Varieties of Central Limit Theorems

Michael Taylor

Contents

0. Introduction

An eye-opening and side-splitting book review, [F], recently raised the interesting question of just what hypotheses on a sequence of IID random variables are needed for the sequence to satisfy a central limit theorem. One answer to this question is that one gets different central limit theorems depending on the specific hypotheses put forth. The purpose of this note is to describe explicitly some of the varieties of central limit theorems that arise. These results have, no doubt, been known for a long time, but it is perhaps useful to collect them.

To set things up, suppose $(\Omega, \mathcal{F}, \mu)$ is a probability space (Ω a set, $\mathcal{F}$ a σ-algebra, μ a probability measure) and that $\{f_j\}$ is a sequence of (real valued) independent, identically distributed random variables on Ω, with mean 0 and variance σ, so

$$f_j \in L^2(\Omega, \mu), \quad \int_\Omega f_j \, d\mu = 0, \quad \int_\Omega f_j^2 \, d\mu = \sigma > 0. \tag{0.1}$$

In such a case, the independence implies

$$(f_i, f_j)_{L^2} = 0, \quad \text{for} \quad i \neq j. \tag{0.2}$$

The weak law of large numbers says that, as $k \to \infty$,

$$S_k = \frac{1}{k} \sum_{j=1}^{k} f_j \longrightarrow 0, \quad \text{in } L^2\text{-norm.} \tag{0.3}$$

The proof is simple:

$$\Big\| \frac{1}{k} \sum_{j=1}^{k} f_j \Big\|_{L^2}^2 = \frac{1}{k^2} \sum_{i,j=1}^{k} (f_i, f_j)_{L^2} = \frac{\sigma}{k}. \tag{0.4}$$

A standard presentation of the weak law says that $S_k \to 0$ in measure, which follows from (0.3) (or better, from (0.4)), via Chebychev's inequality.

Kolmogoroff's strong law of large numbers produces pointwise a.e. convergence, and relaxes the L^2 hypothesis, down to L^1 (and then yields L^1-norm convergence), but we will not be concerned with that here. (Cf. Chapter 15 of [T] for a treatment, making a connection to Birkhoff's ergodic theorem.)

To proceed, each real-valued random variable f on Ω induces a probability measure ν_f on $\mathbb{R}$, given by

$$\nu_f(S) = \mu(f^{-1}(S)), \tag{0.5}$$

when $S \subset \mathbb{R}$ is a Borel set. Note that

$$
\begin{aligned}
f \in L^1(\Omega, \mu) &\Longleftrightarrow \int |x| \, d\nu_f(x) < \infty, \\
\int_\Omega f \, d\mu &= \int_{\mathbb{R}} x \, d\nu_f(x).
\end{aligned} \tag{0.6}
$$

Similarly,

$$
\int_\Omega f^2 \, d\mu = \int_{\mathbb{R}} x^2 \, d\nu_f(x), \tag{0.7}
$$

and, more generally, for $p \in [1, \infty)$,

$$
\int_\Omega |f|^p \, d\mu = \int_{\mathbb{R}} |x|^p \, d\nu_f(x). \tag{0.8}
$$

Given f as above, the function

$$
\begin{aligned}
\chi_f(\xi) &= \int_\Omega e^{-i\xi f} \, d\mu \\
&= \int_{\mathbb{R}} e^{-ix\xi} \, d\nu_f(x) \\
&= \sqrt{2\pi} \hat{\nu}_f(\xi)
\end{aligned} \tag{0.9}
$$

is called the characteristic function of f. If $\{f_j\}$ are independent, then

$$
G_k = \sum_{j=1}^k f_j \Longrightarrow \chi_{G_k}(\xi) = \chi_{f_1}(\xi) \cdots \chi_{f_k}(\xi). \tag{0.10}
$$

A special class of probability distributions on $\mathbb{R}$, called centered Gaussian distributions, has the form

$$
\gamma^\sigma(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-x^2/2\sigma}. \tag{0.11}
$$

One computes

$$
\int x \, \gamma^\sigma(x) \, dx = 0, \quad \int x^2 \, \gamma^\sigma(x) \, dx = \sigma. \tag{0.12}
$$

A random variable f on $(\Omega, \mathcal{F}, \mu)$ is said to be Gaussian if ν_f is Gaussian. A standard Fourier transform calculation gives

$$\sqrt{2\pi}\hat{\gamma}^\sigma(\xi) = e^{-\sigma\xi^2/2}. \tag{0.13}$$

Hence $f : \Omega \to \mathbb{R}$ is Gaussian with mean 0 and variance σ if and only if

$$\chi_f(\xi) = e^{-\sigma\xi^2/2}. \tag{0.14}$$

We note that

$$\gamma^\sigma * \gamma^\tau = \gamma^{\sigma+\tau}, \tag{0.15}$$

and that if f_j are independent, centered Gaussian random variables on Ω, then the sum $G_k = f_1 + \cdots + f_k$ is also Gaussian.

Gaussian distributions are often approximated by distributions of the sum of a large number of IID random variables, suitably rescaled. Theorems to this effect are called Central Limit Theorems. As stated in the opening paragraph, our goal is to present some of these theorems here.

Given that $\{f_j\}$ is IID and satisfies (0.1), the appropriate rescaling of $f_1+\cdots+f_k$ is suggested by the computation (0.4). We have

$$g_k = \frac{1}{\sqrt{k}} \sum_{j=1}^k f_j \Longrightarrow \|g_k\|_{L^2}^2 \equiv \sigma. \tag{0.16}$$

Note that if ν_1 is the probability distribution of f_1 (hence of f_j for all j), then for any Borel set $B \subset \mathbb{R}$,

$$\nu_{g_k}(B) = \nu_k(\sqrt{k}B), \quad \nu_k = \nu_1 * \cdots * \nu_1 \ (k \text{ factors}). \tag{0.17}$$

Note that

$$\int x^2 \, d\nu_1 = \sigma, \quad \int x \, d\nu_1 = 0. \tag{0.18}$$

In §1 we prove the following version of CLT:

Theorem 0.1. *If $\{f_j : j \in \mathbb{N}\}$ is IID on $(\Omega, \mathcal{F}, \mu)$, satisfying (0.1), and g_k is given by (0.16), then*

$$\nu_{g_k} \longrightarrow \gamma^\sigma, \quad \text{weak}^* \text{ in } \mathcal{M}(\widehat{\mathbb{R}}) = C(\widehat{\mathbb{R}})', \tag{0.19}$$

where $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, so

$$C(\widehat{\mathbb{R}}) = \{u \in C(\mathbb{R}) : u(x) \to u_\infty \text{ as } |x| \to \infty\}. \tag{0.20}$$

In §1 we also strengthen the conclusion (0.19) to

$$(1+x^2)\nu_{g_k} \longrightarrow (1+x^2)\gamma^\sigma, \quad \text{weak}^* \text{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{0.21}$$

Remark 1. The weak* convergence (0.19) means

$$\int f\, d\nu_{g_k} \longrightarrow \int f\, d\gamma^\sigma, \tag{0.22}$$

for each $f \in C(\widehat{\mathbb{R}})$. Since ν_{g_k} are finite positive measures, and γ^σ is absolutely continuous on $\widehat{\mathbb{R}}$, it is an automatic consequence that (0.22) holds whenever f is a bounded Borel function that is Riemann integrable on $\widehat{\mathbb{R}} \approx S^1$. See Appendix A for a brief discussion of this fact.

Remark 2. In contrast to the law of large numbers, the central limit theorem does not assert that $\{g_k\}$ converges to a random variable on Ω that is Gaussian with variance σ. In fact, the set $\{\sigma^{-1/2} f_j\}$ forms an orthonormal basis of a Hilbert space $\mathcal{H} \subset L^2(\Omega,\mu)$, and each g_k is an element of $\mathcal{H}$, and so is any limit. But, for each fixed j,

$$\lim_{k\to\infty} (f_j, g_k)_{L^2} = 0, \tag{0.23}$$

so in fact, as $k \to \infty$,

$$g_k \longrightarrow 0, \quad \text{weakly in } \ L^2(\Omega,\mu). \tag{0.24}$$

Remark 3. The review [F] seems to say that the proof of CLT on p. 194 of [GS] requires all the moments of ν_{f_1} to be finite. We can only recommend that the interested reader make an independent assessment of the proof given there. On the other hand, we must acknowledge the gaffe made on line 6, p. 200, of [O], though ignoring this errant phrase leaves a proof that is OK.

In §2 we study the coin toss, for which

$$\nu_{f_j} = \frac{1}{2}(\delta_1 + \delta_{-1}). \tag{0.25}$$

The analysis of ν_{g_k} for this case illustrates the "rough" manner in which the weak* limit (0.19) holds. Indeed, we have

$$\nu_{g_k} = \frac{1}{\sqrt{2\pi}} \widehat{C}_k(x)\, \lambda_k, \tag{0.26}$$

where λ_k is a sum of point masses supported at integer multiples of $k^{-1/2}$ (see (2.12)), and $C_k(\xi)$ is given by (2.5) and (2.8). While this does illuminate rough weak* convergence, we get a much more precise result than (0.19), namely, as $k \to \infty$,

$$\nu_{g_k} - \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \lambda_k \longrightarrow 0 \text{ in TV norm on } \mathcal{M}(\mathbb{R}). \tag{0.27}$$

This is proved as a consequence of the result that

$$\widehat{C}_k(x) \longrightarrow e^{-x^2/2}, \quad \text{uniformly, as } k \to \infty. \tag{0.28}$$

Going further, we show that, for each $\ell \in \mathbb{N}$,

$$\partial_x^\ell \widehat{C}_k(x) \longrightarrow \partial_x^\ell e^{-x^2/2}, \quad \text{uniformly, as } k \to \infty, \tag{0.29}$$

and also that

$$x^\ell \widehat{C}_k(x) \longrightarrow x^\ell e^{-x^2/2}, \quad \text{uniformly, as } k \to \infty, \tag{0.30}$$

where we start the sequence (0.30) at $k = \ell + 1$. We also have quantitative estimates on the rate of convergence, such as

$$\sup_{x \in \mathbb{R}} |\widehat{C}_k(x) - e^{-x^2/2}| \le \frac{C}{k}, \tag{0.31}$$

refining (0.28), and

$$\|\nu_{g_k} - \gamma^1 \, \lambda_k\|_{\mathrm{TV}(\mathbb{R})} \le C \frac{\sqrt{\log k}}{k}, \tag{0.32}$$

refining (0.27).

In §3, we return to more general IID sequences and examine the rate at which ν_{g_k} converges to γ^σ. We establish the following complement to Theorem 0.1.

