is convex and $I(0) = 0$. In addition, $I(\Omega'(\delta_1)) = a\delta_1 - \Omega(\delta_1)$.

Now, we are in a position to prove a Large Deviation Principle for S_N using Theorem 1.3. The following lemma shows that Theorem 1.3 applies in our case.

Lemma 4.2.1. *Suppose* $(B1), (B2)$ *and* $(B3)$ *hold. Then, there exists* $0 < \delta_2 \leq \delta_1$ *such that for* $\theta \in (-\delta_2, \delta_2)$,

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}(e^{\theta S_N}) = \log \mu(\theta)$$

Proof. Because $\ell(\Pi_0 v) > 0$, there exists δ_2 and $m > 0$ such that for $\theta \in [-\delta_2, \delta_2]$

$\ell(\Pi_\theta v) > 2m$. Because $\|\Lambda_\theta^N\| < C\mu(\theta)^N$ for large N, we have that

$$\lim_{N\to\infty} \mu(\theta)^{-N}\ell(\Lambda_\theta^N v) = 0.$$

Hence, there exists N_0 such that for $N > N_0$,

$$m < \ell(\Pi_\theta v) + \mu(\theta)^{-N}\ell(\Lambda_\theta^N v) < 3m.$$

Hence,

$$\lim_{N\to\infty} \frac{1}{N} \ln\left[\ell(\Pi_\theta v) + \mu(\theta)^{-N}\ell(\Lambda_\theta^N v)\right] = 0.$$

Now, for $\theta \in (-\delta_2, \delta_2)$ we can rewrite (4.9) as

$$\frac{1}{N}\log \mathbb{E}(e^{\theta S_N}) = \log\mu(\theta) + \frac{1}{N}\log\left[\ell(\Pi_\theta v) + \mu(\theta)^{-N}\ell(\Lambda_\theta^N v)\right].$$

This implies that,

$$\lim_{N\to\infty} \frac{1}{N}\log \mathbb{E}(e^{\theta S_N}) = \log\mu(\theta).$$

□

Combining this lemma with Theorem 1.3 and the analysis proceeding it, we have the following LDP.

Theorem 4.2.2. *Suppose* $(B1), (B2)$ *and* $(B3)$ *hold. Then, there exists* $\delta_2 \in (0, \delta_1]$ *such that for all* $a \in \left(0, \dfrac{\log\mu(\delta_2)}{\delta_2}\right)$,

$$\lim_{N\to\infty} \frac{1}{N}\log \mathbb{P}(S_N \geq aN) = -I(a) \tag{4.10}$$

where

$$I(a) = \sup_{\theta\in(-\delta_2,\delta_2)} [a\theta - \log\mu(\theta)] = a\theta_a - \log\mu(\theta_a) \tag{4.11}$$

and θ_a *is the unique* θ *solving* $\big(\log\mu(\theta)\big)' = \dfrac{\mu'(\theta)}{\mu(\theta)} = a.$

Remark 4.2.1. *The range of a for which the LDP holds, is constrained by the assumptions $(B1), (B2)$ and $(B3)$. We require a positive top eigenvalue $\mu(\theta)$ to exist, $\log\mu(\theta)$ to be strictly convex and $\ell(\Pi_\theta v) > 0$. Larger the range of θ for which these hold, larger the range of a. In particular, if these hold for all $\theta \in \mathbb{R}$, then by convexity $B = \lim_{\delta\to\infty} \dfrac{\log\mu(\delta)}{\delta}$ exists as an extended real number and for all $a \in (0, B)$ the LDP holds.*

Next, we compute higher order asymptotics of this LDP. To this end, we make two more assumptions about $\mathcal{L}_z$.

(B4) For all $\theta \in (-\delta_2, \delta_2)$, for all real numbers $t \neq 0$, $\mathrm{sp}(\mathcal{L}_{\theta+it}) \subset \{|z| < \mu(\theta)\}$.

(B5) There are positive real numbers r_1, r_2, C, K and N_0 such that for all $\theta \in (-\delta_2, \delta_2)$, for all $N > N_0$ and for all $K < |t| < N^{r_1}$, $\left\|\mathcal{L}^N_{\theta+it}\right\| \leq C\dfrac{\mu(\theta)^N}{N^{r_2}}$.

Remark 4.2.2. *As in Remark 3.1.1 it follows that by slightly decreasing r_1 we can assume r_2 to be as large as required for large enough N.*

Pick $a \in \left(0, \dfrac{\log\mu(\delta_2)}{\delta_2}\right)$. Then,

$$\begin{aligned}
\mathbb{E}(f(S_N - aN))e^{a\theta N} &= \mathbb{E}(e^{\theta S_N} e^{-(S_N - aN)\theta} f(S_N - aN)) \\
&= \frac{1}{2\pi}\int \widehat{f}_\theta(t) e^{-iatN} \ell(\mathcal{L}^N_{\theta+it} v)\, dt
\end{aligned}$$

where $f_\theta(x) = \dfrac{1}{2\pi} e^{-\theta x} f(x)$. Now define, $\overline{\mathcal{L}}_{\theta+it} = \dfrac{e^{-iat}}{\mu(\theta)}\mathcal{L}_{\theta+it}$. Then,

$$\mathbb{E}(f(S_N - aN))e^{a\theta N} = \mu(\theta)^N \int \widehat{f}_\theta(t)\ell(\overline{\mathcal{L}}^N_{\theta+it} v)\, dt.$$

From this we have,

$$\mathbb{E}(f(S_N - aN))e^{[a\theta - \log\mu(\theta)]N} = \int \widehat{f}_\theta(t)\ell(\overline{\mathcal{L}}^N_{\theta+it} v)\, dt.$$

In particular,

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \int \widehat{f}_{\theta_a}(t)\ell(\overline{\mathcal{L}}^N_{\theta+it_a}v)\, dt. \tag{4.12}$$

Note that for $|\theta_a + it| < \delta_0$ the top eigenvalue of $\overline{L}_{\theta_a+it}$ is $\overline{\mu}(\theta_a + it) = \dfrac{e^{-iat}}{\mu(\theta_a)}\mu(\theta_a + it)$. As a function of t, $\overline{\mu}(\theta_a + it)$ is analytic in a neighbourhood of 0 by (4.8). Further,

$$\overline{\mu}(\theta_a) = 1,\ \overline{\mu}'(\theta_a) = \frac{d}{dt}\overline{\mu}(z)\Big|_{t=0} = -ia + i\frac{\mu'(\theta_a)}{\mu(\theta_a)} = 0,\ \overline{\mu}''(\theta_a) = -\frac{\mu''(\theta_a)}{\mu(\theta_a)} = -\sigma_a^2$$

with $\sigma_a > 0$. Thus, there exists $\overline{\delta}$ such that

$$|\overline{\mu}(\theta_a + it)| < e^{-\sigma_a^2 t^2/4},\ |t| < \overline{\delta}. \tag{4.13}$$

We also notice that,

$$\lim_{N\to\infty} \frac{\ell(\Lambda_\theta^N v)}{\mu(\theta)^N} = 0$$

because the spectral radius of Λ_θ is strictly smaller than $\mu(\theta)$. Combining this with $\mathbb{E}(e^{\theta S_N}) = \mu(\theta)^N \ell(\Pi_\theta v) + \ell(\Lambda_\theta^N v)$ we conclude that for all θ,

$$\ell(\Pi_\theta v) = \lim_{N\to\infty} \frac{\mathbb{E}(e^{\theta S_N})}{\mu(\theta)^N}.$$

The following lemma allows us to obtain asymtotics of (4.12). We note that it is analogous to Theorem 3.1.4 where asymptotics of $\mathbb{E}(f(S_N - aN))$ for $f \in F^{q+2}_{r+1}$ are discussed and can be proven using the ideas in the proof of Theorem 3.1.4. One just has to replace $\mathcal{L}_t$ by $\mathcal{L}_{\theta_a+it}$ there and introduce the corresponding changes.

