This is the first of (what is projected to be) a short series of technical reports aimed at articulating certain connections between the broad subject headings of Gaussians, zeros, and linear combinations of $L$-functions within the setting of analytic number theory.

In later installments, we shall find ourselves particularly interested in developing some techniques by means of which it becomes feasible to say something fairly precise (at least near $\text{Re}(s) = \frac{1}{2}$) about both the asymptotic value distribution and the distribution of zeros manifested by run-of-the-mill linear combinations $F(s)$ of the aforementioned sort. A large portion of our results will have underpinnings that go back to work of Atle Selberg from the 1940s and mid-'70s. See [Sel] for an updated perspective on part of this; also [BH].

To set the stage for these things, it is convenient to begin with several more basic reports devoted to discussing some background material that will for the most part be seen to either be standard - or else a natural variant of something that is. Apart from two or three items of the latter type, any novelty in these earlier installments will be confined solely to methodological matters.

§1. Some Interpolation Formulae and Related Heuristics

1.1. The traditional reference for entire functions of Beurling-Selberg type is [Beu] coupled with [Sel2, pp.213–218, 226]. In very loose terms, the central issue here function-theoretically is as follows: Consider the real line $\mathbb{R}$. Let $\chi_E(x)$ denote the indicator function of set $E$ and

$$\text{sgn}(x) = \begin{cases} 0, & x = 0 \\ x/|x|, & x \neq 0 \end{cases}. \quad (1.1)$$

Keeping $\ell \in \mathbb{Z}^+$, let $m(x)$ be either $\text{sgn}(x)$ or $\chi_{[0,\ell]}(x)$. How does one construct entire functions $f(z)$ of exponential type at most $2\pi$ such that, along $\mathbb{R}$, one

---

$^1$Part of the work in this paper was supported by a Bell Companies Fellowship at the Institute for Advanced Study in Princeton, N.J.
has

(i) \( f(x) \in \mathbb{R} \) and \( f(x) \geq m(x) \);

and

(ii) \( \lim_{|x| \to \infty} (f(x) - m(x)) = 0 \)?

In (ii), the faster the vanishing, the better. Can one rig things, for instance, so that one also has

(iii) \( \int_{-\infty}^{\infty} |f(x) - m(x)|^q \, dx < \infty \) for every \( q \in [1, \infty) \)?

An \( L_1 \)-norm of (something close to) minimal size would be especially desirable.

In §1, our goal will simply be to informally identify some natural candidates for \( f(z) \). Substantiation of any proposed \( f \) will be left for §2.

1.2. To get started, we need a few basics. We’ll assume the standard \( L_2 \) theory of Fourier transforms on \( \mathbb{R} \) as being known. The prototypical formulae

\[
\hat{f}(p) = \int_{-\infty}^{\infty} f(x) e^{-2\pi ipx} \, dx
\]

\[
f(x) = \int_{-\infty}^{\infty} \hat{f}(p) e^{2\pi ipx} \, dp
\]

\[
\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{-\infty}^{\infty} |\hat{f}(p)|^2 \, dp
\]

\[
\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle
\]

\[
(f')^\wedge = 2\pi ip \hat{f}(p)
\]

\[
(f * g)^\wedge = \hat{f} \hat{g}
\]

are thus available as need be (in, for instance, Schwartz space \( S \)). Over and beyond this, one readily checks that

\[
f(x) = g(\lambda x) \Rightarrow \hat{f}(p) = \frac{1}{|\lambda|} \hat{g}(p/\lambda)
\]

for \( \lambda \neq 0 \) and that

\[
f(x) = e^{-\pi ax^2} \Rightarrow \hat{f}(p) = \sqrt{\frac{1}{a}} e^{-\pi p^2/a}
\]

\[
f(x) = \max\{0, b - |x|\} \Rightarrow \hat{f}(p) = \frac{\sin^2(\pi pb)}{\pi^2 p^2}
\]

\[
f(x) = \chi_{[-c,c]}(x) \Rightarrow \hat{f}(p) = \frac{\sin(2\pi pc)}{\pi p}
\]

\[
f(x) = \text{sgn}(x) \chi_{[-c,c]}(x) \Rightarrow \hat{f}(p) = -2\pi i p \frac{\sin^2(\pi pc)}{\pi^2 p^2}.
\]

\(^2\text{Note that the analogous constructions for } f \leq m \text{ and non-integral } \ell \text{ are also treated there.}\)
For test functions \( \varphi \) in \( C(\mathbb{R}) \cap L^1(\mathbb{R}) \) whose “periodization”

\[
\Phi(x) = \sum_{k=-\infty}^{\infty} \varphi(x + k)
\]

converges \textit{uniformly} on \([0, 1]\), one knows that the Poisson summation formula

\[
\sum_{k=-\infty}^{\infty} \varphi(k) = \sum_{m=-\infty}^{\infty} \hat{\varphi}(m)
\]

holds anytime the right-hand series is convergent. Cf. [Z1, pp.68, 89(3.4)]; also [Gau, pp.88–89]. In particular, matters are readily seen to hold with absolute convergence throughout anytime \( \varphi \) is \( C^2 \) and \( \varphi^{(j)} \in L^1(\mathbb{R}) \) for \( 0 \leq j \leq 2 \).

1.3. Another result that’s very familiar, albeit on a much deeper level, is the Paley-Wiener theorem, which characterizes \( L^2(\mathbb{R}) \) functions obtained as restrictions of entire functions of exponential type. In this regard, cf. [Boa, p.103 (theorem 6.8.1)], [PaW, p.13 (note the typos in 6.10)], and [Z2, p.272 (theorem 7.2)]. When the type is at most \( 2\pi \alpha \), the representation

\[
f(z) = \int_{-\alpha}^{\alpha} g(v)e^{2\pi i vz} dv
\]

holds with a unique \( g \in L^2[-\alpha, \alpha] \) which, in turn, satisfies

\[
\int_{-\alpha}^{\alpha} |g(v)|^2 dv = \int_{-\infty}^{\infty} |f(x)|^2 dx .
\]

A simple application of Leibnitz’s rule then gives

\[
f^{(m)}(z) = \int_{-\alpha}^{\alpha} (2\pi iv)^m g(v)e^{2\pi i vz} dv
\]

for every \( m \geq 1 \) (the \( L^2 \) norm of \( f^{(m)} \) necessarily being finite in each case).

By applying Parseval’s equation to the \textit{Fourier series} of \( (2\pi iv)^m g(v) \) on \([-\alpha, \alpha] \), one immediately sees that:

\[
\int_{-\alpha}^{\alpha} |g(v)|^2 dv = \frac{1}{2\alpha} \sum_{k=-\infty}^{\infty} \left| f\left( \frac{k}{2\alpha} \right) \right|^2 = \int_{-\infty}^{\infty} |f(x)|^2 dx ; \quad (1.5)
\]

\[
\frac{1}{2\alpha} \sum_{k=-\infty}^{\infty} \left| f^{(m)}\left( \frac{k}{2\alpha} \right) \right|^2 = \int_{-\infty}^{\infty} \left| f^{(m)}(x) \right|^2 dx \leq (2\pi \alpha)^2 m \int_{-\infty}^{\infty} |f(x)|^2 dx . \quad (1.6)
\]

\[3\]At the other extreme, see [Kat, pp.130(problem 15), 12(2.7), 125(1.10)] for a “mass redistribution style” construction of a continuous probability density \( g(t) \) on \( \mathbb{R} \) whose Fourier transform \( \varphi \) belongs to \( C(\mathbb{R}) \cap L^1(\mathbb{R}) \) but, for which, \( \varphi(n) = \delta_{n,0} \) and \( g(n) = 0 \) hold at every \( n \in \mathbb{Z} \). Equation (1.2) clearly \textit{fails} for the pair \( (\varphi, g) \). By either [Z1, p.12(6.2)] or [Kat, p.13], the periodizations of \( \varphi \) and \( g \) must sum to 0 and 1, respectively, almost everywhere. Compare [KL, p.306] and (overly-optimistic) problem 12 in [Fel2, p.648].
Relations (1.5) and (1.6) continue to hold when $k$ is replaced throughout by $k + \Theta \ (0 \leq \Theta < 1)$. At the same time, it pays to keep in mind that, by (1.4),

$$|f^{(m)}(x)| \leq \sqrt{2\alpha} \frac{(2\pi\alpha)^m}{\sqrt{1 + 2m}} \|f\|_2$$

(1.7) holds for every $x \in \mathbb{R}$ and $m \geq 0$.

1.4. To develop an interpolation formula for $f$, one simply returns to (1.3) and substitutes the standard $L_2$ Fourier expansion

$$\sum_{n=-\infty}^{\infty} c_n e^{inx/\alpha}$$

of $g(v)$. Since $c_n = \frac{1}{2\alpha} f\left(\frac{-n}{2\alpha}\right)$, a quick calculation yields

$$f(z) = \frac{\sin(2\pi\alpha z)}{2\pi\alpha} \sum_{k=-\infty}^{\infty} (-1)^k \frac{f(k/2\alpha)}{z - k/2\alpha}$$

(1.8)

just as in [Boa, p.220 (11.5.8)]. Cf. also here [Z2, p.275 (theorem 7.19)] and [Lev, p.150 (theorem 1)]. Expansion (1.8) is often referred to as the “cardinal series” or sampling theorem.

By passing to $(f(z) - f(0))/z$, one can now accommodate milder growth restrictions on $|f(x)|$. References [Boa] and [Z2] (loc. cit.) provide an immediate elaboration on this point; one finds with very little effort that

$$f(z) = f'(0) \frac{\sin(2\pi\alpha z)}{2\pi\alpha} + \sum_{k=0}^{\infty} (-1)^k \frac{f(k/2\alpha)}{z - k/2\alpha}$$

(1.9)

anytime $|f(x)|/(1 + |x|) \in L_2(\mathbb{R})$. The classical relation

$$\frac{\pi}{\sin \pi w} = \frac{1}{w} + \sum_{k \neq 0} (-1)^k \left(\frac{1}{w - k} + \frac{1}{k}\right)$$

(1.10)

turns out to be exactly what is needed to handle the terms arising from $f(0)$ in $(f(z) - f(0))/z$. (Notice that (1.10) is just (1.9) with $f \equiv 1$.)

Formula (1.9) has antecedents going back many years. In chronological order, see [Po1, §5], [Val, especially p.204 (9)(11) with $\mu = 1$], [PSz, problem III.165], [Har, theorem 10], and [Hig1, §1.4 (minus the typo in (10))] for an interesting historical perspective on this. Note that (1.8) follows from (1.9) simply by applying the latter to the product function $zf(z)$.

*which again has exponential type at most $2\pi\alpha$*
1.5. Looked at geometrically, the set of nodes in formula (1.8) can be said to be $\frac{1}{\alpha}Z$ in an obvious sense. If additional information in the form of values of $f'$ happens to be available at the nodes, an elementary reshuffling in (1.3) can be used to “cut the nodal set down” to just $\frac{1}{\alpha}Z$.

To see this, we follow Vaaler [Vaa, p.195 (3.5)(3.6)]. (Compare [Sel2, p.214 (lines 18–19)].) Clearly:

$$f(z) = \int_0^\alpha g(v)e^{2\pi ivz}dv + e^{-2\pi iz} \int_0^\alpha g(t - \alpha)e^{2\pi itz}dt$$

$$f\left(\frac{k}{\alpha}\right) = \int_0^\alpha [g(v) + g(v - \alpha)]e^{2\pi ikv/\alpha}dv;$$

$$\frac{1}{2\pi i}f'(z) = \int_0^\alpha vg(v)e^{2\pi ivz}dv + e^{-2\pi iz} \int_0^\alpha (t - \alpha)g(t - \alpha)e^{2\pi itz}dt$$

$$\frac{1}{2\pi i}f'(\frac{k}{\alpha}) = \int_0^\alpha [vg(v) + (v - \alpha)g(v - \alpha)]e^{2\pi ikv/\alpha}dv.$$

Accordingly, as $L_2$ Fourier series on $[0,\alpha]$,

$$g(v) + g(v - \alpha) \sim \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right)e^{-2\pi ikv/\alpha}$$

$$vg(v) + (v - \alpha)g(v - \alpha) \sim \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha})e^{-2\pi ikv/\alpha}.$$

Denoting the left-hand expressions by $G_1$ and $G_2$, respectively, we can now write

$$g(v) = \left(1 - \frac{v}{\alpha}\right)G_1 + \frac{1}{\alpha}G_2$$

$$g(v - \alpha) = \frac{v}{\alpha}G_1 - \frac{1}{\alpha}G_2$$

and, in this way, obtain

$$g(v) \sim \left(1 - \frac{v}{\alpha}\right)\sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right)e^{-2\pi ikv/\alpha} + \frac{1}{\alpha} \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha})e^{-2\pi ikv/\alpha}$$

$$g(v - \alpha) \sim \frac{v}{\alpha} \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right)e^{-2\pi ikv/\alpha} - \frac{1}{\alpha} \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha})e^{-2\pi ikv/\alpha}$$

in a natural $L_2$ sense over $[0,\alpha]$. By substituting back, we immediately get

$$f(z) = \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right) \int_0^\alpha \left(1 - \frac{v}{\alpha}\right)e^{-2\pi ikv/\alpha}e^{2\pi ivz}dv$$

$$+ e^{-2\pi iz} \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right) \int_0^\alpha \frac{t}{\alpha}e^{-2\pi ikv/\alpha}e^{2\pi itz}dt$$

$$+ \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha}) \int_0^\alpha \frac{1}{\alpha}e^{-2\pi ikv/\alpha}e^{2\pi ivz}dv$$

$$+ e^{-2\pi iz} \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha}) \int_0^\alpha \frac{(-1)}{\alpha}e^{-2\pi ikv/\alpha}e^{2\pi itz}dt.$$

Denoting the left-hand expressions by $G_1$ and $G_2$, respectively, we can now write

$$g(v) = \left(1 - \frac{v}{\alpha}\right)G_1 + \frac{1}{\alpha}G_2$$

$$g(v - \alpha) = \frac{v}{\alpha}G_1 - \frac{1}{\alpha}G_2$$

and, in this way, obtain

$$g(v) \sim \left(1 - \frac{v}{\alpha}\right)\sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right)e^{-2\pi ikv/\alpha} + \frac{1}{\alpha} \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha})e^{-2\pi ikv/\alpha}$$

$$g(v - \alpha) \sim \frac{v}{\alpha} \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right)e^{-2\pi ikv/\alpha} - \frac{1}{\alpha} \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha})e^{-2\pi ikv/\alpha}$$

in a natural $L_2$ sense over $[0,\alpha]$. By substituting back, we immediately get

$$f(z) = \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right) \int_0^\alpha \left(1 - \frac{v}{\alpha}\right)e^{-2\pi ikv/\alpha}e^{2\pi ivz}dv$$

$$+ e^{-2\pi iz} \sum_k \frac{1}{\alpha}f\left(\frac{k}{\alpha}\right) \int_0^\alpha \frac{t}{\alpha}e^{-2\pi ikv/\alpha}e^{2\pi itz}dt$$

$$+ \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha}) \int_0^\alpha \frac{1}{\alpha}e^{-2\pi ikv/\alpha}e^{2\pi ivz}dv$$

$$+ e^{-2\pi iz} \sum_k \frac{1}{2\pi i \alpha}f'(\frac{k}{\alpha}) \int_0^\alpha \frac{(-1)}{\alpha}e^{-2\pi ikv/\alpha}e^{2\pi itz}dt.$$
Writing \( t = \alpha + u \) in the second and fourth sums produces:

\[
\begin{align*}
\int \frac{f(k/\alpha)}{\alpha} \int_0^\alpha (1 - \frac{v}{\alpha}) e^{-2\pi ivk/\alpha} e^{2\pi ivu} dv \\
+ \int \frac{f(k/\alpha)}{\alpha} \int_0^0 (1 + \frac{u}{\alpha}) e^{-2\pi ivk/\alpha} e^{2\pi ivu} dv \\
+ \int \frac{1}{2\pi i \alpha^2} f'(k/\alpha) \int_0^\alpha e^{-2\pi ivk/\alpha} e^{2\pi ivu} dv \\
+ \int \frac{1}{2\pi i \alpha^2} f'(k/\alpha) \int_0^0 e^{-2\pi ivk/\alpha} e^{2\pi ivu} dv \\
= \sum f(k/\alpha) \int_\alpha^\alpha (\frac{\alpha - |v|}{\alpha^2}) e^{2\pi iv(z - k/\alpha)} dv \\
+ \sum f'(k/\alpha) \int_\alpha^\alpha \frac{\text{sgn}(v)}{2\pi i \alpha^2} e^{2\pi iv(z - k/\alpha)} dv \\
= \sum f(k/\alpha) \left( \sin(\pi \alpha p) \right)^2 \\
+ \sum f'(k/\alpha) (-p) \left( \sin(\pi \alpha p) \right)^2 \\
\{ \text{with } p \equiv k/\alpha - z \} \\
= \sum f(k/\alpha) \left( \sin(\pi \alpha z) \right)^2 (z - k/\alpha)^{-2} \\
+ \sum f'(k/\alpha) (z - k/\alpha) \left( \sin(\pi \alpha z) \right)^2 (z - k/\alpha)^{-2}.
\end{align*}
\]

In other words, the alternate representation

\[
f(z) = \left( \frac{\sin(\pi \alpha z)}{\pi \alpha} \right)^2 \left\{ \sum_{k=-\infty}^\infty \frac{f(k/\alpha)}{(z - k/\alpha)^2} + \sum_{k=-\infty}^\infty \frac{f'(k/\alpha)}{z - k/\alpha} \right\} \tag{1.11}
\]

is available to us anytime (1.3) holds. Compare: (1.8), (1.5), (1.6). One readily checks that the right-hand expression behaves like

\[
f\left( \frac{k}{\alpha} \right) + f'\left( \frac{k}{\alpha} \right) \left( z - \frac{k}{\alpha} \right) + O(1) \left( z - \frac{k}{\alpha} \right)^2 \tag{1.11'}
\]

near each nodal point \( k/\alpha \). (See [Hig2, pp. 97–100] for an interesting higher-order generalization; also p.58 in the second volume of this work.)

1.6. Just as with (1.8), once (1.11) is known, one can pass to \( (f(z) - f(0))/z \) to widen the class of admissible \( f \). (See [Vaa].) For present purposes, however, this augmentation is unnecessary and a simple limit trick suffices to give us the heuristic formats that we seek (à la §1.1) in both cases of \( m \).
One reasons as follows. In view of (1.11) and (1.11') [with \( \alpha = 1 \)] and the fact that \( \ell \in \mathbb{Z}^+ \), the interpolatory format

\[
\mathcal{F}_\ell(z) = \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=0}^{\ell} \frac{1}{(z-k)^2} + \frac{A}{z} + \frac{B}{\ell - z} \right\} \quad (A > 0, B > 0) \tag{1.12}
\]

clearly occupies a distinguished position vis-à-vis \( \chi_{[0,\ell]}(x) \) — at least for suitably controlled \( A \) and \( B \).

For the \( L_1 \) norm of \( \mathcal{F}_\ell(x) - \chi_{[0,\ell]}(x) \) to be finite, it is necessary that \( B - A \) reduce to zero. In considering (1.12) as a format, we therefore tacitly assume this to be so.