Proposition 0.2. *In the setting of Theorem 0.1, and under the additional hypothesis that, for some $a > 0$,*

$$\chi_{f_j}(\xi) = e^{-\sigma \xi^2/2 + \xi^2 \beta(\xi)}, \quad \textit{for } |\xi| \le a, \tag{0.33}$$

where $|\beta(\xi)| \le \sigma/4$ on this interval, and

$$|\beta(\xi)| \le b|\xi|^r, \quad \textit{for some } r \in (0, 2], \tag{0.34}$$

we have

$$|\langle \nu_{g_k} - \gamma^\sigma, v \rangle| \le C k^{-r/2} \mathcal{A}(v) + |\langle \psi(k^{-1/2} D) \nu_{g_k}, v \rangle|, \tag{0.35}$$

where

$$\mathcal{A}(v) = \int_{-\infty}^{\infty} |\hat{v}(\xi)| e^{-\sigma\xi^2/8} |\xi|^{2+r}\, d\xi. \tag{0.36}$$

In (0.35), we take

$$\psi \in C^\infty(\mathbb{R}), \quad \psi(\xi) = 0 \text{ for } |\xi| \le \frac{a}{2},\ 1 \text{ for } |\xi| \ge a. \tag{0.37}$$

This result leads to the task of estimating the last term on the right side of (0.35), which we denote $\mathcal{B}_k(v)$. One straightforward estimate, established in §3, is that if $v \in \operatorname{Lip}(\mathbb{R})$, then

$$\mathcal{B}_k(v) \le C k^{-1/2} \operatorname{Lip}(v). \tag{0.38}$$

More generally, if $v \in C(\mathbb{R})$ and $\partial_x^m v \in L^\infty(\mathbb{R})$, we have

$$\mathcal{B}_k(v) \le C \|\partial_x^m v\|_{L^\infty} k^{-m/2}. \tag{0.39}$$

In §4, we consider circumstances under which we can derive a rate at which

$$\Phi_k(y) - G(y) \longrightarrow 0, \tag{0.40}$$

as $k \to \infty$, where

$$\Phi_k(y) = \nu_{g_k}((-\infty, y]), \quad G(y) = \gamma^\sigma((-\infty, y]). \tag{0.41}$$

The magnitude of this difference is of the form (0.35), with $v = v_y$ the indicator function of $(-\infty, y]$, but in this case the estimate of the last term in (0.35) is more difficult than that covered by (0.38). To deal with this, we approximate v_y by smooth functions $w_{y,h}$, equal to 0 for $x \ge y$, to 1 for $x \le y - h$, taking values in $[0,1]$ for $y - h \le x \le y$, and satisfying

$$|\partial_x^m w_{y,h}(x)| \le C_m h^{-m}. \tag{0.42}$$

An elementary argument gives

$$\sup_y |\Phi_k(y) - G(y)| \le \sup_y |\langle \nu_{g_k} - \gamma^\sigma, w_{y,h}\rangle| + Ch, \tag{0.43}$$

hence the left side of (0.43) is dominated by $Ck^{-r/2} + C_m k^{-m/2} h^{-m} + Ch$. Taking $h = k^{-m/2(m+1)}$ yields

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2} + C_m k^{-m/2(m+1)}, \tag{0.44}$$

in the setting of Proposition 0.2. In particular, taking m large enough, we have that the left side of (0.44) is

$$\le Ck^{-r/2}, \quad \text{provided } \ 0 < r < 1. \tag{0.45}$$

Such estimates were established by Liapunov.

In §5 we discuss the Berry-Esseen theorem, which treats the endpoint case of (0.45):

Theorem 0.3. *In the setting of Proposition 0.2, under the hypothesis that (0.33) holds with*

$$|\beta(\xi)| \le b|\xi|, \quad \text{for } \ |\xi| \le a, \tag{0.46}$$

we have

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-1/2}. \tag{0.47}$$

Note that the coin toss satisfies the hypotheses (0.33)–(0.34) with $r = 2$, but, as is clear from the estimate (0.32), comparing ν_{g_k} to the discretized Gaussian, in this case the exponent $-1/2$ in (0.47) cannot be improved.

While the exponent in (0.47) is optimal for the coin toss, there are other interesting cases where it is not. One example occurs when

$$\nu_{f_1} = \frac{1}{2}\mathbf{1}_{[-1,1]}(x), \tag{0.48}$$

in which case

$$\chi(\xi) = \frac{\sin \xi}{\xi}. \tag{0.49}$$

In §6 we estabish the following.

Proposition 0.4. *In the setting of Proposition 0.2, particularly with (0.34) for some $r \in (0, 2]$, and with the additional hypotheses that*

$$\sup_{|\xi| \ge a/2} |\chi(\xi)| \le \delta < 1, \quad \text{and} \quad \int_{-\infty}^{\infty} |\chi(\xi)|^{\ell}\, d\xi < \infty, \tag{0.50}$$

for some $\ell \in \mathbb{N}$, we have

$$|\langle \nu_{g_k} - \gamma^\sigma, v\rangle| \le C\mathcal{A}(v)k^{-r/2} + C\mathcal{S}_k(v)\delta^{k-\ell}k^{1/2}, \tag{0.51}$$

with $\mathcal{A}(v)$ as in (0.36) and

$$\mathcal{S}_k(v) = \sup_{|\xi| \ge (a/2)k^{1/2}} |\tilde{v}(\xi)|. \tag{0.52}$$

This applies to $v = v_y$, the indicator function of $(-\infty, y]$, to yield

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2}, \tag{0.53}$$

under the hypotheses of Proposition 0.4. In particular, we treat the case (0.48), obtaining (0.53) with $r = 2$.

In §7 we turn our attention to *tail estimates.* The first result of this nature is (0.21), which sharpens (0.19), in that it says more about the behavior of $\{\nu_{g_k}\}$ far out in $(-\infty,\infty)$. Going further, we establish the following.

Proposition 0.5. *In the setting of Theorem 0.1, assume also that, for some* $\ell \in \mathbb{N},\ \ell \ge 2$,

$$\int x^{2\ell}\, d\nu_{f_1}(x) < \infty. \tag{0.54}$$

Then

$$(1+x^{2\ell})\nu_{g_k} \longrightarrow (1+x^{2\ell})\gamma^\sigma, \quad \textit{weak}^* \textit{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{0.55}$$

This result is complemented by the following.

Proposition 0.6. *In the setting of Proposition 0.2, particularly including (0.34),*

$$\rho < r+2 \Rightarrow (1+x^2)^{\rho/2}\nu_{g_k} \to (1+x^2)^{\rho/2}\gamma^\sigma, \quad \textit{weak}^* \textit{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{0.56}$$

Furthermore, for such ρ,

$$v \in S^\rho(\mathbb{R}) \Longrightarrow |\langle \nu_{g_k} - \gamma^\sigma, v\rangle| \le C k^{-r/2}. \tag{0.57}$$

Here,

$$S^\rho(\mathbb{R}) = \{v \in C^\infty(\mathbb{R}) : |v^{(\ell)}(x)| \le C_\ell (1+|x|)^{\rho-\ell},\ \forall\, \ell \in \mathbb{Z}^+\}. \tag{0.58}$$

In §8 we expand the scope of CLT beyond results on approximating Gaussians. We look at probability measures on $\mathbb{R}$ arising from fractional diffusion equations:

$$\gamma^t_\alpha(x) = e^{-t(-\partial_x^2)^{\alpha/2}}\delta(x), \tag{0.59}$$

for $t>0,\ \alpha \in (0,2)$, and establish the following:

Theorem 0.7. *Assume* $\{f_j : j \in \mathbb{N}\}$ *is an IID sequence on* $(\Omega,\mathcal{F},\mu)$ *whose characteristic function* $\chi(\xi)$ *satisfies*

$$\chi(\xi) = 1 - t|\xi|^\alpha + r(\xi), \quad r(\xi) = o(|\xi|^\alpha), \quad \textit{as } \xi \to 0, \tag{0.60}$$

for some $t>0,\ \alpha \in (0,2)$. *Define* g_k *by*

$$g_k = k^{-1/\alpha}(f_1 + \cdots + f_k). \tag{0.61}$$

Then

$$\nu_{g_k} \longrightarrow \gamma^t_\alpha, \quad \textit{weak}^* \textit{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{0.62}$$

We close with some appendices. Appendix A discusses the fact that if we have a weak* convergent sequence of probability measures, $\nu_k \to \mu$, so

$$\int\limits_X f\, d\nu_k \longrightarrow \int\limits_X f\, d\mu, \quad \text{as} \quad k \to \infty, \tag{0.63}$$

for continuous f on X, then (0.63) automatically holds for a larger class of functions f, namely bounded Borel functions $f : X \to \mathbb{R}$ such that

$$f \in \mathcal{R}(X, \mu), \tag{0.64}$$

a space of "Riemann integrable" functions. Here X denotes a compact metric space, and μ is a probability measure on X. In the body of the text, this has several applications when $X = \widehat{\mathbb{R}}$, involving matters related to the Levy-Cramér continuity theorem. In Appendix B we pursue this further when $X = \widehat{\mathbb{R}}$ and μ has no atoms, and apply it to results on uniform convergence of $\Phi_k \to G$.

1. General CLT for IID random variables with finite second moments

As advertised in the introduction, our first task in this section is to prove the following.

Theorem 1.1. *Assume* $\{f_j : j \in \mathbb{N}\}$ *is an IID sequence on* $(\Omega, \mathcal{F}, \mu)$*, with mean zero, and satisfying* $\|f_j\|^2_{L^2(\Omega,\mu)} \equiv \sigma$*. Set*

$$g_k = \frac{1}{\sqrt{k}} \sum_{j=1}^{k} f_j, \tag{1.1}$$

and define γ^σ *as in (0.11). Then*

$$\nu_{g_k} \longrightarrow \gamma^\sigma, \quad \text{weak}^* \ \text{in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{1.2}$$

Proof. Applying the Fourier transform to the convolution identity in (0.17) yields

$$\chi_{g_k}(\xi) = \chi(k^{-1/2}\xi)^k, \tag{1.3}$$

where $\chi(\xi) = \sqrt{2\pi}\hat{\nu}_1(\xi)$. By (0.6)–(0.7) applied to (0.1), and the fact that the Fourier transform intertwines multiplication by x and $id/d\xi$, and that the Fourier transform of a finite measure is a bounded, continuous function, we have

$$\chi \in C^2(\mathbb{R}), \quad \chi'(0) = 0, \quad \chi''(0) = -\sigma. \tag{1.4}$$

Hence

$$\chi(\xi) = 1 - \frac{\sigma}{2}\xi^2 + r(\xi), \quad r(\xi) = o(\xi^2), \quad \text{as } \ \xi \to 0. \tag{1.5}$$

Equivalently, there exists $a > 0$ such that, for $|\xi| \le a$,

$$\chi(\xi) = e^{-\sigma\xi^2/2 + \xi^2\beta(\xi)}, \quad \beta(\xi) \to 0, \ \text{ as } \ \xi \to 0. \tag{1.6}$$

Hence

$$\chi_{g_k}(\xi) = e^{-\sigma\xi^2/2 + \xi^2\beta(k^{-1/2}\xi)}, \quad \text{for } \ |\xi| \le ak^{1/2}, \tag{1.7}$$

with

$$\beta(k^{-1/2}\xi) \longrightarrow 0 \ \text{ as } \ k \to \infty, \quad \forall\, \xi \in \mathbb{R}. \tag{1.8}$$

Therefore,

$$\lim_{k\to\infty} \hat{\nu}_{g_k}(\xi) = \hat{\gamma}^\sigma(\xi), \quad \forall\, \xi \in \mathbb{R}. \tag{1.9}$$

Now the functions $\hat{\nu}_{g_k}(\xi)$ are uniformly bounded by $1/\sqrt{2\pi}$. Making use of (1.9), the Parseval identity for the Fourier transform, and the dominated convergence theorem, we obtain for each $v \in \mathcal{S}(\mathbb{R})$ (the Schwartz space of rapidly decreasing functions) that

$$\begin{aligned} \int v\, d\nu_{g_k} &= \int \hat{v}(\xi)\hat{\nu}_{g_k}(\xi)\, d\xi \\ &\to \int \hat{v}(\xi)\hat{\gamma}^\sigma(\xi)\, d\xi \\ &= \int v\, \gamma^\sigma\, dx. \end{aligned} \tag{1.10}$$

An equivalent statement is that

$$\nu_{g_k} \longrightarrow \gamma^\sigma \quad \text{in} \quad \mathcal{S}'(\mathbb{R}), \tag{1.11}$$

where $\mathcal{S}'(\mathbb{R})$ denotes the Schwartz space of tempered distributions. However, since $\{\nu_{g_k} : k \in \mathbb{N}\}$ is bounded in $\mathcal{M}(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$ is dense in

$$C_*(\mathbb{R}) = \{u \in C(\widehat{\mathbb{R}}) : u(\infty) = 0\}, \tag{1.12}$$

we also have

$$\int v\, d\nu_{g_k} \longrightarrow \int v\, \gamma^\sigma\, dx, \tag{1.13}$$

for all $v \in C_*(\mathbb{R})$. Clearly (1.13) also holds for $v = 1$, so we have the conclusion (1.2).

