The Central Limit Theorem: Generalization, Localization, and Application to Random Walks

James Chen

4/9/2021

1 Introduction and Definitions

In the following paper, we examine three forms of the central limit theorem: the classical version, the generalized version with the finite variance assumption weakened, and the discrete localized version. We also apply the localized version to random walks to prove recurrence in dimensions of 2 or less and transience in dimensions of 3 or more.

Definition 1.1 We define $(\Omega, \mathcal{F}, P)$ to be a *probability space*, where Ω is a *sample space*, P is a *probability measure* (i.e. $P(\Omega) = 1$), and $\mathcal{F}$ is the *event space* (defined to be the σ-algebra of some P-measurable sets of Ω). In the following section, we will take Ω to be the open interval $(0, 1)$, $\mathcal{F}$ to be the Borel sets (smallest σ-algebra containing the open intervals of $(0, 1)$), and P to be the Lebesgue measure.

Definition 1.2 A real valued function X defined on Ω is said to be a *random variable* if for every Borel set $B \subseteq \mathbb{R}$ we have

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

Definition 1.3. If X is a random variable, then X induces a probability measure on $\mathbb{R}$ called its *distribution* by setting

$$\mu(B) := P(X^{-1}(B))$$

The distribution of a X is usually described by giving its *(cumulative) distribution function*,

$$F(x) := P\left(X^{-1}((-\infty, x])\right) = P(\{\omega \in \Omega : X(\omega) \leq x\})$$

We donote $\mu(B) = P(X \in B)$ for convenience. For example, we denote $F(x) = P(X \leq x)$ for convenience.

Definition 1.4 If $X \geq 0$ is a random variable on $(\Omega, \mathcal{F}, P)$, then we define its *expected value* to be the the Lebesgue-Stieltjes integral

$$\mathbb{E}[X] := \int_{\Omega} X \; dP(\omega)$$

which always makes sense, but may be ∞. If $\mathbb{E}[X^2] < \infty$ then the *variance* of X is defined to be

$$\mathrm{var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2]$$

Definition 1.5. If X is a random variable, then we define its *characteristic function* by

$$\varphi(t) := \mathbb{E}[\exp(itX)]$$

The characteristic function can thus be seen as the Fourier transform of the probability density function, if X admits one.

Definition 1.6. A sequence of distribution functions $\{F_n\}$ is said to *converge weakly* to a limit F (written $F_n \implies F$)) if $F_n(y) \to F(y)$ for all y that are continuity points of F. A sequence of random variables $\{X_n\}$ is said to *converge weakly* or *converge in distribution* to a limit X_∞ (written $X_n \implies X_\infty$) if their distribution functions converge weakly.

Definition 1.7 Two events A and B are *independent* if

$$P(A \setminus B) = P(A)P(B)$$

Two random variables X and Y are *independent* if for all Borel sets C, D,

$$P(X \in C, Y \in D) = P(X \in C)P(Y \in D)$$

i.e., the events $A = \{X \in C\}$ and $B = \{Y \in D\}$ are independent.

2 The Central Limit Theorem

Theorem 1. *If F is non-decreasing, cadlag (right-continuous), and satisfies both $\lim\limits_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$, then it is the distribution function of some random variable.*

Proof. (Adapted from p. 11 of Durrett [1]) For $\omega \in \Omega$, let

$$X(\omega) = \sup\{y : F(y) < \omega\}$$

Once we show that

$\{\omega : X(\omega) \le x\} = \{\omega : \omega \le F(x)\}$

the desired result follows immediately since $P(\omega : \omega \le F(x)) = F(x)$. We observe that if $\omega \le F(x)$ then $X(\omega) \le x$, since $x \notin \{y : F(y) < \omega\}$. On the other hand if $\omega > F(x)$, then since F is right continuous, there is an $\varepsilon > 0$ such that $F(x+\varepsilon) < \omega$ and $X(\omega) \ge x + \varepsilon > x$.

Theorem 2 (Skorokhod's Representation Theorem). *If $F_n \implies F$ then there are random variables Y_n and Y with distributions F_n and F so that $Y_n \to Y$ a.s.*

Proof. (Adapted from p. 118 of Durrett [1]) Let $Y_n(x) = F_n^{\rightarrow}(x) := \sup\{y : F_n(y) < x\}$ ($n = \infty$ corresponds to F and Y), where $F_n^{\rightarrow}$ is called the "right inverse". By Theorem 1, Y_n has distribution F_n.

We begin by identifying the exceptional set. Let $a_x = \sup\{y : F(y) < x\}$, $b_x = \inf\{y : F(y) > x\}$, and $\Omega_0 = \{x : (a_x, b_x) = \emptyset\}$ where (a_x, b_x) is the open interval with the indicated endpoints. Then $\Omega \setminus \Omega_0$ is countable since the (a_x, b_x) are disjoint and each non-empty interval contains a different rational number.

If $x \in \Omega_0$ then $F(y) < x$ for $y < F^{\rightarrow}(x)$ and $F(z) > x$ for $z > F^{\rightarrow}(x)$. Let $y < F^{\rightarrow}(x)$ be such that F is continuous at y. Since $x \in \Omega_0$, $F(y) < x$ and if n is sufficiently large $F_n(y) < x$, i.e., $F_n^{\rightarrow}(x) \ge y$. Since this holds for all y satisfying the indicated restrictions, we have $\liminf_{n\to\infty} F_n^{\rightarrow}(x) \ge F^{\rightarrow}(x)$. Similarly, $\limsup_{n\to\infty} F_n^{\rightarrow}(x) \le F^{\rightarrow}(x)$, and we have that $F_n^{\rightarrow}(x) \to F^{\rightarrow}(x)$ for $x \in \Omega_0$.

Theorem 3 *$X_n \implies X_\infty$ iff, for every continuous function g of compact support, we have $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X_\infty)]$.*

Proof of $\implies$. (Adapted from p. 119 of Durrett [1]) Consider the Skorokhod representations Y_n (from Skorokhod's Representation Theorem) of the distributions of X_n. That is to say, let Y_n have the same distribution as X_n and converge a.s. Since g is continuous on compact support, g is bounded (by the Extreme Value Theorem) and $g(Y_n) \to g(Y)$ a.s. The bounded convergence theorem then implies

$$\mathbb{E}[g(X_n)] = \mathbb{E}[g(Y_n)] \to \mathbb{E}[g(Y)] = \mathbb{E}[g(X)]$$

Proof of $\impliedby$. (Adapted from p. 63 of Bhattacharya and Waymire [2]) Let f be an arbitrary bounded continuous function, $|f(x)| \le c$ for all x. Given $\varepsilon > 0$ there exists N such that $P(|X| \ge N) < \frac{\varepsilon}{4c}$. Let

$$g_N(x) = \begin{cases} 1 & x \le N \\ 0 & x \ge N+1 \\ \text{linear} & N \le x \le N+1 \end{cases}$$

Then, because $g_N(x) = 1$ for $x \le N$, g_N is continuous, and $g_N(x) = 0$ for $x > N+1$,

$$\liminf_{n\to\infty} P(|X_n| \le N+1) \ge \liminf_{n\to\infty} \mathbb{E}[g_N(X_n)] = \mathbb{E}[g_N(X)] \ge P(|X| \le N) > 1 - \frac{\varepsilon}{4c}$$

so that

$$\limsup_{n\to\infty} P(|X_n| > N+1) = 1 - \liminf_{n\to\infty} P(|X_n| \le N+1) < \frac{\varepsilon}{4c}$$

Now define $f_N := f \cdot g_{N+1}$. Note that $f = f_N$ on $\{x \le N+1\}$ and that on $\{x > N+1\}$ one has $|f(x)| \le c$. Since

$$f = f \cdot \mathbf{1}_{\{x\le N+1\}} + f \cdot \mathbf{1}_{\{x>N+1\}} \qquad \text{and} \qquad \mathbb{E}\left[f_N \cdot \mathbf{1}_{\{x\le N+1\}}\right] = \mathbb{E}[f_N] - \mathbb{E}\left[f_N \cdot \mathbf{1}_{\{N+1<x\le N+2\}}\right]$$

(and similarly for the integral with respect to P), one has

$$\begin{aligned}
\limsup_{n\to\infty} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\le \limsup_{n\to\infty} |\mathbb{E}[f_N(X_n)] - \mathbb{E}[f_N(X)]| \\
&\quad + \limsup_{n\to\infty} 2c \cdot P(|X_n| > N+1) + 2c \cdot P(|x| > N+1) \\
&< 2c\frac{\varepsilon}{4c} + 2c\frac{\varepsilon}{4c} \\
&= \varepsilon
\end{aligned}$$

By an ε of room, $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$.