For bounded \( A \), simple use of the identity

\[
1 = \left( \frac{\sin \pi x}{\pi} \right)^2 \sum_{n=-\infty}^{\infty} \frac{1}{(x-n)^2} \tag{1.13}
\]

shows that \( \| \mathcal{F}_\ell - \chi_{[0,\ell]} \|_p = O(\ell^{1/p}) \) for each \( p > 1 \). (The contribution from \( |x| \geq 3\ell \) is trivially seen to be \( o(1) \).) Consideration of the elementary calculus estimate

\[
\sum_{n=1}^{\infty} \frac{1}{(n+\omega)^2} < \frac{1}{\omega} \quad (\omega > 0)
\]

quickly demonstrates that the earlier exponent \( 1/p \) is replaceable by 0.

In light of this, it now makes eminently good sense to try to approximate \( \text{sgn}(x) \) by letting \( \ell \to \infty \) in the combination \( 2(\mathcal{F}_\ell(x) - \frac{1}{2}) \). Insofar as the choice of \( A \) can be kept bounded, the term with \( B \) automatically drops out on compact subsets of \( \mathbb{R} \) (in this connection, see also (1.7)). The expression

\[
\mathcal{F}^*(z) = 2 \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=0}^{\infty} \frac{1}{(z-k)^2} + \frac{A}{z} \right\} - 1
\]

thus takes on a special significance in regard to \( \text{sgn}(x) \). Notice incidentally that \( \mathcal{F}^*(0) = 1 \) (not 0); also that \( \mathcal{F}^*(x) = O(1) \).

By virtue of (1.13), \( \mathcal{F}^*(z) \) can be written in three different ways:

\[
\mathcal{F}^* = \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=0}^{\infty} \frac{1}{(z-k)^2} + \sum_{k<0} \frac{1}{(z-k)^2} + \frac{2A}{z} \right\}; \tag{1.14_a}
\]

\[
\mathcal{F}^* = -1 + \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ 2 \sum_{k=0}^{\infty} \frac{1}{(z-k)^2} + \frac{2A}{z} \right\}; \tag{1.14_b}
\]

\[
\mathcal{F}^* = 1 + \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ -2 \sum_{k=0}^{\infty} \frac{1}{(z-k)^2} + \frac{2A}{z} \right\}. \tag{1.14_c}
\]

One wants the \( L_1 \) norm of \( \mathcal{F}^*(x) - \text{sgn}(x) \) to be finite. Since the earlier calculus estimate can be strengthened to read

\[
\frac{1}{\omega} - \frac{1}{\omega^2} < \sum_{n=1}^{\infty} \frac{1}{(n+\omega)^2} < \frac{1}{\omega}, \tag{1.15}
\]
equations (1.14b) and (1.14c) immediately show that there is precisely one ad-
missible value of \( A \); viz., \( A = 1 \).

We’ll subsequently consider (1.14) only under this restriction. The obvious hope here, of course, is that, when \( A = 1 \), the norms \( \| \mathcal{F}_\ell - \chi_{[0,\ell]} \|_p \) will remain uniformly bounded for all \( p \geq 1 \) as \( \ell \to \infty \).

With this, our informal determination of plausible \( f(z) \)-formats is complete.

**Remark 1.7.** Though, in this section, the role played by equation (1.9) has mainly been motivational, in other approximation settings (less “constrained” as to \( f' \)), matters change and (1.9), rather than (1.11), becomes the “ansatz-of-choice.” See [Vaa, p.187 ff] – or [Ess, p.16 ff] – for a good example.

§2. Entire Functions of Beurling-Selberg Type

2.1. After the introductory discussion given in §1, it is relatively pedestrian to go back and place matters on a rigorous footing. For this purpose, we first recall (1.11′) and then write

\[
B(z) = \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=0}^{\infty} \frac{1}{(z-k)^2} - \sum_{n=1}^{\infty} \frac{1}{(z+n)^2} + \frac{2}{z} \right\}; \quad (2.1)
\]

\[
b(z) = \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=1}^{\infty} \frac{1}{(z-k)^2} - \sum_{n=1}^{\infty} \frac{1}{(z+n)^2} + \frac{2}{z} \right\}; \quad (2.2)
\]

\[
W(z) = \left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=1}^{\infty} \frac{1}{(z-k)^2} - \sum_{n=1}^{\infty} \frac{1}{(z+n)^2} + \frac{2}{z} \right\}; \quad (2.3)
\]

\[
K(z) = \left( \frac{\sin \pi z}{\pi z} \right)^2 = \int_{-1}^{1} \frac{1}{1-|v|} e^{2\pi i vz} dv. \quad (2.4)
\]

One immediately checks that \( B(z) \) corresponds to \( \mathcal{F}^* (z) \) (à la (1.14)), \( b(z) = -B(-z) \), \( W = B - K = b + K \), and that the function \( 2H(x) \) appearing in [Beu, p.371 (bot)] is simply \( b(x/2\pi) \). According to Selberg, the functions \( B \) and \( H \) were first considered by Beurling around 1940 or so; cf. [Sel2, pp.226 (lines 10–17, 218 (20.17)].

Motivated by equations (20.19), (20.2), and (20.12) in [Sel2, pp.214–218], we also define

\[
S_\ell(z) = \frac{1}{2} \left[ B(z) + B(\ell - z) \right] \quad (2.5)
\]

\[
\sigma_\ell(z) = \frac{1}{2} \left[ b(z) + b(\ell - z) \right] \quad (2.6)
\]

for any positive real \( \ell \). When \( \ell \) is integral, one immediately checks that \( S_\ell(z) \) reduces to \( \mathcal{F}_\ell(z) \) (cf. (1.12)) with \( B = A = 1 \). In a similar way, \( \sigma_\ell(z) \) is seen to reduce to

\[
\left( \frac{\sin \pi z}{\pi} \right)^2 \left\{ \sum_{k=1}^{\ell-1} \frac{1}{(z-k)^2} + \frac{1}{z} + \frac{1}{\ell - z} \right\}.
\]
These reductions and the discussion in the first part of §1.6 make it plausible at least that one will have
\[ \sigma_\ell(x) \leq \chi_{[0,\ell)}(x) \leq S_\ell(x) \]
at every \( x \in \mathbb{R} \). (Compare [Sel2, p.217 (lines 13, 3, 10)].) For later use, notice too that \( 0 \leq K(x) \leq 1 \).

**Theorem 2.2.** The functions \( B, b, W \) are entire and satisfy the following basic properties for real \( x \):

(a) \( B(x) = W(x) + K(x), b(x) = W(x) - K(x) \);
(b) \( W(x) + W(-x) = 0, B(x) + B(-x) = 2K(x) \);
(c) \( b(x) \leq \text{sgn}(x) \leq B(x) \) (strictly if \( x \notin \mathbb{Z} \));
(d) \( W(x) \in [1 - K(x), 1] \) if \( x > 0 \);
(e) \( W(x) \in [-1, -1 + K(x)] \) if \( x < 0 \);
(f) \( |b(x) - \text{sgn}(x)| \leq 2K(x), |B(x) - \text{sgn}(x)| \leq 2K(x) \);
(g) \( \int_{-\infty}^{\infty} (\text{sgn}(x) - b(x)) \, dx = \int_{-\infty}^{\infty} (B(x) - \text{sgn}(x)) \, dx = 1 \).

For \( z \in \mathbb{C} \), the magnitudes of the numbers \( |B(z)|, |b(z)|, \) and \( |W(z)| \) are at most \( O(1) \exp(2\pi |\text{Im}(z)|) \).

**Proof.** Assertions (a) and (b) are trivial; (c) is nearly so upon utilizing relations (1.14) and (1.15). To verify (d), one needs to exploit a refinement of (1.15); viz.,
\[ \frac{1}{\omega} - \frac{1}{2\omega^2} < \sum_{k=1}^{\infty} \frac{1}{(k + \omega)^2} < \frac{1}{\omega} . \]
The left-hand portion of this is immediately recognized as a standard consequence of the Euler-Maclaurin summation formula with \( f(t) = (t + \omega)^{-2} \). Cf. [Ste, pp.124(bot), 132(9)(10), 133(lines 6–20)] with \( m = 2\tau = 2 \). (Alternatively: see [Jor, pp.8(bot), 254(2), 255(6)(7), 261(4)] and [Malm, pp.58(8), 69(47)].)

Assertions (e) and (f) follow from (d) by means of (b) and (a), respectively.

To address (g), one simply notes that:
\[
\int_{-\infty}^{0} (B(x) + 1) \, dx = \int_{0}^{\infty} (B(-y) + 1) \, dy \\
= \int_{0}^{\infty} (2K(y) - B(y) + 1) \, dy \\
= \int_{-\infty}^{\infty} K(y) \, dy - \int_{0}^{\infty} (B(y) - 1) \, dy;
\]
\[
\int_{\mathbb{R}} (\text{sgn}(x) - b(x)) \, dx = \int_{\mathbb{R}} (\text{sgn}(-y) - b(-y)) \, dy \\
= \int_{\mathbb{R}} (B(y) - \text{sgn}(y)) \, dy .
\]
It remains to control things for $z \in \mathbb{C}$. For this, it is enough to look at $|W(z)|$ and, then, only over $\{\operatorname{Re}(z) \geq 0\} \cap \{|z| \geq \frac{1}{2}\}$. By virtue of (1.13),

$$1 - W(z) = 2 \left( \frac{\sin \frac{\pi z}{\pi}}{2 \pi} \right)^2 \left[ \frac{1}{2z^2} + \sum_{n=1}^{\infty} \frac{1}{(z+n)^2} - \frac{1}{z} \right]. \quad (2.7)$$

To estimate the bracketed term, one can either use Euler-MacLaurin with $f(t) = (t + z)^{-2}$ and a suitably big $m = 2\tau$ or else Stirling’s formula coupled with the fact ([Erd, p.15 (2)]) that

$$\frac{d^2}{dz^2} \log \Gamma(z) = -\sum_{n=0}^{\infty} \frac{1}{(z+n)^2}.$$

The bracket is quickly seen to be $O(|z|^{-3})$ and we are done. \[\blacksquare\]

The foregoing ideas are readily combined with a little calculation to yield

$$1 - W(z) = \left\{ \begin{array}{ll}
\left( \frac{\sin \frac{\pi x}{\pi}}{2 \pi} \right)^2 \left[ \sum_{k=1}^{N} \frac{2B_{2k}}{2^{2k+1}} + R_N(x) \right], & x > 0 \\
K(x) \left[ 1 - 2x + 2 \sum_{n=2}^{\infty} (-1)^n \zeta(n)(n-1)x^n \right], & 0 \leq x < 1
\end{array} \right\}, \quad (2.8)$$

wherein $B_\nu$ are the standard Bernoulli numbers ([Jor, Ste]), $N \geq 0$, and the remainder term $R_N(x)$ satisfies

$$R_N(x) = 2\theta B_{2N+2}x^{-2N-3} \quad \text{with} \quad 0 < \theta < 1 \, ^5$$

An analogous formula holds for $1 - W(z)$ on $\mathbb{C} \cap \{|\operatorname{Arg}(z)| \leq \pi - \delta\}$ modulo a minor revision in the part concerning $\theta$. Cf. (2.7) and [Erd, p.47 (1)(7)]. Restricting things to $\{|z| < 1\}$ leads to the familiar expansion of $\csc^2(\pi z)$ near $z = 0$ (cf. [Erd, p.51 (3)]) and to the curious Taylor series development

$$\frac{W(z)}{K(z)} = 2z + \sum_{m=1}^{\infty} 4m \zeta(2m+1)z^{2m+1}. \quad (2.9)$$

**Theorem 2.3.** Given any $\ell > 0$. The functions $\sigma_\ell$ and $S_\ell$ are entire, have magnitude $O(1)$ exp$(2\pi |\operatorname{Im}(z)|)$, and satisfy the following basic properties for real $x$:

(a) $\sigma_\ell(x) \leq \chi[0,\ell](x) \leq S_\ell(x)$;
(b) $\max \{ \chi[0,\ell](x) - \sigma_\ell(x), S_\ell(x) - \chi[0,\ell](x) \} \leq K(x) + K(\ell - x)$;
(c) $\int_{-\infty}^{\infty} (\chi[0,\ell](x) - \sigma_\ell(x)) \, dx = \int_{-\infty}^{\infty} (S_\ell(x) - \chi[0,\ell](x)) \, dx = 1$.

^5When combined with the elementary formula for $W(u) - W(u+m) \, (m \geq 1)$, this estimate leads to a very efficient way of computing $W(x)$ at any $x \in \mathbb{R}$. Compare [Ess, p.18 (24)], where the same idea is readily seen to work after breaking things up into partial fractions.
Proof. It suffices to verify (a) for $x \neq 0, \ell$. Under this restriction,
\[
\chi_{[0,\ell]}(x) = \frac{1}{2} \{	ext{sgn}(x) + \text{sgn}(\ell - x)\}
\]
holds and the desired estimate follows immediately from Theorem 2.2(c). Assertion (b) is proved similarly utilizing Theorem 2.2(f). Equation (c) follows directly from (2.10) and Theorem 2.2(g). ■

Theorem 2.4. The entire functions $\{B, b, W, S_\ell, \sigma_\ell\}$ have type exactly equal to $2\pi$.

Proof. Simply put $z = \beta + iy$ (y large) in the asymptotic development for $W(z)$ cited after (2.8) and keep $\beta$ bounded; likewise for $K(z)$. Any negative powers of $\ell - z$ are best re-expressed as power series in $z^{-1}$. (When $\ell \in \frac{1}{2} + \mathbb{Z}$, a slight anomaly occurs in the $\{\sigma_\ell, S_\ell\}$ asymptotics.) Application of the difference operator $T[f] \equiv f(z + 1) - f(z)$ immediately furnishes an alternate proof in the case of $\{B, b, W\}$. With a bit more effort, the same low-level approach also works for $\sigma_\ell$ and $S_\ell$. ■

2.5. To streamline the development of the next two theorems, it is helpful to preface matters with a general result about entire functions of exponential type.

Lemma 2.6. Given any positive $p$ and $\tau$. Let $f(z)$ be an entire function of exponential type $\leq \tau$. Suppose that $f(x) \in L_p(\mathbb{R})$. Then:

(a) $f(x) \in L_q(\mathbb{R})$ for $q \in [p, \infty)$;
(b) $\int_{-\infty}^{\infty} |f(x + iy)|^p dx \leq e^{p\tau|y|} \int_{-\infty}^{\infty} |f(x)|^p dx$ for $y \neq 0$;
(c) $f(z) = o(1)$ on every horizontal strip $\{|\text{Im}(z)| \leq \Delta\}$;
(d) $\sum_{n=1}^{\infty} |f(z + n)|^p$ converges uniformly on $\{0 \leq \text{Re}(z) \leq 1, |\text{Im}(z)| \leq \Delta\}$;
(e) $f^{(k)}(x) \in L_p(\mathbb{R})$ for every $k \geq 1$.

Proof. Assertion (b) is a classical result of Plancherel and Pólya; see [PIP, pp.120–124] and [Lev, pp.50, 51, 38 (theorem 2)]. (The proof hinges on subharmonicity and some related Phragmén-Lindelöf type estimates.) Straightforward use of the subharmonicity of $|f|^p$ then gives
\[
|f(z)| = O_\Delta(1) \left\{ \int_{-\infty}^{\infty} |f(t)|^p dt \right\}^{1/p}
\]
on every $\{|\text{Im}(z)| \leq \Delta\}$. Cf. [Lev, p.51 (bot)]. Assertion (c) follows by a minor adaptation of the same manipulation; cf. [Lev, p.138 (top)]. Assertion (a) is then obvious. By using a second (more astute!) adaptation, one obtains (d). Cf. [Lev, p.138 (bot)], [PIP, p.126 (top)], and [Boa, p.101 (lines 13–23)].

To establish (e), we follow [PIP, p.127]. It suffices to treat $k = 1$. For $x_0 \in \mathbb{R},$
the function $g(z) \equiv z^{-1}[f(x_0 + z) - f(x_0)]$ is entire. Accordingly:

$$
|g(0)|^p \leq \frac{1}{2\pi} \int_0^{2\pi} |g(\delta e^{i\phi})|^p d\phi
$$

$$
|f'(x_0)|^p \leq \frac{1}{2\pi \delta^p} \int_0^{2\pi} |f(x_0 + \delta e^{i\phi}) - f(x_0)|^p d\phi
\leq \{ |u + v|^p \leq 2^p \max(|u|^p, |v|^p) \}$$

$$
|f'(x_0)|^p \leq \frac{2^p}{2\pi \delta^p} \int_0^{2\pi} \left[ |f(x_0 + \delta e^{i\phi})|^p + |f(x_0)|^p \right] d\phi
$$

To conclude, one simply integrates over $x_0$ and applies (b). Upon taking $\delta = \tau^{-1}$, we find that

$$
\int_{-\infty}^{\infty} |f'(x)|^p dx \leq 2(2\pi \tau)^p \int_{-\infty}^{\infty} |f(x)|^p dx.
$$

Compare [Boa, p.211 (bot)]. ■

**Theorem 2.7.** Let $\mathcal{M}^+$ be the set of all entire functions of exponential type $\leq 2\pi$ which majorize $\text{sgn}(x)$ along $\mathbb{R}$. Let $\mathcal{M}^-$ be the counterpart for $f(x) \leq \text{sgn}(x)$. We then have

$$
\inf_{f \in \mathcal{M}^+} \int_{\mathbb{R}} (f(x) - \text{sgn}(x)) dx = 1 = \inf_{h \in \mathcal{M}^-} \int_{\mathbb{R}} ( \text{sgn}(x) - h(x)) dx.
$$

The extremal functions are respectively $B(z)$ and $b(z)$; they are unique.

**Proof.** Thanks to the transformation $h(z) = -f(-z)$, it is enough to look at $\mathcal{M}^+$. There is clearly no loss of generality if we also restrict ourselves to functions which satisfy $f(x) - \text{sgn}(x) \in L_1(\mathbb{R})$. The differences $f(z) - B(z)$ will then satisfy the hypotheses of Lemma 2.6 with $(p, \tau) = (1, 2\pi)$. By combining this information with the asymptotic of $W(z)$ mentioned after (2.8), we immediately see that $f(z) - B(z) = o(1)$ on every strip $\{|\text{Im}(z)| \leq \Delta\}$ and that

$$
\int_{-\infty}^{\infty} |f'(x)| dx \leq \int_{-\infty}^{\infty} |f' - B'| dx + \int_{-\infty}^{\infty} |K'| dx + \int_{-\infty}^{\infty} |W'| dx < \infty. \quad (2.11)
$$

Likewise for $f^{(k)}(x)$ with $k \geq 2$.

We now follow Beurling ([Beu]) and look at matters in the framework of a judiciously chosen integration by parts. To set the stage, we first write

$$
\theta(x) = f(x) - \text{sgn}(x), \quad \theta^*(x) = B(x) - \text{sgn}(x), \quad k(x) = x - \left\lfloor x \right\rfloor - \frac{1}{2}
$$

and recall that
Entire Functions of Beurling-Selberg Type

\[ k(x) = -\sum_{n=1}^{\infty} \frac{\sin(2\pi nx)}{\pi n} \quad \text{for} \quad x \notin \mathbb{Z}. \]

The decomposition \( \theta = \theta^* + (f - B) \) assures us that \( \theta(x) = o(1) \). (Cf. Theorem 2.2(f) à propos \( \theta^*(x) \).)