Lemma 4.2.3. *Suppose (B1) through (B5) hold. Let $r \in \mathbb{N}$. Then, there exist $\delta_2 \in (0, \delta)$ such that for all $a \in \left(0, \dfrac{\log \mu(\delta_2)}{\delta_2}\right)$ there are polynomials $P_p(z)$ such that for $g \in F^{q+1}_{r+1}$ with $q > \dfrac{r+1}{2r_1}$,*

$$\int \widehat{g}(t)\ell(\overline{\mathcal{L}}^N_{\theta_a+it}v)\, dt = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+1/2}} \int P_p(z)g(z)dz + C^{q+2}_{r+1}(g) \cdot o_{r,\theta_a}\left(\frac{1}{N^{\frac{r+1}{2}}}\right)$$

where θ_a is as in (4.11).

Proof. We state how to estimate LHS away from 0. The rest of the proof, which contains the construction of polynomials P_p, is identical to that of Theorem 3.1.5 with it replaced by $\theta_a + it$.

Fix $\delta > 0$ as in (4.13). By (B4), for $\delta \le |t| \le K$, there exists $c_0 \in (0,1)$ such that $\|\overline{\mathcal{L}}^n_{\theta_a+it}\| \le c_0^n$. Thus, $\left| \int_{\delta<|t|<K} \widehat{g}(t)\ell(\overline{\mathcal{L}}^n_{\theta_a+it}v)\, dt \right| \le C\|g\|_1 c_0^n$.

WLOG assuming $r_2 > r_1 + (r+1)/2$,

$$\left| \int_{K<|t|<n^{r_1}} \widehat{g}(t)\ell(\overline{\mathcal{L}}^n_{\theta_a+it}v)\, dt \right| \le C\|g\|_1 \int_{K<|t|<n^{r_1}} \|\mathcal{L}^n_{\theta_a+it}\|\, dt \le \frac{C\|g\|_1}{n^{r_2-r_1}}$$
$$= \|g\|_1 o(n^{-(r+1)/2}).$$

Since, $g \in F^{q+2}_{r+1}$, we have that $t^q\widehat{g}(t) = (-i)^q\widehat{g^{(q)}}(t)$ and $\widehat{g^{(q)}}$ is integrable. Integrability of $\widehat{g^{(q)}}$ along with $q > \dfrac{r+1}{2r_1}$ implies,

$$\begin{aligned}\left| \int_{|t|>n^{r_1}} \widehat{g}(t)\ell(\overline{\mathcal{L}}^n_{\theta_a+it}v)\, dt \right| &\le \int_{|t|>n^{r_1}} |\widehat{g}(t)|\, dt \le \int_{|t|>n^{r_1}} \left| \frac{\widehat{g^{(q)}}(t)}{t^q} \right| dt \\ &\le \frac{\|\widehat{g^{(q)}}\|_1}{n^{r_1 q}} = \|\widehat{g^{(q)}}\|_1 o(n^{-(r+1)/2}).\end{aligned} \tag{4.14}$$

Therefore,

$$\left| \int_{|t|>\delta} \widehat{g}(t)\ell(\overline{\mathcal{L}}^n_{\theta_a+it}v)\, dt \right| = o(n^{-(r+1)/2}). \tag{4.15}$$

□

Remark 4.2.3.

1. *The proof is almost identical to the proof of Theorem 3.1.4 and hence, the coefficients of polynomials P_p can be computed as shown in Section 3.3. In particular, they depend on exponential moments of S_N.*

2. *Since θ_a depends on a, the coeffients of the polynomials P_p also depend on a.*

As a direct consequence of Lemma 4.2.3 and equation (4.12), we have the following theorem.

Theorem 4.2.4. *Suppose (B1) through (B5) hold. Let $r \in \mathbb{N}$. Then, for $a \in \left(0, \dfrac{\log \mu(\delta_2)}{\delta_2}\right)$ there exist $\theta_a \in (0, \delta_2)$ and polynomials $P_p(z)$ such that for $f \in F^{q+2}_{r+1,\alpha}$ with $q > \dfrac{r+1}{2r_1}$ and $\alpha > \delta_2$,*

$$\mathbb{E}(f(S_N - aN))e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{1}{N^{p+1/2}} \int P_p(z) f_{\theta_a}(z) dz + C^{q+2}_{r+1}(f_{\theta_a}) \cdot o_{r,\theta_a}\left(\frac{1}{N^{\frac{r+1}{2}}}\right)$$

where $f_\theta(x) = \dfrac{1}{2\pi} e^{-\theta x} f(x)$, I and θ_a as in (4.11).

Remark 4.2.4. *In particular, the theorem holds for all $f \in C^\infty_c(\mathbb{R})$.*

This is the weak asymptotic expansion which gives us the required higher order asymptotics for (4.10), the LDP in Theorem 4.2.2.

Next, we replace (B5) by the following stronger assumption which allows us to conclude existence of strong expansions for the LDP. Compare this assumption with assumption (A5) in Chapter 3.

$\widetilde{\text{(B5)}}$ There are positive real numbers r_1, r_2, r_3, C, K and N_0 such that for all $\theta \in (-\delta_2, \delta_2)$, for all $N > N_0$ and for all $|t| > K$, $\left\|\mathcal{L}^N_{\theta+it}\right\| \le C \dfrac{\mu(\theta)^N}{N^{r_2}|t|^{r_3}}$.

As in the case of (B5), we can assume r_2 and r_3 to be large after slightly reducing r_1. Therefore we have the following theorem.

Theorem 4.2.5. *Suppose (B1) through (B4) and $\widetilde{(B5)}$ hold. Let $r \in \mathbb{N}$. Then, there exists $0 < \delta_2 \le \delta$ such that S_N admits a weak asymptotic expansions for the LDP in the range $\left(0, \dfrac{\log \mu(\delta_2)}{\delta_2}\right)$ for $f \in F^1_{r+1,\alpha}$ with $\alpha > \delta_2$.*

In particular, for all $a \in \left(0, \frac{\log \mu(\delta_2)}{\delta_2}\right)$ *there exist constants* $C_p(a)$ *such that*

$$\mathbb{P}(S_N \geq aN)e^{I(a)N} = \sum_{p=0}^{\lfloor r/2 \rfloor} \frac{C_p(a)}{N^{p+1/2}} + C_{r,\theta_a}\ o\left(\frac{1}{N^{\frac{r+1}{2}}}\right).$$

where

$$C_p(a) = \frac{1}{2\pi} \int_0^\infty e^{-\theta_a z} P_p(z) dz$$

for some polynomials $P_0(z), \dots, P_r(z)$ *depending on* a *and unique* $\theta_a \in (0, \delta_2)$ *such that*

$$I(a) = \sup_{\theta \in (-\delta_2, \delta_2)} [a\theta - \log \mu(\theta)] = a\theta_a - \log \mu(\theta_a).$$

Proof. The proof of the first part is similar to that of Theorem 4.2.4. The only difference is the estimate (4.14).