We can strengthen the conclusion of Theorem 1.1, by using

$$\int x^2\, d\nu_{g_k}(x) = \|g_k\|^2_{L^2} \equiv \sigma. \tag{1.14}$$

In particular,

$$\{(1 + x^2)\nu_{g_k} : k \in \mathbb{N}\} \text{ is bounded in } \mathcal{M}(\widehat{\mathbb{R}}), \tag{1.15}$$

and we have from (1.11) that

$$(1 + x^2)\nu_{g_k} \longrightarrow (1 + x^2)\gamma^\sigma, \tag{1.16}$$

in $\mathcal{S}'(\mathbb{R})$, hence in $C_*(\mathbb{R})'$, and then, by (1.14), in $C(\widehat{\mathbb{R}})'$. This gives:

Proposition 1.2. *In the setting of Theorem 1.1, we have*

$$(1 + x^2)\nu_{g_k} \longrightarrow (1 + x^2)\gamma^\sigma, \quad \textit{weak}^* \textit{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{1.17}$$

2. Coin toss

To model a fair coin toss, one takes $X = \{1, -1\}$, each point having measure 1/2, and forms the probability space

$$\Omega = \prod_{j\in\mathbb{N}} \{1, -1\}, \tag{2.1}$$

with product Borel field and product measure. The random variables f_j, given by

$$f_j(\omega_1, \omega_2, \omega_3, \dots) = \omega_j, \tag{2.2}$$

are independent and satisfy (0.1), with $\sigma = 1$. We have

$$\nu_{f_j} = \nu = \frac{1}{2}(\delta_1 + \delta_{-1}), \quad \chi_{f_j}(\xi) = \chi(\xi) = \cos\xi, \tag{2.3}$$

and g_k, given by (0.16), has characteristic function

$$\chi_{g_k}(\xi) = \chi(k^{-1/2}\xi)^k, \tag{2.4}$$

as in (1.3).

To analyze this, we set

$$\begin{aligned} C(\xi) = \cos\xi \quad &\text{for} \quad |\xi| \le \frac{\pi}{2}, \\ 0 \quad &\text{otherwise}, \end{aligned} \tag{2.5}$$

so

$$\chi(\xi) = \sum_{n\in\mathbb{Z}} (-1)^n C(\xi + n\pi), \tag{2.6}$$

hence

$$\chi_{g_k}(\xi) = \sum_{n\in\mathbb{Z}} (-1)^{kn} C_k(\xi + k^{1/2}n\pi), \tag{2.7}$$

where we have set

$$C_k(\xi) = C(k^{-1/2}\xi)^k. \tag{2.8}$$

Note that the series (2.7) converges in $\mathcal{S}'(\mathbb{R})$. Applying the Fourier transform gives

$$\sqrt{2\pi}\,\nu_{g_k} = \widehat{C}_k(x)\lambda_k, \tag{2.9}$$

where

$$\text{(2.10/11)} \qquad \begin{aligned} \lambda_k &= \sum_{n\in\mathbb{Z}} (-1)^{kn} e^{ink^{1/2}\pi x} \\ &= \sum_{n\in\mathbb{Z}} e^{in\pi k^{1/2}(x+k^{1/2})}, \end{aligned}$$

convergence also holding in $\mathcal{S}'(\mathbb{R})$, on which $\widehat{C}_k$ acts as a multiplier. The Poisson summation formula gives

$$\text{(2.12)} \qquad \begin{aligned} \lambda_k = 2k^{-1/2} &\sum_{\ell\in\mathbb{Z}} \delta_{2\ell k^{-1/2}}, \qquad && k \text{ even}, \\ 2k^{-1/2} &\sum_{\ell\in\mathbb{Z}} \delta_{(2\ell+1)k^{-1/2}}, && k \ \text{ odd}. \end{aligned}$$

Thanks to (2.9), the task of producing a detailed asymptotic analysis of the behavior of ν_{g_k} is reduced to that of analyzing $\widehat{C}_k(x)$. For this, we can use techniques similar to those brought to bear in §1. These will yield stronger conclusions on $\widehat{C}_k$ than we obtained there for ν_{g_k}. Parallel to (1.6), we can write

$$\text{(2.13)} \qquad C(\xi) = e^{-\xi^2/2+\xi^2\beta(\xi)}, \quad \text{for } \ |\xi| < \frac{\pi}{2},$$

with

$$\text{(2.14)} \qquad \beta \in C^\infty(I), \quad \beta(\xi) = O(\xi^2), \quad I = \Big(-\frac{\pi}{2}, \frac{\pi}{2}\Big).$$

We also have

$$\text{(2.15)} \qquad 0 \le C(\xi) \le e^{-a\xi^2}, \quad \forall\, \xi \in \mathbb{R},$$

for some $a > 0$. It follows that

$$\text{(2.16)} \qquad C_k(\xi) = e^{-\xi^2/2+\xi^2\beta(k^{-1/2}\xi)}, \quad \text{for } \ |\xi| < \frac{\pi}{2}k^{1/2},$$

and furthermore

$$\text{(2.17)} \qquad 0 \le C_k(\xi) \le e^{-a\xi^2}, \quad \forall\, \xi \in \mathbb{R}.$$

Parallel to (1.9), we have from (2.16) and (2.14) that

$$\text{(2.18)} \qquad C_k(\xi) \longrightarrow e^{-\xi^2/2}, \quad \forall\, \xi \in \mathbb{R}.$$

The additional uniform bound (2.17) allows us to use the dominated convergence theorem to deduce that

$$\text{(2.19)} \qquad C_k \longrightarrow e^{-\xi^2/2} \ \text{ in } \ L^1(\mathbb{R}), \text{as } k \to \infty.$$

Hence

$$\text{(2.20)} \qquad \widehat{C}_k(x) \longrightarrow e^{-x^2/2} = \sqrt{2\pi}\gamma^1(x), \quad \text{uniformly, as } \ k \to \infty.$$

We are now in a position to establish the following, giving a much more precise analysis of ν_{g_k} than Theorem 1.1 does.

Proposition 2.1. *For Ω and f_j given by (2.1)–(2.2), λ_k by (2.12), we have*

$$\nu_{g_k} - \gamma^1(x)\lambda_k \longrightarrow 0 \ \ \text{in} \ \ \mathcal{M}(\mathbb{R}), \quad \text{in total variation norm.} \tag{2.21}$$

Proof. By (2.9), our conclusion is equivalent to the assertion that

$$\Big((2\pi)^{-1/2}\widehat{C}_k(x) - \gamma^1(x)\Big)\lambda_k \longrightarrow 0, \quad \text{in total variation norm.} \tag{2.22}$$

We can deduce this from (2.20) in concert with the facts that

$$(2\pi)^{-1/2}\widehat{C}_k\lambda_k = \nu_{g_k} \ \ \text{are probability measures on } \mathbb{R}, \tag{2.23}$$

and

$$\gamma^1(x)\lambda_k \ \ \text{are positive measures with mass } m_k \to 1. \tag{2.24}$$

To see this, pick $\varepsilon > 0$. Pick $A \in (0,\infty)$ so that, for all $k \in \mathbb{N}$, the total mass of $\gamma^1\lambda_k$ outside $[-A, A]$ is $\le \varepsilon$. Then pick $K \in \mathbb{N}$ so that

$$\begin{aligned} k \ge K \Longrightarrow |m_k - 1| \le \varepsilon, \ \text{and} \\ \max_{|x|\le A} \big|(2\pi)^{-1/2}\widehat{C}_k(x) - \gamma^1(x)\big| \le \frac{\varepsilon}{2A}. \end{aligned} \tag{2.25}$$

It follows that, for $k \ge K$, the total mass of the measure in (2.22) is $\le 4\varepsilon$, and we deduce the asserted result.

To complement the results (2.20)–(2.21), let us note that (2.17)–(2.18) imply

$$\begin{aligned} &\xi^\ell C_k(\xi) \longrightarrow \xi^\ell e^{-\xi^2/2}, \\ &0 \le |\xi|^\ell C_k(\xi) \le |\xi|^\ell e^{-a\xi^2}, \quad \forall\, \xi \in \mathbb{R}, \ \ell \in \mathbb{N}, \end{aligned} \tag{2.26}$$

hence

$$\partial_x^\ell \widehat{C}_k(x) \longrightarrow \partial_x^\ell e^{-x^2/2}, \quad \text{uniformly, as } k \to \infty, \ \forall\, \ell \in \mathbb{N}. \tag{2.27}$$

To proceed, we analyze the behavior of derivatives of $C_k(\xi)$. Note that

$$C_k \in C^\ell(\mathbb{R}), \quad \forall\, \ell < k. \tag{2.28}$$

Now (2.14) implies that, for each $m \in \mathbb{N}$,

$$\{\beta(k^{-1/2}\xi) : k \ge m\} \longrightarrow 0 \ \ \text{in} \ \ C^\infty(I_m), \tag{2.29}$$

as $k \to \infty$, where

$$I_m = \Big\{ \xi \in \mathbb{R} : |\xi| < \frac{\pi}{2} m^{1/2} \Big\}. \tag{2.30}$$

We deduce from (2.16) that

$$\{C_k : k \geq m\} \longrightarrow e^{-\xi^2/2} \;\text{ in }\; C^\infty(I_m), \tag{2.31}$$

as $k \to \infty$, and consequently, for each $\ell \in \mathbb{N}$,

$$\{C_k^{(\ell)}(\xi) : k > \ell\} \longrightarrow \partial_\xi^\ell e^{-\xi^2/2}, \quad \forall\, \xi \in \mathbb{R}. \tag{2.32}$$

Having this extension of (2.18), we seek uniform estimates on $\{C_k^{(\ell)} : k > \ell\}$, parallel to (2.17). Indeed, differentiating

$$C_k(\xi) = C(k^{-1/2}\xi)^k, \tag{2.33}$$

we have

$$\begin{aligned} C_k'(\xi) &= k^{1/2} C'(k^{-1/2}\xi) C(k^{-1/2}\xi)^{k-1} \\ &= \big[-k^{1/2}\sin(k^{-1/2}\xi)\big] C(k^{-1/2}\xi)^{k-1} \\ &= -\frac{\sin(k^{-1/2}\xi)}{k^{-1/2}\xi}\, \xi\, C(k^{-1/2}\xi)^{k-1}, \end{aligned} \tag{2.34}$$

so, by (2.17),

$$\begin{aligned} |C_k'(\xi)| &\leq |\xi| e^{-a(1-1/k)\xi^2} \\ &\leq |\xi| e^{-a\xi^2/2}, \quad \text{for } \; k \geq 2. \end{aligned} \tag{2.35}$$

Next,

$$\begin{aligned} C_k''(\xi) &= C''(k^{-1/2}\xi) C(k^{-1/2}\xi)^{k-1} \\ &\quad + (k-1) C'(k^{-1/2}\xi)^2 C(k^{-1/2}\xi)^{k-2}, \end{aligned} \tag{2.36}$$

and the analysis of $k^{1/2} C'(k^{-1/2}\xi)$ used in (2.34) yields

$$\begin{aligned} |C_k''(\xi)| &\leq C(k^{-1/2}\xi)^{k-1} + \xi^2 C(k^{-1/2}\xi)^{k-2} \\ &\leq (1+\xi^2) e^{-a(1-2/k)\xi^2} \\ &\leq (1+\xi^2) e^{-a\xi^2/3}, \quad \text{for } \; k \geq 3. \end{aligned} \tag{2.37}$$

From (2.32), (2.35), (2.37), and the dominated convergence theorem, we have

$$C_k^{(\ell)} \longrightarrow \partial_\xi^\ell e^{-\xi^2/2} \quad \text{in} \quad L^1(\mathbb{R}), \qquad \text{as} \quad k \to \infty, \tag{2.38}$$

for $\ell = 1, 2$, hence, complementing (2.20),

$$x^\ell \widehat{C}_k(x) \longrightarrow x^\ell e^{-x^2/2}, \quad \text{uniformly, as } k \to \infty, \tag{2.39}$$

for $\ell = 1, 2$. This is enough to give an alternative proof of (2.22), hence of Proposition 2.1.