Fix any \( y \notin \mathbb{Z} \). For large \( N \) and small \( \varepsilon \), we then have:

\[
\begin{align*}
\int_{y+\varepsilon}^{N-\varepsilon} k(u)d\theta(y-u) + \int_{y+\varepsilon}^{N-\varepsilon} \theta(y-u)dk(u) &= \left[ k(u)\theta(y-u) \right]_{y+\varepsilon}^{N-\varepsilon} \\
\int_{y+\varepsilon}^{N-\varepsilon} k(u)f'(y-u)(-1)du + \int_{y+\varepsilon}^{N-\varepsilon} \theta(y-u)du &= \sum_{y+\varepsilon < n < N} \theta(y-n) + \left[ k(u)\theta(y-u) \right]_{y+\varepsilon}^{N-\varepsilon}.
\end{align*}
\]

Upon letting \( N \to \infty \),

\[
-\int_{y+\varepsilon}^{\infty} k(u)f'(y-u)du + \int_{y+\varepsilon}^{\infty} \theta(y-u)du = -k(y+\varepsilon)\theta(-\varepsilon) + \sum_{n>y} \theta(y-n).
\]

In a similar way,

\[
-\int_{-\infty}^{y-\varepsilon} k(u)f'(y-u)du + \int_{-\infty}^{y-\varepsilon} \theta(y-u)du = k(y-\varepsilon)\theta(\varepsilon) + \sum_{n<y} \theta(y-n).
\]

Accordingly:

\[
\int_{|u-y|>\varepsilon} \left( -k(u)f'(y-u) + \theta(y-u) \right)du = \left[ k(y-v)\theta(v) \right]_{-\varepsilon}^{\varepsilon} + \sum_{-\infty}^{\infty} \theta(y-n).
\]

By passing to the limit in \( \varepsilon \), we conclude that

\[
2k(y) = \sum_{-\infty}^{\infty} \theta(y-n) + \int_{\mathbb{R}} k(u)f'(y-u)du - \int_{\mathbb{R}} \theta(v)dv.
\]

(The format of the final integral clearly depends crucially on the fact that \( k'(u) = 1 \) almost everywhere.)

Observe now that the Paley-Wiener theorem applies to \( f'(z) \). Cf. (1.3) with \( \alpha = 1 \). Since \( f'(x) \in L_2 \cap L_1 \), its Fourier transform is continuous on all of \( \mathbb{R} \). As such, the “\( g \)” in (1.3) necessarily vanishes at the endpoints. By virtue of the Poisson summation formula (1.2), we therefore have

\[
\sum_{m=-\infty}^{\infty} f'(x + m) = \int_{-\infty}^{\infty} f'(y)dy + 0 = 2,
\]

\(^6\text{Cf. [Z2, pp.249 (para 4), 250 (2.17)]; also [Z1, p.26 (11.6)].}\)
from which it follows that

$$\int_{\mathbb{R}} k(u)f'(y-u)du = \int_{0}^{1} k(u)\left[ \sum_{m=-\infty}^{\infty} f'(y-u+m) \right] du = 0.$$ 

In other words:

$$2k(y) = \sum_{-\infty}^{\infty} \theta(y-n) - \int_{\mathbb{R}} \theta(v)dv, \; y \notin \mathbb{Z}. \quad (2.12)$$

Since $\theta(v) \geq 0$, making $y \to 0^+$ yields

$$-1 = \left[ \text{non-negative number} \right] - \int_{\mathbb{R}} \theta(v)dv;$$

i.e., $\int_{\mathbb{R}} \theta(v)dv \geq 1$. Equality holds only if

$$\lim_{y \to 0^+} \sum_{-N}^{N} \theta(y-n) = 0$$

for each large $N$. This necessitates that $f(n) = B(n)$ for each fixed $n$. Since $f(x) - \text{sgn}(x) = \theta(x) \geq 0$, we would also need to have $f'(n) = 0, n \neq 0$, and $f'(0) \geq 0$. Utilizing (1.11)\footnote{or, even better, the \{G_1, G_2\} formalism in §1.5 for $f(z) - B(z)$}, we finally deduce that

$$f(z) - B(z) = \left(\frac{\sin \pi z}{\pi}\right)^{2} \frac{c}{z}$$

with $c = f'(0) - B'(0) = f'(0) - 2$. Since $f - B \in L_1(\mathbb{R})$, the constant $c$ must be zero; hence $f(z) \equiv B(z)$. $\blacksquare$

Though Theorem 2.7 is due to Beurling, the following one is perhaps best attributed to Selberg. (Cf. [Sel2, p.226 (lines 5–10)].)

**Theorem 2.8.** Let $\ell$ be a positive integer and $\mathcal{M}_\ell$ be the counterpart of $\mathcal{M}_G$ for $\chi_{[0,\ell]}(x)$. We then have

$$\inf_{f \in \mathcal{M}_\ell} \int_{\mathbb{R}} (f(x) - \chi_{[0,\ell]}(x))dx = 1 = \inf_{h \in \mathcal{M}_\ell} \int_{\mathbb{R}} (\chi_{[0,\ell]}(x) - h(x))dx.$$ 

The functions $S_\ell(z)$ and $\sigma_\ell(z)$ are extremal, but they are not unique.

**Proof.** We begin with $\mathcal{M}_\ell^+$ and again restrict ourselves to $f$ which also satisfy $\|f - \chi_{[0,\ell]}\|_1 < \infty$. Just as in the proof of Theorem 2.7 (but with $f' \hookrightarrow f$),

$$\sum_{m=-\infty}^{\infty} f(x + m) = \hat{f}(0) + 0.$$
Putting \( x = 0 \) immediately gives \( \hat{f}(0) \geq \ell + 1 \). By Theorem 2.3(c), equality holds for \( f = S_\ell \). Any other extremal function \( F \) necessarily satisfies
\[
F(m) = \chi_{[0, \ell]}(m) \quad \text{and} \quad F'(k) = 0 \text{ if } k \neq 0, \ell.
\]
By (1.11) and \( F \in L_1 \), we then get
\[
F(z) = S_\ell(z) + \eta \left( \frac{\sin \pi z}{\pi} \right)^2 \frac{\ell}{z(\ell - z)},
\]
wherein \( \eta = F'(0) - S_\ell'(0) = F'(0) - 1 \). Nonuniqueness stems from the fact that the right-hand side of (2.13) continues to lie in \( \mathcal{M}_\ell^+ \) for any sufficiently small \( \eta \in \mathbb{R} - \{0\} \). Cf. [Sel2, p.217 (line 3)]. Observe too that, for \( \ell \in \mathbb{Z} \),
\[
\int_{-\infty}^{\infty} \left( \frac{\sin \pi x}{\pi} \right)^2 \frac{x}{x(\ell - x)} \, dx = \lim_{R \to \infty} \int_{-R}^{R} \left( \frac{\sin \pi x}{\pi} \right)^2 \left[ \frac{1}{x} + \frac{1}{\ell - x} \right] \, dx = 0.
\]
The analysis for \( \mathcal{M}_\ell^- \) runs similarly; see [Sel2, p.217 (line 10)] for the pertinent \( \eta \)-condition. ■

If \( \ell \not\in \mathbb{Z} \), the qualitative reasoning in [Sel2, pp.218 (bot), 219 (top)] immediately shows that the functions \( S_\ell \) and \( \sigma_\ell \) are no longer extremal: one already does better, in fact, with certain multiples \((1 - \varepsilon)S_\ell(z)\) and \((1 + u)\sigma_\ell(z)\). (Cf. the second clause of Theorem 2.2(c). Take \( u = -1 \) if \( 0 < \ell < 1 \).)

For further information on this, consult [DL, Log]\(^8\) and the two graphs given in [Sel2, p.219]. The fact that \( S_\ell \) and \( \sigma_\ell \) are suboptimal in regard to \( L_1 \) will turn out to be of relatively little consequence for subsequent purposes.

**Remark 2.9.** In constructive function theory, approximants (of various \( L_p \) types) for a given function \( m(x) \) along \( \mathbb{R} \) are often obtained simply by taking convolutions with a rescaled Fejér kernel \( \alpha K(\alpha x) \) or similar. See [Tim, Ach] and, e.g., [Kat, p.125 (theorem 1.10)]. Since \( \{B, b, W, S_\ell, \sigma_\ell\} \) are basically defined via (1.11), they are not really amenable to being re-interpreted in a natural way as convolutions over \( \mathbb{R} \) of the aforementioned special type. The formal \( L_2 \) expansions for \( g(v) \) and \( g(v - \alpha) \) derived in §1.5 offer valuable insight on this last point.

**2.10.** We close §2 by determining what the analog of representation (1.3) is for the function \( W(z) \). Since
\[
\text{sgn}(x) = \int_{-\infty}^{\infty} \frac{1}{\pi v} \sin(2\pi xv) \, dv,
\]
there is a natural suspicion that
\[
W(z) = \int_{-1}^{1} \frac{Q(v)}{v} \sin(2\pi vz) \, dv
\]
holds for some nice, even, continuous function \( Q \) satisfying \( Q(0) = 1/\pi \).

\(^8\)in the case of \( \mathcal{M}_\ell^+ \)
Following the hint provided by the proof of Theorem 2.7, we first pass to \( W'(z) \) and write

\[
W'(z) = \int_{-\infty}^{\infty} A(v) \cos(2\pi z v) dv
\]

with an even continuous \( A(v) \) having support \( \subseteq [-1, 1] \). The asymptotics of \( W(z) \) cited near (2.8) assure us that \( W'(z) = O((1 + |z|)^{-3}) \) on every strip \( \{|\text{Im}(z)| \leq \Delta\} \). By Fourier inversion,

\[
A(v) = \int_{-\infty}^{\infty} W'(x) \cos(2\pi x v) dx.
\]

Letting \( v = 0 \) gives \( A(0) = 2 \). Since

\[
A'(v) = -2\pi \int_{-\infty}^{\infty} x W'(x) \sin(2\pi x v) dx,
\]

the function \( A \) necessarily belongs to \( C^1(\mathbb{R}) \). In view of the restriction on \( \text{supp}(A) \),

\[
A(1) = A'(1) = 0.
\]

We now write \( W(x) = \int_0^x W'(t) dt \) and exploit Fubini’s theorem. This gives

\[
W(x) = \int_0^x \int_{-1}^{1} A(v) \cos(2\pi t v) dv dt
\]

\[
= \int_{-1}^{1} \int_0^x A(v) \cos(2\pi t v) dt dv
\]

\[
= \int_{-1}^{1} A(v) \frac{\sin(2\pi x v)}{2\pi v} dv,
\]

from which it is evident that

\[
Q(v) = \frac{1}{2\pi} A(v)
\]

works in (2.15). (That \( Q \) is unique in (2.15) follows immediately from Leibnitz’s rule and the Plancherel theorem.)

**Theorem 2.11.** The function \( W(z) \) is representable in the form (2.15) with

\[
Q(v) = \frac{1}{\pi} |v| + (1 - |v|) \text{ctn}(\pi v).
\]

Extending \( Q \) to be zero off \([−1, 1]\) produces an even \( C^1(\mathbb{R}) \) function (call it \( q \)) which satisfies

\[
q(0) = \frac{1}{\pi}, \quad q(1) = 0, \quad q'(1) = 0
\]

\[
q''(0\pm) = \frac{2}{3}\pi, \quad q''(1-) = \frac{3}{2}\pi, \quad q''(1+) = 0
\]

\[
q'(v) < 0 \quad \text{for} \quad 0 < v < 1
\]
Proof. Consider $W(z)$ on the disk $\{ |z| \leq R \}$. One has
\[ W(z) = \lim_{N \to \infty} \left( \sin \frac{\pi z}{N} \right) \left\{ \sum_{m=1}^{N} \frac{1}{(z-m)^2} - \sum_{m=1}^{N} \frac{1}{(z+m)^2} + \frac{2}{z} \right\} \]
as a uniform limit. But,
\[ K(p-m) = \int_{-1}^{1} (1-|v|) e^{2\pi iv} e^{-2\pi i sv} dv \]
and
\[ pK(p) = \int_{-1}^{1} \frac{i}{2\pi} sgn(v) e^{-2\pi iv} dv \]
via §1.2. Accordingly (note the “−z”):\[ W(z) = \lim_{N \to \infty} \int_{-1}^{1} \left( \sum_{m=1}^{N} (1-|v|) \operatorname{sgn}(m) e^{2\pi iv} + \frac{i}{\pi} \operatorname{sgn}(v) \right) e^{-2\pi iv} dv \]
\[ = \lim_{N \to \infty} i \int_{-1}^{1} \left[ \frac{2(1-|v|)}{2\sin(\pi v)} \Re(e^{\pi iv} - e^{2\pi i(N+\frac{1}{2})}) + \frac{1}{\pi} \operatorname{sgn}(v) \right] e^{-2\pi iv} dv \]
\[ = \int_{-1}^{1} \left[ (1-|v|) \operatorname{ctn}(\pi v) + \frac{1}{\pi} \operatorname{sgn}(v) \right] \sin(2\pi zv) dv \]
\[ - \lim_{N \to \infty} \int_{-1}^{1} \frac{1-|v|}{\sin(\pi v)} \cos(2\pi v(N+\frac{1}{2})) \sin(2\pi zv) dv . \]
Since $\frac{1-|v|}{\sin(\pi v)} \sin(2\pi zv)$ is a continuous function of $(v, z)$, the final $N$-limit is 0 (uniformly w.r.t. $z$) by the Riemann-Lebesgue lemma. Representation (2.15) follows at once.

To conclude the proof, one simply uses the relation $q = \frac{1}{2\pi} A$ and a bit of elementary calculus. Observe that $q(v) + q(1-v) = \frac{1}{2}$ holds on $[0,1]$. ■

§3. Esseen’s Lemma for Probability Distributions over $\mathbb{R}$

3.1. Our primary focus will now shift to probability distributions $F(x)$ and their characteristic functions
\[ \varphi(\xi) \equiv \int_{\mathbb{R}} e^{i\xi x} dF(x) , \]
as discussed, for instance, in [Fel2, chap. 15], [Z2, p.262], and [Bil, §26].\[ ^9 \]

Lemma 3.2. Let $F$ and $G$ be any two probability distributions on $\mathbb{R}$ having continuous densities $f(x)$ and $g(x)$, respectively. Assume that $0 \leq g(x) \leq m$ and that
\[ \int_{\mathbb{R}} |x|^a dF(x) < \infty , \quad \int_{\mathbb{R}} |x|^a dG(x) < \infty \]

\[ ^9 \text{We make the tacit assumption that all probability distributions on } \mathbb{R} \text{ are taken to be right continuous. Similarly for } \mathbb{R}^k . \]
for some positive $\alpha$. Take $\Omega > 0$ and put
\[
\varphi(\zeta) = \int_{\mathbb{R}} e^{i \xi x} dF(x), \quad \psi(\zeta) = \int_{\mathbb{R}} e^{i \xi x} dG(x).
\]
The a priori inequality
\[
|F(t) - G(t)| \leq c_1 \int_{-\Omega}^{\Omega} \left| \frac{\varphi(\zeta) - \psi(\zeta)}{\zeta} \right| d\zeta + c_2 \frac{m}{\Omega} \tag{3.1}
\]
will then hold on $\mathbb{R}$ for certain universal constants $c_1$ and $c_2$. (The $d\zeta$-integral is automatically convergent as an improper Riemann integral.)

Proof. By considering $F(x_0 + u)$ and $G(x_0 + u)$, it suffices to treat the case $t = 0$. To expedite matters, we refer to Theorem 2.11 and set
\[
T(v) = \begin{cases} 
0, & v = 0 \\
\frac{Q(v) - Q(0)}{v}, & 0 < |v| \leq 1 
\end{cases} \in C[-1, 1];
\]
\[
R_B(v) = \frac{1}{i} T(v) + (1 - |v|) \in C[-1, 1];
\]
\[
R_b(v) = \frac{1}{i} T(v) - (1 - |v|) \in C[-1, 1];
\]
\[
\lambda = \sup_{0 \leq \xi \leq 1} \left| -i Q(\xi) + \xi (1 - \xi) \right|.
\]
For $A > 0$ and $h \in C\{0 < |v| \leq A\}$, we also set
\[
\int_{-A}^{A} h(v)dv = \lim_{\varepsilon \to 0^+} \int_{\varepsilon \leq |v| \leq A} h(v)dv \tag{3.2}
\]
whenever the right-hand limit exists.