Theorem 4 (Levy's Continuity Theorem). *Let $\{X_n\}$ be a sequence of random variables and X be an additional random variable. Then φ_{X_n} converges pointwise to φ_X iff X_n converges in distribution to X.*

Proof. (Adapted from p. 99 of Tao [3]) The reverse implication is immediate from the definitions of the characteristic function and convergence in distribution, since the function $x \mapsto \exp(sit \cdot x)$ is bounded and continuous.

Now suppose that $\varphi_{X_n}(t) \to \varphi_X(t)$ pointwise. We wish to show that $X_n \implies X$. By Theorem 3, it suffices to show that

$$\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X_n)]$$

for continuous, compactly supported functions $g : \mathbb{R} \to \mathbb{R}$. Using techniques similar to those of homework 3 (by convolving g with a Gaussian family of good kernels and then approximating the convolution with Riemann sums), we can estimate g uniformly by a linear combination of Gaussians f. Hence, it suffices to show the result for f (by the three-term estimate). From homework 3, we have (for linear combinations of Gaussians) the Fourier inversion formula

$$f(X_n) = \int_{\mathbb{R}} \widehat{f}(t) \exp(sit \cdot X_n)\ dt$$

where $\widehat{f}$ is a Schwartz function, and is in particular absolutely integrable. From the Fubini-Tonelli theorem, we thus have

$$\mathbb{E}f(X_n) = \int_{\mathbb{R}} \widehat{\varphi}(t) F_{X_n}(t)\ dt$$

and similarly for X. The claim now follows from the Lebesgue dominated convergence theorem.

Lemma 5. *We have the following, more potent estimate for the Lagrange remainder:*

$$\left| \exp(ix) - \sum_{m=0}^{n} \frac{(ix)^m}{m!} \right| \leq \min\left(\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right)$$

The first term on the right is the usual order of magnitude we expect in the correction term. The second is better for large $|x|$ and will help us prove the central limit theorem without assuming finite third moments.

Proof. (Adapted from p. 134 of Durrett [1]) Integrating by parts gives

$$\int_0^x (x-s)^n \exp(is)\ ds = \frac{x^{n+1}}{n+1} + \frac{i}{n+1} \int_0^x (x-s)^{n+1} \exp(is)\ ds$$

When $n = 0$, this says

$$\int_0^x \exp(is)\ ds = x + i \int_0^x (x-s) \exp(is)\ ds$$

The left-hand side evaluates to $\frac{(\exp(ix)-1)}{i}$, so rearranging gives

$$\exp(ix) = 1 + ix + i^2 \int_0^x (x-s) \exp(is)\ ds$$

Using the result for $n = 1$ now gives

$$\exp(ix) = 1 + ix + \frac{i^2 x^2}{2} + \frac{i^3}{2} \int_0^x (x-s)^2 \exp(is)\ ds$$

and iterating we arrive at **(a)**:

$$\exp(ix) - \sum_{m=0}^{n} \frac{(ix)^m}{m!} = \frac{i^{n+1}}{n!} \int_0^x (x-s)^n \exp(is)\ ds$$

Since $|\exp(is)| \leq 1$ for all s, we have **(b)**:

$$\left|\frac{i^{n+1}}{n!}\int_0^x (x-s)^n \exp(is)\ ds\right| \le \frac{|x|^{n+1}}{(n+1)!}$$

The last estimate is good when x is small. The next is designed for large x. Integrating by parts

$$\frac{i}{n}\int_0^x (x-s)^n \exp(is)\ ds = -\frac{x^n}{n} + \int_0^x (x-s)^{n-1}\exp(is)\ ds$$

Noticing $\frac{x^n}{n} = \int_0^x (x-s)^{n-1}\ ds$ now gives

$$\frac{i^{n+1}}{n!}\int_0^x (x-s)^n \exp(is)\ ds = \frac{i^n}{(n-1)!}\int_0^x (x-s)^{n-1}(\exp(is)-1)\ ds$$

and since $|\exp(ix)-1| \le 2$, it follows that **(c)**:

$$\left|\frac{i^{n+1}}{n!}\int_0^x (x-s)^n \exp(is)\ ds\right| \le \left|\frac{2}{(n-1)!}\int_0^x (x-s)^{n-1}\ ds\right| \le \frac{2|x|^n}{n!}$$

Combining **(a)**, **(b)**, and **(c)**, we have the desired result.

Theorem 6. *If $E[|X|^2] < \infty$ then*

$$\varphi(t) = 1 + it\mathbb{E}[X] - \frac{t^2}{2}\mathbb{E}[X^2] + o(t^2)$$

Proof. (Adapted from p. 136 of Durrett [1]) Taking expected values, using Jensen's inequality, and applying Lemma 4 to $x = tX$, gives

$$\left|\mathbb{E}[\exp(itX)] - \sum_{m=0}^{2}\mathbb{E}\left[\frac{(itX)^m}{m!}\right]\right| \le \mathbb{E}\left[\left|\exp(itX) - \sum_{m=0}^{2}\frac{(itX)^m}{m!}\right|\right] \le \mathbb{E}[\min(|tX|^3, 2|tX|^2)]$$

where in the second step we have dropped the denominators to make the bound simpler. The error term is thus $\le t^2\mathbb{E}[\min(|t|\cdot|X|^3, 2|X|^2)]$. The variable in parentheses is smaller than $2|X|^2$ and converges to 0 as $t \to 0$, so the desired conclusion follows from the dominated convergence theorem.

Lemma 7. *Let $z_1, \ldots, z_n$ and $w_1, \ldots, w_n$ be complex numbers of modulus $\le \theta$. Then*

$$\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| \le \theta^{n-1}\sum_{m=1}^{n}|z_m - w_m|$$

Proof. (Adapted from p. 144 of Durrett [1]) The result is true for $n = 1$. To prove it for $n > 1$ observe that

$$\begin{aligned}\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| &\le \left|z_1 \prod_{m=2}^{n} z_m - z_1 \prod_{m=2}^{n} w_m\right| + \left|z_1 \prod_{m=2}^{n} w_m - w_1 \prod_{m=2}^{n} w_m\right| \\ &\le \theta \left|\prod_{m=2}^{n} z_m - \prod_{m=2}^{n} w_m\right| + \theta^{n-1}|z_1 - w_1|\end{aligned}$$

and use induction.

Lemma 8. *If b is a complex number with $|b| \le 1$ then $|\exp(b) - (1 + b)| \le |b|^2$.*

Proof. (Adapted from p. 145 of Durrett [1])

$$\exp(b) - (1 + b) = \frac{b^2}{2!} + \frac{b^3}{3!} + \frac{b^4}{4!} + \dots$$

so if $|b| \le 1$ then

$$|\exp(b) - (1 + b)| \le \frac{|b|^2}{2}\left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) = |b|^2$$

Theorem 9. *If $c_n \to c \in \mathbb{C}$ then $(1 + \frac{c_n}{n})^n \to e^c$.*

Proof. (Adapted from p. 145 of Durrett [1]) Let $z_m = (1 + \frac{c_n}{n})$, $w_m = \exp(\frac{c_n}{n})$, and $\gamma > |c|$. For large n, $|c_n| < \gamma$. Since $1 + \frac{\gamma}{n} \le \exp(\frac{\gamma}{n})$, it follows from Lemmas 7 and 8 that

$$\left|\left(1 + \frac{c_n}{n}\right)^n - \exp(c_n)\right| \le \exp\left(\frac{\gamma}{n}\right)^{n-1} n \left|\frac{c_n}{n}\right|^2 \le \exp(\gamma)\frac{\gamma^2}{n} \to 0$$

as $n \to \infty$.