Since $f \in F^1_{r+1,\alpha}$, we have $g = f_\theta \in F^1_{r+1}$. So $t\widehat{g}(t) = (-i)\widehat{g'}(t)$. WLOG assume $r_3 > \frac{r+1}{2r_1}$. Then,

$$\begin{aligned} \left| \int_{|t|>n^{r_1}} \widehat{g}(t) \ell(\overline{\mathcal{L}}^n_{\theta_a+it} v)\, dt \right| \leq C \int_{|t|>n^{r_1}} |\widehat{g}(t)| \|\overline{\mathcal{L}}^n_{\theta_a+it}\|\, dt &\leq C \int_{|t|>n^{r_1}} \left| \frac{\widehat{g'}(t)}{t^{1+r_3}} \right| dt \\ &\leq \frac{C\|g'\|_1}{n^{r_1 r_3}} \\ &= \|g'\|_1 o(n^{-(r+1)/2}) \end{aligned}$$

Now, the existence of the strong expansion follows from the first part of the theorem and Theorem 4.1.4. □

As in the i.i.d. case, $C_p(a)$ does not depend r because θ_a and P_p do not. Also, there are two ways the coefficients $C(a)$ depend on a. First note θ_a depends on the choice of a. Also, from Section 3.3, we know exactly how the coefficients of P_p depend on the derivatives of the $\mu(z)$ and $\ell(\Pi_z(\cdot))$ at θ_a and thus, on the exponential moments of S_N. Since this dependence of $C(a)$ on a is explicit, one can compute these coefficients.

4.3 An application to Markov Chains.

Take x_n to be a time homogeneous Markov process on a compact connected manifold $\mathcal{M}$ with smooth transition density $p(x, y)$ which is bounded away from 0, and $X_n = h(x_{n-1}, x_n)$ for smooth function $h : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$. We assume that $h(x, y)$ can not be written in the form

$$h(x, y) = H(y) - H(x) + c(x, y) \tag{4.16}$$

where $c(x, y)$ is piece-wise constant. (An equivalent condition is given in Lemma 3.5.1). This is exactly the setting we worked in Section 3.5.3.1.

We need the following lemma to establish (B1) through (B5).

Lemma 4.3.1. *Let $K(x, y)$ be a smooth positive function on $\mathcal{M} \times \mathcal{M}$. Let P be an operator on $L^\infty(\mathcal{M})$ given by $Pu(x) = \int_{\mathcal{M}} K(x, y)u(y)\,dy$. Then, P has a simple leading eigenvalue $\lambda > 0$ and the corresponding eigenfunction g is positive and smooth.*

Proof. From the Weierstrass theorem, $K(x, y)$ is a uniform limit of functions of the form $\sum_{r \le n} J_r(x)L_r(y)$. Therefore, P can be approximated by finite rank operators. So P is compact. Since P is an operator which leaves the cone of positive functions invariant, by a direct application of Birkhoff Theory (see [2]), P has a leading eigenvalue λ which is positive and simple. The corresponding eigenfunction g is also positive.

Because P is compact, there is $l \in (0, \lambda)$ such that $\mathrm{sp}_{L^\infty}(P) \cap \{|z| > r\} = \{\lambda\}$. Next, we consider P acting on $C^1(\mathcal{M})$. Observe that,

$$\frac{d}{dx}(Pu)(x) = \int_{\mathcal{M}} \frac{\partial K}{\partial x}(x, y)u(y)\, dy.$$

So, $\|Pu(x)\|_{C^1} \leq C\|u\|_\infty$ for some C. Since $\|\cdot\|_\infty \leq \|\cdot\|_{C^1}$ unit ball with respect to $\|\cdot\|_{C^1}$ is relatively compact with respect to $\|\cdot\|_\infty$. Therefore the essential spectral radius is 0 by [24, Lemma 2.2]. This gives us, $\mathrm{sp}_{C^1}(P) \cap \{|z| > r\} \subseteq \{\lambda\}$.

To see that equality holds, note that the constant function $1 \in C^1(\mathcal{M})$. By positivity of P,

$$1 \geq \frac{g}{\sup g} \implies P^n 1 \geq \frac{P^n g}{\sup g} \implies P^n 1 \geq \frac{\lambda^n g}{\sup g} \implies |||P^n||| \geq \lambda^n \|\frac{g}{\sup g}\|_{C^1} \geq \lambda^n$$

where $|||\cdot|||$ is the operator norm of P acting on $C^1(\mathcal{M})$. Therefore, the spectral radius of P is $\geq \lambda$. This establishes that $g \in C^1$. We can repeat the argument and show $g \in C^r$ for $r \in \mathbb{N}$. $\square$

Take $\mathbb{B} = L^\infty(\mathcal{M})$ and consider the family of integral operators,

$$(\mathcal{L}_z u)(x) = \int_{\mathcal{M}} p(x, y)e^{zh(x,y)}u(y)\, dy,\ z \in \mathbb{C}.$$

Let μ be the initial distribution of the Markov chain. Then, using the Markov property, we have $\mathbb{E}_\mu[e^{zS_n}] = \mu(\mathcal{L}_z^N 1)$. Now, we check conditions (B1) through (B5).

Conditions (B1) and (B2) coincide with the conditions (A1) and (A2) in Chapter 3 and we verify them in Section 3.5.3.1. In particular, (B1) holds with $\delta = \infty$. Note that, for all θ, $\mathcal{L}_\theta$ is of the form P in Lemma 4.3.1. Therefore, (B3) holds for all θ. Take $\lambda(\theta)$ be the top eigenvalue and g_θ to be the corresponding eigenfunction. Then, g_θ is smooth.

To show (B4) and (B5) we define a new operator Q_θ as follows.

$$(Q_\theta u)(x) = \frac{1}{\lambda(\theta)} \int_{\mathcal{M}} e^{\theta h(x,y)} p(x,y) u(y) \frac{g_\theta(y)}{g_\theta(x)} \, d(y).$$

It is easy see to that $p_\theta(x,y) = \dfrac{e^{\theta h(x,y)} p(x,y)}{g_\theta(x)\lambda(\theta)}$ and $dm_\theta(y) = g_\theta(y)\, d(y)$ defines a new Markov chain x_n^θ with the associated Markov operator Q_θ. That is, Q_θ is a positive operator and $Q_\theta 1 = \dfrac{1}{\lambda(\theta)} \displaystyle\int_{\mathcal{M}} e^{\theta h(x,y)} p(x,y) \frac{g_\theta(y)}{g_\theta(x)} \, dy = 1$ because g_θ is the eigenfunction corresponding to eigenvalue $\lambda(\theta)$ of $\mathcal{L}_\theta$.

Now, we can repeat the argument in Section 3.5.3.1 to establish properties of the perturbed operator given by

$$(Q_{\theta+it}) u(x) = \int_{\mathcal{M}} e^{ith(x,y)} p_\theta(x,y) \, dm_\theta(y)$$

Since (4.16) does not hold we conclude that $\mathrm{sp}(\mathcal{L}_{\theta+it}) \subset \{|z| < 1\}$.