By combining Theorems 2.2 and 2.11, it is virtually self-evident that
\[
B(x) = \int_{-1}^{1} \left[ \frac{Q(v)}{iv} + (1 - |v|) \right] e^{2\pi i xv} dv
\]
\[
b(x) = \int_{-1}^{1} \left[ \frac{Q(v)}{iv} - (1 - |v|) \right] e^{2\pi i xv} dv
\]
\[
B(x) = \int_{-1}^{1} \left[ \frac{1}{\pi iv} + R_B(v) \right] e^{2\pi i xv} dv \tag{3.3a}
\]
\[
b(x) = \int_{-1}^{1} \left[ \frac{1}{\pi iv} + R_b(v) \right] e^{2\pi i xv} dv. \tag{3.3b}
\]
One knows that
\[
B(x) = \text{sgn}(x) + \theta(x), \quad 0 \leq \theta(x) \leq 2K(x);
\]
\[
b(x) = \text{sgn}(x) + \sigma(x), \quad -2K(x) \leq \sigma(x) \leq 0.
\]
Observe now that:

\[ 2(F(0) - G(0)) = 2(1 - G(0)) - 2(1 - F(0)) \]

\[ = \int_{-\infty}^{\infty} (g(x) - f(x))[1 + \text{sgn}(\Omega x)] \, dx \]

\[ = \int_{-\infty}^{\infty} (g(x) - f(x)) \text{sgn}(\Omega x) \, dx \]

\[ = \int_{-\infty}^{\infty} (g(x) - f(x))B(\Omega x) \, dx - \int_{-\infty}^{\infty} g(x)\theta(\Omega x) \, dx \]

\[ + \int_{-\infty}^{\infty} f(x)\theta(\Omega x) \, dx \]

\[ \geq \int_{-\infty}^{\infty} (g(x) - f(x))B(\Omega x) \, dx - m \int_{-\infty}^{\infty} \theta(\Omega x) \, dx \]

by Theorem 2.2(g). A trivial substitution for \( B(\Omega x) \) gives

\[ \int_{-\infty}^{\infty} (g(x) - f(x))B(\Omega x) \, dx = \int_{-\infty}^{\infty} [g(x) - f(x)] \left( \int_{-\Omega}^{\Omega} e^{2\pi i w \frac{w}{\Omega}} \, dw \right) \, dx \]

\[ + \int_{-\infty}^{\infty} [g(x) - f(x)] \left( \int_{-\Omega}^{\Omega} R_{B}(\frac{w}{\Omega}) e^{2\pi i w \frac{w}{\Omega}} \, dw \right) \, dx. \]

Notice, however, that one has

\[ \int_{\varepsilon \leq |w| \leq \Omega} e^{2\pi i w \frac{w}{\Omega}} \, dw = \int_{\varepsilon \leq |w| \leq \Omega} \frac{\sin(2\pi x w)}{\pi w} \, dw \]

\[ = 2 \int_{\varepsilon \leq |w| \leq \Omega} \frac{\sin(2\pi x w)}{w} \, dw \]

\[ = 2 \int_{\varepsilon \leq |w| \leq \Omega} \frac{\sin u}{u} \, du = O(1) \]

with an implied constant that is absolute (since \( \int_{0}^{\infty} u^{-1} \sin(u) \, du \) converges to \( \pi/2 \)). By the Lebesgue dominated convergence theorem and Fubini, it follows that

\[ \int_{-\infty}^{\infty} (g(x) - f(x))B(\Omega x) \, dx = \int_{-\Omega}^{\Omega} 1_{\Omega} \int_{-\infty}^{\infty} [g(x) - f(x)] e^{2\pi i w \frac{w}{\Omega}} \, dx \, dw \]

\[ + \int_{-\Omega}^{\Omega} 1_{\Omega} R_{B}(\frac{w}{\Omega}) \int_{-\infty}^{\infty} [g(x) - f(x)] e^{2\pi i w \frac{w}{\Omega}} \, dx \, dw \]

\[ = \int_{-\Omega}^{\Omega} r(w) [\psi(2\pi w) - \varphi(2\pi w)] \, dw \]
wherein
\[
    r(w) = \frac{1}{\pi iw} + \frac{1}{\Omega} R_B\left(\frac{w}{\Omega}\right)
\]
\[
    = \frac{1}{\Omega} \left( \frac{1}{\pi iv} + R_B(v) \right) \quad \text{with } v = w/\Omega
\]
\[
    = \frac{1}{\Omega} \left[ \frac{Q(v)}{iw} + (1 - |v|) \right] = \frac{1}{\Omega} \left[ \frac{-iQ(v) + v(1 - |v|)}{v} \right].
\]
Since \( Q \) is even, we clearly have
\[
    |r(w)| \leq \frac{\lambda}{|w|} \quad \text{for } 0 < |w| \leq \Omega.
\]

Putting everything together, we find that:
\[
    2(F(0) - G(0)) \geq \int_{-\infty}^{\infty} (g(x) - f(x)) B(\Omega x) dx - \frac{m}{\Omega}
\]
\[
    = \oint_{\Omega} r(w) \left[ \psi(2\pi w) - \varphi(2\pi w) \right] dw - \frac{m}{\Omega}
\]
\[
    \geq - \oint_{-\Omega} \left| r(w) \right| \left| \psi(2\pi w) - \varphi(2\pi w) \right| dw - \frac{m}{\Omega}
\]
\[
    \geq - \lambda \oint_{-\Omega} \left| \frac{\varphi(2\pi w) - \varphi(2\pi w)}{|w|} \right| dw - \frac{m}{\Omega}
\]
\[
    \geq - \lambda \oint_{-2\pi\Omega} \left| \frac{\varphi(w) - \varphi(w)}{|w|} \right| dw - \frac{m}{\Omega}
\]
where, in the last three lines, a natural convention is made if the \( dw \)-integral diverges. A parallel manipulation with \( \text{sgn}(\Omega x) = b(\Omega x) - \sigma(\Omega x) \) gives
\[
    2(F(0) - G(0)) \leq \lambda \oint_{-2\pi\Omega} \left| \frac{\varphi(w) - \varphi(w)}{|w|} \right| dw + \frac{m}{\Omega}
\]
Dividing by 2 and replacing \( \Omega \) by \( \Omega/2\pi \) finally yields
\[
    |F(0) - G(0)| \leq \frac{\lambda}{2} \oint_{-2\pi\Omega} \left| \frac{\varphi(w) - \varphi(w)}{|w|} \right| dw + \frac{\pi m}{\Omega}
\]
To finish up, we set \( \tilde{\alpha} = \min\{\alpha, 1\} \) and simply note that
\[
    \varphi(w) - 1 = \int_{\mathbb{R}} (e^{ixw} - 1) dF(x) = O(1)|w|^{\tilde{\alpha}} \int_{\mathbb{R}} |x|^{\tilde{\alpha}} dF(x),
\]
the implied constant being at most 2. Similarly for \( \psi(w) - 1 \). ■

Estimate (3.1) corresponds to [Ess, p.32] and [Fel2, p.538 (3.13)] and is often referred to (together with Theorem 3.3 below) as the Esseen smoothing lemma. The traditional proof of (3.1) makes use of an auxiliary convolution. The approach adopted here, based on the Beurling function (i.e., eq. (3.3)), is some-
what more direct. Since
\[
\sqrt{0 + \frac{1}{\pi^2}} \leq \lambda \leq \sqrt{\frac{1}{10} + \frac{1}{\pi^2}} < \frac{1}{2},
\]
taking \((c_1, c_2) = (\frac{1}{4}, \pi)\) will certainly be admissible.\(^{10}\)

**Theorem 3.3.** Inequality (3.1) remains true without the simplifying hypothesis that \(F\) have a continuous density function on \(\mathbb{R}\). It is also valid when the initial \(C^1\) function \(G\) is only known to be real and to satisfy
\[
G(-\infty) = 0, \quad G(\infty) = 1, \quad |G'(x)| \leq m, \quad \int_{-\infty}^{\infty} |x|^\alpha |G'(x)| dx < \infty. \quad (3.4)
\]

**Proof.** The extension to a non-monotonic \(G\) is self-evident upon reviewing the earlier manipulations. With that augmentation in hand, widening the class of admissible \(F\) is most easily achieved by way of approximation. To this end, one forms the (Riemann-Stieltjes) convolution \(N_\varepsilon * F\) with
\[
N_\varepsilon(x) = \frac{1}{\varepsilon \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2\varepsilon^2} du
\]
and then passes to the \(\varepsilon = 0\) limit in (3.1). Compare [Fel2, pp.507 (bottom), 146 (theorem 4)]. The pertinent density and characteristic functions are
\[
\frac{1}{\varepsilon \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-u)^2}{2\varepsilon^2}} dF(u) \quad \text{and} \quad e^{-\frac{\varepsilon^2\zeta^2}{2}} \varphi(\zeta),
\]
respectively. The \(dF\)-integral is manifestly continuous w.r.t. \(x\); one also knows that
\[
\mathbb{E}(|Q + X|^\alpha) \leq 2^\alpha \mathbb{E}(|Q|^\alpha) + 2^\alpha \mathbb{E}(|X|^\alpha)
\]
for general random variables \(Q\) and \(X\). That being said, to finish the proof, one simply notes that
\[
\lim_{\varepsilon \to 0^+} \left[ (N_\varepsilon * F)(t) - F(t) \right] = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}} \left[ F(t - v) - F(t) \right] dN_\varepsilon(v) = 0
\]
holds at every point of continuity of \(F\). Since such \(t\) are everywhere dense along \(\mathbb{R}\), inequality (3.1) follows with the same constants \(c_j\) as before. \(\blacksquare\)

Esseen’s lemma admits an important (and very well-known) corollary in connection with convergence of probability distributions.

**Corollary 3.4.** Given \(\{G, g, \alpha, \psi\}\) as in Lemma 3.2. Let \(\{F_n\}_{n=1}^\infty\) be any sequence of probability distributions on \(\mathbb{R}\) for which
\begin{enumerate}
  \item \(\int_{\mathbb{R}} |x|^\alpha dF_n(x) = O(1)\);
  \item \(\int_{\mathbb{R}} e^{i\zeta x} dF_n(x) \to \psi(\zeta)\) pointwise on \(\mathbb{R}\) as \(n \to \infty\).
\end{enumerate}

\(^{10}\) Readers having computer experience will readily check that \(\lambda = 0.3263598\) to 7 decimal places. Those familiar with [Ess, chap. II] may also wish to note that \(\lambda/2 \approx (1.025)/2\pi\).
We then have

\[ F_n(x) \to G(x) \]

uniformly on \(\mathbb{R}\). An analogous result holds when \(n\) is replaced by a continuous variable \(\xi\).

Proof. An immediate consequence of Theorem 3.3 and the Lebesgue dominated convergence theorem. Cf. (i) and the last three lines in the proof of Lemma 3.2. A trivial modification of this shows that the characteristic functions \(\{\varphi_n(\zeta)\}_{n=1}^{\infty}\) satisfy a uniform Hölder \(\tilde{\alpha}\)-condition on \(\mathbb{R}\). The convergence hypothesized in (ii) will thus be uniform on every \(\zeta\)-interval \([-\Omega, \Omega]\).

To treat \(F_\xi\), one can either exploit the Hölder \(\tilde{\alpha}\)-condition or simply reason by contradiction. ■

3.5. In terms of applications, the single most common one for Corollary 3.4 is undoubtedly its use in establishing the central limit theorem for sums of \(N\) identically distributed, independent random variables \(X_j\). See, for instance, [Fel2, p.515]. In Corollary 5.5 below, we look at a somewhat more involved case. (See [Fel2, pp.542–543] for a prototypical effective example.)

3.6. Hypothesis (i) in Corollary 3.4 is of course very mild. By a change in method, it can simply be expunged.

To appreciate this, it suffices to recall two standard facts from basic probability theory. First: that (ii) is tantamount to \(F_n\) approaching \(G\) weakly, i.e., in distribution. Second: that distributional convergence \(F_n \to P\) is automatically upgradable to uniform convergence on \(\mathbb{R}\) anytime the limiting probability distribution \(P\) is everywhere continuous. (Cf. [Fel2, p.285 (problem 5)] or [KeS, §4.11] as regards the latter assertion.) Insofar as the “target” distribution \(G\) is continuous, hypothesis (ii) will thus furnish both a necessary and sufficient condition for the uniform \(x\)-limit articulated in Corollary 3.4.

The downside to switching over to this much more rudimentary viewpoint [based ultimately on just the Helly selection principle and Fourier-Stieltjes inversion] is that the rate of convergence will typically not be very transparent.

§4. The Esseen Smoothing Lemma in Several Variables

4.1. It is only natural to wonder if there exists a reasonably simple counterpart of (3.1) and Theorem 3.3 in higher dimensions. We examine this question in the present section, placing a premium on retaining some measure of formal similarity. Because of certain algebraic difficulties, the matter is not as straightforward as one might initially expect.

To streamline our exposition, it is helpful to begin with several preliminaries and a bit of notation.

\[11\] Consideration of \(N_\epsilon \ast G\) provides a good illustration of what can go wrong if \(G\) has atoms.

\[12\] It is worth noting that the technique used in [Fel2, pp.257–260(top)] runs into similar issues. The second line in (4.7) should be deleted there.
4.2. When $k > 1$, the one-variable inequalities $f_j(x_j) \leq \chi_{A_j}(x_j)$ do not in general concatenate w.r.t. $j$ to produce the relation

$$\prod_{j=1}^{k} f_j(x_j) \leq \chi_A(x_1, \ldots, x_k)$$

with $A = A_1 \times \cdots \times A_k$ over $\mathbb{R}^k$. Fortunately, an elementary ring-theoretic lemma [whose form I owe to a 1990 conversation with A. Selberg] enables one to circumvent this difficulty with minimal fuss in a wide variety of technical settings.

Lemma 4.3. Given any integer $k \geq 2$. Define symbolic functions

$$f_j = \chi_j - \delta_j, \quad g_j = \chi_j + \varepsilon_j$$

for $1 \leq j \leq k$. We then have

$$(1-k) g_1 \cdots g_k + f_1 g_2 \cdots g_k + g_1 f_2 \cdots g_k + \ldots + g_1 \cdots g_{k-1} f_k \quad (4.1)$$

$$= \chi_1 \cdots \chi_k - S,$$

wherein $S$ is a nonempty sum of monomials $\omega_1 \cdots \omega_k$ satisfying the conditions

(i) $\omega_j \in \{\chi_j, \varepsilon_j, \delta_j\}$ for each index $j$; 
(ii) $\omega_\tau \neq \chi_\tau$ for at least one $\tau$.

When $k > 2$, certain of the products $\omega_1 \cdots \omega_k$ will appear in $S$ with a multiplicity larger than one; e.g., $\varepsilon_1 \cdots \varepsilon_k$.

Proof. Put

$$G_j(m) = \left\{ \begin{array} {cl} \chi_j, & m = 0 \\ \varepsilon_j, & m = 1 \end{array} \right\}$$

and then keep $m_j \in \{0, 1\}$. Since the LHS of (4.1) is just

$$g_1 \cdots g_k + (f_1 - g_1) g_2 \cdots g_k + \ldots + g_1 \cdots g_{k-1} (f_k - g_k),$$

one is free to re-express things as

$$\sum_{\{1, \ldots, k\}} G_1(m_1) \cdots G_k(m_k) - (\delta_1 + \varepsilon_1) \sum_{\{2, \ldots, k\}} G_2(m_2) \cdots G_k(m_k) - \cdots$$

$$- (\delta_k + \varepsilon_k) \sum_{\{1, \ldots, k-1\}} G_1(m_1) \cdots G_{k-1}(m_{k-1}),$$

$$\{a, \ldots, b\}$$

serving here as an obvious shorthand for $(m_a, \ldots, m_b)$. But:

$$\sum_{\{1, \ldots, k\}} G_1(m_1) \cdots G_k(m_k) \quad (4.2)$$

$$= \chi_1 \cdots \chi_k + \sum_{m_1 + \cdots + m_k > 0} G_1(m_1) \cdots G_k(m_k).$$

\[13\] that of §2 being fairly typical
The last sum is clearly “subsumed” algebraically by
\[
G_1(1) \sum_{\{2, \ldots, k\}} G_2(m_2) \cdots G_k(m_k) + \cdots + G_k(1) \sum_{\{1, \ldots, k-1\}} G_1(m_1) \cdots G_{k-1}(m_{k-1})
\]
\[
\equiv \varepsilon_1 \sum_{\{2, \ldots, k\}} G_2(m_2) \cdots G_k(m_k) + \cdots + G_k(1) \sum_{\{1, \ldots, k-1\}} G_2(m_1) \cdots G_{k-1}(m_{k-1}) .
\]
Inserting a “redacted” form of the above into (4.2) and then backtracking, it quickly becomes apparent that \( S \) has the properties we’ve posited for it. ■

For later use, we’ll abbreviate equation (4.2) as
\[
g_1 \cdots g_k = \chi_1 \cdots \chi_k + \tilde{S}, \tag{4.1bis}
\]
wherein \( \tilde{S} \) is now viewed as an \( \{\chi_j, \varepsilon_j\} \) counterpart of the expression \( S \) that occurs in (4.1).

4.4. Moving onward, suppose next that \( n \geq 1 \) and that the variables \( \{x, v\} \cup \{x_j, v_j\}_{j=1}^n \) are all real. Keep \( \Delta \) and \( \Delta_j \) positive, and write
\[
E = [-\Delta, \Delta] , \ E_j = [-\Delta_j, \Delta_j].
\]
Also set
\[
\mathcal{E} = E_1 \times \cdots \times E_n
\]
and let \( d\vec{v} \) be an abbreviation for the standard Euclidean volume element \( dv_1 \cdots dv_n \).

For functions \( f \in C(E) \), we introduce a difference operator \( D \) by writing
\[
(Df)(v) = \frac{1}{2} [f(v) - f(-v)]. \tag{4.3}
\]
For \( G \in C(\mathcal{E}) \), we then let \( D_j G \) denote the obvious partial difference. It is easily seen that the operators \( D_j \) commute; moreover, \( D_j^2 = D_j \). The vector space \( C(\mathcal{E}) \) therefore splits into a direct sum of \( 2^n \) simultaneous \( D_j \)-eigenspaces \( C(\mathcal{E})_\sigma \), wherein \( \sigma \) corresponds to \( \prod_{j=1}^n \{0, 1\} \) in an obvious fashion.

On a related note, recall that a function \( h \in C(E) \) is said to be Hermitian when \( h(-x) = \overline{h(x)} \). Similarly for \( H \in C(\mathcal{E}) \); here
\[
H(-x_1, \ldots, -x_n) = \overline{H(x_1, \ldots, x_n)}.
\]
By the symmetry properties of \( d\vec{v} \), one immediately sees that
\[
\int_{\mathcal{E}} H(v_1, \ldots, v_n) d\vec{v} \in \mathbb{R}.
\]
Similarly for \( h \).

A moment’s thought shows that functions in \( C(\mathcal{E}) \) are Hermitian precisely when their \( \sigma \)– components in \( C(\mathcal{E}) \) (\( G_\sigma \), say) are respectively either real or purely imaginary depending on the parity of \( \|\sigma\| \equiv \sigma_1 + \cdots + \sigma_n \).

In line with this, it is also helpful to observe that:
(i) the characteristic function \( \varphi(v_1, \ldots, v_n) \) of any probability distribution \( F \) (or totally finite signed measure \( \eta \)) on \( \mathbb{R}^n \) is automatically Hermitian;
(ii) likewise for any \( \mathbb{R} \)-linear combination \( (D_{i_1} \cdots D_{i_r}) \varphi \) with \( \iota_m \uparrow \) and \( 1 \leq r \leq n \);
(iii) and – yet again – for the functions \( \frac{1}{k} R_B(\frac{x}{k}) \) and \( \frac{1}{k} R_b(\frac{x}{k}) \) in \( C(\mathcal{E}) \) that arose in the proof of (3.1).

4.5. Our subsequent continuity considerations and crude estimates for functions of types (i) and (ii) will generally rest either implicitly or explicitly on some combination of three basic inequalities; viz.,

\[
|e^{2i(\tau + \phi)} - e^{2i\phi}| = |e^{2i\tau} - 1| = 2|\sin \tau| \leq 2 \min(1, |\tau|) \leq 2 \min(1, |\tau|^{\omega}),
\]

\[
\prod_{j=1}^n (|x_j| + |y_j|)^{\beta_j} \leq (\|X\|_\infty + \|Y\|_\infty)^{\beta_1 + \cdots + \beta_n},
\]

\[
|\prod_{j=1}^n w_j - \prod_{j=1}^n \xi_j| \leq \sum_{j=1}^n |w_j - \xi_j| \quad \text{whenever} \quad |w_j|, |\xi_j| \leq 1.
\]

It is understood herein that \( \phi \) and \( \tau \) are real, \( 0 < \omega < 1 \), and \( \beta_j \geq 0 \).

4.6. Finally, in connection with both \( \mathbb{R} \) and \( \mathbb{R}^n \), it is convenient to have at one’s disposal a symbolic Dirac delta function \( \delta(\cdot) \) such that, for any \( r \in [1, n] \) and \( G \in C(\mathcal{E}) \),

\[
\int_{\mathcal{E}} \delta(v_1) \cdots \delta(v_r) G(v_1, \ldots, v_n) \, d\vec{v} = \left\{ \begin{array}{c} G(0, \ldots, 0) & \text{if } r = n \\ \int_{E_{r+1} \times \cdots \times E_n} G(0, \ldots, 0, v_{r+1}, \ldots, v_n) \prod_{j=r+1}^n dv_j & \text{if } r < n \end{array} \right\}.
\]

Similarly for \( f \in C(\mathcal{E}) \) and permuted sets of indices. Also for functions \( \tilde{G} \) whose continuity is assumed only over, e.g.,

\[
\tilde{\mathcal{E}} \equiv \mathcal{E} \cap \{v_{r+1} \cdots v_n \neq 0\},
\]

the behavior elsewhere being such that one at least has

\[
\sup_{E_1 \times \cdots \times E_r} |\tilde{G}(\xi_1, \ldots, \xi_r; v_{r+1}, \ldots, v_n)| \in L_1(E_{r+1} \times \cdots \times E_n),
\]

the latter condition serving to ensure that the \( dv_{r+1} \cdots dv_n \) “cross-sectional integral” of \( \tilde{G} \) remains nicely continuous as the level \( (\xi_1, \ldots, \xi_r) \) varies.
4.7. We are now able to state a prototypical version of our $k$-variable counterpart of (3.1). Matters are facilitated in this by letting $P$ denote a generic partition of $\{1, 2, ..., k\}$ into three disjoint (possibly empty!) subsets $\langle B, C, D \rangle$ and then writing

$$D_C G \equiv \left( \prod_{j \in C} D_j \right) G.$$  

(Here and below, we follow the standard convention that empty products are understood to be either 1 or the identity operator. When the context is clear, we also agree that “$\sim$” atop anything signifies expungement.)