Theorem 10. *If X_1 and X_2 are independent and have characteristic function's φ_1 and φ_2 then $X_1 + X_2$ has characteristic function $\varphi_1(t)\varphi_2(t)$.*

Proof. (Adapted from p. 126 of Durrett [1])

$$\mathbb{E}[\exp(it(X_1 + X_2))] = \mathbb{E}[\exp(itX_1)\exp(itX_2))] = \mathbb{E}[\exp(itX_1)]\mathbb{E}[\exp(itX_2)]$$

since $\exp(itX_1)$ and $\exp(itX_2)$ are independent.

Theorem 11 (Central Limit Theorem). *Let $X_1, X_2, \dots$ be independent and identically distributed with $\mathbb{E}[X_i] = \mu$, var$(X_i) = \sigma^2 \in (0, \infty)$. If $S_n = X_1 + \cdots + X_n$ then*

$$\frac{(S_n - n\mu)}{\sigma\sqrt{n}} \implies N(0, 1)$$

where $N(0,1)$ denotes the standard Gaussian distribution with mean 0 and variance 1.

Proof. (Adapted from p. 134 of Durrett [1]) By considering $X_i' = X_i - \mu$, it suffices to prove the result when $\mu = 0$. From Theorem 6

$$\varphi(t) = \mathbb{E}[\exp(itX_1)] = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$$

so, by Theorem 10,

$$\mathbb{E}\left[\frac{\exp(itS_n)}{\sigma n^{1/2}}\right] = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

From Lemma 8 it should be clear that the last quantity $\to \exp(-\frac{t^2}{2})$ as $n \to \infty$, which, with Levy's Continuity Theorem, completes the proof.

3 A Generalization

The classical central limit theorem imposes three assumptions on the random variables to be aggregated: identical distribution, independence, and finite variance. The first can be relaxed under the Lyapunov and Lindeberg-Feller central limit theorems. We will not cover these theorems in this paper; treatments can be found on p. 148 in Durrett [1] and p. 99 in Bhattacharya and Waymire [2]. As of the time of writing, the second is a delicate condition that's spawned an active field of research with a long history. There are many contextual results, some of which are treated on p. 328 in Chow and Teicher [4] as well as on p. 334 in Ibragimov and Linnik [5].

We will be focusing on what happens when the third condition is relaxed. We will not prove the actual generalized central limit theorem, as it does not rely on techniques taught in this course and thus does not allow us to behold the power of Fourier methods. Instead, we will illuminate the form that the characteristic function of the weak limit of (properly shifted and scaled) sums of random variables with possibly infinite variance takes on under different circumstances.

Definition 3.1 A random variable X has a *stable distribution (function)* if, for all $a_1, a_2 > 0$, there exist constants $a > 0$ and b such that its characteristic function satisfies

$$\varphi\left(\frac{t}{a_1}\right) \cdot \varphi\left(\frac{t}{a_2}\right) = \varphi\left(\frac{t}{a}\right) \exp(-ibt)$$

Theorem 12 (Generalized Central Limit Theorem). *Let $X_1, X_2, \ldots$ be independent and identically distributed random variables and $\{C_n\}, \{D_n\}$ be two families of constants. Then the random variable χ has a stable distribution iff*

$$\frac{X_1 + X_2 + \cdots + X_n}{D_n} - C_n \implies \chi$$

Futhermore, unless F is degenerate, $D_n = n^{1/\alpha} h(n)$, where $0 < \alpha \leq 2$ and $h(n)$ is a slowly varying funciton in the sense of Karamata.

Proof. A treatment can be found on p. 38 of Ibragimov and Linnik [5].

Theorem 13. *If φ is a characteristic function, then $\varphi(-z) = \overline{\varphi}(z)$.*

Proof. (Adapted from p. 126 of Durrett [1]). Simply note that

$$\varphi(-t) = \mathbb{E}[\cos(-tX) + i\sin(-tX)] = \mathbb{E}[\cos(tX) - i\sin(tX)] = \overline{\varphi}(z)$$

Theorem 14 (Characteristic Functions of Stable Distributions). *χ (as in the previous theorem) has a stable distribution iff the characteristic function of χ has the form*

$$\varphi(t) = \exp\left(-a|t|^\alpha \left[1 - i\beta \mathrm{sgn}(t)\phi(t)\right]\right)$$

where a, α, and β are (to be determined) constants and

$$\phi(t) = \begin{cases} \tan\left(\frac{\pi}{2}\alpha\right) & \alpha \neq 1 \\ \frac{-2}{\pi}\log|t| & \alpha = 1 \end{cases}$$

Proof. A rigorous proof can be found on p. 43 of Ibragimov and Linnik [5].

Discussion. (Adapted from p. 6 of Amir [6]) We take a break from the standard ask-and-answer style to proceed with an exploration via "renormalization group" (popularized by the physicists). Denoting the partial sums $s_n := X_1 + X_2 + \cdots + X_n$ (with X_n as in the previous theorem), the general scaling one may consider is

$$\xi_n := \frac{s_n - b_n}{a_n}$$

Let us denote the characteristic function of each ξ_n by $\varphi(t)$ (assumed to be approximately independent of n for large n). Consider the variable $y_n = \xi_n a_n$. Its characteristic function is

$$\varphi_y(t) := \varphi(a_n t)$$

Consider next the distribution of the sum s_n. We have that $s_n = y_n + b_n$. It is straightforward to see that shifting a distribution by b_n implies multiplying the characteristic function by $\exp(itb_n)$. Therefore the characteristic function of the sum is

$$\Phi_N(t) := \exp(ib_N t) \cdot \varphi_y(t) = \exp(ib_N t) \cdot \varphi(a_N t)$$

This form will be the basis for the rest of the derivation, where we emphasize that the characteristic function φ is N-independent.

Consider $N = n \cdot m$, where n, m are two large numbers (sorry for the abrubt change in index). The important, "renormalization" insight is to realize that one may compute s_N in two ways: as the sum of N of the original variables, or as the sum of m variables, each one being the sum of n of the original variables. The sum of n variables drawn from the original distribution is given by

$$\Phi_n(t) = \exp(ib_n t) \cdot \varphi(a_n t)$$

If we take a sum of m variables drawn from that distribution (i.e., the one corresponding to the sum of n's), then, by Theorem 10, its characteristic function will be

$$f(t) = \exp(imb_n t) \cdot (\varphi(a_n t))^m = \exp(imb_n t) \exp(m \log[\varphi(a_n t)])$$

with a distribution corresponding to the sum of $n \cdot m = N$ variables drawn from the original distribution. Hence, this total sum does not depend on n or m separately but only on their product N. Assuming that n is suffciently large such that we may treat it as a "continuous" variable, we have that

$$\frac{\partial}{\partial n} \exp\left(i\frac{N}{n} b_n t + \frac{N}{n} \log[\varphi(a_n t)]\right) = 0$$

Defining $d_n := \frac{b_n}{n}$, we find, by the chain and product rule,

$$\implies \qquad iNt\frac{\partial d_n}{\partial n} - \frac{N}{n^2}\log(\varphi) + \frac{N}{n}\frac{\varphi'}{\varphi}\frac{\partial a_n}{\partial n}t = 0 \qquad \implies \qquad \frac{\varphi'(a_n t)t}{\varphi(a_n t)} = \frac{\log(\varphi(a_n t))}{n\frac{\partial a_n}{\partial n}} - it\frac{\partial d_n}{\partial n}\frac{n}{\frac{\partial a_n}{\partial n}}$$

Multiplying both sides by a_n and defining $z := a_n t$, we find that

$$\frac{\varphi'(z)z}{\varphi(z)} - \log(\varphi(z))\frac{a_n}{n\frac{\partial a_n}{\partial n}} + iz\frac{\partial d_n}{\partial n}\frac{n}{\frac{\partial a_n}{\partial n}} = 0$$

Since this equation should hold (with the same function $\varphi(z)$) as we vary n, we expect that $\frac{a_n}{n\frac{\partial a_n}{\partial n}}$ and $\frac{\partial d_n}{\partial n}\frac{n}{\frac{\partial a_n}{\partial n}}$ should be nearly independent of n for large values of n. This differential equation for $\varphi(z)$ has the mathematical structure

$$\frac{\varphi'}{\varphi} - \frac{C_1 \log(\varphi(z))}{z} = iC_2$$

with constants C_1, C_2.