From here, an inductive argument gives, for general $\ell \in \mathbb{N}$,

$$\begin{aligned} |C_k^{(\ell)}(\xi)| &\le A_\ell (1 + |\xi|^\ell) C(k^{-1/2}\xi)^{k-\ell} \\ &\le A_\ell (1 + |\xi|^\ell) e^{-a(1-\ell/k)\xi^2} \\ &\le A_\ell (1 + |\xi|^\ell) e^{-a\xi^2/(\ell+1)}, \quad \text{for} \quad k > \ell. \end{aligned} \tag{2.40}$$

From (2.32), (2.40), and the dominated convergence theorem, we have (2.38) for all $\ell \in \mathbb{N}$, and applying the Fourier transform yields the following result.

Proposition 2.2. *For each integer $\ell \ge 0$,*

$$x^\ell \widehat{C}_k(x) \longrightarrow x^\ell e^{-x^2/2}, \quad \textit{uniformly, as} \quad k \to \infty, \tag{2.41}$$

where we start the sequence (2.41) at $k = \ell + 1$.

We next investigate the rate at which the uniform convergence (2.20) holds, and its implications for an estimate for the rate at which norm convergence in (2.21) holds. We start with a more hands-on approach to (2.19), estimating

$$\int_{-\infty}^{\infty} |C_k(\xi) - e^{-\xi^2/2}|\, d\xi. \tag{2.42}$$

To start, we use the estimate (2.17) to dominate the integrand in (2.42) by $2e^{-a\xi^2}$, and use

$$\begin{aligned} \int\limits_{|\xi|\ge r} e^{-a\xi^2}\, d\xi &= 2\int_r^\infty e^{-a\xi^2}\, d\xi \\ &\le \frac{2}{r} \int_r^\infty e^{-a\xi^2} \xi\, d\xi \\ &= \frac{1}{ar} e^{-ar^2}, \end{aligned} \tag{2.43}$$

to estimate the integral (2.42) over $|\xi| \ge r$ (a quantity to be chosen below). To estimate the integral over $|\xi| \le r$, we use (2.16) (and (2.14)). We have

$$C_k(\xi) - e^{-\xi^2/2} = e^{-\xi^2/2}\Bigl(e^{\xi^2 \beta(k^{-1/2}\xi)} - 1\Bigr), \tag{2.44}$$

with

$$|\xi^2\beta(k^{-1/2}\xi)| \le Ck^{-1}\xi^4, \quad \text{for} \quad |\xi| \le \frac{\pi}{4}k^{1/2}, \tag{2.45}$$

hence

$$\left|e^{\xi^2\beta(k^{-1/2}\xi)} - 1\right| \le Ck^{-1}\xi^4, \quad \text{for} \quad |\xi| \le k^{1/4} \tag{2.46}$$

We deduce that, with

$$r(k) = k^{1/4}, \tag{2.47}$$

we have

$$\begin{aligned} \int\limits_{|\xi|\le r(k)} |C_k(\xi) - e^{-\xi^2/2}|\, d\xi &\le \frac{C}{k} \int\limits_{|\xi|\le r(k)} e^{-\xi^2/2}\xi^4\, d\xi \\ &\le \frac{C}{k}\int_{-\infty}^{\infty} e^{-\xi^2/2}\xi^4\, d\xi \\ &= \frac{C'}{k}. \end{aligned} \tag{2.48}$$

Hence, if we take $r = r(k)$ in (2.43), we have

$$\|C_k - e^{-\xi^2/2}\|_{L^1(\mathbb{R})} \le \frac{C}{k}. \tag{2.49}$$

This refines (2.20) to

$$\sup_{x\in\mathbb{R}} |\widehat{C}_k(x) - e^{-x^2/2}| \le \frac{C}{k}. \tag{2.50}$$

With this estimate in hand, we can tackle the quantitative refinement of Proposition 2.1, and estimate the total variation norm of (2.21). Let's start by considering

$$m_k = \|\gamma^1\lambda_k\|_{\mathrm{TV}}. \tag{2.51}$$

We can deduce from Jacobi's formula,

$$\begin{aligned} \sum_{\ell\in\mathbb{Z}} e^{-\varepsilon\ell^2} &= \Big(\frac{\pi}{\varepsilon}\Big)^{1/2} \sum_{n\in\mathbb{Z}} e^{-n^2\pi^2/\varepsilon} \\ &= \Big(\frac{\pi}{\varepsilon}\Big)^{1/2} \Big(1 + O(e^{-\pi^2/\varepsilon})\Big), \end{aligned} \tag{2.52}$$

that

$$|1 - m_k| \le C e^{-bk^{1/2}}, \tag{2.53}$$

for some $b > 0,\ C < \infty$. It will be convenient to bring in the sequence of probability measures

$$\mu_k = m_k^{-1} \gamma^1(x) \lambda_k. \tag{2.54}$$

Now to the total variation estimate. By (2.22) and (2.50),

$$\|\nu_{g_k} - \gamma^1\, \lambda_k\|_{\mathrm{TV}(I_k)} \le \frac{C}{k} \ell(I_k), \tag{2.55}$$

where

$$I_k = [-s(k), s(k)], \tag{2.56}$$

with $s(k)$ to be selected shortly. Meanwhile, parallel to (2.43),

$$\|\gamma^1\, \lambda_k\|_{\mathrm{TV}(\mathbb{R}\setminus I_k)} \le C e^{-s(k)^2/2}. \tag{2.57}$$

It is hence tempting to take

$$s(k) = \sqrt{2 \log k}. \tag{2.58}$$

In light of (2.53)–(2.54), we have

$$\|\nu_{g_k} - \mu_k\|_{\mathrm{TV}(I_k)} \le C \frac{\sqrt{\log k}}{k}, \quad \|\mu_k\|_{\mathrm{TV}(\mathbb{R}\setminus I_k)} \le \frac{C}{k}. \tag{2.59}$$

Also, since ν_{g_k} and μ_k are both probability measures on $\mathbb{R}$, we have

$$\begin{aligned} \|\nu_{g_k}\|_{\mathrm{TV}(\mathbb{R}\setminus I_k)} &= 1 - \|\nu_{g_k}\|_{\mathrm{TV}(I_k)} \\ &= 1 - \|\mu_k\|_{\mathrm{TV}(I_k)} + O\Big(\frac{\sqrt{\log k}}{k}\Big) \\ &= \|\mu_k\|_{\mathrm{TV}(\mathbb{R}\setminus I_k)} + O\Big(\frac{\sqrt{\log k}}{k}\Big). \end{aligned} \tag{2.60}$$

Putting together (2.55)–(2.60), we have:

Proposition 2.3. *In the setting of Proposition 2.1,*

$$\|\nu_{g_k} - \gamma^1(x)\, \lambda_k\|_{\mathrm{TV}(\mathbb{R})} \le C \frac{\sqrt{\log k}}{k}, \tag{2.61}$$

for $k \ge 2$.

3. Estimates on rate of approach of ν_{g_k} to γ^σ

Here we derive some estimates on the rate at which

$$\langle \nu_{g_k} - \gamma^\sigma, v\rangle \longrightarrow 0, \tag{3.1}$$

as $k \to \infty$, for ν_{g_k} as in (0.17) and γ^σ as in (0.11). We retain the hypothesis (0.1). We take v in various function spaces, and impose various conditions on ν_{f_j}, beyond having a finite second moment. For example, we consider the condition $f_j \in L^p(\Omega, \mu)$ for $p = 2 + r > 2$, or equivalently

$$\int |x|^{2+r}\, d\nu_{f_j}(x) < \infty. \tag{3.2}$$

This implies that

$$\chi = \chi_{f_j} \in C^{2+r}(\mathbb{R}). \tag{3.3}$$

In such a case, we can refine (1.6) to

$$\chi(\xi) = e^{-\sigma\xi^2/2 + \xi^2\beta(\xi)}, \quad \text{for } \ |\xi| \le a, \tag{3.4}$$

where

$$|\beta(\xi)| \le b|\xi|^r, \quad \text{provided } \ r \in (0,1]. \tag{3.5}$$

If by chance (3.2) holds with $r \ge 1$ and

$$\int x^3\, d\nu_{f_j} = 0, \tag{3.6}$$

we can expand the scope of (3.5) to

$$|\beta(\xi)| \le b|\xi|^r, \quad \text{provided } \ r \in (0,2]. \tag{3.7}$$

To start the estimate of (3.1), we have

$$\begin{aligned} \sqrt{2\pi}\langle \nu_{g_k} - \gamma^\sigma, v\rangle &= \sqrt{2\pi}\langle \hat{\nu}_{g_k} - \hat{\gamma}^\sigma, \tilde{v}\rangle \\ &= \int \Bigl[\chi_{g_k}(\xi) - e^{-\sigma\xi^2/2}\Bigr]\overline{\tilde{v}(\xi)}\, d\xi. \end{aligned} \tag{3.8}$$

Now

$$\chi_{g_k}(\xi) - e^{-\sigma\xi^2/2} = e^{-\sigma\xi^2/2}\Bigl(e^{\xi^2\beta(k^{-1/2}\xi)} - 1\Bigr), \quad \text{for } \ |\xi| \le ak^{1/2}, \tag{3.9}$$

and (3.5) (or (3.7)) implies

$$|\xi^2\beta(k^{-1/2}\xi)| \le bk^{-r/2}|\xi|^{2+r}, \quad \text{for} \;\; |\xi| \le ak^{1/2}. \tag{3.10}$$

It follows that

$$\left|e^{\xi^2\beta(k^{-1/2}\xi)} - 1\right| \le \tilde{b}k^{-r/2}|\xi|^{2+r}, \tag{3.11}$$

for $k^{-r/2}|\xi|^{2+r} \le 1$, or equivalently for

$$|\xi| \le k^{e(r)}, \quad e(r) = \frac{r}{2(2+r)}. \tag{3.12}$$

Shrinking a if necessary, we also arrange that

$$|\beta(k^{-1/2}\xi)| \le \frac{\sigma}{4}, \quad \text{for} \;\; |\xi| \le ak^{1/2}, \tag{3.13}$$

so

$$\left|e^{\xi^2\beta(k^{-1/2}\xi)} - 1\right| \le 2e^{\sigma\xi^2/4}, \quad \text{for} \;\; k^{e(r)} \le |\xi| \le ak^{1/2}. \tag{3.14}$$

We will make do with the estimate

$$|\chi_{g_k}(\xi)| \le 1, \quad \text{for} \;\; |\xi| \ge ak^{1/2}. \tag{3.15}$$

We therefore divide the range of integration $\mathbb{R}$ on the right side of (3.8) into three pieces:

$$|\xi| \le k^{e(r)}, \quad k^{e(r)} \le |\xi| \le ak^{1/2}, \quad |\xi| \ge ak^{1/2}, \tag{3.16}$$

and obtain the following result.