Take G_θ to be the operator on $L^\infty(\mathcal{M})$ that corresponds to multiplication by g_θ. Then, $\mathcal{L}_{\theta+it} = \lambda(\theta) G_\theta Q_{\theta+it} G_\theta^{-1}$. Therefore, $\mathrm{sp}(\mathcal{L}_{\theta+it})$ is the $\mathrm{sp}(Q_{\theta+it})$ scaled by $\lambda(\theta)$. This implies $\mathrm{sp}(\mathcal{L}_{\theta+it}) \subset \{|z| < \lambda(\theta)\}$ as required.

Since (4.16) does not hold, the asymptotic variance σ_θ^2 of $X_n^\theta = h(x_{n-1}^\theta, x_n^\theta)$ is positive. Taking $\gamma(\theta+it)$ to be the top eignevalue of $Q_{\theta+it}$, $\lambda(\theta+it) = \lambda(\theta)\gamma(\theta+it)$. Thus, $(\log \lambda(\theta))'' = -\dfrac{d^2}{dt^2} \log \lambda(\theta+it)\Big|_{t=0} = -\dfrac{d^2}{dt^2} \log \gamma(\theta+it)\Big|_{t=0} = -\dfrac{\gamma''(\theta)}{\gamma(\theta)} + \left(\dfrac{\gamma'(\theta)}{\gamma(\theta)}\right)^2 = -\gamma''(\theta) + \gamma'(\theta)^2$ ($\because$ $\gamma(\theta) = 1$). Put $S_N^\theta = X_1^\theta + \cdots + X_N^\theta$. Since, $\mathbb{E}(e^{itS_N^\theta}) = \displaystyle\int Q_{\theta+it}^N 1 \, d\mu$, from (3.37), we have that $\gamma'(\theta)^2 - \gamma''(\theta) = \sigma_\theta^2$. Thus, $(\log \lambda(\theta))'' = \sigma_\theta^2 > 0$. Therefore, $\log \lambda(\theta)$ is a strictly convex function.

Note that, $\mathcal{L}_\theta = \lambda(\theta)\Pi_\theta + \Lambda_\theta$ where Π_θ is the projection onto the top eigenspace. From [27, Chapter III], $\Pi_\theta = g_\theta \otimes \varphi_\theta$ where φ_θ is the top eigenfunction of Q_θ^*, the adjoint of Q_θ. Because Q_θ^* itself is a positive compact operator acting on $(L^\infty)^*$ (the space of finitely additive finite signed measures), φ_θ is a finite positive measure. Hence, $\mu(\Pi_\theta 1) = \varphi_\theta(1)\mu(g_\theta) > 0$ for all θ.

As a result, Lemma 4.2.1 holds with δ_2 arbitrary large and hence, Theorem 4.2.2 holds with δ_2 arbitrary large. So the rate function $I(a)$ in Theorem 4.2.2 is finite for $a \in (0, B)$ where $B = \lim_{\theta\to\infty} \dfrac{\log\lambda(\theta)}{\theta}$. We observe that $B < \infty$ because h is bounded i.e. $\dfrac{S_N}{N} \le \|h\|_\infty$ surely. In fact, we claim $B = \lim_{N\to\infty} \dfrac{B_N}{N}$ where $B_N = \sup_{x_0,\ldots,x_N} \sum_{j=1}^{N} h(x_{j-1}, x_j)$ (the supremum taken over all possible realizations of the Markov chain x_n).

First note that B_N is subadditive. So $\lim_{N\to\infty} \dfrac{B_N}{N}$ exists and is equal to $\inf_N \dfrac{B_N}{N}$. Given, $a > B$ there exists N_0 such that for all $N > N_0$, $\dfrac{S_N}{N} \le \dfrac{B_N}{N} < a$. Therefore, $\mathbb{P}(S_N \ge aN) = 0$ for all $N > N_0$ and hence, $I(a) = \infty$. Next, given $a < B$, for all N, $B_N > aN$. Fix N. Then, there exists a realization $x_1, \ldots, x_N$ such that $aN < \sum_{j=1}^{N} h(x_{j-1}, x_j) \le B$. Since h is uniformly continuous on $\mathcal{M} \times \mathcal{M}$, there exists $\delta > 0$ such that by choosing y_j from a ball of radius δ centred at x_j i.e. $y_j \in \mathbb{B}(x_j, \delta)$, we have $aN < \sum_{j=1}^{N} h(y_{j-1}, y_j) \le B$. We estimate the probability of choosing such a realization $y_1, \ldots, y_N$ and obtain a lower bound for $\mathbb{P}(S_N \ge aN)$:

$$\begin{aligned}
\mathbb{P}(S_N \ge aN) &\ge \int_{\mathbb{B}(x_N,\delta)} \cdots \int_{\mathbb{B}(x_1,\delta)} \int_{\mathbb{B}(x_0,\delta)} p(y_{N-1}, y_N) \ldots p(y_0, y_1)\, d\mu(y_0)\, dy_1 \ldots dy_N \\
&\ge \mu(\mathbb{B}(x_0,\delta)) \Big(\min_{x,y\in\mathcal{M}} p(x,y) \Big)^N \operatorname{vol}(\mathbb{B}_\delta)^N
\end{aligned}$$

Therefore, $I(a) < \infty$ as required.

Also, because g_θ is smooth we can repeat the argument in Section 3.5.3.1 to obtain (3.45) for $Q_{\theta+it}$. That is, there is ϵ_θ and r_θ such that $\|Q^2_{\theta+it}\| \le (1 - \epsilon_\theta)$ for all $|t| > r_\theta$. Therefore,

$$\|\mathcal{L}_{\theta+it}^N\| = \lambda(\theta)^N \|G_\theta Q_{\theta+it}^N G_\theta^{-1}\| \le \lambda(\theta)^N \|G_\theta\| \|Q_{\theta+it}^N\| \|G_\theta^{-1}\| \le C\lambda(\theta)^N (1-\epsilon_\theta)^{\lfloor N/2 \rfloor}.$$

This establishes (B5).

Since the rate in (B5) is exponential and Theorem 4.2.2 holds for $(0, B)$, we conclude that for all $r \in \mathbb{N}$, these Markov chains admit weak expansions for large deviations of order r in the range $(0, B)$ for $F_{r+1,B+}^3$ where $B+ = \infty$, if $B = \infty$ and $B+ > B$, if $B < \infty$.

We need a stronger assumption on h to establish $(\widetilde{\text{B5}})$. Suppose,

$$\text{For all } x, y \text{ critical points of } z \mapsto (h(x,z) + h(z,y)) \text{ are non-degenerate.} \tag{4.17}$$

Since critical points of $z \mapsto (h(x,z) + h(z,y))$ are non-degenerate we can use the stationary phase asymptotics in [48, Chapter VIII.2], to obtain,

$$\Big| \int_{\mathcal{M}} e^{it(h(x,z)+h(z,y))} p(x,z) p(z,y) e^{\theta(h(x,z)+h(z,y))} \, dz \Big| \le \frac{M}{|t|^{d/2}}$$

where M is a constant and d is the dimension of $\mathcal{M}$. Therefore, $\|\mathcal{L}_{\theta+it}^2\| \le \dfrac{M}{|t|^{d/2}}$. Choose $K = (2M)^{2/d}$. Then for all $|t| > K$, $\|Q_{\theta+it}^2\| \le \dfrac{1}{2}$ and hence,

$$\|\mathcal{L}_{\theta+it}^N\| \le \|\mathcal{L}_{\theta+it}^{N-2}\| \|\mathcal{L}_{\theta+it}^2\| \le \Big(\frac{1}{2}\Big)^{\lfloor (N-2)/2 \rfloor} \frac{M}{t^{d/2}}, \ |t| > K.$$

By convexity, $\lambda(\theta) > 1$. Thus,

$$\|\mathcal{L}_{\theta+it}^N\| \le M \Big(\frac{1}{2}\Big)^{\lfloor (N-2)/2 \rfloor} \frac{\lambda(\theta)^N}{t^{d/2}}, \ |t| > K.$$

This establishes $(\widetilde{\text{B5}})$.