If $C_2 = 0$, we can write the diff. eq. in the form

$$\frac{C_1}{z} = \frac{\varphi'}{\varphi \log(\varphi)} = \frac{(\log(\varphi))'}{\log(\varphi)} = [\log(\log(\varphi))]' \qquad \Longrightarrow \qquad \log(\log(\varphi)) = C_1 \log|z| + \text{const}$$

$$\Longrightarrow \varphi(z) = \exp(A \cdot |z|^{C_1})$$

This is the general solution to the homogeneous equation. Guessing a particular solution to the inhomogeneous equation in the form $\varphi(z) = \exp(Dz) \iff \log(\varphi) = Dz$ leads to

$$D - C_1 D = iC_2 \qquad \longrightarrow \qquad D = \frac{iC_2}{1 - C_1}$$

As long as $C_1 \neq 1$, we have found a solution. Combining the solutions to the homogeneous and inhomogeneous equations, we obtain the general solution to the diff. eq. when $C_1 \neq 1$,

$$\varphi(z) = \exp(A|z|^{C_1} + Dz)$$

The D term is associated with a trivial shift of the distribution (related to the linear scaling of b_n) and can be eliminated. We will therefore not consider it in the following analysis. By Theorem 12, the characteristic function must take the form

$$\varphi(z) = \begin{cases} \exp(Az^{C_1}) & t > 0 \\ \exp(A^*|z|^{C_1}) & t < 0 \end{cases}$$

Letting $\alpha := C_1 \neq 1$, this may be rewritten as

$$\varphi(z) = \exp\left(-a|z|^\alpha \left[1 - i\beta \text{sgn}(z) \tan\left(\frac{\pi\alpha}{2}\right)\right]\right)$$

When $C_1 = 1$, we can guess a solution of the form $\varphi(z) = \exp(Dz\log(z)) \iff \log(\varphi) = Dz\log(z)$ to the inhomogeneous equation and find:

$$D\log(z) + D - D\log(z) = iC_2$$

Hence, we have a solution when $D = iC_2$. Again combining the solutions to the homogeneous and inhomogeneous equations, we obtain the general solution to the diff. eq. when $C_1 = 1$:

$$\varphi(z) = \exp(A|z| + iC_2 z\log(z))$$

Repeating the logic we used before to establish the coefficients A, D, C_2 for $z > 0$ and $z < 0$ leads to:

$$\varphi(z) = \exp\left(-a|z|\left[1 - i\beta sgn(z)\left(-\frac{2}{\pi}\log|z|\right)\right]\right)$$

Heuristic Remark *We are about to introduce the notion of Tauberian relations, the idea that the asymptotic nature of a function can be determined by examining the behavior of the transform close to the origin. Before proceeding, we note that Tauberian-theoretic methods are legitimized by the Heisenberg uncertainty principle discussed in class: as we zoom into smaller regions of the frequency domain, we will glean the asymptotic behavior of the original function in the time domain, as we are forced by the lower bound. Unfortunately, I have decided to denote inputs from the frequency space with t.*

Theorem 15 (Karamata's Tauberian Theorem). *The constants from the previous theorem satisfy $a \geq 0$, $0 < \alpha \leq 2$, and $|\beta| \leq 1$. Furthermore, the probability density function f corresponding to χ (as in the last theorem) has the following asymptotic behavior:*

$$f(x) \sim \frac{A}{x^{1+\alpha}}$$

for some constant A.

Proof. Again, a rigorous proof may be found on p. 115 of Bhattacharya and Waymmire [2].

Discussion. (Adapted from p. 10 in Amir [6]) Consider a probability density function, $f(x)$, which decays, for $|x| \geq x_0 > 0$ bounded away from 0, as

$$f(x) \sim \frac{A}{x^{1+\alpha}}$$

with $\alpha > 0$. This is what it means for f to be of moderate decrease; from class, we know that we can apply the Fourier transform to f. In doing so, we shall derive a Tauberian relation: we will see that the Fourier transform of $f(x)$ takes on a form similar to that of solutions to the differential equation derived from the "renormalization group" near the origin. This, in turn, allows us to "characterize" the asymptotic behavior of all solutions to the differential equation. For small, positive t we find

$$\Phi(t) := \int_{-\infty}^{\infty} \frac{A}{x^{1+\alpha}} \exp(itx)\, dx = A_+ \int_{tx_0}^{\infty} \frac{\exp(itx)}{x^{1+\alpha}}\, dx + A_- \int_{-\infty}^{-tx_0} \frac{\exp(itx)}{x^{1+\alpha}}\, dx$$

where we substituted $m = tx$ and split the integral over right-tail-dominant and left-tail-dominant power-law distributions. Focusing in on one of the integrals (the one with the left tail decaying much faster than the right),

$$I_+(t) := \int_{tx_0}^{\infty} \frac{\exp(itx)}{x^{1+\alpha}}\, dx = \int_{tx_0}^{\infty} \exp(im)\frac{t^{1+\alpha}}{m^{1+\alpha}} \cdot \frac{1}{t}\, dm$$

we integrate by parts to obtain

$$I_+(t) = \left[-t^\alpha \frac{m^{-\alpha}}{\alpha}\exp(im)\right]_{m=tx_0}^{\infty} + t^\alpha \int_{tx_0}^{\infty} \frac{im^{-\alpha}}{\alpha}\exp(im)\ dm$$

For $\alpha < 1$, we may approximate the integral by replacing the lower limit of integration by 0, to find

$$I_+(t) \approx \frac{x_0^{-\alpha}}{\alpha}\exp(ix_0 t) + \frac{t^\alpha}{\alpha} i \int_0^\infty \frac{\exp(im)}{m^\alpha}\ dm$$

and $\int_0^\infty \frac{\exp(im)}{m^\alpha}\ dm = i\Gamma(1-\alpha)\exp(-i\frac{\pi}{2}\alpha)$ (this can easily be evaluated using contour integration). Thus, for small t we have

$$I_+(t) \approx (\text{const.}) - C_+ t^\alpha$$

with $C_+ := \frac{\Gamma(1-\alpha)\exp(-i\frac{\pi}{2}\alpha)}{\alpha}$. If we look at the integral corresponding to the distribution with a rapidly decaying right tail, similar analysis leads to

$$I_-(t) \approx (\text{const.}) - C_-(-t)^\alpha \qquad \text{with} \qquad C_- := \frac{\Gamma(1-\alpha)\exp(i\frac{\pi}{2}\alpha)}{\alpha}$$

where a is real. $\Phi(t)$ is technically not the characteristic function, since the near-zero bounds of the integrals were actually tx_0. However, since the integrals correspond to tails that decay fast, we can bound the rest of the integral to be at most linear in t for small t (up to a constant). Therefore the characteristic function near the origin is approximated by

$$\varphi(t) \approx 1 - C|t|^\alpha \approx \exp\left(-C|t|^\alpha\right)$$

where the second approximation comes from Taylor series and

$$C := \frac{\Gamma(1-\alpha)}{\alpha}\left(A_+ \exp\left(i\alpha\frac{\pi}{2}\right) + A_- \exp\left(-i\alpha\frac{\pi}{2}\right)\right)$$
$$\implies \frac{\Im(C)}{\Re(C)} = -\frac{\sin(\frac{\pi}{2}\alpha)}{\cos(\frac{\pi}{2}\alpha)}\left(\frac{A_+ - A_-}{A_+ + A_-}\right) = -\tan\left(\frac{\pi}{2}\alpha\right)\beta$$

with

$$\beta := \frac{A_+ - A_-}{A_+ + A_-}$$

Thus, the characteristic function is approximate near the origin by

$$\varphi(t) \approx \exp\left(-c|t|^{\alpha}\left[1 - \tan\left(\frac{\pi}{2}\alpha\right)\beta\right]\right)$$

with c a real constant. In the case that $\alpha = 1$, note that the logic following up to the integration-by-parts step is still intact for the case $\alpha = 1$. However, we can no longer replace the lower limit of integration by 0. Instead, note that the real part of the integral can be evaluated by parts, and diverges as $-\log|t|$, while the imaginary part of the integral does not suffer from such a divergence and we can well approximate it by replacing the lower limit of integration with 0.