Proposition 3.1. *In the setting of Theorem 1.1, and with the additional hypothesis that (3.4) holds, with*

$$|\beta(\xi)| \le b|\xi|^r, \quad \textit{for some} \;\; r \in (0,2], \tag{3.17}$$

we have

$$\sqrt{2\pi}\left|\langle \nu_{g_k} - \gamma^\sigma, v\rangle\right| \le A_k(v) + B_k(v) + C_k(v), \tag{3.18}$$

where

$$\begin{aligned} A_k(v) &= \tilde{b}k^{-r/2} \int\limits_{|\xi|\le k^{e(r)}} |\tilde{v}(\xi)|e^{-\sigma\xi^2/2}|\xi|^{2+r}\,d\xi, \\ B_k(v) &= 2 \int\limits_{k^{e(r)}\le|\xi|\le ak^{1/2}} |\tilde{v}(\xi)|e^{-\sigma\xi^2/4}\,d\xi, \\ C_k(v) &= 2 \int\limits_{|\xi|\ge ak^{1/2}} |\tilde{v}(\xi)|\,d\xi. \end{aligned} \tag{3.19}$$

Note that

$$
\begin{aligned}
&A_k(v) \le \widetilde{A}_k(v) = \tilde{b} k^{-r/2} \int_{-\infty}^{\infty} |\tilde{v}(\xi)| e^{-\sigma \xi^2/2} |\xi|^{2+r}\, d\xi, \\
&B_k(v) \le \widetilde{B}_k(v) = 2 e^{-(\sigma/8) k^{2e(r)}} \int\limits_{|\xi| \ge k^{e(r)}} |\tilde{v}(\xi)| e^{-\sigma \xi^2/8}\, d\xi, \\
&C_k(v) \le \widetilde{C}_k(v) = \frac{4}{a} k^{-1/2} \sup_{\xi} \xi^2 |\tilde{v}(\xi)|.
\end{aligned} \tag{3.20}
$$

Clearly the seminorms $\widetilde{A}_k$ and $\widetilde{B}_k$ are quite nicely behaved on rather wild functions v. However, the seminorms C_k and $\widetilde{C}_k$ are not finite on a number of test functions v we would like to use. This provides motivation to modify the frequency cutoffs. We hence bring in the functions φ and ψ, satisfying the following conditions:

$$
\varphi, \psi \in C^\infty(\mathbb{R}), \quad \varphi(\xi) = 1 \text{ for } |\xi| \le \frac{a}{2},\ 0 \text{ for } |\xi| \ge a, \quad \psi = 1 - \varphi. \tag{3.21}
$$

We toss in the conditions

$$
0 \le \varphi \le 1, \quad \varphi(-\xi) = \varphi(\xi). \tag{3.22}
$$

Now we have

$$
\langle \nu_{g_k} - \gamma^\sigma, v \rangle = \langle \varphi(k^{-1/2} D)(\nu_{g_k} - \gamma^\sigma), v \rangle + \langle \psi(k^{-1/2} D)(\nu_{g_k} - \gamma^\sigma), v \rangle, \tag{3.23}
$$

and estimates arising in the proof of Proposition 3.1 imply

$$
|\langle \varphi(k^{-1/2} D)(\nu_{g_k} - \gamma^\sigma), v \rangle| \le C k^{-r/2} \mathcal{A}(v), \tag{3.24}
$$

where

$$
\mathcal{A}(v) = \int_{-\infty}^{\infty} |\tilde{v}(\xi)| e^{-\sigma \xi^2/8} |\xi|^{2+r}\, d\xi. \tag{3.25}
$$

We also have

$$
|\langle \psi(k^{-1/2} D) \gamma^\sigma, v \rangle| = \frac{1}{\sqrt{2\pi}} |\langle e^{-\sigma \xi^2/2}, \psi(k^{-1/2} \xi) \tilde{v}(\xi) \rangle| \le C e^{-b k^{1/2}} \mathcal{A}(v). \tag{3.26}
$$

This gives the following.

Proposition 3.2. *In the setting of Proposition 3.1,*

$$|\langle \nu_{g_k} - \gamma^\sigma, v\rangle| \le Ck^{-r/2}\mathcal{A}(v) + |\langle \psi(k^{-1/2}D)\nu_{g_k}, v\rangle|. \tag{3.27}$$

Other ways to present the last term arise via the identities

$$\begin{aligned} \langle \psi(k^{-1/2}D)\nu_{g_k}, v\rangle &= \langle \nu_{g_k}, \psi(k^{-1/2}D)v\rangle \\ &= \langle \psi(2k^{-1/2}D)\nu_{g_k}, \psi(k^{-1/2}D)v\rangle, \end{aligned} \tag{3.28}$$

the latter via

$$\psi(2\xi)\psi(\xi) = \psi(\xi). \tag{3.29}$$

We now have the task of estimating

$$\mathcal{B}_k(v) = |\langle \nu_{g_k}, \psi(k^{-1/2}D)v\rangle|. \tag{3.30}$$

Here is one straightforward result.

Proposition 3.3. *Assume v is Lipschitz continuous, with Lipschitz constant* $\operatorname{Lip}(v) = L$*:*

$$|v(x) - v(y)| \le L|x-y|, \quad \forall\, x, y \in \mathbb{R}. \tag{3.31}$$

Then

$$\mathcal{B}_k(v) \le Ck^{-1/2}\operatorname{Lip}(v). \tag{3.32}$$

Proof. Clearly

$$\mathcal{B}_k(v) \le \sup_x |\psi(k^{-1/2}D)v(x)|. \tag{3.33}$$

With $f = \sqrt{2\pi}\hat{\varphi}$, an element of $\mathcal{S}(\mathbb{R})$ that integrates to 1, we have, for all $x \in \mathbb{R}$,

$$\begin{aligned} |\psi(k^{-1/2}D)v(x)| &= \Bigl| \int k^{1/2} f(k^{1/2}y) v(x-y)\,dy - v(x) \Bigr| \\ &= \Bigl| \int k^{1/2} f(k^{1/2}y) \Bigl[v(x-y) - v(x) \Bigr]\, dy \Bigr| \\ &\le \operatorname{Lip}(v) \int k^{1/2} |f(k^{1/2}y)y|\, dy \\ &= k^{-1/2}\operatorname{Lip}(v) \int |f(y)y|\, dy. \end{aligned} \tag{3.34}$$

This gives (3.32).

The following result is a useful extension of Proposition 3.3.

Proposition 3.4. *Let* $m \in \mathbb{N}$*. Take* $v \in C(\mathbb{R})$ *and assume*

$$\partial_x^m v \in L^\infty(\mathbb{R}). \tag{3.35}$$

Then

$$\mathcal{B}_k(v) \le C_m k^{-m/2} L_m(v), \quad L_m(v) = \|\partial_x^m v\|_{L^\infty(\mathbb{R})}. \tag{3.36}$$

Proof. Set

$$\psi_m(\xi) = \xi^{-m}\psi(\xi), \tag{3.37}$$

and note that

$$\hat{\psi}_m \in L^1(\mathbb{R}), \quad \text{for} \ \ m \in \mathbb{N}. \tag{3.38}$$

We have

$$\psi(k^{-1/2}D)v(x) = k^{-m/2}\psi_m(k^{-1/2}D)(i\partial_x)^m v(x), \tag{3.39}$$

so

$$\sup_x |\psi(k^{-1/2}D)v(x)| \le C\|\hat{\psi}_m\|_{L^1(\mathbb{R})}\|\partial_x^m v\|_{L^\infty(\mathbb{R})} k^{-m/2}, \tag{3.40}$$

and (3.36) follows.

4. Convergence of distribution functions – Liapunov estimates

In this section we study the rate of convergence of

$$\Phi_k(y) \longrightarrow G(y), \tag{4.1}$$

as $k \to \infty$, where

$$\Phi_k(y) = \nu_{g_k}((-\infty, y]), \quad G(y) = \gamma^\sigma((-\infty, y]). \tag{4.2}$$

We retain the hypotheses on g_k in effect in Theorem 1.1, supplemented by those in Proposition 3.1, especially that (3.4) and (3.17) hold, i.e., the characteristic function $\chi(\xi)$ satisfies

$$\chi(\xi) = e^{-\sigma\xi^2/2+\xi^2\beta(\xi)}, \quad \text{for} \ \ |\xi| \le a, \tag{4.3}$$

and

$$|\beta(\xi)| \le b|\xi|^r, \quad \text{with} \ \ r \in (0, 2]. \tag{4.4}$$

Recall that

$$\chi_{g_k}(\xi) = \chi(k^{-1/2}\xi)^k. \tag{4.5}$$

To put the desired analysis in the framework of Proposition 3.2, we have

$$\Phi_k(y) - G(y) = \langle \nu_{g_k} - \gamma^\sigma, v_y \rangle, \tag{4.6}$$

where

$$\begin{aligned} v_y(x) = 1, &\quad \text{if} \ \ x \le y, \\ 0, &\quad \text{if} \ \ x > y. \end{aligned} \tag{4.7}$$

Proposition 3.2 is applicable, and we have

$$|\Phi_k(y) - G(y)| \le Ck^{-r/2}\mathcal{A}(v_y) + \mathcal{B}_k(v_y), \tag{4.5}$$

where

$$\begin{aligned} \mathcal{A}(v) &= \int_{-\infty}^{\infty} |\tilde{v}(\xi)| e^{-\sigma\xi^2/8} |\xi|^{2+r} \, d\xi, \\ \mathcal{B}_k(v) &= |\langle \nu_{g_k}, \psi(k^{-1/2}D)v \rangle|. \end{aligned} \tag{4.9}$$

Note that, with v_y given by (4.7), the inverse Fourier transform $\tilde{v}_y$ is a principal value distribution, with $1/\xi$ type blowup as $\xi \to 0$, but this singularity is cancelled out by the factor $|\xi|^{2+r}$. We have $\tilde{v}_y = e^{iy\xi}\tilde{v}_0$, so there is a uniform bound

$$\mathcal{A}(v_y) \le A_0 < \infty, \quad \forall\, y \in \mathbb{R}. \tag{4.10}$$

A direct estimate of $\mathcal{B}_k(v_r)$ seems not so simple. Instead, we follow [V] and sneak up on the problem of estimating (4.6) by bringing in

$$w_{y,h}(x) = \begin{array}{ll} 0, & \text{if } x \ge y, \\ \frac{h-(x-y)}{h}, & \text{if } y-h \le x \le y, \\ 1, & \text{if } x \le y-h. \end{array} \tag{4.11}$$

For $h \ge 0$, $v_{y-h} \le w_{y,h} \le v_y$, so

$$\langle \nu_{g_k}, v_{y-h}\rangle \le \langle \nu_{g_k}, w_{y,h}\rangle \le \langle \nu_{g_k}, v_y\rangle, \tag{4.12}$$

and

$$-\langle \gamma^\sigma, v_y\rangle \le -\langle \gamma^\sigma, w_{y,h}\rangle \le -\langle \gamma^\sigma, v_{y-h}\rangle, \tag{4.13}$$

hence

$$\langle \nu_{g_k} - \gamma^\sigma, w_{y,h}\rangle \le \langle \nu_{g_k} - \gamma^\sigma, v_y\rangle + \langle \gamma^\sigma, v_y - v_{y-h}\rangle, \tag{4.14}$$

and

$$\langle \nu_{g_k} - \gamma^\sigma, v_{y-h}\rangle - \langle \gamma^\sigma, v_y - v_{y-h}\rangle \le \langle \nu_{g_k} - \gamma^\sigma, w_{y,h}\rangle. \tag{4.15}$$

Since $0 \le \langle \gamma^\sigma, v_y - v_{y-h}\rangle \le Ch$, we have

$$\sup_y |\langle \nu_{g_k} - \gamma^\sigma, v_y\rangle| \le \sup_y |\langle \nu_{g_k} - \gamma^\sigma, w_{y,h}\rangle| + Ch. \tag{4.16}$$

Estimates parallel to (4.10) apply to $\mathcal{A}(w_{y,h})$:

$$\mathcal{A}(w_{y,h}) \le A_1 < \infty, \quad \forall\, y \in \mathbb{R}, h > 0. \tag{4.17}$$

Since also $\mathrm{Lip}(w_{y,h}) = 1/h$, Propositions 3.2–3.3 apply, giving

$$|\langle \nu_{g_k} - \gamma^\sigma, w_{y,h}\rangle| \le Ck^{-r/2}\mathcal{A}(w_{y,h}) + Ck^{-1/2}h^{-1}, \tag{4.18}$$

Hence (4.16) yields

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2} + Ck^{-1/2}h^{-1} + Ch, \tag{4.19}$$

for all $h > 0$. We choose $h = k^{-1/4}$ to balance the last two terms on the right side of (4.19), and obtain the following.

Proposition 4.1. *For ν_{g_k} as in Proposition 3.1, in particular with (4.3)–(4.4) holding, we have*

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2} + Ck^{-1/4}. \tag{4.20}$$

Another way to represent the right side of (4.20) is as

$$\le Ck^{-\delta(r)}, \quad \delta(r) = \min\Bigl(\frac{r}{2}, \frac{1}{4}\Bigr). \tag{4.21}$$

The exponent in (4.21) is sharp if $r \in (0, 1/2]$, but for larger r, one can do better.