In particular, when h depends only on one variable, i.e. $h(x, y) = H(x)$ for some H, we have that $h(x, z) + h(z, y) = H(x) + H(z)$. Then, the condition (4.17) reduces to critical points of H being non-degenerate.

Again, because Theorem 4.2.2 hold for all $(0, B)$ and the rate in $(\widetilde{\mathrm{B5}})$ is exponential, we conclude that these strongly ergodic Markov chains admit strong expansions for large deviations of all orders in the range $(0, B)$.

Chapter A: Appendix

A.1 Convergence of $\mathcal{X}$.

We need some background information. Given a piecewise smooth function $g : \mathbb{R}^d \to \mathbb{R}$ of compact support its Siegel transform is a function on the space of lattices defined by

$$\mathcal{S}(g)(\mathcal{L}) = \sum_{\mathbf{w}\in\mathcal{L}\setminus\{\mathbf{0}\}} g(\mathbf{w}).$$

We need an identity of Siegel, see ([38, Section 3.7] or [46, Lecture XV]) saying that

$$\mathbf{E}_{\mathcal{L}}(\mathcal{S}(g)) = \int_{\mathbb{R}^d} g(\mathbf{w})d\mathbf{w}. \tag{A.1}$$

In particular, if B is a set in $\mathbb{R}^d$ with piecewise smooth boundary not containing $\mathbf{0}$ then

$$\mathbf{P}_{\mathcal{L}}(\mathcal{L}\cap B \neq \emptyset) \leq \mathbf{P}(\mathcal{S}(\mathbb{1}_B)(\mathcal{L}) \geq 1) \leq \mathbf{E}_{\mathcal{L}}(\mathcal{S}(\mathbb{1}_B)) = \mathrm{Vol}(B). \tag{A.2}$$

Proof of Lemma 2.1.2. Let $\mathcal{L}^+ = \{\mathbf{w} \in \mathcal{L} : y(\mathbf{w}) > 0\}$. Since $\dfrac{\sin(2\pi\chi(\mathbf{w}))}{y(\mathbf{w})}$ is even it is enough to restrict the attention to $\mathbf{w} \in \mathcal{L}^+$.

Throughout the proof we fix two numbers $\varepsilon > 0, \tau < 1$ such that $\varepsilon \ll (1-\tau) \ll 1$. It is easy to see using (A.2) and Borel-Cantelli Lemma that for almost every lattice

$\mathcal{L}$, there exists C and β such that $y(\mathbf{w}) > \dfrac{C}{\|\mathbf{w}\|^\beta}$. It follows that

$$\sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\geq||\mathbf{w}||^\varepsilon} \frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})} e^{-||x(\mathbf{w})||^2} \leq \sum_{\mathbf{w}\in\mathcal{L}^+} C||\mathbf{w}||^\beta e^{-||\mathbf{w}||^{2\varepsilon}}$$

converges absolutely. Hence it suffices to establish the convergence of

$$\bar{\mathcal{X}} := \sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\leq||\mathbf{w}||^\varepsilon<R^\varepsilon} \frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})} e^{-||x(\mathbf{w})||^2}.$$

Let $R_{j,k} = 2^k + j2^{\tau k}$, $j = 0, \ldots 2^{(1-\tau)k}$. To prove the convergence of $\bar{\mathcal{X}}$ we will show that for all $\mathcal{L}$ almost all χ satisfy two estimates below

$$\forall \text{ sequence } \{j_k\}\ \bar{\mathcal{X}}_{R_{j_k,k}} \text{ converges as } k\to\infty, \tag{A.3}$$

$$\max_j \sup_{R_{j,k}\leq R\leq R_{j+1,k}} \left|\bar{\mathcal{X}}_R - \bar{\mathcal{X}}_{j,k}\right| \to 0 \text{ as } k\to\infty. \tag{A.4}$$

To prove (A.3) let

$$S_{j,k} = \sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\leq||\mathbf{w}||^\varepsilon,\ R_{j,k}\leq||\mathbf{w}||\leq R_{j+1,k}} \frac{\sin 2\pi\chi(\mathbf{w})}{y(\mathbf{w})} e^{-||x(\mathbf{w})||^2}.$$

Using that $\mathbf{E}_\chi(\sin(2\pi(\chi(\mathbf{w})))) = 0$ and for $\mathbf{w}_1 \neq \pm\mathbf{w}_2$ we have

$$\mathbf{E}_\chi(\sin(2\pi(\chi(\mathbf{w}_1)))\sin(2\pi(\chi(\mathbf{w}_2)))) = 0$$

we see that $\mathbf{E}_\chi(S_{j,k}) = 0$ and

$$\begin{aligned}
\mathrm{Var}_\chi(S_{j,k}) &= \sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\leq||\mathbf{w}||^\varepsilon, R_{j,k}\leq||\mathbf{w}||\leq R_{j+1,k}} \frac{e^{-2\|\mathbf{x}(\mathbf{w})\|^2}}{2y^2(\mathbf{w})} \\
&\leq \frac{1}{2^{2k+1}}\mathrm{Card}(\mathbf{w} : ||x(\mathbf{w})|| \leq ||\mathbf{w}||^\varepsilon, R_{j,k} \leq ||\mathbf{w}|| \leq R_{j+1,k}) \\
&\leq \frac{C(\mathcal{L})}{2^{2k}}\mathrm{Vol}(\mathbf{w} : ||x(\mathbf{w})|| \leq ||\mathbf{w}||^\varepsilon, R_{j,k} \leq ||\mathbf{w}|| \leq R_{j+1,k}) \\
&\leq C(\mathcal{L})2^{(\tau+\varepsilon(d-1)-2)k}.
\end{aligned}$$

Hence by Chebyshev inequality for each j

$$\mathbf{P}_\chi\left(S_{j,k} \geq 2^{-(1-\tau+\varepsilon)k}\right) \leq C(\mathcal{L})2^{(\varepsilon d-\tau)k}$$

and so

$$\mathbf{P}_\chi\left(\exists j:\ S_{j,k} \geq 2^{-(1-\tau+\varepsilon)k}\right) \leq C(\mathcal{L})2^{(1+\varepsilon d-2\tau)k}.$$

Thus if ε is sufficiently small and τ is sufficiently close to 1 then Borel-Cantelli Lemma shows that for almost every χ, if k is large enough, then for all j $S_{j,k} \leq 2^{-(1-\tau+\varepsilon)k}$ and thus $\sum_j S_{j,k} \leq 2^{-\varepsilon k}$ proving (A.3). Likewise,

$$\begin{aligned}
&\sup_{R_{j,k}\leq R\leq R_{j+1,k}} \left|\bar{\mathcal{X}}_R - \bar{\mathcal{X}}_{j,k}\right| \\
&\qquad \leq \sum_{\mathbf{w}\in\mathcal{L}^+:\ ||x(\mathbf{w})||\leq||\mathbf{w}||^\varepsilon, ||\mathbf{w}||\in[R_{j,k},R_{j+1,k}]} \frac{1}{|y(\mathbf{w})|}\, e^{-||x(\mathbf{w})||^2} \\
&\qquad \leq C(\mathcal{L})2^{-2k}\mathrm{Vol}(\mathbf{w} : ||x(\mathbf{w})|| \leq ||\mathbf{w}||^\varepsilon, R_{j,k} \leq ||\mathbf{w}|| \leq R_{j+1,k}) \\
&\qquad \leq \bar{C}(\mathcal{L})2^{\tau+\varepsilon(d-1)-1}
\end{aligned}$$

proving (A.4). Lemma 2.1.2 is established. □

A.2 Hierarchy of Expansions.