**Theorem 4.8.** Given $k \geq 1$ and any list of positive numbers $\{\Omega_1, \ldots, \Omega_k\}$. Let $F$ be any probability distribution on $\mathbb{R}^k$ and $dG = g(x_1, \ldots, x_k) d\vec{x}$ any totally finite signed measure having $g \in C(\mathbb{R}^k)$. Suppose that

$$\int_{\mathbb{R}^k} \left( \max_{1 \leq j \leq k} |x_j| \right)^\alpha dF < \infty, \quad \int_{\mathbb{R}^k} \left( \max_{1 \leq j \leq k} |x_j| \right)^\alpha |g(x_1, \ldots, x_k)| d\vec{x} < \infty \quad (4.9)$$

for some $\alpha > 0$ and that, in addition,

$$\int_{\mathbb{R}^{k-1}} |g(x_1, \ldots, x_k)| dx_1 \cdots dx_{\ell} \cdots dx_k \leq m_\ell \quad (4.10)$$

for every $\ell \in [1, k]$ (the integral sign simply being absent if $k = 1$). Let $\varphi(\vec{v})$ and $\psi(\vec{v})$ be the usual multivariate characteristic functions of $F$ and $dG$. (Cf. §3.1.) With

$$G(y_1, \ldots, y_k) \equiv \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} g(\xi_1, \ldots, \xi_k) d\vec{\xi}, \quad (4.11)$$

there then exist positive constants $c_1$ and $c_2$ depending solely on $k$ such that one has the a priori bound

$$|F(t_1, \ldots, t_k) - G(t_1, \ldots, t_k)| \leq c_1 \sum_p \left| \prod_{j = 1}^k \left( \frac{1}{\Omega_j} \prod_{j \in C} \frac{1}{|\Omega_j|} \prod_{j \in D} \left( \frac{1}{|\Omega_j|} + \frac{|\sin(t_jv_j)|}{|v_j|} \right) \prod_{j \in B} \delta(v_j) \right) d\vec{v} \right|$$

$$+ c_2 \sum_{\ell = 1}^k \frac{m_\ell}{\Omega_\ell}$$

at every point $(t_1, \ldots, t_k) \in \mathbb{R}^k$. (Though (4.12) does not require $\psi(\vec{0}) = 1$, cases having $\psi(\vec{0}) \neq 1$ will generally be of at most subsidiary interest from a probabilistic standpoint; cf. the term with $C = D = \phi$.)

**Proof.** Apart from some additional bookkeeping, the basic procedure is largely a mimic of the approach that was used for Lemma 3.2 and Theorem 3.3. We start by taking $\alpha < 1$ w.l.o.g and writing $\beta = \alpha/k$. By (4.4) and a trivial manipulation,
Esseen’s Lemma over \( \mathbb{R}^k \)

\[
|\sin u| \leq \min\{1, |u|\} \leq \min\{1, |u|^\beta\};
\]

\[
D_{\epsilon\varphi} = \int_{\mathbb{R}^k} \prod_{j \in \mathcal{E}} (i \sin(v_j x_j)) \prod_{j \in \mathcal{B}, \mathcal{D}} e^{iv_j x_j} dF(x).
\]  
(4.13)

(Similarly for \( D_{\epsilon\psi} \).) Relation (4.8) is thus fulfilled by a large margin for every \( \mathcal{P} \)-summand in (4.12); cf. (4.9) and §4.5. Notice too that there is no loss of generality if a large positive constant \( N \) is added to each “max” in (4.9).

It will be convenient to first prove (4.12) in the case where \( (t_j) = (0) \) and \( F \) has a continuous density function \( f(x_1, \ldots, x_k) \). To this end, we write

\[
\chi(x) = \chi(-\infty, 0)(x)
\]

and then use Theorem 2.2 to get

\[
\frac{1}{2} [1 - B(\Omega x)] \leq \chi(x) \leq \frac{1}{2} [1 - b(\Omega x)]
\]

for each \( \Omega > 0 \). This quickly leads to

\[
\tilde{f}_\Omega(x) \leq \chi(x) \leq \tilde{g}_\Omega(x)
\]  
(4.14)

with

\[
\tilde{f}_\Omega(x) = - \frac{1}{2} \int_{\Omega} \frac{D_w(e^{2\pi i w x})}{\pi i w} dw + \frac{1}{2} \int_{-\Omega} \frac{e^{2\pi i x w}}{w} \left[ \delta(w) - \frac{1}{\Omega} R_B(w) \right] dw
\]

\[
\tilde{g}_\Omega(x) = - \frac{1}{2} \int_{-\Omega} \frac{D_w(e^{2\pi i w x})}{\pi i w} dw + \frac{1}{2} \int_{\Omega} \frac{e^{2\pi i x w}}{w} \left[ \delta(w) - \frac{1}{\Omega} R_B(w) \right] dw
\]

thanks to equation (3.3). By Theorem 2.2(a), one knows that

\[
\tilde{g}_\Omega(x) - \tilde{f}_\Omega(x) = K(\Omega x).
\]

The situation of Lemma 4.3 is thus applicable with

\[
\left\{ \begin{array}{l}
\chi_j = \chi(x_j), f_j = \tilde{f}_\Omega_j(x_j), g_j = \tilde{g}_\Omega_j(x_j) \\
0 \leq \delta_j \leq K(\Omega_j x_j), 0 \leq \epsilon_j \leq K(\Omega_j x_j)
\end{array} \right. \}
\]  
(4.15)

To continue, we let

\[
T^-(x_1, \ldots, x_k) = \text{the LHS of (4.1)}
\]

\[
T^+(x_1, \ldots, x_k) = g_1 \cdots g_k
\]

and then exploit the fact that

\[
T^- + S = \chi_1 \cdots \chi_k = T^+ - \tilde{S}
\]  
(4.16)

in accordance with (4.1) and (4.1bis). Since

\[
F(\tilde{0}) - G(\tilde{0}) = \int_{\mathbb{R}^k} (f - g) \chi_1 \cdots \chi_k d\tilde{z}
\]
and $f \geq 0$, one immediately obtains

$$F(\vec{0}) - G(\vec{0}) \geq \int_{\mathbb{R}^k} (f - g)T^- d\vec{x} - \int_{\mathbb{R}^k} |g|S d\vec{x}$$  \hspace{1cm} (4.17a)

$$F(\vec{0}) - G(\vec{0}) \leq \int_{\mathbb{R}^k} (f - g)T^+ d\vec{x} + \int_{\mathbb{R}^k} |g|\tilde{S} d\vec{x}.$$  \hspace{1cm} (4.17b)

Each $g_j$ or $f_j$ appearing in $T^\pm$ can and will be viewed as the sum of three obvious chunks; cf. the formulae following (4.14).

Prior to pushing onward with this, we note that, by (4.10), (4.15), and the algebraic format of $S$ and $\tilde{S}$, the final terms in (4.17a) and (4.17b) are manifestly dominated by

$$\sum_{\ell=1}^k c_3 \int_{-\infty}^{\infty} K(\Omega_{\ell}x_\ell) m_\ell dx_\ell = c_3 \sum_{\ell=1}^k \frac{m_\ell}{\Omega_{\ell}}$$  \hspace{1cm} (4.18)

for some choice of $c_3 \neq c_3(k)$.

The central issue now becomes one of substitution and straightforward justification of the applicability of Fubini’s theorem. For the latter, it suffices to combine (4.9), (4.5), and our earlier observations about $|\sin u|$ and $\mathfrak{M}$. Cf. also here the discussion following (3.3), especially as regards the uniform boundedness of the integral

$$\int_{\gamma_1}^{\gamma_2} \frac{\sin(2\pi xw)}{\pi w} dw$$

and related use of the Lebesgue dominated convergence theorem.

On the basis of these remarks, it quickly becomes apparent that all relevant manipulations for “case $\vec{0}$” are easily carried out — in fact, separately so for both $\phi$ and $\psi$, and with good majorants throughout. One clearly obtains (4.12) with $\phi(2\pi \vec{w})$, $\psi(2\pi \vec{w})$ in place of $\phi(\vec{v})$, $\psi(\vec{v})$. To finish up format-wise, one simply replaces each $\Omega_j$ by $\Omega_j/2\pi$ and writes $v_j = 2\pi w_j$.

Keeping $dF = fd\vec{x}$, we turn next to the case of an arbitrary point $\{t_j\} \in \mathbb{R}^k$. One freezes $\{t_j\}$ and introduces the new densities

$$f^*(x_1, \ldots, x_k) = f(t_1 + x_1, \ldots, t_k + x_k), \quad g^*(x_1, \ldots, x_k) = g(t_1 + x_1, \ldots, t_k + x_k).$$

Conditions (4.9) and (4.10) are still satisfied. Working separately with the analogs of (4.17a) and (4.17b) in accordance with the first two sentences of the preceding paragraph [cf. also here (4.18)], it is immediately evident that, up to trivial $Q$-coefficients, one obtains a specific sum of $d\vec{w}$–integrals over

$$\mathcal{E} = [-\Omega_1, \Omega_1] \times \cdots \times [-\Omega_k, \Omega_k]$$

whose integrands have the form

$$\prod_{j \in \mathcal{C}} \frac{1}{\pi w_j} \prod_{j \in \mathcal{D}} \frac{1}{\Omega_j} R_s \left( \frac{w_j}{\Omega_j} \right) \prod_{j \in \mathcal{B}} \delta(w_j)$$

$$\cdot \int_{\mathbb{R}^k} \prod_{j \in \mathcal{C}} \sin(2\pi w_j x_j) \prod_{j \notin \mathcal{C}} e^{2\pi i w_j x_j} (f^* - g^*) d\vec{x}.$$
Esseen’s Lemma over $\mathbb{R}^k$

(Cf. (4.14).) Writing $x_j = y_j - t_j$, each such integrand then becomes, again up to trivial $Q$-coefficients, a specific sum of expressions like

$$
\prod_{j \in D} \frac{1}{\Omega_j} R_* \left( \frac{w_j}{\Omega_j} \right) \prod_{j \in B} \delta(w_j) \prod_{j \notin C} e^{-2\pi i w_j t_j}
$$

$$
\cdot \int_{\mathbb{R}^k} \prod_{j \in C_1} \left( \frac{\sin(2\pi w_j y_j)}{\pi w_j} \right) \cos(2\pi w_j t_j) \prod_{j \in C_2} \left( \frac{\sin(2\pi w_j t_j)}{\pi w_j} \right) \cos(2\pi w_j y_j)
$$

$$
\cdot \prod_{j \notin C} e^{2\pi i w_j y_j} (f - g) d\vec{y},
$$

wherein $(C_1, C_2)$ is a partition of $C$. Since $\max\{|y_j|, |t_j|\} \leq |y_j| + |t_j|$ for $j \in C$ and

$$
\int_{\mathbb{R}^k} (1 + \|\vec{y}\|_\infty + \|\vec{t}\|_\infty)^{\alpha} (f + |g|) d\vec{y} < +\infty,
$$

each of these new integrands clearly admits a good majorant over the set $E$. Observe now that one can successively “flip” each $\cos(2\pi w_j y_j)$ (with $j \in C_2$) into $\exp(2\pi i w_j y_j)$ without changing the numerical value of the corresponding iterated $d\vec{y}$–integral over $E$. Similarly for $\cos(2\pi w_j t_j)$ with $j \in C_1$.

The upshot of course is that each expression in (4.19) can thus be replaced by

$$
\prod_{j \in D} \frac{1}{\Omega_j} R_* \left( \frac{w_j}{\Omega_j} \right) \prod_{j \in B} \delta(w_j) \prod_{j \notin C} e^{-2\pi i w_j t_j}
$$

$$
\cdot \int_{\mathbb{R}^k} \prod_{j \in C_1} \left( \frac{i \sin(2\pi w_j y_j)}{i\pi w_j} \right) \prod_{j \notin C} \left( \frac{\sin(2\pi w_j t_j)}{\pi w_j} \right) \prod_{j \in C_2} e^{2\pi i w_j y_j} (f - g) d\vec{y},
$$

It makes sense to rewrite this as

$$
\prod_{j \in D} \frac{1}{\Omega_j} R_* \left( \frac{w_j}{\Omega_j} \right) \prod_{j \in B} \delta(w_j) \prod_{j \notin C} e^{-2\pi i w_j t_j} \prod_{j \in C_2} \left( \frac{\sin(2\pi w_j t_j)}{\pi w_j} \right)
$$

$$
\cdot \int_{\mathbb{R}^k} \prod_{j \in C_1} \left( \frac{i \sin(2\pi w_j y_j)}{i\pi w_j} \right) \prod_{j \notin C} e^{2\pi i w_j y_j} (f - g) d\vec{y},
$$

the $d\vec{y}$–integral herein simply being

$$
D_{C_1}(\varphi - \psi)(2\pi \vec{w}) \prod_{j \in C_1} (i\pi w_j).
$$

Notice incidentally that (4.20) is manifestly Hermitian w.r.t. $(w_1, \ldots, w_k)$.

Writing $D_a = D \cup C_2$, the corresponding contribution to the overall estimate for $|F(t_j) - G(t_j)|$ thus becomes some real number having absolute value
\[
\leq c_4 \int_{\mathcal{E}} \prod_{j \in \mathcal{D}_a} \left( \frac{1}{2\pi \Omega_j} + \frac{\left| \sin(2\pi w_j t_j) \right|}{2\pi |w_j|} \right) \prod_{j \in \mathcal{B}} \delta(w_j) \cdot \frac{|D_{\mathcal{E}_j}(\varphi - \psi)(2\pi w)|}{\prod_{j \in \mathcal{E}_j} |2\pi w|} \, dw.
\]

Needless to say: $c_4 = c_4(k)$ and $\{1, 2, \ldots, k\} = \mathcal{C}_1 \cup \mathcal{D}_a \cup \mathcal{B}$.

Upon taking $v_j = 2\pi w_j$, temporarily setting $\Omega_j = E_j/2\pi$, and looking at
\[
\mathcal{P}^* = \langle \mathcal{B}, \mathcal{C}_1, \mathcal{D}_a \rangle
\]
in lieu of $\mathcal{P}$, relation (4.12) follows\(^\text{15}\)

To establish (4.12) for a perfectly general $F$, one simply convolves $F$ with a multivariate Gaussian ($N_{\varphi}^{(k)}$, say) having density
\[
\left( \frac{1}{\varepsilon \sqrt{2\pi}} \right)^k \prod_{j=1}^k \exp \left( - \frac{u_j^2}{2 \varepsilon^2} \right)
\]
and then lets $\varepsilon \to 0$ in obvious analogy to what was done earlier for $k = 1$ and Theorem 3.3. Since
\[
\prod_{j=1}^k \exp \left( - \frac{1}{2} \varepsilon^2 v_j^2 \right)
\]
is even w.r.t. each $v_j$, the action of $D_{\mathcal{E}}$ is trivially visualized and there is no difficulty securing a good majorant for each $\mathcal{P}$–summand over $\mathcal{E}$. \[\blacksquare\]

4.9. When $k = 1$, (4.12) is readily seen to provide a very slight improvement in (3.1) [apart from choice of constants]. For $k \geq 2$, however, matters are less satisfactory. There (4.12) turns out to have two features that tend to be somewhat problematic for purposes of applications.

The most egregious of these is the fact that, when $k > 1$, the RHS of (4.12) typically tends to blow up anytime one or more of the entries in $(t_1, \ldots, t_k)$ diverges to $\pm \infty$.

The difficulty stems from the “sine” portion of the $D$-terms. It suffices to look at the partition for which $\mathcal{D} = \{1, 2, \ldots, k\}$. Assume, for simplicity, that $\alpha$ can be taken very large in (4.9). After utilizing the well-known relation
\[
e^{iy} = \sum_{j=0}^N \frac{(iy)^j}{j!} + \Theta \frac{|y|^{N+1}}{(N+1)!} \quad \text{(for $y \in \mathbb{R}$)} \quad (4.21)
\]
to form obvious Taylor polynomials and rescaling things a bit, the operative “inflationary” mechanism is most easily appreciated [in a model computation

\(^{15}\) Note that the form of the product for $j \in \mathcal{D}_a$ already accounts for all choices of $\mathcal{C}_2$ associated with a given $\mathcal{D}_a$.

\(^{16}\) Cf. equation (5.1) below.
with \( k = 3 \) by contemplating the fact that one has
\[
\int_{1/T_1}^{1} \int_{1/B_2}^{1} \int_{1/B_3}^{1} \left[ \frac{1}{T_1} \frac{1}{B_2} \frac{1}{B_3} \left( u \ell + v \ell + w \ell - \lambda u \ell - \lambda v \ell - \lambda w \ell \right) \right] dwdvdudvdu
\]
\[
\geq \int_{1/T_1}^{1} \int_{1/B_2}^{1} \int_{1/B_3}^{1} \left[ u \ell + v \ell + w \ell \right] \frac{u \ell + v \ell + w \ell}{uvw} dwdvdudvdu
\]
\[
\geq \frac{1 - 2^{-\ell}}{2\ell} \left[ (\log B_2)(\log B_3) + (\log T_1)(\log B_3) + (\log T_1)(\log B_2) \right]
\]
anytime \( T_1 \geq 2, B_2 \geq 2, B_3 \geq 2, 1 \leq c \leq \ell, \lambda \in [0, 1/2] \). Here \( T_1 \) corresponds essentially to \( |t_1| \).

The second snag is more subtle and – as will be seen momentarily – has a predominantly “operational” nature. It arises in connection with those entries

\[
D_C(\varphi - \psi) \prod_{j \in \mathcal{C}} v_j (\equiv Q_C)
\]
in (4.12) for which \( \text{card}(\mathcal{C}) \geq 2 \). We take \( \mathcal{C} = \{1, \ldots, m\} \) w.l.o.g. (and continue to assume \( k > 1 \)). To explicate matters\(^{17}\), it is helpful to first observe that

\[
(D_1 \cdots D_m)f = \frac{1}{2^m}(v_1 \cdots v_m) \int_{[-1,1]^m} D[m]f(v_1 u_1, \ldots, v_m u_m) d\vec{u} \quad (4.22)
\]
holds whenever \( f \in C^m \). Here \( D[m]f \) is the obvious mixed partial. (There is clearly no harm in suppressing, as we do, any variables that are inactive. Notice too that the \( d\vec{u} \)-integral is automatically even w.r.t. each \( v_j \).)