For this, we want to replace the mollification $w_{y,h}$ of v_y by the following. Take

$$\zeta \in C_0^\infty(-1, 0), \quad \zeta \ge 0, \quad \int \zeta(x)\, dx = 1, \tag{4.22}$$

set $\zeta_h(x) = h^{-1}\zeta(h^{-1}x)$, and then set

$$w_{y,h} = \zeta_h * v_y. \tag{4.23}$$

In common with (4.11), we have

$$\begin{aligned} w_{y,h}(x) = 0&, \quad \text{if } \ x \ge y, \\ 1&, \quad \text{if } \ x \le y - h, \end{aligned} \tag{4.24}$$

and

$$0 \le w_{y,h}(x) \le 1, \quad \text{if } \ y - h \le x \le y, \tag{4.25}$$

but now $w_{y,h} \in C^\infty(\mathbb{R})$, and, for $m \in \mathbb{N}$,

$$\|\partial_x^m w_{y,h}\|_{L^\infty(\mathbb{R})} = A_m h^{-m}. \tag{4.26}$$

Estimates of the form (4.12)–(4.17) continue to hold. This time, we use (4.26) in concert with Propositions 3.2 and 3.4 to obtain the following variant of (4.18):

$$|\langle \nu_{g_k} - \gamma^\sigma, w_{y,h}\rangle| \le Ck^{-r/2}\mathcal{A}(w_{y,h}) + C_m k^{-m/2} h^{-m}, \tag{4.27}$$

which in concert with (4.16) gives the following variant of (4.19):

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2} + Ck^{-m/2}h^{-m} + Ch, \tag{4.28}$$

for all $h \in (0, 1]$. This time we choose h to make $k^{-m/2}h^{-m} = h$, i.e.,

$$h = k^{-m/2(m+1)}, \tag{4.29}$$

and we get the following extension of Proposition 4.1.

Proposition 4.2. *In the setting of Proposition 4.1, we have, for each* $m \in \mathbb{N}$*,*

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2} + C_m k^{-m/2(m+1)}. \tag{4.30}$$

Consequently, as long as (4.3)–(4.4) hold with

$$0 < r < 1, \tag{4.31}$$

we have

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-r/2}. \tag{4.32}$$

One interesting corollary arises by writing

$$\nu_{g_k}([y, y + k^{-r/2}]) = \Phi_k(y + k^{-r/2}) - \Phi_k(y), \tag{4.33}$$

using (4.32), and estimating $G(y + k^{-1/2}) - G(y)$. We obtain the following.

Corollary 4.3. *In the setting of Proposition 4.2, particularly assuming (4.3)-(4.4) hold and* $r \in (0, 1)$*, there exists* $C < \infty$ *such that*

$$\nu_{g_k}([y, y + k^{-r/2}]) \le Ck^{-r/2}, \quad \forall\, y \in \mathbb{R}. \tag{4.34}$$

5. The Berry-Esseen theorem

The Berry-Esseen theorem treats the endpoint case of the results established in §4. Here is a statement.

Theorem 5.1. *Assume f_j are IID random variables satisfying (0.1), and define g_k as in (0.16), and Φ_k and G as in (0.41). Assume in addition that*

$$\int_{\Omega} |f_j|^3 \, d\mu = \rho < \infty. \tag{5.1}$$

Then there exists $C < \infty$ such that

$$\sup_y |\Phi_k(y) - G(y)| \le C k^{-1/2}. \tag{5.2}$$

To start the proof, we have the setting of Proposition 3.2, with $r = 1$. Hence (4.8)–(4.10) hold, with $r = 1$ and v_y given by (4.7). That is to say,

$$|\Phi_k(y) - G(y)| \le C A_0 k^{-1/2} + \mathcal{B}_k(v_y), \tag{5.3}$$

and, recall,

$$\mathcal{B}_k(v_y) = |\langle \nu_{g_k}, \psi(k^{-1/2} D) v_y \rangle|, \tag{5.4}$$

with ψ as in (3.21).

Tp proceed, we take an approach to the estimate of $\mathcal{B}_k(v_y)$ rather different from that used in §4. Note that

$$\psi(k^{-1/2} D) v_y(x) = \psi(k^{-1/2} D) v_0(x - y), \tag{5.5}$$

and

$$\psi(k^{-1/2} D) v_0(x) = v_0(x) - \varphi(k^{-1/2} D) v_0(x) = V(k^{1/2} x), \tag{5.6}$$

where $V \in L^\infty(\mathbb{R}) \cap C^\infty(\mathbb{R} \setminus 0)$ has a simple jump at $x = 0$ and $V(x)$ is rapidly decreasing as $|x| \to \infty$. Then

$$\psi(k^{-1/2} D) v_y(x) = V(k^{1/2}(x - y)), \tag{5.7}$$

with

$$|V(x)| \le C_n \langle x \rangle^{-n}, \quad \forall\, n \in \mathbb{N}. \tag{5.8}$$

The next key ingredient in the proof of Theorem 5.1 is the following useful extension of the estimates (4.34) on ν_{g_k}.

Proposition 5.2. *Assume f_j are IID random variables satisfying (0.1) and define g_k as in (0.12). Then there exists $C < \infty$ such that*

$$\nu_{g_k}([y, y+k^{-1/2}]) \le C k^{-1/2}, \quad \forall\, y \in \mathbb{R}, k \in \mathbb{N}. \tag{5.9}$$

Once we have this, then we get

$$\mathcal{B}_k(v_y) \le \int |V(k^{1/2}(x-y))|\, d\nu_{g_k}(x), \tag{5.10}$$

and (5.8)–(5.9) imply this is $\le C k^{-1/2}$, as stated in (5.2). It remains to give the

Proof of Proposition 5.2. Pick ϕ satisfying

$$\phi \in C_0^\infty((-a,a)), \quad \phi \ge 0, \quad \phi(0) = 1. \tag{5.11}$$

We desire to estimate

$$\phi(k^{-1/2}D)\nu_{g_k}(x). \tag{5.12}$$

Note that its Fourier transform is

$$\phi(k^{-1/2}\xi)\chi_{g_k}(\xi) = \phi(k^{-1/2}\xi) e^{-\sigma\xi^2/2 + \xi^2\beta(k^{-1/2}\xi)}. \tag{5.13}$$

As in (3.13), we can assume

$$|\beta(\xi)| \le \frac{\sigma}{4} \ \text{ for } \ |\xi| \le a, \tag{5.14}$$

so

$$|\phi(k^{-1/2}\xi)\chi_{g_k}(\xi)| \le C e^{-\sigma\xi^2/8}, \quad \forall\, \xi \in \mathbb{R}, k \in \mathbb{N}. \tag{5.15}$$

This gives an L^1-norm bound that implies

$$|\phi(k^{-1/2}D)\nu_{g_k}(x)| \le C, \quad \forall\, x \in \mathbb{R}, k \in \mathbb{N}. \tag{5.16}$$

Note that

$$\phi(k^{-1/2}D)\nu_{g_k}(x) = c k^{1/2} \int \hat{\phi}(k^{1/2}(x-y))\, d\nu_{g_k}(y). \tag{5.17}$$

We can pick ϕ satisfying (5.11) and also

$$\hat{\phi}(x) \ge 0, \quad \forall\, x \in \mathbb{R}. \tag{5.18}$$

Then $\hat{\phi}(x)$ is bounded away from 0 on some neighborhood of 0, so (5.16)–(5.17) yield (5.9).

The proof of Theorem 5.1 is complete.

6. Faster convergence for more regular ν_{f_1}

The Berry-Esseen theorem gives the optimal rate of convergence of Φ_k to G for general IID random variables $f_j \in L^p(\Omega,\mu)$, satisfying (0,1), for each $p \ge 3$, namely

$$\sup_y |\Phi_k(y) - G(y)| \le Ck^{-1/2}. \tag{6.1}$$

As we have noted, this estimate is optimal for the coin toss. However, one does have faster convergence for lots of natural cases. Consider for example a case where ν_{f_j} is Lebesgue measure on $\mathbb{R}$ times

$$\begin{aligned} F(x) = \frac{1}{2} &\quad \text{for} \ \ |x| \le 1, \\ 0 &\quad \text{otherwise}. \end{aligned} \tag{6.2}$$

We have

$$\chi(\xi) = \frac{\sin \xi}{\xi}, \tag{6.3}$$

and $\chi(k^{-1/2}\xi)^k$ tends to $e^{-\sigma\xi^2/2}$ (with $\sigma = 1/3$) much more nicely than does its counterpart for the coin toss. The following result distills features that lead to improvements of (6.1).

Proposition 6.1. *Take an IID sequence $\{f_j\}$ as in Theorem 1.1. As in Proposition 3.1, assume $\chi = \chi_{f_j}$ satisfies (for some $a > 0$)*

$$\chi(\xi) = e^{-\sigma\xi^2/2+\xi^2\beta(\xi)}, \quad \textit{for} \ \ |\xi| \le a, \tag{6.4}$$

where, for ξ in this interval,

$$|\beta(\xi)| \le \frac{\sigma}{4}, \quad \textit{and} \ \ |\beta(\xi)| \le C|\xi|^r, \quad \textit{for some} \ \ r \in (0,2]. \tag{6.5}$$

Add the following hypotheses:

$$\sup_{|\xi|\ge a/2} |\chi(\xi)| \le \delta < 1, \tag{6.6}$$

and, for some $\ell \in \mathbb{N}$,

$$\int_{-\infty}^{\infty} |\chi(\xi)|^\ell \, d\xi < \infty. \tag{6.7}$$

Then, for $k \geq \ell$,

$$|\langle \nu_{g_k} - \gamma^\sigma, v\rangle \leq C\mathcal{A}(v)k^{-r/2} + C\mathcal{S}_k(v)\delta^{k-\ell}k^{1/2}, \tag{6.8}$$

with $\mathcal{A}(v)$ as in (3.25), and

$$\mathcal{S}_k(v) = \sup_{|\xi|\geq(a/2)k^{1/2}} |\tilde{v}(\xi)|. \tag{6.9}$$

Proof. By Proposition 3.2, it remains to estimate

$$\langle \psi(k^{-1/2}D)\nu_{g_k}, v\rangle = \int \psi(k^{-1/2}\xi)\chi(k^{-1/2}\xi)^k \tilde{v}(\xi)\, d\xi. \tag{6.10}$$

If $k \geq \ell$, this is bounded in absolute value by

$$\begin{aligned} &\int\limits_{|\xi|\geq(a/2)k^{1/2}} |\chi(k^{-1/2}\xi)|^k \, d\xi \cdot \mathcal{S}_k(v) \\ &\leq \delta^{k-\ell} \int |\chi(k^{-1/2}\xi)|^\ell \, d\xi \cdot \mathcal{S}_k(v) \\ &\leq C\delta^{k-\ell}k^{1/2}\mathcal{S}_k(v), \end{aligned} \tag{6.11}$$

as desired.

We can apply Proposition 6.1 to $v = v_y$, where

$$\begin{aligned} v_y(x) = 1 &\quad \text{for} \ \ x \leq y, \\ 0 &\quad \text{otherwise}. \end{aligned} \tag{6.12}$$

Then $\tilde{v}_y$ is a PV type distribution with $1/\xi$ type blowup at $\xi = 0$, and $|\tilde{v}(\xi)| \leq C/|\xi|$ on $\mathbb{R} \setminus 0$. Thus we have $\mathcal{A}(v_y) \leq A < \infty$, uniformly in y, and also

$$k^{1/2}\mathcal{S}_k(v_y) \leq S < \infty, \quad \text{uniformly in} \ \ y. \tag{6.13}$$

We deduce that, when ν_{f_1} satisfies the hypotheses of Proposition 6.1, then

$$\sup_y |\Phi_k(y) - G(y)| \leq Ck^{-r/2}, \tag{6.14}$$

and this works whenever (6.5) holds and $r \in (0, 2]$.

For example, when ν_{f_1} is given by (6.2), then (6.14) holds with $r = 2$.