In the discussion below, we do not assume the abstract setting introduced in section 3.1. Therefore the hierarchy of asymptotic expansions provided here holds true in general.

We observe that the classical Edgeworth expansion is the strongest form of asymptotic expansion among the expansions for non-lattice random variables. The following proposition and remark A.2.1 establish this fact.

Proposition A.2.1. *Suppose S_N admits order r Edgeworth expansions, then it also admits order r weak global expansion for $f \in F_0^1$ and order r averaged expansions for $f \in L^1$. Further, if the polynomials P_p in the Edgeworth expansion has opposite parity as p then S_N admits order $r-1$ weak local expansion for $f \in F_r^1$.*

Remark A.2.1. *Section 3.5.2 contains examples for which the weak and averaged forms of expansions exist but the strong expansion does not. Therefore none of the above implications are reversible.*

Proof of Proposition A.2.1. Suppose $f \in F_0^1$. Let $F_n = \mathbb{P}\Big(\dfrac{S_n - nA}{\sqrt{n}} \leq x\Big)$ and put

$$\mathcal{E}_{r,n}(x) = \mathfrak{N}(x) + \sum_{p=1}^{r} \frac{P_p(x)}{n^{p/2}} \mathfrak{n}(x).$$

Observe that $F_n(x) - \mathcal{E}_n(x) = o(n^{-r/2})$ uniformly in x and,

$$d\mathcal{E}_{r,n}(x) = \mathfrak{n}(x)\, dx + \sum_{p=1}^{r} \frac{1}{n^{p/2}} \left[P_p'(x)\, \mathfrak{n}(x) + P_p(x)\mathfrak{n}'(x)\right] dx = \sum_{p=0}^{r} \frac{R_p(x)}{n^{p/2}} \mathfrak{n}(x)\, dx$$

where R_p are polynomials given by $R_p = P_p' + P_p Q$ and Q is such that $\mathfrak{n}'(x) =$

$Q(x)\mathfrak{n}(x)$. Next, we observe that,

$$
\begin{aligned}
\mathbb{E}(f(S_n - nA)) = \mathbb{E}\Big(f\Big(\frac{S_n - nA}{\sqrt{n}}\sqrt{n}\Big)\Big) &= \int f(x\sqrt{n})\,dF_n(x) \\
&= \int f(x\sqrt{n})\,d\mathcal{E}_{r,n}(x) + \int f(x\sqrt{n})\,d(F_n - \mathcal{E}_{r,n})(x).
\end{aligned}
$$

Now we integrate by parts and use $\mathcal{E}_{r,n}(\infty) = F_n(\infty) = 1$ and $\mathcal{E}_{r,n}(-\infty) = F_n(-\infty) = 0$ to obtain,

$$
\begin{aligned}
\mathbb{E}(f(S_n - nA)) &= \int f(x\sqrt{n})\,d\mathcal{E}_{r,n}(x) + (F_n - \mathcal{E}_{r,n})(x) f(x\sqrt{n})\Big|_{-\infty}^{\infty} \\
&\qquad - \int (F_n - \mathcal{E}_{r,n})(x)\sqrt{n} f'(x\sqrt{n})\,dx \\
&= \int \sum_{p=0}^{r} \frac{1}{n^{p/2}} R_p(x)\mathfrak{n}(x)\, f(x\sqrt{n})dx + o\left(n^{-r/2}\right) \int \sqrt{n} f'(x\sqrt{n})\,dx \\
&= \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int R_p(x)\mathfrak{n}(x)\, f(x\sqrt{n})dx + o\left(n^{-r/2}\right).
\end{aligned}
$$

This is the order r weak global Edgeworth expansion. The existence of the order $r-1$ weak local expansion follows from this. This is our next theorem. So we postpone its proof.

For $f \in L^1$ substituting x by $x + \dfrac{y}{\sqrt{n}}$ in the Edgeworth expansion for S_n we have

$$
\begin{aligned}
\mathbb{P}\left(\frac{S_n - nA}{\sqrt{n}} \le x + \frac{y}{\sqrt{n}}\right) - \mathfrak{N}\left(x + \frac{y}{\sqrt{n}}\right) \\
= \sum_{p=1}^{r} \frac{1}{n^{p/2}} P_p\left(x + \frac{y}{\sqrt{n}}\right) \mathfrak{n}\left(x + \frac{y}{\sqrt{n}}\right) + o\left(n^{-r/2}\right).
\end{aligned}
$$

For fixed x, the error is uniform in y. Therefore, multiplying the equation by $f(y)$ and then integrating we can conclude that the order r averaged expansion exists. $\square$

Remark A.2.2. *We have seen from the derivation of the Edgeworth expansion in section 3.2 that $P_p(x)$ and p have opposite parity in the weakly dependent case. This* *implies that $P_{p,g}$ has the same parity as p. This is true in the i.i.d. case as well. Even though this assumption may look artificial in the general case, it is reasonable. When using characteristic functions to derive the expansions, one is likely to end up with Hermite polynomials which is the reason behind the parity relation.*

Next, we compare the the relationships among the weak and averaged forms of Edgeworth expansions.

Proposition A.2.2. *Suppose S_N admits order r weak global Edgeworth expansion for $f \in F_r^{q+1}$ for some $q \geq 0$. If the polynomials $P_{p,g}$ in the global Edgeworth expansion has the same parity as p then S_N admits order $r-1$ weak local expansion for f.*

Proof. Assume, $f \in F_r^1$. Then, from the Plancherel formula,

$$\int_{\mathbb{R}} \sqrt{n} f\big(x\sqrt{n}\big) P_{p,g}(x)\mathfrak{n}(x)\, dx = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big) A_p(t) e^{-\frac{\sigma^2 t^2}{2}}\, dt$$

where $A_p(t)$ are polynomials constructed using the following relation,

$$P_{p,g}(t)e^{-\frac{t^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} A_p\left(-i\frac{d}{dt}\right)\left[e^{-\frac{t^2}{2\sigma^2}}\right].$$

By construction $P_{p,g}$ and A_p has the same parity. This means A_p has the same parity as p.