Therefore we find:

$$\frac{I}{|t|} \approx -\log|t| + i\int_0^{\infty} \frac{\sin|x|}{|x|}\, dx = -\log|t| + i\frac{\pi}{2}$$

This leads to a characteristic function that is approximated near the origin by

$$\varphi(t) \approx \exp\left(-c|t|\left[1 + i\frac{2}{\pi}\log|t|\right]\right)$$

for some real constant c.

Notice that the assumption $\alpha < 1$ was needed in order to approximate the near-zero bounds of the integral. Along with the $\alpha = 1$ case, we can extend our analysis to $1 < \alpha \leq 2$ but no further.

Let $f(x) := g'(x)$, where g is a probability density function that has tails that decay according to a power-law with an index $0 < \widetilde{\alpha} < 1$, as derived in the last few paragraphs. Note that g could decay to a finite constant that may be non-zero, as $g'(x) = (g(x) + a_1)'$. Its Fourier transform will thus, with similar steps taken, be approximated by the following form for small t:

$$\widehat{g}(t) \approx \frac{a_1}{t} + a_2 - |t|^{\widetilde{\alpha}}$$

where the $\frac{1}{t}$ divergence comes from the constant value a_1 of g at infinity. From class, we know that the Fourier transform sends the derivative to a multiplication by a linear factor, like so: $\widehat{f}(t) = -it\widehat{g}(t)$. Therefore, for small t, we have

$$\widehat{f}(t) \approx 1 - (-it)C_+|t|^{\widetilde{\alpha}} + iBt$$

where the constant B must be equal to the expectation value of x (by the definition of the characteristic function), and C_+ is as before.

For the characteristic function approximation, first note that the linear term will drop out due to the shift in the scaling ξ_n. Then let $\alpha := \widetilde{\alpha}$ (this will be the index of f). We can simplify the rest using the trigonometric identities $\tan(\frac{\pi}{2}\alpha) = \tan(\frac{\pi}{2}\widetilde{\alpha} + \frac{\pi}{2}) = -\cot(\frac{\pi}{2}\widetilde{\alpha})$:

$$-iC_+ \propto -i\left[1 - i\tan\left(\frac{\pi}{2}\widetilde{\alpha}\right)\right] = -\tan\left(\frac{\pi}{2}\widetilde{\alpha}\right) - i = c\left[1 + i\cot\left(\frac{\pi}{2}\widetilde{\alpha}\right)\right] = c\left[1 - i\tan\left(\frac{\pi}{2}\alpha\right)\right]$$

for some real constant c. Thus the form of $\varphi(t)$ for $f = g'$ is.

$$\varphi(t) \approx 1 - iC_+|t|^{\widetilde{\alpha}+1} \approx \exp\left(-c|t|^{\alpha}\left[1 - \tan\left(\frac{\pi}{2}\alpha\right)\right]\right)$$

where the second approximation is due to Taylor series. We've just shown that, for $0 < \alpha < 2$, the characteristic function of a function of moderate decrease take on the canonical stable form near the origin. Thus, by our Tauberian theorem, we have that probability density functions of stable distributions with index $0 < \alpha < 2$ decrease moderately as $|x| \to \infty$.

Finally, as $\alpha \to 2$, $\tan(\frac{\pi}{2}\alpha) \to \tan(\pi) = 0$, making

$$\varphi(t) \approx \exp\left(-c|t|^2\right)$$

The corresponding probability density function has asymptotic behavior like

$$f(x) \sim \exp\left(-c|t|^2\right)$$

The Fourier inversion of a Gaussian is still a Gaussian. As Gaussians are functions of moderate decrease, our claim follows.

4 A Local Limit Theorem

Let d be a natural number and $V = \{x_1, \ldots, x_l\} \subseteq \mathbb{Z}^d \setminus \{0\}$ be a generating set of $\mathbb{Z}^d$ for the rest of this paper.

Definition 4.1. A probability measure P (or its associated distribution function) is said to be *discrete* if there is a countable set S with $P(S^c) = 0$. A *discrete random variable* is simply a random variable whose domain probability space has a discrete measure. For the rest of this paper, we will take the probability measure P to be the (discrete) counting measure with $P(X = y) = 0$ for all $\mathbb{R}^d \ni y \notin \mathbb{Z}^d$ and $\sum_{y \in \mathbb{Z}^d} P(X = y) = 1$.

Definition 4.2. A *random walk* p is the sum of discrete, identically distributed random variables X. This sum is written as

$$S_n := S_0 + X_1 + X_2 + \cdots + X_n$$

Where S_0 is given the trivial distribution $P(S_0 = 0) = 1$ (the random walk starts at the origin). For the rest of this paper we will use the following notation for convenience:

$$p(x) := P(X = x) \qquad \text{and} \qquad p_n(x) := P(S_n = x)$$

Note that the random walk can also be defined as a Markov chain with

$$P(S_{n+1} = z | S_n = y) = P(X = (z - y))$$

The *simple random walk* has the distribution

$$P(X_j = e_k) = P(X_j = -e_k) = \frac{1}{2d}$$

for $k = 1, \ldots, d$ and e_k the unit vectors in $\mathbb{Z}^d$. We define a few conditions on the random walk:

- A random walk is *aperiodic* if $P(S_n = 0) > 0$ for all n.
- A random walk is *rreducible* if, for all $y \in \mathbb{Z}^d$, there exists N such that if $n \geq N$ then $P(S_n = y) > 0$.
- A random walk is *symmetric* if, $P(X = y) = P(X = -y)$ for all $y \in \mathbb{Z}^d$.

Let there be function $\mathcal{K} : V \to (0, 1)$. An aperiodic, irreducible, symmetric random walk can be characterized by the vector set V and function $\mathcal{K}$ like so:

$$\mathcal{K}(X_1) + \cdots + \mathcal{K}(X_l) < 1 \qquad \mathcal{K}(0) = 1 - \sum_{j=1}^{l} \mathcal{K}(x_j) \qquad P(X = x_j) = P(X = -x_j) = \frac{1}{2}\mathcal{K}(x_j)$$

We let $\mathcal{P}$ denote the set of all random walks of this form.

Definition 4.3. We define the $(j_1, \ldots, j_d)$ *moment* of X as:

$$\mathbb{E}\left[(X^1)^{j_1} \ldots (X^d)^{j_d}\right] := \sum_{x \in \mathbb{Z}^d} (x^1)^{j_1} \ldots (x^d)^{j_d} p(x)$$

The *covariance matrix* of X, called Γ, is a $d \times d$ matrix where the (k, l)th entry of the matrix is given by:

$$\Gamma_{kl} = \mathbb{E}[(X^k)(X^l)] = \sum_{x \in \mathbb{Z}^d} (x^k)(x^l) p(x)$$

This corresponds to the moment that has $j_k = j_l = 1$ and all other moment values equal to 0.

Lemma 16. *For all $p \in \mathcal{P}$ with covariance matrix Γ, and for all $v \in \mathbb{Z} \setminus \{0\}$, the following are true:*

(1) $\mathbb{E}[(X \cdot v)^m] = 0$ *if m is odd*

(2) $\mathbb{E}[(X \cdot v)^2] = v \cdot \Gamma v > 0$

Proof. (Adapted from p. 3 of Hill [7]). To prove (1), note that for

$$\mathbb{E}[(X \cdot v)^m] = \sum_{x \in \mathbb{Z}^d} (x \cdot v)^m p(x)$$

we only need to examine $x \in \mathbb{Z}^d$ where both $p(x) \neq 0$ and $x \cdot v \neq 0$, because otherwise $(x \cdot v)^m p(x) = 0$. Since $p(-x) = p(x)$ (by the definition of P), we have that

$$(x \cdot v)^m p(x) + ((-x) \cdot v)^m p(-x) = 0$$

for each $x \in \mathbb{Z}^d$ if m is odd.