Suppose now that \( \vec{v} \) is situated in that portion of \( \prod_{j=1}^{k} [-\Omega_j, \Omega_j] \) on which

\[\{ |v_1| < \tau, \ldots, |v_m| < \tau \}.\]

Here \( \tau \) is some suitably small constant. Think of \( \|\vec{t}\| \) as being bounded. The need to avoid spurious divergences in (4.12) growing out of

\[
\int_{0}^{\tau} \cdots \int_{0}^{\tau} \frac{\lambda(v_1^{\ell+1} + \cdots + v_m^{\ell+1})}{v_1 \cdots v_m} dv_m \cdots dv_1 = +\infty \quad (\lambda > 0)
\]
when relation (4.21) is applied crudely over \( \mathbb{R}^m \) strongly suggests that any Taylor approximations for \( D_C(\varphi - \psi) \) w.r.t. \( (v_1, \ldots, v_m) \) that one proposes to exploit be created in a “pre-factored” format by means of either \( (4.22)+(4.21) \) or else \( (4.13) + \) a variant of (4.21) [call it (4.21)\(^\sharp\)] which refers to the product function

\[\sin(y_1) \cdots \sin(y_m)\]

\[\frac{y_1}{y_1} \cdots \frac{y_m}{y_m} .\]

The remainder term in relation (4.21)\(^\sharp\) is readily seen (after a bit of manipulation with multinomial expansions) to have magnitude at most

\(^{17}\) (some readers may prefer to merely skim the details that follow)
for \( y_\nu \in \mathbb{R} \) once the total \( \vec{y} \)–degree pushes past \( N \). A closer analysis reveals that any fixed number \( \tilde{D} \geq 1 \) can be inserted as a denominator in \( \sum_{j=1}^{m} |y_j| \).

To ensure that everything remains well-defined here [in the first approach as well], one tacitly assumes in the foregoing that \((4.21)\) holds with some \( \alpha \geq m + N + 1 \).

Matters receive a useful clarification when Taylor’s theorem with remainder written in integral form is brought into the picture. (Recall that the \( \mathbb{R}^m \)-version of Taylor’s theorem is an easy consequence of the corresponding result over \( \mathbb{R}^1 \).) By writing things out for the function \( G = \prod_{j=1}^{m} \exp(i\xi_j) \), putting \( \xi_j = y_j u_j \), and then averaging w.r.t. \( \vec{u} \) over \([-1,1]^m \), an explicit form of \((4.21)\) immediately ensues. The essential point here, of course, is that

\[
\left( \frac{\sin y}{y} \right)^{(a)} = \frac{1}{2} \int_{-1}^{1} (iu)^a e^{i\mu u} du \quad \text{for } a \geq 0. \tag{4.23}
\]

The equation for \( D_{\vec{y}}\varphi/(v_1 \cdots v_m) \) produced by this version of \((4.21)\) is clearly just the earlier Taylor manipulation repeated for the augmented function \( \prod_{j=1}^{m} \exp(i\xi_j) \cdot \Omega \), wherein \( \xi_j = (v_j x_j) u_j \) and

\[
\Omega = \prod_{j=m+1}^{k} e^{i v_j x_j} \cdot (ix_1) \cdots (ix_m) dF,
\]

followed by an integration w.r.t. \((\vec{u}, \vec{x})\), the part involving \( \vec{x} \) being done last.

Since \( D^{[m]}\varphi(w_1, \ldots, w_m; v_{m+1}, \ldots, v_k) \) can be viewed as a continuous superposition of \( \prod_{j=1}^{m} \exp(iw_j x_j) \cdot \Omega \), it is hardly surprising that formula \((4.22)\) with \( f = \varphi \) appears when exactly the same procedure is followed, but the integration w.r.t. \((\vec{u}, \vec{x})\) is done in reverse order. Integrals having the form

\[
\int_{\mathbb{R}^k} (ix_1)^{a_1} e^{i w_1 x_1} \cdots (ix_m)^{a_m} e^{i w_m x_m} \cdot \Omega
\]

are merely partial derivatives of \( D^{[m]}\varphi \); as such, it is easily seen that the Taylor developments of \( D_{\vec{y}}\varphi/(v_1 \cdots v_m) \) obtained by first integrating the one attached to \((4.21)\) and, then, that of \( D^{[m]}\varphi \) (à la \((4.22)\)) actually have summands that are numerically identical term-by-term. Methods 1 and 2 for handling \( Q_\varphi \) over the given portion of \( \prod_{j=1}^{k} [-\Omega_j, \Omega_j] \) are thus equivalent, at least to the extent that the sharper version of \((4.21)\) is used.

When \( F \) is fixed, none of this presents any serious difficulty. In settings, however, where \( \varphi \) is basically given as a product of many (scaled down, more basic) characteristic functions \( \varphi_\nu \), making any effective use of this multiplicity vis à vis \( D_{\vec{y}}\varphi/(v_1 \cdots v_m) \) requires a bit of thought, lest unwelcome spurious divergences re-appear.

On a practical level, for run-of-the-mill \( \varphi_\nu \), one finds two main options. First: viewing \( F \) as an \( M \)-fold convolution, one can seek to obtain some sort of \emph{a priori} hold on the magnitudes of appropriately large \((4.9)\)-style moments of \( dF \). In certain settings, such information may be relatively easy to obtain by means of a bit of algebra. (With bounds of any reasonable quality here, use of even the cruder version of \((4.21)\) may already turn out to be sufficient for one’s needs.)
The second option would be to exploit method 1; i.e., come in by way of (4.22). Writing \( f = \varphi = \prod \varphi_\nu \), one would simply expand \( D^{[m]} \varphi \) via a rule of Leibnitz type. This is not unreasonable since \( m \leq k \). (We tacitly assume here that each \( \varphi_\nu \) corresponds to a \( dF_\nu \) having finite moments out to order at least \( m + N + 1 \).) Proceeding in this way enables one to achieve fairly quickly at least some orderly use of the multiplicativity “without any division worries.” Depending on the form of \( \varphi_\nu \), keeping \( \tau \) sufficiently small may also facilitate the calculation of any relevant derivatives of \( \log \varphi_\nu \).

In situations where \( \varphi_\nu \) and \( \varphi = \prod \varphi_\nu \) are nicely-behaved analytic functions of \( \vec{v} \), a third line-of-attack would be to express the linear combination \( D_{\vec{v}} \varphi \) as a Cauchy-type integral w.r.t. \( \{v_j\}_{j=1}^m \). Doing this produces a factorization formula akin to (4.22).

In all three of these approaches, it pays to observe that those portions of \( \prod_{j=1}^m [-\Omega_j, \Omega_j] \) having, e.g.,

\[
\{ |v_1| < \tau, \ldots, |v_\ell| < \tau; |v_{\ell+1}| \geq \tau, \ldots, |v_m| \geq \tau \} \quad (1 \leq \ell \leq m)
\]

can be easily treated by first expressing \( D_{\vec{v}} \varphi \) as \( D_1 \cdots D_\ell g \), wherein \( g \) is the explicit linear combination \( D_{\ell+1} \cdots D_m \varphi \), and then simply repeating the earlier ideas with \( R_{\ell} \) in place of \( R_m \).

What is evident from all this is that, in seeking to gain some sensible control on \( Q_C \) in (4.12) by way of Taylor approximations when \( \varphi = \prod \varphi_\nu \), there is invariably an element of context that enters the picture even if, say, \( \alpha \geq 1000k \).

4.10. This fact, which is not entirely surprising, prompts one to ponder the possibility of making a fundamental change in mindset that would enable one to simply sidestep the bulk of any contextual issues with \( Q_C \).

To that end, the following corollary of (4.22) [taken now with any \( m \geq 1 \)] is highly suggestive, particularly when specialized to \( \varphi - \psi \). Namely: for any \( f \in C^h \) and \( \ell \in [1, m] \), one always has

\[
|D_1 \cdots D_m f| \leq |v_1|^{h/\ell} \cdots |v_\ell|^{h/\ell} \left( \prod_{\sigma} M_\sigma f \right)^{1/(\ell)} ,
\]

wherein \( h \in [1, \ell] \), \( \sigma \) ranges over all subsets of \( [1, m] \) for which

\[
\text{card}(\sigma) = h \quad \text{and} \quad \sigma \subseteq \{1, 2, \ldots, \ell\} ,
\]

and one defines, e.g.,

\[
M_{\{1,2,\ldots,h\}} f = \max_{\varepsilon_{|s|} \leq 1} \sup_{|u_i| \leq 1} \left| \frac{\partial^h f}{\partial \xi_1 \cdots \partial \xi_h} (u_1 v_1, \ldots, u_h v_h; \varepsilon_{h+1} v_{h+1}, \ldots, \varepsilon_m v_m) \right| .
\]

The index \( \gamma \) ranges over \( \sigma \), while the “"" serves to stress that only \( \sigma \subseteq [1, \ell] \) are included in the product.

The case \( h = 0 \) is readily seen to hold trivially if the \( M_\sigma \) - product is given an obvious interpretation. At the same time, it is also worth noting that, as a technical artifice, taking \( h = 1 \) will already be sufficient for most purposes.

\[\text{Compare [Fel2, p.528 (problem 15)].}\]
To prove (4.24) per se, we first put \( g = D_{h+1} \cdots D_m f \) and then think of \( g \) as a linear combination. The sum of the absolute values of its coefficients is \( 2^{m-h}/2^{m-h} = 1 \). With \( \{v_{h+1}, \ldots, v_m\} \) frozen, we now apply (4.22) but with \( m \) replaced by \( h \). This gives:

\[
|D_1 \cdots D_m f| = |(D_1 \cdots D_h) g| \leq |v_1| \cdots |v_h| (M_{(1, \ldots, h)} g) \leq |v_1| \cdots |v_h| (M_{(1, \ldots, h)} f).
\]

(The first \( M \) is taken on \( \mathbb{R}^h \).) Similarly for any other \( \sigma \subseteq [1, \ell] \) having cardinality \( h \). Multiplying over the collection of all relevant \( \sigma \), one clearly gets

\[
|D_1 \cdots D_m f| \bigg|_{\sigma} \leq \prod_{j=1}^{\ell} |v_j| (\ell-1) \cdot (\prod_{\sigma} M_{\sigma} f).
\]

Since

\[
\binom{\ell - 1}{h - 1} = \frac{h}{\ell} \binom{\ell}{h}
\]

whenever \( 1 \leq h \leq \ell \), inequality (4.24) follows immediately.

There exists a slight extension of (4.24) that will also turn out to be useful in connection with \( Q_\mathcal{C} \). To explain it, we’ll continue to work in the setting of §4.4 with \( \mathbb{R}^n \) replaced by \( \mathbb{R}^m \). In addition to the operators \( \{D_1, \ldots, D_m\} \), we propose to look at

\[
E_j = I - D_j, \quad (P_j f)(\bm{v}) = f[((1 - \delta_{hj})v_h)_{h=1}^m], \quad \Delta_j = I - P_j
\]

for \( 1 \leq j \leq m \). By straightforward symbol-pushing, each of these operators on \( C(\mathcal{E}) \) is readily seen to commute with the other \( 4m - 1 \), and to be an idempotent (i.e., satisfy \( T^2 = T \)). One also sees that

\[
D_j E_j = 0, \quad D_j P_j = 0, \quad D_j \Delta_j = D_j,
\]

\[
E_j P_j = P_j, \quad P_j \Delta_j = 0;
\]

\[
I = P_j + D_j + E_j \Delta_j. \tag{4.25}
\]

Besides being an idempotent, each summand in (4.25) annihilates the other two. The associated direct sum decomposition of \( C(\mathcal{E}) \) can be viewed as providing a kind of “Boolean” Taylor development w.r.t. \( v_j \). Corresponding to (4.22), one easily checks that

\[
(\Delta_1 \cdots \Delta_m) f = (v_1 \cdots v_m) \int_{[0,1]^m} (D^{(m)} f)(t_1 v_1, \ldots, t_m v_m) d\mathbf{t}
\]

\[
(\prod_{j=1}^m E_j \Delta_j) f = \prod_{j=1}^m \left( \frac{1}{2} v_j^2 \right) \int_{[0,1]^m \times [-1,1]^m} (D^{(m,m)} f)((t_j u_j v_j)) t_1 \cdots t_m d\mathbf{u} d\mathbf{t}
\]

\[
(\prod_{j=1}^m E_j \Delta_j) f = \prod_{j=1}^m \left( \frac{1}{2} v_j^2 \right) \int_{[-1,1]^m} \prod_{j=1}^m (1 - |\xi_j|) \cdot (D^{(m,m)} f)((\xi_j v_j)) d\mathbf{\xi}
\]

\[\text{When } f \text{ is the restriction of a characteristic function } \varphi \text{ to } \mathbb{R}^m, \text{ there are of course bounds similar to } (4.24) \text{ having exponent } \lambda/\ell \text{ (} 0 < \lambda \leq \ell \text{) in place of } h/\ell. \text{ These follow immediately from (4.4) and were already referenced, albeit implicitly, in the proof of Theorem 4.8. The } M_{\sigma} \text{-portion is best written as } E(\|\mathbf{\bar{x}}\|_\lambda^2), \text{ where } \|\mathbf{\bar{x}}\|_\lambda \equiv \max\{|x_j| : 1 \leq j \leq \ell\}.\]
for $f \in C^m$ and $C^{2m}$, respectively. The notations $D^{[m,m]}$ and $((t_j u_j v_j))$ are obvious abbreviations.

Let $1 \leq n \leq m$ and $r = m - n$. Writing $S_j = \partial / \partial \xi_j$, we now define

$$T_j = \begin{cases} S_j^2, & \text{if } j \leq n \\ S_j, & \text{if } j > n \end{cases}.$$  

Fix any $\ell \in [1,n]$ and $\delta \in [0,r]$. Put $L = \ell + \delta$. In line with $\{1,\ldots,\ell\} \cup \{n + 1,\ldots,n + \delta\}$ and the actions of both $I = P_j + \Delta_j$ and (4.25), for any $h \in [1,L]$ and $f \in C^h, C^{2h}$ (resp.), one has

$$|\Delta_1 \cdots \Delta_n D_{n+1} \cdots D_m f| \leq 2^{m-h} (\prod_{j=1}^{\ell} |v_j| \cdot \prod_{j=1}^{\delta} |v_{n+j}|)^{h/L} (\prod_{\sigma \subseteq \{1,\ldots,\ell\}} \mathcal{M}_{\sigma} f)^{1/(\ell)} (4.26)$$

$$|\prod_{j=1}^{\ell} E_j \Delta_j \prod_{j=n+1}^{m} D_j f| \leq 2^{m-h} (\prod_{j=1}^{\ell} |v_j| \cdot \prod_{j=1}^{\delta} |v_{n+j}|)^{h/L} (\prod_{\sigma \subseteq \{1,\ldots,\ell\}} \mathcal{N}_{\sigma} f)^{1/(\ell)} (4.27)$$

$$|\prod_{j=1}^{\ell} E_j \Delta_j \prod_{j=n+1}^{m} D_j f| \leq 2^{m-h} (\prod_{j=1}^{\ell} |v_j|^2 \cdot \prod_{j=1}^{\delta} |v_{n+j}|)^{h/L} (\prod_{\sigma \subseteq \{1,\ldots,\ell\}} \mathcal{N}_{\sigma} f)^{1/(\ell)} (4.28)$$

wherein $\sigma$ ranges over all subsets of $[1,n+r] (=[1,m])$ for which

$$\text{card}(\sigma) = h, \quad \sigma \subseteq \{1,\ldots,\ell\} \cup \{n + 1,\ldots,n + \delta\}$$

and one defines, e.g.,

$$\mathcal{M}_{(1,\ldots,h)} f = \max_{\varepsilon_\beta \geq 0,1} \sup \left| (S_1 \cdots S_h f)(u_1 v_1,\ldots,u_h v_h; \varepsilon_{h+1} v_{h+1},\ldots,\varepsilon_m v_m) \right|$$

$$\mathcal{N}_{(1,\ldots,h)} f = \max_{\varepsilon_\beta \geq 0,1} \sup \left| (T_1 \cdots T_h f)(u_1 v_1,\ldots,u_h v_h; \varepsilon_{h+1} v_{h+1},\ldots,\varepsilon_m v_m) \right|$$

when $h \leq \ell$. Similarly for a split set $\{1,\ldots,h_1\} \cup \{n + 1,\ldots,n + h_2\}$. In each instance, the index $\gamma$ ranges over $\sigma$ and “$\uparrow$” means “omit $\varepsilon_\beta = 0$ if $\beta > n$.”

Estimate (4.27) follows immediately from (4.26). The proofs of (4.26) and (4.28) are easy mimics of that of (4.24). To keep things simple, one makes an intentional overshoot in the coefficient $2^{m-h}$.

Just as before, the case “$h = 0$” holds trivially when the product over $\sigma$ is given a natural interpretation.

In regard to any applications with $f = \varphi - \psi$, the key point in (4.24) and (4.26)–(4.28) is that $h$ can be taken equal to 1. This imposes only the mild restriction that hypothesis (4.9) be true with $\alpha = 1$ or 2 (as opposed to something more like $\alpha \geq m + N + 1$, in our earlier notation).

**4.11.** Before opting to place any further restrictions on $\alpha$, it seems only prudent to ask if there might not exist some easy variant of (4.12) [applicable for any $\alpha$] in which the structural hassles that we identified earlier (in §4.9) are simply absent.