First replace

$$\int P_{p,g}(x)\mathfrak{n}(x)\, f(x\sqrt{n})dx$$

by

$$\frac{1}{2\pi\sqrt{n}} \int_{\mathbb{R}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big) A_p(t) e^{-\frac{\sigma^2 t^2}{2}}\, dt$$

in the weak global expansion to obtain,

$$\sqrt{n}\mathbb{E}(f(S_n - nA)) = \frac{1}{2\pi}\sum_{p=0}^{r}\frac{1}{n^{p/2}}\int_{\mathbb{R}}\widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)A_p(t)e^{-\frac{\sigma^2 t^2}{2}}\,dt + o\left(n^{-(r-1)/2}\right).$$

Then substituting for $\widehat{f}$ with its order $r-1$ Taylor expansion,

$$\sqrt{n}\mathbb{E}(f(S_n - nA)) \;=\; \frac{1}{2\pi}\sum_{p=0}^{r}\sum_{j=0}^{r-1}\frac{\widehat{f}^{(j)}(0)}{j!n^{(j+p)/2}}\int_{\mathbb{R}} t^j e^{-\sigma^2 t^2/2}A_p(t)\,dt + o\left(n^{-(r-1)/2}\right).$$

Put

$$a_{pj} = \int_{\mathbb{R}} t^j e^{-\sigma^2 t^2/2}A_p(t)\,dt = 0 \quad\text{and}\quad f^{(j)}(0) = \int (-it)^j f(t)\,dt$$

to get,

$$\sqrt{n}\mathbb{E}(f(S_n - nA)) = \frac{1}{2\pi}\sum_{p=0}^{r}\sum_{j=0}^{r-1}\frac{a_{pj}}{j!n^{(j+p)/2}}\int_{\mathbb{R}}(-it)^j f(t)\,dt + o\left(n^{-(r-1)/2}\right)$$

Since p and A_p are of the same parity, when $j+p$ is odd. $a_{pj} = 0$. So we collect terms such that $p+j = 2k$ where $k = 0, \dots, r-1$ and write,

$$P_{k,w} = \sum_{p+j=2k}\frac{a_{pj}}{j!}(-it)^j$$

Then, rearranging, simplifying and absorbing higher order terms to the error, we obtain,

$$\sqrt{n}\mathbb{E}(f(S_n - nA)) = \frac{1}{2\pi}\sum_{k=0}^{\lfloor (r-1)/2\rfloor}\frac{1}{n^k}\int_{\mathbb{R}}P_{k,w}(t)f(t)\,dt + o\left(n^{-(r-1)/2}\right)$$

which is the order $r-1$ weak local Edgeworth expansion. $\square$

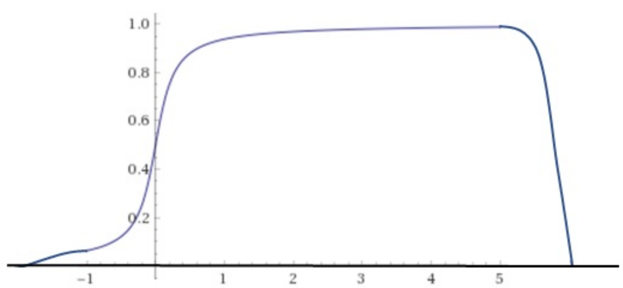
A.3 Construction of $\{f_k\}$.

For each k, let $f_k(x) = \dfrac{1}{\pi}\tan^{-1}(kx) + \dfrac{1}{2}$ for $x \in [-1, k]$. Extend f_k to $[-2, k+1]$ in such a way that $f_k(-2) = f_k(k+1) = 0$, f_k is continuously differentiable and satisfying the following conditions.

1. f_k is increasing on $[-2, k]$ with derivative on $[-2, -1]$ is bounded above by 1.
2. f_k is decreasing on $[k + 1/2, k + 1]$ with derivative bounded below by -5.
3. $|f_k'| \leq 5$ on $[k, k + 1]$.
4. $0 \leq f_k \leq 1$ on $[-2, k + 1]$ and $f_k = 0$ elsewhere.

Then, f_k is supported on $[-2, k + 1]$. Here our choice of bounds 1 and -5 in some sense arbitrary. As long as they are large enough and independent of k, we obtain an appropriate sequence of functions.

As an example, when $k = 5$, the graph of f_5 looks like:



For all $\gamma > 0$,

$$\int |(f_k)_\gamma(x)|\,dx = \int |e^{-\gamma x} f_k(x)|\,dx \leq \int_{-2}^{\infty} e^{-\gamma x}\,dx = C_{\gamma,1} < \infty$$

because $0 \leq f_k \leq 1$.

Since $|f_k'| \le 5$ on $[k, k+1]$, $0 \le f_k \le 1$ and f_k is increasing on $[-2, k]$,

$$\begin{aligned}
\int |((f_k)_\gamma)'(x)| dx &= \int_{-2}^{k+1} |\gamma e^{-\gamma x} f_k(x) + e^{-\gamma x} f_k'(x)|\, dx \\
&\le \int_{-2}^{k+1} \left(\gamma e^{-\gamma x} f_k(x) + e^{-\gamma x} |f_k'(x)|\right) dx \\
&\le \int_{-2}^{k} \gamma e^{-\gamma x}\, dx + \int_{-1}^{k} f_k'(x)\, dx + \int_{k}^{k+1} (\gamma e^{-\gamma x} + 5 e^{-\gamma x})\, dx \\
&\le 1 + \int_{-2}^{k+1} (5+\gamma) e^{-\gamma x}\, dx = C_{\gamma,2} < \infty
\end{aligned}$$

Also, note that $|x^l f_k(x)| \le x^l e^{-\gamma x}$ for all $x \in [-2, k+1]$. Hence,

$$\int |x^l f_k(x)|\, dx \le \int_{-2}^{\infty} x^l e^{-\gamma x}\, dx = J_{\gamma,l} < \infty$$

Put $J_r(\gamma) = \max\limits_{1 \le l \le r} J_{\gamma,l}$ and $C_\gamma(r) = \max\{J_r(\gamma), C_{\gamma,1}, C_{\gamma,2}\}$. Then, $C_\gamma(r)$ is finite and depends only on γ and r.

Now, we have the following,

1. $C^1_{r+1}((f_k)_\gamma) \le C_\gamma(r)$ for all k.
2. Since, $\dfrac{1}{\pi} \tan^{-1}(kx) + \dfrac{1}{2}$ converges pointwise to $1_{[0,\infty)}(x)$, it is easy to see that $f_k \to 1_{[0,\infty)}$ pointwise.
3. Since for each p, $e^{-\gamma z} P_p(z) f_k(z)$ converges pointwise to $e^{-\gamma z} P_p(z) 1_{[0,\infty)}(z)$, $e^{-\gamma z} |P_p(z)| 1_{[-2,\infty)}$ is integrable and $|e^{-\gamma z} P_p(z) f_k(z)| \le e^{-\gamma z} |P_p(z)| 1_{[-2,\infty)}$, we can apply the LDCT to conclude,

$$\int P_p(z) g_k(z)\, dz = \int_{-2}^{\infty} e^{-\gamma z} P_p(z) f_k(z)\, dz \to \int_0^{\infty} e^{-\gamma z} P_p(z)\, dz.$$

Bibliography

[1] Bahadur, R. R., Ranga Rao R.; *On Deviations of the Sample Mean*. Ann. Math. Statist. **31** (1960), no. 4, 1015-1027.

[2] Birkhoff, G. *Extensions of Jentzschs theorem*. Trans. Amer. Math. Soc. **85** (1957), no. 1, 219–227.

[3] Bhattacharya, R. N., Ranga Rao R.; Normal Approximation and Asymptotic Expansions, first edition, John Wiley and Sons, 1976, xiv+274 pp.