For (2), $\mathbb{E}[(X \cdot v)^2] > 0$ is clear from the fact that $0 \notin \mathbb{Z}^d \setminus \{0\}$, which implies there must be some $x_j \in \mathbb{Z}^d$ with $x_j \cdot v \neq 0$. For the equivalence,

$$\mathbb{E}[(X \cdot v)^2] = \sum_{x \in \mathbb{Z}^d} (x \cdot v)^2 p(x) = \sum_{x \in \mathbb{Z}^d} (x^1 v^1 + \cdots + x^d v^d)^2 p(x) = \sum_{x \in \mathbb{Z}^d} \left(\sum_{j=1}^{d} \sum_{k=1}^{d} x^j v^j x^k v^k \right) p(x)$$

Thus,

$$v \cdot \Gamma v = \sum_{j=1}^{d} v^j (\Gamma v)^j = \sum_{j=1}^{d} v^j \sum_{k=1}^{d} v^k \mathbb{E}[X^j X^k] = \sum_{j=1}^{d} v^j \sum_{k=1}^{d} v^k \sum_{x \in \mathbb{Z}^d} x^j x^k p(x) = \mathbb{E}[(X \cdot v)^2]$$

where rearrangement is allowed due to the finiteness of all but one sum.

Lemma 17. *The following are useful properties of the characteristic function.*

Proof. (Adapted from p. 4 of Hill [7])

$$\varphi(0) = \mathbb{E}[\exp(ix \cdot 0)] = \mathbb{E}[1] = 1$$

$$|\varphi(\theta)| = \left| \sum_{x \in \mathbb{Z}^d} p(x) \exp(ix \cdot \theta) \right| \leq \sum_{x \in \mathbb{Z}^d} |p(x) \exp(ix \cdot \theta)| \leq \sum_{x \in \mathbb{Z}^d} |p(x)| = 1$$

For continuity, notice that

$$
\begin{aligned}
|\varphi(\theta+\theta_1)-\varphi(\theta)| &= \left|\sum_{x\in\mathbb{Z}^d}\exp(i(\theta+\theta_1)\cdot x)p(x)-\sum_{x\in\mathbb{Z}^d}\exp(i\theta\cdot x)p(x)\right| \\
&= \left|\sum_{x\in\mathbb{Z}^d}\exp(ix\cdot\theta)(\exp(ix\cdot\theta_1-1))p(x)\right| \\
&\leq \sum_{x\in\mathbb{Z}^d}|(\exp(ix\cdot\theta_1-1)p(x)|
\end{aligned}
$$

Since $\lim_{\theta_1\to 0}|\exp(ix\theta_1-1)|=0$ and $|p(x)|>0$ for only a finite number of x,

$$
\lim_{\theta_1\to 0}|\varphi(\theta+\theta_1)-\varphi(\theta)|=0 \qquad \forall\theta
$$

Let $\varphi_u(s)$ be the one-dimensional characteristic function of the random variable $X\cdot u$, with $|u|=1$ and $\theta=su$. Then its mth derivative is

$$
\varphi_u^{(m)}(s)=i^m\mathbb{E}[(X\cdot u)^m\exp(i(X\cdot u)s)]
$$

by the chain rule. Taylor expanding around the origin, we get the estimate

$$
\left|\varphi_u(s)-\sum_{j=0}^{m}\frac{i^j\mathbb{E}[(X\cdot u)^j]}{j!}s^j\right| \leq \left|\frac{\mathbb{E}[(X\cdot u)^{m+1}]s^{m+1}}{(m+1)!}\right|
$$

Redefining $\varphi_u(s)$ with $\varphi(\theta)$, we expand $\varphi(\theta)$ about the origin as follows, making use of Lemma 16:

$$
\begin{aligned}
\varphi(0) &= 1 \\
\varphi'(0) &= i\mathbb{E}[X\cdot u]=0 \\
\varphi''(0) &= -\mathbb{E}[(X\cdot u)^2] \\
\varphi'''(0) &= -i\mathbb{E}[(X\cdot u)^3]=0
\end{aligned}
$$

These give us the second order Taylor polynomial:

$$
\varphi(\theta)=1-\frac{\mathbb{E}[(X\cdot u)^2]}{2}s^2+h(\theta)
$$

$h(\theta)$ is bounded by the Lagrange remainder with $m=3$. Again, by Lemma 16,

$$
\mathbb{E}[(X\cdot u)^2]s^2=\mathbb{E}[(X\cdot\theta)^2]=\theta\cdot\Gamma\theta
$$

and so

$$\varphi(\theta) = 1 - \frac{\theta \cdot \Gamma\theta}{2} + h(\theta)$$

Theorem 18. *If $X = (X^1, \ldots, X^d)$ is a $\mathbb{Z}^d$-valued random variable with characteristic function $\varphi(\theta)$, then the following holds:*

$$p(x) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \varphi(\theta) \exp(-ix \cdot \theta)\ d\theta$$

Proof. (Adapted from p. 6 of Hill [7]) Since

$$\varphi(\theta) = \sum_{y \in \mathbb{Z}^d} \exp(iy \cdot \theta) p(y)$$

is clearly periodic (as $\exp(it)$ is period) we have that

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(\theta) \exp(-ix \cdot \theta)\ d\theta &= \int_{[-\pi,\pi]^d} \left(\sum_{y \in \mathbb{Z}^d} \exp(iy \cdot \theta) p(y) \right) \exp(-ix \cdot \theta)\ d\theta \\ &= \sum_{y \in \mathbb{Z}^d} p(y) \int_{[-\pi,\pi]^d} \exp i(y - x) \cdot \theta)\ d\theta\end{aligned}$$

When $y \neq x$, we have that $y - x \neq 2\pi k$ for $k \in \mathbb{Z}^d \setminus \{0\}$, since π is not an integer. Thus,

$$\int_{[-\pi,\pi]^d} \exp(i(y - x) \cdot \theta)\ d\theta = 0$$

Thus,

$$\int_{[-\pi,\pi]^d} \varphi(\theta) \exp(-ix \cdot \theta)\ d\theta = p(x) \int_{[-\pi,\pi]^d} d\theta = p(x)(2\pi)^d$$

which, when rearranged, gives the theorem.

Lemma 19. *For all θ in $[-\pi, \pi]^d$ and $n > 0$, there exists $b > 0$ such that*

$$|\varphi(\theta)|^r < \exp\left(-br|\theta|^2\right)$$

Proof. (Adapted from p. 7 of Hill [7]) By Lemma 17, $|\varphi(\theta)| \leq 1$. Consider the set of all θ such that $|\varphi(\theta)| = 1$, Then $|\varphi(\theta)^n| = 1$ for all n, making

$$\left| \sum_{x \in \mathbb{Z}^d} p_n(x) \exp(ix \cdot \theta) \right| = \sum_{x \in \mathbb{Z}^d} |p_n(x) \exp(ix \cdot \theta)| = 1$$

Recall that, for complex numbers $w_1, \ldots, w_k$, if $|w_1 + \cdots + w_k| = |w_1| + \cdots + |w_k| = 1$, then there exists $\gamma \in \mathbb{R}$ such that $w_j = r_j \exp(i\gamma)$ with each $r_j \geq 0$. Thus, each

$$p_n(x) \exp(ix \cdot \theta) = r_j \exp(i\gamma)$$

for some γ that depends on n. By the way a random walk $p \in \mathcal{P}$ was defined, for all x, there is an N such that $p_n(x) > 0$. From this we know that, for all $x \in \mathbb{Z}^d$, there exists γ and N with

$$p_N(x) \exp(ix \cdot \theta) = r \cdot \exp(i\gamma)$$

$p_N(x)$ must equal r, so we end up with

$$2k\pi + x \cdot \theta = \gamma$$

The only $\theta \in [-\pi, \pi]^d$ that satisfies this equation for all $x \in \mathbb{Z}^d$ is 0, so $|\varphi(\theta)| < 1$ for $\theta \in [-\pi, \pi]^d \setminus \{0\}$.