The following result, based on truncation, achieves this goal in a relatively cheap way.
Theorem 4.8A. Given the situation of Theorem 4.8 and any number $\Delta > 1$. Assume further that $\min(\Omega_1, \ldots, \Omega_k) > 1$. Putting $v^* \equiv v/\min\{1, \Delta |v|\}$ (and, say, $0^* = 1/\Delta$), we then have

$$|F(t_1, \ldots, t_k) - G(t_1, \ldots, t_k)|$$

$$\leq c_5 \int_{\Omega_1}^{\Omega_k} \cdots \int_{\Omega_k}^{\Omega_k} |\varphi(v) - \psi(v)| |v_1^*| \cdots |v_k^*| dv + c_6 \sum_{\ell=1}^{k} \frac{m_\ell}{\Omega_\ell}$$

$$(k + 1)\Delta^{-\alpha} \int_{\mathbb{R}^k} (\max_{1 \leq j \leq k} |x_j|)^\alpha (dF + |g|d\bar{x})$$

for every $(t_1, \ldots, t_k) \in \mathbb{R}^k$, wherein $c_5$ and $c_6$ are certain positive constants that depend solely on $k$. (Note the absence of $i$ on the right hand side.) Each factor $1/|v_j^*|$ in (4.29) can be replaced, whenever convenient, by the larger value

$$\frac{\Delta}{|v_j^*|}, \text{ wherein } v^* \equiv \frac{v}{\min\{1, |v|\}}.$$

Proof. It is useful to begin with an intermediate result. Treating $\Delta$ as a natural cut-off, we first write

$$t_j^* = \begin{cases} 
  t_j, & |t_j| \leq \Delta \\
  \Delta, & t_j > \Delta \\
  -\Delta, & t_j < -\Delta
\end{cases}$$

for every index $j$. Since

$$|F(t_1, \eta_2, \ldots, \eta_k) - F(t_1^*, \eta_2, \ldots, \eta_k)| \leq \Delta^{-\alpha} \int_{\mathbb{R}^k} |x_1|^\alpha dF(x)$$

$$\vdots$$

$$|F(\eta_1, \ldots, \eta_{k-1}, t_k) - F(\eta_1, \ldots, \eta_{k-1}, t_k^*)| \leq \Delta^{-\alpha} \int_{\mathbb{R}^k} |x_k|^\alpha dF(x),$$

we immediately see that

$$|F(t_1, \ldots, t_k) - F(t_1^*, \ldots, t_k^*)| \leq k\Delta^{-\alpha} \int_{\mathbb{R}^k} (\max_{1 \leq j \leq k} |x_j|)^\alpha dF(x).$$

(4.30)

Similarly for $G$ and $|g|d\bar{x}$. Observe, however, that one has

$$|F(t_j) - G(t_j)| \leq |F(t_j) - F(t_j^*)| + |F(t_j^*) - G(t_j^*)| + |G(t_j^*) - G(t_j)|$$

in an obvious shorthand. Since $|\sin u| \leq \min\{1, |u|\}$, by applying (4.12) to $|F(t_j^*) - G(t_j^*)|$ and then noting that, on $\{v_j \leq \Omega_j\}$,

$$\frac{1}{\Omega_j} + \min\{|t_j^*|, \frac{1}{|v_j|}\} = \min\left\{\frac{1}{\Omega_j}, |t_j^*|, \frac{1}{|v_j|}\right\} \leq \min\left\{2\Delta, \frac{2}{|v_j|}\right\} = \frac{2}{|v_j^*|},$$
one immediately finds [as a natural relative of (4.12)] that

\[ |F(t_1, \ldots, t_k) - G(t_1, \ldots, t_k)| \]

\[ \leq 2^k c_1 \sum_p \int_{\Omega_1}^{\Omega_k} \cdots \int_{-\Omega_1}^{-\Omega_k} |D_{\mathcal{C}}(\varphi - \psi)| \prod_{j \in \mathcal{C}} \frac{1}{|v_j|} \prod_{j \in \mathcal{D}} \frac{1}{|v_j|} \prod_{j \in \mathcal{B}} \delta(v_j) d\bar{v} \]

\[ + c_2 \sum_{\ell=1}^k \frac{m_{\ell}}{\Omega_\ell} + k \Delta^{-\alpha} \int_{\mathbb{R}^k} (\max_{1 \leq j \leq k} |x_j|)^\alpha (dF + |g|d\bar{x}) \]

for every \((t_1, \ldots, t_k) \in \mathbb{R}^k\). Each \(\mathcal{D}\)-entry can be improved slightly since, a couple lines earlier, we could have just as easily written

\[ \frac{1}{\Omega_j} + \min \left( \frac{|t_j^*|}{|v_j|} \right) = \min \left( \frac{1}{\Omega_j} + |t_j|, \frac{1}{\Omega_j} + \Delta, \frac{1}{\Omega_j} + \frac{1}{|v_j|} \right) \]

\[ \leq \min \left( \frac{2}{\Omega_j} + 2|t_j|, 2\Delta, \frac{2}{|v_j|} \right) \]

\[ = 2 \min \left( \frac{1}{\Omega_j} + |t_j|, \frac{1}{|v_j|} \right). \]

Letting \(\Delta \to \infty\) in this improved version of things basically reproduces (4.12).

Going back to (4.31), the issue is now one of trying to restructure things so that all partitions but one \((\mathcal{D} = \{1, \ldots, k\})\) somehow disappear. For this purpose, we need to return to the proof of Theorem 4.8 and make several key modifications aimed at mollifying the \(1/\pi iw\)-singularity in \(\tilde{f}_\Omega\) and \(\tilde{g}_\Omega\) (cf. (4.14)).

To simplify matters, it is helpful to first treat the case in which \(F\) has a continuous density function \(f(\bar{x})\). We put

\[ L = 2\Delta \]

and initially work with any numbers \(\{\Omega_j\}_{j=1}^k\) situated within \((1/2\pi, \infty)\). We also set

\[ F_L(\tilde{\xi}) = \int_A f(\bar{x})d\bar{x} \quad \text{and} \quad G_L(\tilde{\xi}) = \int_A g(\bar{x})d\bar{x}, \]

wherein

\[ A = (-L, \xi_1] \times \cdots \times (-L, \xi_k] \]

and each \(\xi_j\) is understood to exceed \(-L\). Since

\[ \prod_{j=1}^k (-\infty, \xi_j] - A \subseteq \{ \bar{x} \in \mathbb{R}^k : x_\ell \leq -L \text{ for some index } \ell \}, \]

it is self-evident that

\[ \begin{cases} 0 \leq F(\tilde{\xi}) - F_L(\tilde{\xi}) \leq L^{-\alpha} \int_{\mathbb{R}^k} (\max_{1 \leq j \leq k} |x_j|)^\alpha f d\bar{x} \\ |G(\tilde{\xi}) - G_L(\tilde{\xi})| \leq L^{-\alpha} \int_{\mathbb{R}^k} (\max_{1 \leq j \leq k} |x_j|)^\alpha |g|d\bar{x} \end{cases}. \]
The plan is to now bound $|F_L(\tilde{\xi}) - G_L(\tilde{\xi})|$ by mimicking the considerations utilized near equation (4.16) in the proof of Theorem 4.8. With $\xi$ temporarily frozen, we define

$$\chi_j(x) = \chi(-L,\xi_j)(x) \quad (1 \leq j \leq k)$$

and observe that

$$\chi_k(\vec{x}) = \chi_1(x_1) \cdots \chi_k(x_k)$$

and

$$F_L(\tilde{\xi}) - G_L(\tilde{\xi}) = \int_{\mathbb{R}_k} (f - g) \chi_1(x_1) \cdots \chi_k(x_k) d\vec{x}.$$ 

Dropping the subscript $j$, we also note (cf. (2.10)) that

$$\frac{b[\Omega(x + L)] - B[\Omega(x - \xi)]}{2} \leq \chi(-L,\xi)(x) \leq \frac{B[\Omega(x + L)] - b[\Omega(x - \xi)]}{2}$$

for all $x \in \mathbb{R}$ (the points $-L$ and $\xi$ being handled via trivial limits). By Theorem 2.2(a), the extreme members of (4.33) have a difference equal to $K_1[\Omega(x + L)] + K_2[\Omega(x - \xi)]$.

Following (4.14), we express the minorant portion of (4.33) as a Fourier transform

$$\tilde{f}_\Omega(x) = \frac{1}{2} \int_{-\Omega}^{\Omega} \left\{ \frac{e^{2\pi i L w}}{\pi i w} - \frac{e^{-2\pi i \xi w}}{\pi i w} \right\} e^{2\pi i x w} dw$$

and proceed similarly with the majorant (calling it $\tilde{g}_\Omega$). Each “brace” is clearly Hermitian w.r.t. $w$. In marked contrast, however, to what happened previously, the first brace is now a continuous function on $[-\Omega, \Omega]$: its value is simply

$$2 \frac{\sin[\pi(L + \xi)w]}{\pi w} e^{\pi i (L - \xi)w}.$$ 

At the same time, it is also natural to write

$$\mathcal{R}^-(w) = \frac{e^{2\pi i L w}}{\Omega} R_b\left(\frac{w}{\Omega}\right) - \frac{e^{-2\pi i \xi w}}{\Omega} R_B\left(\frac{w}{\Omega}\right);$$

$$\mathcal{R}^+(w) = \frac{e^{2\pi i L w}}{\Omega} R_B\left(\frac{w}{\Omega}\right) - \frac{e^{-2\pi i \xi w}}{\Omega} R_b\left(\frac{w}{\Omega}\right).$$

One clearly has

$$|\mathcal{R}^\pm(w)| = O(\frac{1}{\Omega})$$

with an implied constant that is absolute (the value $15/4$ certainly works here as a bit of elementary calculus shows).

With the essential components all in place, the earlier reasoning begins
at (4.16)] is readily repeated with \( P = \langle \phi, \mathcal{E}, \mathcal{D} \rangle \) and leads to
\[
|F_L(\xi) - G_L(\xi)| \\
\leq cT \int_{-\Omega_i}^{\Omega_i} \cdots \int_{-\Omega_k}^{\Omega_k} |\varphi(2\pi w) - \psi(2\pi w)| \prod_{j=1}^{k} \left( \frac{|\sin \pi w_j (L + \xi_j)|}{\pi |w_j|} \right) + \frac{1}{\Omega_j} d\bar{w} \\
+ 2^k c_3 \sum_{\ell=1}^{k} \frac{m_\ell}{\Omega_\ell}
\]

The coefficient \( 2^k c_3 \) arises from (4.18), the fact that
\[
\int_{-\infty}^{\infty} \{ K[\Omega_t(x + L)] + K[\Omega_t(x - \xi)] \} dx = \frac{2}{\Omega_t},
\]
and the possibility that each \( \delta_j, \varepsilon_j \) [in Lemma 4.3] can theoretically become as large as 2 (cf. just after (4.33)).

In situations where \( |\xi_j| \leq \Delta \), we clearly have
\[
\frac{|\sin \pi w_j (L + \xi_j)|}{\pi |w_j|} \leq \min \left( \frac{1}{\Omega_j}, |L + \xi_j| \right) \leq \min \left( \frac{1}{|w_j|}, \frac{3}{2}L \right)
\]
\[
\frac{1}{\Omega_j} + \frac{|\sin(\ast \ast \ast)|}{\pi |w_j|} \leq \min \left( \frac{1}{\Omega_j}, \frac{1}{|w_j|}, \frac{1}{\Omega_j} + \frac{3}{2}L \right) \leq \min \left( \frac{2}{|w_j|}, 5L \right)
\]
on \( \{|w_j| \leq \Omega_j\} \), since \( 1/\Omega_j < 2\pi < \pi L \). Here, then,
\[
|F_L(\xi) - G_L(\xi)| \leq 5^k c_7 \int_{-\Omega_i}^{\Omega_i} \cdots \int_{-\Omega_k}^{\Omega_k} |\varphi(2\pi w) - \psi(2\pi w)| \prod_{j=1}^{k} \min \left( \frac{1}{|w_j|}, L \right) d\bar{w} \\
+ 2^k c_3 \sum_{\ell=1}^{k} \frac{m_\ell}{\Omega_\ell}
\]

Putting \( v_j = 2\pi w_j \) and temporarily writing \( \Omega_t = E_t/2\pi \) now gives
\[
|F_L(\xi) - G_L(\xi)| \leq 5^k c_7 \int_{-\Omega_i}^{\Omega_i} \cdots \int_{-\Omega_k}^{\Omega_k} |\varphi(\vec{v}) - \psi(\vec{v})| \prod_{j=1}^{k} \min \left( \frac{1}{|v_j|}, \Delta \right) d\bar{v} \\
+ 2^k (2\pi c_3) \sum_{\ell=1}^{k} \frac{m_\ell}{\Omega_\ell}
\]
for the original \( \Omega_j \) satisfying \( \min(\Omega_1, \ldots, \Omega_k) > 1 \). (Observe that \( 2\pi c_3 \) is simply the coefficient \( c_2 \) that occurred earlier in (4.12); cf. (4.18).) Relation (4.29) follows upon taking \( \xi_j = t_j^* \), noticing that
\[
|F(t_j) - G(t_j)| \leq |F(t_j) - F(t_j^*)| + |F(t_j^*) - F_L(t_j^*)| \\
+ |F_L(t_j^*) - G_L(t_j^*)| + |G_L(t_j^*) - G(t_j)| + |G(t_j) - G(t_j)|,
\]
and finally substituting (4.30) + (4.32).
The case of an arbitrary $F$ is handled by the usual convolution argument. In this connection, it may be worthwhile to point out that, while
\[
E[(\max_j |U_j\pm Y_j|)^\alpha]^{1/\alpha} \leq E[(\max_j |U_j| + \max_j |Y_j|)^\alpha]^{1/\alpha} 
\]
\[
\leq E[(\max_j |U_j|)^\alpha + E[(\max_j |Y_j|)^\alpha]]^{1/\alpha}
\]
is completely standard for $\alpha \geq 1$, the elementary inequality
\[
(|x_1| + |x_2|)^\omega \leq |x_1|^\omega + |x_2|^\omega \quad \text{for} \quad 0 < \omega < 1
\]
leads to
\[
E[(\max_j |U_j\pm Y_j|)^\alpha] \leq E[(\max_j |U_j| + \max_j |Y_j|)^\alpha] 
\]
\[
\leq E[(\max_j |U_j|)^\alpha + (\max_j |Y_j|)^\alpha] 
\]
\[
= E[(\max_j |U_j|)^\alpha] + E[(\max_j |Y_j|)^\alpha]
\]
when $0 < \alpha < 1$. In either situation, taking a multivariate Gaussian $N_2^{(k)}$ for $Y$ and letting $\varepsilon \to 0$ clearly leads to no change in constants in (4.29). ■

By making a slight adaptation in the foregoing argument, the following result parallel to Theorem 4.8A is obtained virtually immediately.

**Theorem 4.8B.** Given the situation of Theorem 4.8 and any number $D > 0$. Assume further that $\min(\Omega_1, \ldots, \Omega_k) > 1$. Writing $v^* = v/\min\{1, \Delta |v|\}$ with $\Delta = 1 + D$, and taking
\[
A_{ab} = \prod_{j=1}^k (a_j, b_j) \subseteq \mathbb{R}^k,
\]
we then have, in the sense of measures,
\[
|F\{A_{ab}\} - G\{A_{ab}\}| \leq c_8 \int_{-\Omega_1}^{\Omega_1} \cdots \int_{-\Omega_k}^{\Omega_k} \frac{|\varphi(\vec{v}) - \psi(\vec{v})|}{|\vec{v}\_1| \cdots |\vec{v}\_k|} d\vec{v} + c_9 \sum_{\ell=1}^k m\ell \Omega\ell \quad (4.35)
\]
whenever $0 \leq b_j - a_j \leq D$. The coefficients $c_8$ and $c_9$ will depend solely on $k$.

**Proof.** The crux of the matter begins just prior to (4.33); one simply replaces $-L$ by $a_j$ and $\xi_j$ by $b_j$, and proceeds from there. Taking $\Omega_j > 1/2\pi$ as before, the key observation is that, on $\{|w_j| \leq \Omega_j\}$,
\[
\frac{1}{\Omega_j} + \frac{\sin \pi w_j(b_j - a_j)}{\pi |w_j|} \leq \frac{1}{\Omega_j} + \min \left( \frac{1}{\pi |w_j|}, |b_j - a_j| \right) \leq 2 \min \left( \frac{1}{|w_j|}, \pi + D \right).
\]
This quickly produces (4.35) with $c_8 = 2^k c_7$ and $c_9 = 2^k (2\pi c_3) = 2^k c_2$. ■
As a \( k \)-dimensional counterpart to Corollary 3.4, one now has

**Corollary 4.12.** Given probability distributions \( F_n \) and \( G \) on \( \mathbb{R}^k \) having characteristic functions \( \varphi_n \) and \( \psi \). Assume that \( G \) has a continuous density function \( g(\bar{x}) \) which satisfies (4.9) and (4.10) of Theorem 4.8. Assume further that

(i) \[ \int_{\mathbb{R}^k} (\|\bar{x}\|)^{\alpha} \, dF_n(\bar{x}) = O(1); \]

(ii) \[ \varphi_n(\bar{v}) \to \psi(\bar{v}) \text{ pointwise on } \mathbb{R}^k \text{ as } n \to \infty. \]

We then have

\[ F_n(\bar{x}) \to G(\bar{x}) \]

uniformly on \( \mathbb{R}^k \). An analogous result holds when \( n \) is replaced by a continuous variable \( \xi \).

Proof. An easy consequence of either Theorem 4.8A or 4.8B as the the cut-off parameter \( \Delta \) is moved upward incrementally. In the case of 4.8B, one exploits the double cut-off idea implicit in (4.30)+(4.32) with, say, \( \frac{1}{2} \Delta \) in place of \( \Delta \) and keeps \( \xi \in [-\frac{1}{2} \Delta, \frac{1}{2} \Delta] \) in relation (4.32). The essential observation is then (4.34). See (4.6)+(4.4) concerning the necessary Hölder estimates. (Recall too §3.6 and our earlier comment about continuous distributions \( \mathcal{P} \).) \( \blacksquare \)

If (i) holds with some \( \alpha > \ell \) (\( \ell \in \mathbb{Z}^+ \)), and \( \mathcal{D} \) is any partial derivative w.r.t. \( \bar{v} \) of order \( \leq \ell \), the assertion \( F_n \to G \) readily implies \( \mathcal{D} \varphi_n \Rightarrow \mathcal{D} \psi \) on compacta. Similarly for the associated moments and absolute moments of \( dF_n \).

**4.13.** Our third, and final, variant of Theorem 4.8 will arise from taking \( \alpha = 1 \) in (4.9), and then revisiting the proof of Theorem 4.8 aided by the circle-of-ideas looked at in §4.10. Though its statement over \( \mathcal{P} \) will be simple enough (see (4.40) infra), the argument used in proving it will entail a bit of bookkeeping. It is helpful to lay the groundwork for this side of things first.

To that end, we begin by noticing, in connection with (4.25), that the functions \( D_j f, E_j f, \Delta_j f, P_j f, \) and \( E_j \Delta_j f \) are all Hermitian anytime \( f \) is; cf. §4.4. At the same time, we recall that

\[ \int_{\gamma_1}^{\gamma_2} \frac{\sin(tv)}{v} \, dv = O(1) \text{ for any } \gamma_2 > \gamma_1, \ t \in \mathbb{R} \quad (4.36) \]

and

\[ \delta(v) = a \delta(av) \quad \text{whenever} \quad a > 0. \quad (4.37) \]

For \( \mathcal{C} \subseteq \{1, 2, \ldots, k\} \) and \( \bar{v} \in \mathbb{R}^k \), we now write

\[ \mathcal{C}_0(\bar{v}) = \{ j \in \mathcal{C} : |v_j| \geq 1 \}, \quad \mathcal{C}_s(\bar{v}) = \{ j \in \mathcal{C} : |v_j| < 1 \}; \quad (\sim \text{big/small}) \]

\[ E_{\mathcal{C}}(\bar{v}) = \{ \xi : \xi_j = v_j \text{ if } j \notin \mathcal{C}, \xi_j = \pm v_j \text{ if } j \in \mathcal{C}_b(\bar{v}), |\xi_j| \leq |v_j| \text{ if } j \in \mathcal{C}_s(\bar{v}) \}, \]

\[ E_{\mathcal{C}}(\bar{v}) = \{ \xi : \xi_j = v_j \text{ if } j \notin \mathcal{C}, \xi_j = \pm v_j \text{ if } j \in \mathcal{C}_b(\bar{v}), |\xi_j| \leq 1 \text{ if } j \in \mathcal{C}_s(\bar{v}) \}. \]

We also select any numbers \( F_1, \ldots, F_k \) in \( (1, \infty) \) and then restrict \( \bar{v} \) to lie in \( \prod_{j=1}^{k} [-F_j, F_j] \). By applying the partition \( [-F_j, F_j] = (-1, 1) \cup \{1 \leq |v_j| \leq F_j \} \),
the box $\prod_{j=1}^{k}[-F_j, F_j]$ clearly splits into $2^k$ symmetrically-shaped (generally disconnected) subregions. The set of interior points of each such subregion will be succinctly referred to as a “slab.” Taking the union of all $2^k$ of these slabs produces $\prod_{j=1}^{k}[-F_j, F_j]$ apart from a set of ($k$-dimensional) measure 0. The resulting “grid” furnishes a natural underpinning for the obvious electrical short-circuit interpretation of the sets $E_c(\vec{v})$ and $E_c(\vec{v})$. Finally, let $S_j = \partial/\partial \xi_j$ as in §4.10 and $\vec{v} = 2\pi \vec{w}$. We thus have

$$\vec{v} \in \prod_{j=1}^{k}[-F_j, F_j] \quad \text{and} \quad \vec{w} \in \prod_{j=1}^{k}[-\omega_j, \omega_j],$$

wherein $\omega_j = \frac{1}{2\pi} F_j$.