7. Tail estimates

As seen in Proposition 1.2, we can sharpen the result $\nu_{g_k} \to \gamma^\sigma$, weak* in $\mathcal{M}(\widehat{\mathbb{R}})$, to

$$(1+x^2)\nu_{g_k} \longrightarrow (1+x^2)\gamma^\sigma, \quad \text{weak}^* \text{ in } \mathcal{M}(\widehat{\mathbb{R}}), \tag{7.1}$$

under the hypotheses of Theorem 1.1, especially $\int x^2\, d\nu_{f_1}(x) = \sigma < \infty$. Then general results discussed in Appendix A yield

$$\Phi_{2,k}(y) \longrightarrow G_2(y), \quad \text{as} \;\; k \to \infty, \; \forall\, y \in \mathbb{R}, \tag{7.2}$$

where, complementing (0.41), we set

$$\begin{aligned} \Phi_{2,k}(y) &= \int_{-\infty}^{y} x^2\, d\nu_{g_k}(x), \\ G_2(y) &= \int_{-\infty}^{y} x^2 \gamma^\sigma(x)\, dx. \end{aligned} \tag{7.3}$$

Such results constitute *tail estimates.* Here we seek further tail estimates when we have higher moments that are finite, i.e.,

$$\int |x|^p\, d\nu_{f_1}(x) < \infty, \quad p > 2. \tag{7.4}$$

We concentrate on the cases $p = 2\ell,\ \ell \in \mathbb{N},\ \ell > 1$. In such a case, taking

$$\chi(\xi) = \int_{\mathbb{R}} e^{-ix\xi}\, d\nu_{f_1}(\xi), \tag{7.5}$$

we have that, if (7.4) holds with $p = 2\ell$, then $\chi \in C^{2\ell}(\mathbb{R})$ and

$$\chi^{(2\ell)}(0) = (-1)^\ell \int_{\mathbb{R}} x^{2\ell}\, d\nu_{f_1}(x). \tag{7.6}$$

Conversely, if $\chi \in C^{(2\ell)}(\mathbb{R})$, then (7.4) holds, with $p = 2\ell$, and we have (7.6).

Now, to obtain tail estimates, we start with the following observation.

Proposition 7.1. *Assume f_j are IID random variables satisfying (0.1), and define g_k as in (0.16). Fix $\ell \in \mathbb{N}$, $\ell > 1$. If*

$$\int x^{2\ell}\, d\nu_{f_1}(x) < \infty, \tag{7.7}$$

then there exists $A < \infty$, independent of k, such that

$$\int x^{2\ell}\, d\nu_{g_k}(x) \le A, \quad \forall\, k. \tag{7.8}$$

Proof. As in (1.6), there exists $a > 0$ such that, for $|\xi| \le a$,

$$\chi(\xi) = e^{\Psi(\xi)}, \quad \Psi(0) = \Psi'(0) = 0. \tag{7.9}$$

If (7.7) holds, then $\chi \in C^{2\ell}(\mathbb{R})$, hence

$$\Psi \in C^{2\ell}((-a,a)). \tag{7.10}$$

Now, as in (1.7), for $|\xi| \le ak^{1/2}$,

$$\chi_{g_k}(\xi) = e^{\Psi_k(\xi)}, \quad \Psi_k(\xi) = k\Psi(k^{-1/2}\xi). \tag{7.11}$$

We have

$$\Psi_k^{(j)}(\xi) = k^{1-j/2}\Psi^{(j)}(k^{-1/2}\xi), \tag{7.12}$$

for $j \le 2\ell$, hence

$$\Psi_k^{(j)}(0) = k^{1-j/2}\Psi^{(j)}(0), \quad 0 \le j \le 2\ell. \tag{7.13}$$

Note that the exponent in $k^{1-j/2}$ is > 0 if and only if $j = 0$ or 1, and in these cases the right side of (7.13) vanishes. It readily follows that there exists $A < \infty$ such that

$$|\chi_{g_k}^{(2\ell)}(0)| \le A, \quad \forall\, k, \tag{7.14}$$

and this gives (7.8).

We can now extend Proposition 1.2.

Proposition 7.2. *Under the hypotheses of Proposition 7.1,*

$$(1+x^{2\ell})\nu_{g_k} \longrightarrow (1+x^{2\ell})\gamma^\sigma, \quad \textit{weak}^* \ \textit{in } \mathcal{M}(\widehat{\mathbb{R}}), \tag{7.15}$$

as $k \to \infty$.

Proof. We know from Theorem 1.1 that

$$\langle (1+x^{2\ell})\nu_{g_k}, v\rangle \longrightarrow \langle (1+x^{2\ell})\gamma^\sigma, v\rangle, \tag{7.16}$$

as $k \to \infty$, for all continuous v on $\mathbb{R}$ with compact support, hence, thanks to (7.8), for all $v \in C(\widehat{\mathbb{R}})$ satisfying $v(\infty) = 0$. To get (7.15), it remains to obtain (7.16) for $v \equiv 1$, hence to obtain

$$\int_{\mathbb{R}} x^{2\ell}\, d\nu_{g_k}(x) \longrightarrow \int_{\mathbb{R}} x^{2\ell}\gamma^\sigma(x)\, dx, \quad \text{as } \ k \to \infty. \tag{7.17}$$

This is equivalent to

$$\chi_{g_k}^{(2\ell)}(0) \longrightarrow \Bigl(\frac{d}{d\xi}\Bigr)^{2\ell}\gamma^\sigma(0), \quad \text{as } \ k \to \infty. \tag{7.18}$$

In turn, (7.18) follows from (7.9)–(7.13), supplemented by the identity

$$\Psi''(0) = -\sigma, \tag{7.19}$$

which follows from (0.1).

Results of Appendix A then yield the following.

Corollary 7.3. *In the setting of Proposition 7.2, if* $v : \widehat{\mathbb{R}} \to \mathbb{R}$ *is bounded, Borel, and Riemann integrable on* $\widehat{\mathbb{R}}$*, then*

$$\int_{\mathbb{R}} v(x)(1+x^{2\ell})\, d\nu_{g_k}(x) \longrightarrow \int_{\mathbb{R}} v(x)(1+x^{2\ell})\gamma^\sigma(x)\, dx, \tag{7.20}$$

as $k \to 0$.

Our next tail estimates will make use of results of §3. Recall from Proposition 3.2 that, if f_j are IID random variables satisfying (0.1), and if (7.9) holds, with

$$\begin{aligned} &\Psi(\xi) = -\frac{\sigma}{2}\xi^2 + \xi^2\beta(\xi), \\ &|\beta(\xi)| \le b|\xi|^r, \quad |\beta(\xi)| \le \frac{\sigma}{4}, \end{aligned} \tag{7.21}$$

for $|\xi| \le a$, and for some $r \in (0,2]$, then

$$|\langle \nu_{g_k} - \gamma^\sigma, v\rangle| \le Ck^{-r/2}\mathcal{A}(v) + \|\psi(k^{-1/2}D)v\|_{L^\infty}, \tag{7.22}$$

with

$$\mathcal{A}(v) = \int_{-\infty}^{\infty} e^{-\sigma\xi^2/8}|\xi|^{2+r}|\hat{v}(\xi)|\,d\xi, \tag{7.23}$$

and ψ as in (0.37). Hence

$$|\langle \nu_{g_k}, v\rangle| \le |\langle \gamma^\sigma, v\rangle| + Ck^{-r/2}\mathcal{A}(v) + \|\psi(k^{-1/2}D)v\|_{L^\infty}. \tag{7.24}$$

To state our next result, we bring in the following spaces of functions, for $\rho \in \mathbb{R}$:

$$S^\rho(\mathbb{R}) = \{v \in C^\infty(\mathbb{R}) : |v^{(\ell)}(x)| \le C_\ell(1+|x|)^{\rho-\ell},\ \forall\, \ell \in \mathbb{Z}^+\}. \tag{7.25}$$

Then (cf. Proposition 2.4 in [T1], Chapter 7, but note the roles of x and ξ are switched), we have

$$|\hat{v}(\xi)| \le \begin{array}{ll} C|\xi|^{-\rho-1}, & \text{for } |\xi| \le 1 \quad (\text{provided } \rho > -1), \\ C_\nu|\xi|^{-\nu}, & \text{for } |\xi| \ge 1. \end{array} \tag{7.26}$$

We see that

$$v \in S^\rho(\mathbb{R}),\ \rho < r+2 \Longrightarrow \mathcal{A}(v) < \infty, \tag{7.27}$$

and

$$\begin{aligned} v \in S^\rho(\mathbb{R}),\ \rho \in \mathbb{R} \Longrightarrow |\langle \gamma^\sigma, v\rangle| < \infty,\ &\text{and} \\ \|\psi(k^{-1/2}D)v\|_{L^\infty} &\le C'_\nu k^{-\nu/2}. \end{aligned} \tag{7.28}$$

Note that, for each $\rho \in \mathbb{R}$,

$$(1+x^2)^{\rho/2} \in S^\rho(\mathbb{R}). \tag{7.29}$$

We now have the following.

Proposition 7.4. *Assume f_j are IID random variables, satisfying (0.1), (7.9), and (7.21), for $|\xi| \le a$, and some $r \in (0,2]$. Then*

$$\rho < r+2 \Rightarrow (1+x^2)^{\rho/2}\nu_{g_k} \to (1+x^2)^{\rho/2}\gamma^\sigma, \quad \textit{weak}^* \textit{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{7.30}$$

Furthermore, for such ρ,

$$v \in S^\rho(\mathbb{R}) \Longrightarrow |\langle \nu_{g_k} - \gamma^\sigma, v\rangle| \le Ck^{-r/2}. \tag{7.31}$$

Remark. When Proposition 7.2 applies, the result (7.15) is stronger than its counterpart in (7.30), whose hypotheses hold with $r = 2$ if $\int x^3\, d\nu_{f_1} = 0$, and with $r = 1$ otherwise. On the other hand, (7.31) provides useful additional information.

8. CLT associated with a fractional diffusion

For $0 < \alpha \le 2$, the semigroups

$$P_\alpha^t = e^{-t(-\partial_x^2)^{\alpha/2}}, \quad t \ge 0, \tag{8.1}$$

consist of positivity-preserving operators with the property that

$$\int_{\mathbb{R}} P_\alpha^t u(x)\, dx = \int_{\mathbb{R}} u(x)\, dx, \tag{8.2}$$

for $u \in L^1(\mathbb{R})$. They are convolution operators,

$$P_\alpha^t u(x) = \gamma_\alpha^t * u(x), \tag{8.3}$$

where each γ_α^t is a probability measure on $\mathbb{R}$, whose characteristic function is

$$\chi_{t,\alpha}(\xi) = \int e^{-ix\xi} \gamma_\alpha^t(x)\, dx = e^{-t|\xi|^\alpha}. \tag{8.4}$$

If $\alpha < 2$, the measures γ_α^t do not have finite second moments, and if $\alpha \le 1$ they do not have finite first moments.

For $\alpha = 2$, the operators $P_2^t = e^{t\partial_x^2}$ form the diffusion semigroup. For $\alpha < 2$, these are fractional diffusions. They give rise to stochastic processes belonging to the family of Levy processes. For material on this, see [T3], which also treats the higher dimensional case.

Here we formulate and prove a version of CLT associated with such fractional diffusion semigroups.

To begin, suppose $f_j : \Omega \to \mathbb{R}$ are IID random variables on a probability space $(\Omega, \mathcal{F}, \mu)$, inducing the probability measure ν on $\mathbb{R}$, as in (1.6), with characteristic function

$$\chi(\xi) = \int_{\Omega} e^{-i\xi f_j}\, d\mu = \int_{\mathbb{R}} e^{-ix\xi}\, d\nu(\xi). \tag{8.5}$$

Extending the setting of Theorem 1.1, involving (1.5), we will fix $t > 0$, $\alpha \in (0, 2)$, and make the hypothesis that

$$\chi(\xi) = 1 - t|\xi|^\alpha + r(\xi), \quad r(\xi) = o(|\xi|^\alpha), \quad \text{as} \quad \xi \to 0, \tag{8.6}$$

or, equivalently, there exists $a > 0$ such that, for $|\xi| \le a$,

$$\chi(\xi) = e^{-t|\xi|^\alpha + |\xi|^\alpha \beta(\xi)}, \quad \beta(\xi) \to 0 \quad \text{as} \quad \xi \to 0. \tag{8.7}$$

An example of (8.6) (with $t=1$) is

$$\chi(\xi) = (1+|\xi|^\alpha)^{-1} = \int_0^\infty e^{-s(1+|\xi|^\alpha)}\, ds, \tag{8.8}$$

the second identity implying that χ is the characteristic function of a probability measure on $\mathbb{R}$.