By the previous statement and the Taylor expansion of the characteristic function, we know we can find a $b > 0$ with $|\varphi(\theta)| < |1 - b|\theta|^2|$, since θ is restricted to a bounded set. By the same reasoning, we know from the Taylor expansion of $\exp(-b|\theta|^2)$ that $|1 - b|\theta|^2| < \exp(-b|\theta|^2)$. Combining these inequalities and raising to the $r > 0$ power, we get

$$|\varphi(\theta)|^r < \exp\left(-b|\theta|^2\right)^r$$

which proves the lemma.

Lemma 20. *Suppose that we have $p \in \mathcal{P}$ with covariance matrix Γ and characteristic function φ. Then there exists $\varepsilon > 0$ such that if $|\theta| < \varepsilon\sqrt{n}$ then*

$$\varphi\left(\frac{\theta}{\sqrt{n}}\right)^n = \exp\left(-\frac{\theta \cdot \Gamma\theta}{2}\right)(1 + F_n(\theta))$$

and

$$|F_n(\theta)| < \exp\left(\frac{\theta \cdot \Gamma\theta}{4}\right) + 1$$

Proof. (Adapted from p. 9 of Hill [7]) We find $0 < \delta < 1$ such that if $|\theta| < \delta$ then $|1 - \varphi(\theta)| < 1$. From the Taylor expansion of the characteristic function, we know that $1 - \varphi(\theta) = \frac{\theta \cdot \Gamma\theta}{2} - h(\theta)$, and using the Taylor expansion of the complex logarithm, we can write

$$\log(\varphi(\theta)) = \log(1-(1-\varphi(\theta))) = -\frac{\theta\cdot\Gamma\theta}{2} + h(\theta) - \frac{(\theta\cdot\Gamma\theta)^2}{8} + \frac{\theta\cdot\Gamma\theta}{2}h(\theta) - \frac{1}{2}h(\theta)^2 + q(\theta)$$

and by Taylor's Theorem, there exists a $c > 0$ with

$$|q(\theta)| < c|1-\varphi(\theta)|^3 = c\left|\frac{\theta\cdot\Gamma\theta}{2} - h(\theta)\right|^3$$

Similarly,

$$n\log\left(\varphi\left(\frac{\theta}{\sqrt{n}}\right)\right) = -\frac{\theta\cdot\Gamma\theta}{2} + n\cdot h\left(\frac{\theta}{\sqrt{n}}\right) - \frac{(\theta\cdot\Gamma\theta)^2}{8n} + \frac{\theta\cdot\Gamma\theta}{2}h\left(\frac{\theta}{\sqrt{n}}\right) - \frac{1}{2}h\left(\frac{\theta}{\sqrt{n}}\right)^2 + q\left(\frac{\theta}{\sqrt{n}}\right)$$

with q bounded as above. Define $g(\theta, n)$ as follows:

$$g(\theta, n) = n\log\left(\varphi\left(\frac{\theta}{\sqrt{n}}\right)\right) + \frac{\theta\cdot\Gamma\theta}{2}$$

so that

$$\varphi\left(\frac{\theta}{\sqrt{n}}\right)^n = \exp\left(-\frac{\theta\cdot\Gamma\theta}{2}\right)\exp(g(\theta, n))$$

Being the Lagrange remainder, we know that $|h(\theta)|$ is bounded by $c_1|\theta|4$ for some finite c_1. Because of this and the fact that $|\theta| < \delta < 1$, $\frac{1}{n}\frac{(\theta\cdot\Gamma\theta)^2}{8}$ becomes the dominant error term for $g(\theta, n)$, so there exists c_2 such that

$$|g(\theta, n)| < n\cdot h\left(\frac{\theta}{\sqrt{n}}\right) + \frac{c_2|\theta|^4}{n}$$

Both of these terms can be made arbitrarily small by choosing a small enough θ, so there exists ε where $0 < \varepsilon < \delta$ such that if $|\theta| < \varepsilon\sqrt{n}$, then

$$|g(\theta, n)| < \frac{\theta\cdot\Gamma\theta}{4}$$

Let $F_n(\theta) = \exp(g(\theta, n)) - 1$. It is clear that $|F_n(\theta)| = |\exp(g(\theta, n)) - 1| < \exp(\frac{\theta\cdot\Gamma\theta}{4}) + 1$, which proves the lemma.

Lemma 21 *The following estimate holds for $r \in \mathbb{R}^d$, with $c > 0$ and $\beta > 0$.*

$$\left|\int_{|x|>r} \exp(-t^2)\ dt\right| < c\cdot\exp\left(-\beta|r|^2\right)$$

Proof. We proceed with formal techniques. By iterated integration by parts, we see that

$$\begin{aligned}
\int_x^\infty \exp(-t^2)\ dt &= \int_x^\infty \left(\frac{-1}{2t}\right)(-2t\cdot\exp(-t^2))\ dt \\
&= -\frac{1}{2}\left[\frac{(\exp(-t^2)}{t}\right]_x^\infty - \frac{1}{2}\int_x^\infty \frac{\exp(-t^2)}{t^2}\ dt \\
&= \left(\frac{\exp(-x^2)}{2x}\right) + \frac{1}{4}\int_x^\infty \left(\frac{1}{t^3}\right)(-2t\cdot\exp(-t^2))\ dt \\
&= \frac{\exp(-x^2)}{2x} - \frac{\exp(-x^2)}{4x^3} - \dots \\
&= \frac{\exp(-x^2)}{2x} - \sum_{j=1}^\infty \frac{\exp(-x^2)}{2(j+1)X^{2j+1}}
\end{aligned}$$

As the real exponential function is always positive, we are subtracting something positive. The estimate thus holds for formal integration. Translation of the estimate to multiple variables should also hold.

Theorem 22. *Let*

$$\overline{p}_n(x) := \frac{1}{(2\pi n)^{d/2}\sqrt{\det\Gamma}}\exp\left(\frac{-(x\cdot\Gamma^{-1}x)^2}{2n}\right) = \frac{1}{(2\pi n)^{d/2}}\int_{\mathbb{R}^d}\exp\left(\frac{ips\cdot x}{\sqrt{n}}\right)\exp\left(-\frac{s\cdot\Gamma s}{2}\right)\ ds$$

Then for $0 \le r \le \varepsilon\sqrt{n}$, there exists a $c > 0$ and a $\beta > 0$ such that:

$$\left|p_n(x) - \left(\overline{p}_n(x) + \frac{1}{(2\pi)^d n^{d/2}}\int_{|\theta|\le r}\exp(-ipx\cdot\theta)^n\exp\left(\frac{\theta\cdot\Gamma\theta}{2}\right)F_n(\theta)\ d\theta\right)\right| < c{\cdot}n^{-d/2}\exp\left(-\beta r^2\right)$$

Proof. (Adapted from p. 10 of Hill [7]) Recalling the inversion formula from Theorem 18, we use the substitution $\theta = \frac{s}{\sqrt{n}}$ to write $p_n(x)$ as:

$$p_n(x) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}\varphi(\theta)^n\exp(-ix{\cdot}\theta)\ d\theta = \frac{1}{(2\pi)^d n^{d/2}}\int_{\left[-\sqrt{n}\pi,\sqrt{n}\pi\right]^d}\varphi\left(\frac{s}{\sqrt{n}}\right)^n\exp(-iz{\cdot}s)\ ds$$

with $z = \frac{x}{\sqrt{n}}$. Using Lemma 19, we know that there exists a $b > 0$ such that

$$\left|\varphi\left(\frac{s}{\sqrt{n}}\right)\right|^n < \exp\left(-b|s|^2\right)$$

Combining this with Lemma 21, we get:

$$\begin{aligned}\left|p_n(x) - \frac{1}{(2\pi)^d n^{d/2}} \int_{|s|<\varepsilon\sqrt{n}} \varphi\left(\frac{s}{\sqrt{n}}\right)^n \exp(-iz\cdot s)\ ds\right| &< \frac{1}{(2\pi)^d n^{d/2}} \int_{|s|>\varepsilon\sqrt{n}} \exp\left(-b|s|^2\right)\ ds \\ &< c_1 n^{-d/2} \exp\left(-\beta_1 n\right)\end{aligned}$$