[4] Breuillard, E. *Distributions diophantiennes et theoreme limite local sur* $\mathbb{R}^d$. Probab. Theory Related Fields **132** (2005), no. 1, 39–73.

[5] Butterley, O., Eslami, P. *Exponential mixing for skew products with discontinuities*. Trans. Amer. Math. Soc. **369** (2017), no. 2, 783–803.

[6] Bougerol, P., Lacroix, J.; Products of random matrices with applications to Schrödinger operators, Progress in Probability and Statistics, first edition, Birkhäuser Basel, Boston, 1985, xi+284 pp.

[7] Chaganty, N. R., Sethuraman, J., *Strong Large Deviation and Local Limit Theorems*, Ann. Probab. **21** (1993), no. 3, 1671-1690.

[8] Chebyshev, P. L. *Sur le développement des fonctions à une seule variable*. Bull, de l'Acad. Imp. des Sci. de St. Petersbourg **3(1)** (1860), 193-200.

[9] Coelho, Z., Parry, W. *Central limit asymptotics for shifts of finite type*. Israel J. Math. **69** (1990), no. 2, 235–249.

[10] Cramér, H. *On the composition of elementary errors*. Skand. Aktuarietidskr. **1** (1928), 13–74; 141–180.

[11] Cramér, H.; Random variables and probability distributions. Cambridge Tracts in Mathematics no. 36, Cambridge, 1937, 122 pp.

[12] Dolgopyat, D. *A Local Limit Theorem for sum of independent random vectors,* Electronic J. Prob. **21** (2016) paper 39.

[13] Dolgopyat, D. *On mixing properties of compact group extensions of hyperbolic systems.* Israel J. Math. **130** (2002), 157–205.

[14] Dolgopyat D., Fayad B. *Deviations of ergodic sums for toral translations: Convex bodies,* GAFA **24** (2014) 85–115.

[15] Dolgopyat D., Fayad B. *Limit theorems for toral translations,* Proc. Sympos. Pure Math **89** (2015) 227–277.

[16] Dolgopyat, D., Fernando, K. *An error term in the Central Limit Theorem for sums of discrete random variables.* preprint.

[17] Dembo, A., Zeitouni O.; Large Deviations Techniques and Applications, second edition, Springer–Verlag Berlin Heidelberg, 2010, XVI+396.

[18] Eskin A., McMullen C. *Mixing, counting, and equidistribution in Lie groups,* Duke Math. J. **71** (1993) 181–209.

[19] Esséen, C.–G. *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law,* Acta Math. **77** (1945) 1–125.

[20] Feller, W.; An introduction to probability theory and its applications Vol. II., Second edition, John Wiley & Sons, Inc., New York-London-Sydney, 1971, xxiv+669.

[21] Fernando, K., Liverani, C. *Edgeworth expansions for weakly dependent random variables.* `arXiv:1803.07667 [math.PR]`.

[22] Götze F., Hipp C. *Asymptotic Expansions for sums of Weakly Dependent Random Vectors*, Z. Wahrscheinlickeitstheorie verw., **64** (1983) 211-239.

[23] Gnedenko B.V., Kolmogorov A.N.; Limit distributions for sums of independent random variables, Trans. K.L. Chung, Revised edition, Addison-Wesley, 1968, ix+264 pp.

[24] Gouëzel, S. *Limit theorems in dynamical systems using the spectral method. Hyperbolic dynamics, fluctuations and large deviations*, Proc. Sympos. Pure Math., 89 (2015) 161–193, AMS, Providence, RI.

[25] Guivarc'h, Y., Hardy J. *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov*, Annales de l'I.H.P. Probabilités et statistiques, **24 (1)** (1988) 73-98.

[26] Hall, P. *Contributions of Rabi Bhattacharya to the Central Limit Theory and*

Normal Approximation. In Manfred Denker & Edward C. Waymire (Eds.), Rabi N. Bhattacharya Selected Papers, (pp 3–13). Birkhäuser Basel, 2016.

[27] Hennion, H., Hervé, L.; Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness, Lecture Notes in Mathematics, first edition, Springer-Verlag, Berlin Heidelberg, 2001, viii+125 pp.

[28] Hebbar, P., Nolen, J. *The asymptotics of solutions to parabolic PDE with periodic coefficients*, preprint.

[29] Hervé, L., Pène, F. *The Nagaev-Guivarc'h method via the Keller-Liverani theorem*, Bull. Soc. Math. France **138** (2010) no. 3, 415–489.

[30] den Hollander, F.; Large Deviations, Fields Institute Monographs 14, American Mathematical Society, Providence, RI, 2000, x+142 pp.

[31] Ibragimov, I. A., Linnik, Y. V.; Independent and stationary sequences of random variables. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov. Translation from the Russian edited by J. F. C. Kingman. Wolters-Noordhoff Publishing, Groningen, 1971, 443 pp.

[32] Joutard, C. *Strong large deviations for arbitrary sequences of random variables*, Ann. Inst. Stat. Math. **65** (2013) no. 1, 49-67.

[33] Kato, T.; Perturbation theory for linear operators, Classics in Mathematics, Reprint of the 1980 edition, Springer-Verlag, Berlin, 1995, xxii+619.

[34] Kesten H. *Uniform distribution* mod 1, part I: Ann. of Math. **71** (1960) 445–471, part II: Acta Arith. **7** (1961/1962) 355–380.

[35] Kleinbock D. Y., Margulis G. A. *Bounded orbits of nonquasiunipotent flows on homogeneous spaces,* AMS Transl. **171** (1996) 141–172.

[36] Kleinbock D. Y., Margulis G. A. *Logarithm laws for flows on homogeneous spaces,* Invent. Math. **138** (1999) 451–494.

[37] Liverani, C. *Decay of correlations for piecewise expanding maps.* J. Statist. Phys. **78** (1995), no. 3-4, 1111–1129.

[38] Marklof J. *The n-point correlations between values of a linear form,* Erg. Th., Dynam. Sys. **20** (2000) 1127–1172.

[39] Marklof J., Strombergsson A. *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems,* Ann. Math. **172** (2010) 1949–2033.

[40] Nagaev S. V. *More Exact Statement of Limit Theorems for Homogeneous Markov Chain*, Theory Probab. Appl., **6(1)** (1961) 62–81.

[41] Nagaev S. V. *Some Limit Theorems for Stationary Markov Chains*, Theory Probab. Appl., **2(4)** (1959) 378–406.

[42] Pène F. *Mixing and decorrelation in infinite measure: the case of periodic Sinai billiard*, `arXiv:1706.04461v1 [math.DS]`.

[43] Rubin H., Sethuraman J. *Probabilities of moderate deviations*, Sankhya Ser. A, **27** (1965) 325–346.

[44] Rubin H., Sethuraman J. *Bayes risk efficiency*, Sankhya Ser. A, **27** (1965) 347–356.

[45] Shah N. A. *Limit distributions of expanding translates of certain orbits on homogeneous spaces,* Proc. Indian Acad. Sci. Math. Sci. **106** (1996) 105–125.

[46] Siegel C. L.; Lectures on the geometry of numbers, Springer, Berlin, 1989. x+160 pp.

[47] Sprindzuk V. G.; Metric theory of Diophantine approximations, Scripta Series in Math. V. H. Winston & Sons, Washington, D.C. 1979. xiii+156 pp.

[48] Stein E. M.; Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton University Press, first edition, 1993. xiv+695 pp.