To proceed, we see that the characteristic function of $f_1+\cdots+f_k$ is

$$\begin{aligned}\int\limits_\Omega e^{-i\xi(f_1+\cdots+f_k)}\, d\mu &= \chi(\xi)^k \\ &= e^{-tk|\xi|^\alpha + k|\xi|^\alpha \beta(\xi)}, \ \text{ for } \ |\xi| \le a.\end{aligned} \tag{8.9}$$

This formula tells us how to normalize the sum $f_1+\cdots+f_k$. In place of (0.16), we set

$$g_k = k^{-1/\alpha}(f_1+\cdots+f_k), \tag{8.10}$$

yielding

$$\begin{aligned}\chi_{g_k}(\xi) &= \int\limits_\Omega e^{-i\xi k^{-1/\alpha}(f_1+\cdots+f_k)}\, d\mu \\ &= \chi(k^{-1/\alpha}\xi)^k \\ &= e^{-t|\xi|^\alpha + |\xi|^\alpha \beta(k^{-1/\alpha}\xi)},\end{aligned} \tag{8.11}$$

the last identity holding for

$$|\xi| \le a k^{1/\alpha}. \tag{8.12}$$

Having this, we can formulate the following variant of Theorem 1.1.

Theorem 8.1. *Assume $\{f_j : j \in \mathbb{N}\}$ is an IID sequence on $(\Omega, \mathcal{F}, \mu)$ whose characteristic function $\chi(\xi)$ satisfies (8.6), for some $t>0$, $\alpha \in (0,2)$. Define g_k by (8.10). Then*

$$\nu_{g_k} \longrightarrow \gamma_\alpha^t, \quad \textit{weak}^* \ \textit{in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{8.13}$$

Proof. We see from (8.11)–(8.12) that

$$\lim_{k\to\infty} \hat{\nu}_{g_k}(\xi) = \hat{\gamma}_\alpha^t(\xi), \quad \forall\, \xi \in \mathbb{R}. \tag{8.14}$$

Arguing as in (1.10) yields

$$\int v\, d\nu_{g_k} \longrightarrow \int v \gamma_\alpha^t\, dx, \tag{8.15}$$

for all $v \in \mathcal{S}(\mathbb{R})$. Since ν_{g_k} and γ_α^t are probability measures, this gives (8.15) for all $v \in C_*(\mathbb{R})$, and also for $v \equiv 1$, hence for all $v \in C(\widehat{\mathbb{R}})$, giving the asserted result (8.13).

Here is an illustration of Theorem 8.1, with $\alpha = 1$. Define $\chi \in C(\mathbb{R})$ by

$$\begin{aligned} \chi(\xi) &= 1 - \frac{2}{\pi}|\xi|, \quad |\xi| \le \pi, \\ &= \chi(\xi + 2\pi), \quad \forall\, \xi \in \mathbb{R}. \end{aligned} \tag{8.16}$$

Then

$$\chi(\xi) = \frac{4}{\pi^2} \sum_{k \in \mathbb{Z},\ \text{odd}} \frac{1}{k^2} e^{ik\xi}, \tag{8.17}$$

so χ is the characteristic function of a random variable f satisfying

$$\nu_f = \frac{4}{\pi^2} \sum_{k \in \mathbb{Z},\ \text{odd}} \frac{1}{k^2} \delta_k. \tag{8.18}$$

It follows from Theorem 8.1 that if f_j are IID random variables on $(\Omega, \mathcal{F}, \mu)$ for which ν_{f_j} satisfy (8.18), and we form

$$g_k = \frac{1}{k}(f_1 + \cdots + f_k), \tag{8.19}$$

then

$$\nu_{g_k} \longrightarrow \gamma_1^{2/\pi}, \quad \text{weak}^* \text{ in } \mathcal{M}(\widehat{\mathbb{R}}). \tag{8.20}$$

Note that

$$\gamma_1^t(x) = \frac{1}{\pi} \frac{t}{x^2 + t^2}. \tag{8.21}$$

A. Natural extension of weak* convergence of measures

Let X be a compact metric space, μ a finite positive Borel measure on X. If $f : X \to \mathbb{R}$ is a bounded function, we say $f \in \mathcal{R}(X, \mu)$ provided that, for each $\varepsilon > 0$, there exist

$$u, v \in C(X) \text{ such that } u \le f \le v, \quad \text{and} \quad \int\limits_X (v - u)\, d\mu < \varepsilon. \tag{A.1}$$

If $X = S^1$, the unit circle, and μ is Lebesgue measure, this class coincides with the standard notion of Riemann integrable functions. See [T2] for some basic results on this class of functions. The following is a useful result.

Proposition A.1. *Take X, μ as above, and let ν_k be finite, positive Borel measures on X. Assume*

$$\nu_k \longrightarrow \mu, \ \ \textit{weak}^* \ \textit{in} \ \ \mathcal{M}(X) = C(X)'. \tag{A.2}$$

Then, if $f : X \to \mathbb{R}$ is a bounded, Borel function,

$$f \in \mathcal{R}(X, \mu) \Longrightarrow \int f\, d\nu_k \to \int f\, d\mu. \tag{A.3}$$

Proof. Given $f \in \mathcal{R}(X, \mu)$, take $\varepsilon > 0$ and pick u, v such that (A.1) holds. Then

$$\int f\, d\nu_k \le \int v\, d\nu_k \to \int v\, d\mu < \int f\, d\mu + \varepsilon, \tag{A.4}$$

so

$$\limsup_{k\to\infty} \int f\, d\nu_k \le \int f\, d\mu. \tag{A.5}$$

Similarly

$$\liminf_{k\to\infty} \int f\, d\nu_k \ge \int f\, d\mu, \tag{A.6}$$

so we have (A.3).

Example. Let $X = \widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, and let ν_k and μ be probability measures on $\mathbb{R}$, naturally extended to $\widehat{\mathbb{R}}$, so that $\mu(\{\infty\}) = 0$. Let

$$f : \mathbb{R} \longrightarrow \mathbb{R} \ \text{ be a bounded, continuous function.} \tag{A.7}$$

Then f extends to a bounded function on $\widehat{\mathbb{R}}$, with only ∞ as a point of discontinuity. Hence $f \in \mathcal{R}(\widehat{\mathbb{R}}, \mu)$, and (A.3) applies, so if (A.2) holds,

$$\int f\, d\nu_k \longrightarrow \int f\, d\mu, \tag{A.8}$$

for all f satisfying (A.7). The fact that (A.2) and (A.7) imply (A.8) is part of the Levy-Cramér continuity theorem. See [V], p. 25.

B. Weak* convergence of measures and uniform convergence of distribution functions

let ν_k and μ be probability measures on $\mathbb{R}$. The conditions

$$\begin{aligned}
&\nu_k \to \mu \ \text{ in } \ \mathcal{D}'(\mathbb{R}), \\
&\nu_k \to \mu \ \text{ in } \ \mathcal{S}'(\mathbb{R}), \qquad \text{(B.1)} \\
&\nu_k \to \mu \text{ weak}^* \text{ in } \ \mathcal{M}(\widehat{\mathbb{R}}) = C(\widehat{\mathbb{R}})'
\end{aligned}$$

are all equivalent. They say

$$\int f\, d\nu_k \longrightarrow \int f\, d\mu, \tag{B.2}$$

for $f \in C_0^\infty(\mathbb{R}),\ f \in \mathcal{S}(\mathbb{R})$, and $f \in C(\widehat{\mathbb{R}})$, respectively. Let us now assume

$$\mu \text{ has no atoms.} \tag{B.3}$$

Then, by Proposition A.1, (B.2) holds for $f = \chi_{(-\infty,x]}$, for each $x \in \mathbb{R}$. In other words, if we set

$$\Phi_k(x) = \nu_k((-\infty, x]), \quad G(x) = \mu((-\infty, x]), \tag{B.4}$$

we have

$$\Phi_k(x) \longrightarrow G(x), \quad \forall\, x \in \mathbb{R}. \tag{B.5}$$

We note the following useful (and well known) refinement.

Proposition B.1. *If ν_k and μ are probability measures on $\mathbb{R}$ satisfying (B.1) and (B.3), then*

$$\Phi_k \longrightarrow G, \quad \textit{uniformly on } \ \mathbb{R}. \tag{B.6}$$

Proof. If not, there exist $\varepsilon > 0,\ k_n \to \infty$, and $x_{k_n} \in \mathbb{R}$ such that

$$|\Phi_{k_n}(x_{k_n}) - G(x_{k_n})| \ge \varepsilon. \tag{B.7}$$

If $G(y_0) = \varepsilon/4$ and $G(y_1) - 1 - \varepsilon/4$, then only finitely many x_{n_k} can lie outside $[y_0, y_1]$. Hence there is a subsequence (which we merely denote j) of (k_n) such that

$$x_j \to y \in [y_0, y_1], \quad |\Phi_j(x_j) - G(x_j)| \ge \varepsilon. \tag{B.8}$$

Then there is either a further subsequence satisfying $x_j \nearrow y$ or one satisfying $x_j \searrow y$. Let's deal with the first possibility; a similar argument will handle the second.

To start, pick N so large that

$$|\Phi_j(y) - G(y)| < \frac{\varepsilon}{4}, \quad \text{and} \quad |G(x_j) - G(y)| < \frac{\varepsilon}{4}, \quad \forall\, j \geq N. \tag{B.9}$$

It follows that

$$|\Phi_j(y) - G(x_j)| < \frac{\varepsilon}{2}, \quad \forall\, j \geq N, \tag{B.10}$$

hence, if (B.8) holds,

$$|\Phi_j(x_j) - \Phi_j(y)| > \frac{\varepsilon}{2}, \quad \forall\, j \geq N, \tag{B.11}$$

hence $\nu_j([x_j, y]) > \varepsilon/2$ for $j \geq N$, and a fortiori

$$\nu_j([x_N, y]) \geq \frac{\varepsilon}{2}, \quad \forall\, j \geq N. \tag{B.12}$$

Now we take $j \to \infty$ to conclude that

$$\mu([x_N, y]) \geq \frac{\varepsilon}{2}, \tag{B.13}$$

i.e., $G(y) - G(x_N) \geq \varepsilon/2$, contradicting (B.9). This finishes the proof.

Remark. Coming full circle, we can apply d/dx to (B.6) and obtain (B.1).

References

[F1] W. Feller, On the Berry-Esseen Theorem, Z. Wahrsch. 10 (1968), 261–268.

[F2] W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, J. Wiley, New York, 1972.

[F] C. Fox, Review of "Lectures on the Fourier transform and its applications," by B. Osgood, SIAM Review 62 (2020), 731–735.

[GS] G. Grimmett and D. Stirzaker, Probability and Random Processes, Oxford Univ. Press, 2001.

[O] B. Osgood, Lectures on the Fourier Transform and its Applications, Undergraduate Texts #33, AMS, Providence RI, 2019.

[T] M. Taylor, Measure Theory and Integration, GSM #76, AMS, Providence RI, 2006.

[T1] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer, New York, 1996 (2nd ed., 2011).

[T2] M. Taylor, Riemann integrable functions on a compact, measured metric space, Notes, available at http://mtaylor.web.unc.edu/notes, item #11, "Functional Analysis."

[T3] M. Taylor, Levy Processes, Notes, available at http://mtaylor.web.unc.edu /notes, item #6, "Diffusion processes and other random processes."

[V] S. Varadhan, Probability Theory, Courant Lecture Notes #7, AMS, Providence RI, 2001.