Lemma 21 and the definition of $\overline{p}_n(x)$ gives us:

$$\begin{aligned}\left|\overline{p}_n(x) - \frac{1}{(2\pi)^d n^{d/2}} \int_{|s|<\varepsilon\sqrt{n}} \exp\left(-\frac{s\cdot\Gamma s}{8}\right) \exp(-iz\cdot s) ds\right| &< \frac{1}{(2\pi)^d n^{d/2}} \int_{|s|>\varepsilon\sqrt{n}} \exp\left(-b|s|^2\right)\ ds \\ &< c_2 n^{-d/2} \exp\left(-\beta_2 n\right)\end{aligned}$$

By Lemma 20, we know that, for $|s| < \varepsilon\sqrt{n}$, we can write

$$\varphi\left(\frac{s}{\sqrt{n}}\right)^n = \exp\left(-s\cdot\Gamma s\right)\left(1 + F_n(s)\right)$$

Thus, combining the inequalities gives us

$$\begin{aligned}&\left|p_n(x) - \left(\overline{p}_n(x) + \frac{1}{(2\pi)^d n^{d/2}} \int_{|\theta|<\varepsilon\sqrt{n}} \exp(-iz\cdot s) \exp\left(-\frac{\theta\cdot\Gamma\theta}{2}\right) F_n(\theta) d\theta\right)\right| \\ &\qquad < c_1 n^{-d/2} \exp\left(-\beta_1 n\right) + c_2 n^{-d/2} \exp\left(-\beta_2 n\right)\end{aligned}$$

This equation shows the result holds for $r = \varepsilon\sqrt{n}$. For smaller values of r, we use the fact that $|F_n(s)| < \exp(-\frac{s\cdot\Gamma s}{4}) + 1$ to show

$$\left|\int_{r<|s|<\varepsilon\sqrt{n}} \exp(-iz\cdot s) \exp\left(-\frac{s\cdot\Gamma s}{2}\right) F_n(s) ds\right| < 2\int_{|s|>r} \exp\left(-\frac{s\cdot\gamma s}{8}\right)\ ds < c_3 \exp\left(-\beta_3 r^2\right)$$

Since in this case $r < \varepsilon\sqrt{n}$, this becomes the dominant error term, proving the theorem.

Lemma 23. *There exists a $c > 0$ such that if $|\theta| < n^{1/8}$, then*

$$|F_n(s)| < c\frac{|\theta|^4}{n}$$

Proof. (Adapted from p. 10 of Hill [7]) Recall, from Lemma 20, that we have, for some $c > 0$,

$$|g(\theta, n)| < n \cdot h(\theta\sqrt{n}) + c\frac{|\theta|^4}{n}$$

and that $|h(\theta)| < c|\theta|^4$ by Taylor's theorem. Therefore, for a different $c > 0$, we have

$$|g(\theta, n)| < c\frac{|\theta|^4}{n}$$

If $|\theta| < n^{1/8}$, then $|g(\theta, n)| < \frac{c}{\sqrt{n}} \leq c$ for all n. By the Taylor expansion of the complex exponential, we know that if z is restricted to a bounded set, $|\exp(z) - 1| < cz$, and since $g(\theta, n)$ is restricted to a bounded set for $|\theta| < n^{1/8}$, we can write

$$|F_n(\theta)| = |\exp(g(\theta, n)) - 1| < c_1|g(\theta, n)| < c_1 \left| n \cdot h\left(\frac{\theta}{\sqrt{n}}\right) + c_2\frac{|\theta|^4}{n}\right| < c_3\frac{|\theta|^4}{n}$$

Theorem 24 (Local Central Limit Theorem). *For all $p \in P$, there exists a $c > 0$ with*

$$|p_n(x) - \overline{p}_n(x)| < \frac{c}{n^{(d+2)/2}}$$

Proof. (Adapted from p. 11 of Hill [7]) Let $r = \min(\varepsilon\sqrt{n}, n^{1/8})$. By the previous lemma, we know that for such an r,

$$\left|\int_{|\theta|\leq r} \exp\left(-\frac{ipx \cdot \theta}{\sqrt{n}}\right) \exp\left(-\frac{\theta \cdot \Gamma\theta}{2}\right) F_n(\theta)\ d\theta\right| < \frac{c_1}{n} \int_{\mathbb{R}^d} |\theta|^4 \exp\left(-\frac{\theta \cdot \Gamma\theta}{2}\right)\ d\theta < \frac{c_2}{n}$$

Letting

$$\frac{c}{n^{(d+2)/2}} = \frac{c_2}{n}\frac{1}{(2\pi)^d n^{d/2}}$$

and using Theorem 22, we can now write

$$|p_n(x) - \overline{p}_n(x)| < \frac{c}{n^{(d+2)/2}} + c_3 \exp\left(-\beta n^{1/4}\right)$$

Since $\exp(-\beta n^{1/4})$ is a reciprocal exponential, it decays faster than any polynomial. Thus, the theorem holds.

5 Application to Random Walks

Definition 5.1. A random walk is called *recurrent* if it visits every point in the lattice infinitely often. In other words, $\sum_{j=1}^{n} p_j(x)$ increases without bound as n goes towards ∞ for all $x \in \mathbb{Z}^d$. A random walk is called *transient* if $\sum_{j=N}^{n} p_j(x)$ is finite for all x.

Theorem 25. *For any Markov chain, if*

$$P(S_{n+j} = z | S_n = y]) > 0 \qquad \text{and} \qquad P(S_{n+k} = y | S_n = z) > 0$$

for all n and for some $j, k \geq 0$, then if one of those states is recurrent, the other is as well. The same applies for transience.

Proof. This paper is on the central limit theorem. A proof of this theorem is on p. 282 of Durrett [1].

Corollary 26. *Since we defined random walks $p \in \mathcal{P}$ so that the previous theorem is true for all $y, z \in \mathbb{Z}^d$, any random walk must be either recurrent or transient.*

Theorem 27. *Consider random walk $p \in \mathcal{P}$. Then for $d = 1, 2$, p is recurrent. For $d \geq 3$, p is transient.*

Proof. (Adapted from p. 12 of Hill [7]) Note that it suffices to prove the theorem only for $x = 0$ (since we can always shift the starting point back to origin). Recalling the definition of $\overline{p}_n$, we note that $\overline{p}_n(0) = \frac{1}{(2\pi n)^{d/2}\sqrt{det\Gamma}}$. By the local central limit theorem, we also know that $\lim_{n\to\infty} \frac{p_n(0)}{\overline{p}_n(0)} = 1$. This implies that there exists $0 < c_1 < c_2$ and a sufficiently large N such that if $n > N$, then $c_1 n^{d/2} < p_n(0) < c_2 n^{d/2}$. Therefore,

$$\sum_{j=N}^{n} c_1 j^{d/2} < \sum_{j=N}^{n} p_j(0) < \sum_{j=N}^{n} c_2 j^{d/2}$$

If $d = 1, 2$, $\lim_{n\to\infty} \sum_{j=N}^{n} p_j(0) > \lim_{n\to\infty} \sum_{j=N}^{n} c_1 j^{d/2}$, which diverges, proving that p is recurrent. If $d \geq 3$,

$$\lim_{n\to\infty} \sum_{j=N}^{n} p_j(0) < \lim_{n\to\infty} \sum_{j=N}^{n} c_2 j^{d/2}$$

which converges, proving that p is transient.

References

[1] R. Durrett *Probability: Theory and Examples.* Cambridge, 2019.

[2] R. Bhattacharya and E. Waymire *A Basic Course in Probability Theory.* Springer, 2007.

[3] T. Tao *Topics in Random Matrix Theory.* American Mathematical Society, 2011.

[4] Y. Chow and H. Teicher *Probability Theory. Independence, Interchangeability, Martingales.* Springer,

[5] I. Ibragimov and Y. Linnik *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, 1971.

[6] A. Amir *An elementary renormalization-group approach to the generalized central limit theorem and extreme value distributions.* Institute of Physics, 2017.

[7] M. Hill *Approximating the Random Walk Using the Central Limit Theorem.* https://math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Hill.pdf, 2011.