Associated with each $f \in C^1(\prod_{j=1}^{k}[-F_j, F_j])$ and $\mathcal{C} \subseteq \{1, 2, \ldots, k\}$ are two \textit{piecewise continuous} functions, namely, $|f|_\mathcal{C}$ and $\|f\|_\mathcal{C}$, that one introduces by writing

$$|f|_\mathcal{C}(\vec{v}) = \begin{cases} \max \{|f(\vec{\xi})| : \vec{\xi} \in E_\mathcal{C}(\vec{v})\}, & \text{if } \mathcal{C}_\mathcal{C}(\vec{v}) = \phi \\ \sup\{|(S_j f)(\vec{\xi})| : \vec{\xi} \in E_\mathcal{C}(\vec{v}), j \in \mathcal{C}_\mathcal{C}(\vec{v})\}, & \text{if } \mathcal{C}_\mathcal{C}(\vec{v}) > \phi \end{cases}$$

(4.38)

and

$$\|f\|_\mathcal{C}(\vec{v}) = \begin{cases} \max \{|f(\vec{\xi})| : \vec{\xi} \in E_\mathcal{C}(\vec{v})\}, & \text{if } \mathcal{C}_\mathcal{C}(\vec{v}) = \phi \\ \sup\{|(S_j f)(\vec{\xi})| : \vec{\xi} \in E_\mathcal{C}(\vec{v}), j \in \mathcal{C}_\mathcal{C}(\vec{v})\}, & \text{if } \mathcal{C}_\mathcal{C}(\vec{v}) > \phi \end{cases}$$

(4.39)

The functions $|f|_\mathcal{C}$ and $\|f\|_\mathcal{C}$ are clearly continuous on each slab.

Since $\|f\|_\mathcal{C}$ typically manifests at least a partial dependence on $\vec{v}$, there is a slight abuse of notation when we write (4.39). Needless to say,

$$|f|_\mathcal{C} \leq \|f\|_\mathcal{C} \quad \text{and} \quad |f|_\phi = \|f\|_\phi = |f(\vec{v})|. $$

In working with $|f|_\mathcal{C}$ and $\|f\|_\mathcal{C}$ slab-by-slab, those variables $v_j$ for which $j \in \mathcal{C}$ should naturally be viewed as primary (or “active”). Those having $j \notin \mathcal{C}$ are best treated as auxiliaries.

**Theorem 4.8C.** Consider the situation of Theorem 4.8. Assume that (4.9) holds with $\alpha = 1$ and that the quantities $\Omega_j$ satisfy $\min(\Omega_1, \ldots, \Omega_k) > 1$. Let $v^\bullet = v/\min\{1, |v|\}$ (and $0^\bullet = 1$) as in Theorem 4.8A. We then have

$$|F(t_1, \ldots, t_k) - G(t_1, \ldots, t_k)|$$

(4.40)

$$\leq \hat{c}_1 \sum_{j} \int_{-\Omega_j}^{\Omega_j} \cdots \int_{-\Omega_j}^{\Omega_j} \|\varphi - \psi\|_\mathcal{C} \prod_{j \in \mathcal{C}} \frac{1}{\Omega_j} \prod_{j \in \mathcal{C}} \delta(v_j) d\vec{v} + c_2 \sum_{\ell=1}^{k} \frac{m_\ell}{\Omega_\ell}$$

at every point $(t_1, \ldots, t_k) \in \mathbb{R}^k$, wherein $\hat{c}_1$ and $c_2$ are certain [explicitly computable] positive constants that depend solely on $k$. The constant $c_2$ is identical to the one that occurs in (4.12).
Proof. The basic idea is very simple: one merely repeats the proof of Theorem 4.8 up to (4.20) with \( \omega_j \) in place of \( \Omega_j \), and then, after making a trivial “switch over” to \( \bar{v} = 2\pi \bar{w} \) and \((\omega_j, F_j)\), applies (4.25), (4.27), and (4.36) in a natural way to bound the value of the associated d\( \bar{v} \)-integral on a slab-by-slab basis.

The essential steps in the argument go as follows. Let \( T_{BDE_1 \varepsilon_2} \) signify the updated integrand in (4.20), including any trivial \( Q \)-coefficients. Clearly

\[
T_{BDE_1 \varepsilon_2} = \hat{A} \prod_{j \in B} \delta(v_j) \prod_{j \in D} \frac{1}{F_j} R_s \left( \frac{v_j}{F_j} \right) \prod_{j \in B \cup D} e^{-iv_j t_j} \prod_{j \in \varepsilon_1} e^{-iv_j t_j} \prod_{j \in \varepsilon_2} \frac{D_{\varepsilon_1} (\varphi - \psi) (\bar{v})}{\prod_{j \in \varepsilon_1} (iv_j)} \prod_{j \in \varepsilon_2} \left( \frac{\sin(v_j t_j)}{v_j} \right),
\]

where \( \hat{A} \) is some real constant depending solely on \( k, B, D, \varepsilon_1, \varepsilon_2 \) and whether the RHS of (4.41) is attached to either the majorant or minorant track. Note that \( T_s \) is Hermitian w.r.t. \( \bar{v} \); cf. §4.4. After fixing any number \( \eta \) in \((0,1/k] \), say, 1/k, we set \( |v|_* = \max\{ |v|, |v|^{-\eta} \} \) and form the quantity

\[
Q \equiv \hat{C} \sum_{\varphi} \int_{-F_1}^{F_1} \cdots \int_{-F_k}^{F_k} \frac{|\varphi - \psi| e^v}{|v|_*} \prod_{j \in \varepsilon_1} \frac{1}{F_j} \prod_{j \in \varepsilon_2} \delta(v_j) d\bar{v},
\]

wherein \( \varphi' = \langle B', \varepsilon', D' \rangle \) à la §4.7, and \( \hat{C} \) is some sufficiently large constant depending solely on \( k \).

Let \( S \) be any slab. The plan is to show that the real number \( \int_S T_{BDE_1 \varepsilon_2} d\bar{v} \) splits into at most \( 2^k \) summands, each of which has absolute value majorized by some summand of \( Q \) (restricted, of course, to \( S \)). No specific relation is implied here between \( \{ B, D, \varepsilon_1, \varepsilon_2 \} \) and \( \{ B', D', \varepsilon' \} \). Since

\[
|\varphi - \psi| e^v \leq \| \varphi - \psi \| e^v \quad \text{and} \quad \int_{-1}^{1} \frac{1}{|v|_*} dv < +\infty,
\]

once this inclusion property is shown, estimate (4.40) follows immediately.

With \( S \) and \( \{ B, D, \varepsilon_1, \varepsilon_2 \} \) now viewed as fixed, there is obviously no loss of generality if we assume that things have been relabelled so that:

(i) the portion of \( \varepsilon_2 \) “having” \( |v_j| < 1 \) w.r.t. \( S \) is \( \{ 1, 2, \ldots, N \} \);

(ii) \( \varepsilon_1 = \{ N + 1, \ldots, m \} \);

(iii) \( \varepsilon_2 - \{ 1, \ldots, N \} = \{ m + 1, \ldots, M \} \).

The integers \( N, m, M \) are adapted to any empty sets in the obvious way.

Letting \( \mathfrak{h} = \varphi - \psi \), we next write \( \mathfrak{h} = I \mathfrak{h} \), where

\[
I = \prod_{j=1}^{N} (P_j + D_j + E_j \Delta_j)
\]

à la (4.25). This decomposes the function \( \mathfrak{h} \) into \( 3^N \) summands of the form \( W \mathfrak{h} \), wherein each \( W \) is a “word” in (i.e., element of) \( \prod_{j=1}^{N} \{ P_j, D_j, E_j \Delta_j \} \). The number \( \int_S T_{BDE_1 \varepsilon_2} d\bar{v} \) (= \( T \mathcal{S} \), for short) decomposes similarly. We stress here that \( \varepsilon_1 \cap \varepsilon_2 = \phi \).
Suppose for a moment that $W$ contains some $D_i$. The function $D_{c_1}(W\mathcal{h})$ is odd w.r.t. $v_i$. On the other hand, $\sin(v_it_i)/v_i$ is even. Since $S$ is symmetric, the corresponding “chunk” of $TS$, viewed as an iterated integral (cf. (4.41)), is therefore 0. In other words: we need only look at $W \in \prod_{j=1}^{N} \{ P_j, E_j \Delta_j \}$.

At this point, one is finally free to take, w.l.o.g.,

$$W = (\prod_{j=1}^{n} E_j \Delta_j) \prod_{j=n+1}^{N} P_j \quad \text{and} \quad D_{c_1} W\mathcal{h} = \prod_{j=n+1}^{N} P_j (D_{c_1} \prod_{j=1}^{n} E_j \Delta_j)\mathcal{h}.$$  

Referring back to (4.41), one readily sees that the corresponding summand of $TS$ is real and factors into

$$I_N[n+1] \prod_{j=n+1}^{N} \left( \int_{-1}^{1} \frac{\sin(t_j v_j)}{v_j} dv_j \right),$$

wherein (4.36) applies and $I_N[n+1]$ is a certain iterated integral in the variables \{ $v_j : j \in [1,k]$, $j \notin [n+1,N]$ \} having a “level parameter” \( (v_{n+1}, \ldots, v_N) \) that has been set equal to $(0, \ldots, 0)$. The integrand in this more general expression is simply (4.41) with its \{ $n+1 \leq j \leq N$ \} (sine quotient) terms replaced by 1 and its numerator entry $D_{c_1} (\varphi - \psi)$ replaced by

$$D_{c_1} (E_1 \Delta_1) \cdots (E_n \Delta_n)\mathcal{h}. \quad (4.43)$$

The interpretation of \( (v_{n+1}, \ldots, v_N) \) as a level parameter makes eminently good sense near $(0, \ldots, 0)$ and fits with §4.6. The $C_2$ entries “still standing” in the integrand have value

$$\prod_{j=1}^{n} \left( \frac{\sin(t_j v_j)}{v_j} \right) \prod_{j=n+1}^{M} \left( \frac{\sin(t_j v_j)}{v_j} \right).$$

Note that, in the second half of this product, one has $|v_j| > 1$; cf. (iii).

With $W$ fixed, we now form the absolute value of our $TS$-summand and apply relations (4.41), (4.36), (4.43), and (4.27) with $\ell = n$ and $h = 1$ (or possibly 0). See also (4.24). After reviewing definition (4.38), dividing $C_1$ into two chunks (à la s/b), and doing a bit of slow but elementary bookkeeping, one finds an obvious inclusion w.r.t. $Q$ over $S$, wherein

$$B' = B \cup \{n + 1, \ldots, N\}, \quad D' = D, \quad C' = C_1 \cup \{1, \ldots, n\} \cup \{m + 1, \ldots, M\}.$$ 

The key point is that $L \leq k$ in (4.27); note too that use of $|h|\mathcal{v}$ in (4.42) actually corresponds to making a convenient overshoot in the $N_{\alpha}$-terms.

This completes the proof of (4.40) with $F_j$ in place of $\Omega_j$. 

By restricting one’s attention to those $T$, having $C_2 = \phi$, it is immediately seen that the $B'D'C'$-process “visits” every portion of $Q$ as $S$ varies. It also goes without saying here that the “1” appearing in $C_{\mathcal{h}}(\mathcal{v}), C_{\mathcal{c}}(\mathcal{v}), C_{\mathcal{e}}(\mathcal{v})$, etc., can be replaced by any other positive constant $\tau$, as need be.

4.14. Just as in the case of $\mathbb{R}^1$, the single most common application “of all this” is a result that ensues nearly immediately from Corollary 4.12; viz., the central limit theorem ([Cr, p.112]) for a sum of $N$ identically distributed, inde-
pendent, vector-valued random variables $\vec{X}_j$.

One simply looks at $\varphi_n(\vec{v}) = \varphi(\vec{v}/\sqrt{n})^n$, where $\varphi$ is the characteristic function of the common distribution function $F(\vec{x})$. We assume herein that

$$\int_{\mathbb{R}^k} \vec{x} dF(\vec{x}) = \vec{0} \quad \text{and} \quad \text{rank} \left( \int_{\mathbb{R}^k} x_j x_\ell dF \right) = k.$$  \hspace{1cm} (4.44)

(Cf. (4.10) and [Cr, p.109 (145)].) The key technical ingredient in verifying hypothesis (ii) in Corollary 4.12 is the rudimentary fact that

$$e^{it} = 1 + it + \frac{1}{2}(it)^2 + O(|t|^2) \min\{1, |t|\} \quad \text{for} \quad t \in \mathbb{R}.$$  

In condition (i), one naturally takes $\alpha$ equal to 2.

4.15. From the standpoint of modern probability theory, estimates like (4.12), (4.29), (4.35), and (4.40) [with genesis in Lemma 4.3] have a structural style that is slightly dated. In §5.13, we’ll cite a few references that help to connect things to the more recent literature - and to smoothing ideas of other sorts.

§5. More on The Central Limit Theorem

5.1. With its rich history and multitude of formulations, the central limit theorem ([Cr, Fel1, Fel2, Usp, Po2]) is aptly regarded as one of the main results in classical probability theory.

In the first part of this section, we’ll use Corollary 4.12 to quickly establish two [low-level] variants of the C.L.T. having particular relevance for our later work with logarithms of $L$-functions. To keep matters simple, we’ll formulate things in a Liapounov style and restrict ourselves to contexts where, after rescaling, the underlying sequence of independent random variables $\{X_j\}_{j=1}^\infty$ has a common distribution function $F(z)$ on $\mathbb{C}$ ($\cong \mathbb{R}^2$).

5.2. It is helpful to begin by recalling several very basic inequalities. For this purpose, let $n \geq 0$, $N \geq 1$, $m \geq 2$, $\lambda \geq 1$, and $x_j \in [0, \infty)$. Put

$$s_N = (x_1^2 + \cdots + x_N^2)^{1/2}, \quad B_N = \max\{x_1, \ldots, x_N\},$$

and let $\Theta$ signify a complex number (not always the same) having modulus no bigger than 1. We then have:

$$e^{it} = \sum_{k=0}^n \frac{(it)^k}{k!} + \Theta \frac{|t|^{n+1}}{(n+1)!} \quad \text{for} \quad t \in \mathbb{R} \hspace{1cm} (5.1)$$

$$\log(1 + z) = z + \Theta|z|^2 \quad \text{for} \quad |z| \leq \frac{1}{2} \hspace{1cm} (5.2)$$

$$B_m \leq \left( \sum_{j=1}^m x_j \right)^{\frac{1}{\lambda}} \leq \sum_{j=1}^m x_j \leq m^{1-\frac{1}{\lambda}} \left( \sum_{j=1}^m x_j^\lambda \right)^{\frac{1}{\lambda}} \leq m B_m \hspace{1cm} (5.3)$$
Relation (5.1) follows from the identity
\[ e^{it} = 1 + \int_0^t ie^{iu} du \]
by a trivial induction. At the same time, one also sees that
\[ e^{it} = \sum_{k=0}^n \frac{(it)^k}{k!} + \Theta \frac{2^{1-\omega}|t|^{n+\omega}}{(1)(1+\omega)...(n+\omega)} \]
for any \( \omega \in [0,1) \). The remainder term in (5.1) thus admits the alternate format
\[ \Theta \min \left\{ \frac{|t|^{n+1}}{(n+1)!}, \frac{2|t|^n}{n!} \right\}. \]  

Notice too that, in relation (5.3), the first and last terms simply correspond to taking \( \lambda = \infty \) for each fixed \( m \).

5.3. Let \( \{X_j\}_{j=1}^\infty \) now be any sequence of complex-valued, mutually independent random variables having a common distribution function \( F(z) \) on \( \mathbb{C} \). Assume that \( F \) satisfies
\[ \int_{\mathbb{C}} |z|^3 dF(z) = \rho^3 < \infty \]  
in addition to
\[ \int_{\mathbb{C}} z dF = 0, \quad \int_{\mathbb{C}} |z|^2 dF = 2\beta^2, \quad \int_{\mathbb{C}} z^2 dF = 0 \quad \text{with} \quad \beta > 0 . \]  
(Cf. Remark 5.10 concerning the \( z^2 \)-integral.) Put \( z = x + iy \) as usual and then define the characteristic function of \( F \) by writing
\[ \varphi(\xi) = \int_{\mathbb{C}} \exp[i \text{Re}(\xi z)] dF(z) \quad \text{for} \quad \xi \in \mathbb{C} . \]  
Taking \( \xi = v_1 + iv_2 \) gives the standard \( \mathbb{R}^2 \)-version. Similarly for any other distribution function \( F_{arb} \) on \( \mathbb{C} \). The RHS of (5.8) is, of course, nothing other than \( \mathbb{E}[\exp(i \text{Re}(\xi Y))] \) with \( Y = X_j \).

**Theorem 5.4.** With \( F \) and \( X_j \) as above, let \( \{b_j\}_{j=1}^\infty \) be any sequence of complex numbers such that \( b_j \neq 0 \) and
\[ \lim_{N \to \infty} \sum_{j=1}^N \frac{|b_j|^3}{s_N^3} = 0, \]  
wherein \( s_N = (|b_1|^2 + \cdots + |b_N|^2)^{1/2} \). For \( t \in \mathbb{R} \), let \( A_t = (-\infty, t] \). In the limit of large \( N \), the random variable
\[ T_N \equiv \frac{1}{s_N} \sum_{j=1}^N b_j X_j \]  
...
will then become distributed like a complex Gaussian; i.e.,

\[
\Pr\left\{ T_N \in A_x \times A_y \right\} \to \left( \frac{1}{\beta \sqrt{2\pi}} \right)^2 \int_x^y \int_x^y \exp \left( -\frac{u^2 + v^2}{2\beta^2} \right) \, dv \, du \quad (5.10)
\]

for every \((x, y) \in \mathbb{R}^2\). The convergence will be uniform w.r.t. \(x\) and \(y\).

**Proof.** By Hölder’s inequality, \(E(|T_N|) \leq (2\beta^2)^{1/2} \leq \rho\). In view of Corollary 4.12, it suffices to check that one has

\[
\varphi_N(\xi) \Rightarrow \exp\left( -\frac{1}{2\beta^2}|\xi|^2 \right)
\]

on every finite disk \(|\xi| \leq A\). Here \(\varphi_N\) is the characteristic function of \(T_N\). Since the \(X_j\) are independent, one immediately gets

\[
\varphi_N(\xi) = \prod_{j=1}^N \varphi(\mathcal{U}_j), \quad \mathcal{U}_j \equiv \frac{b_j}{s_N} \xi.
\]

Using (5.1) and conditions (5.6)+(5.7), we quickly see that

\[
\varphi_N(\xi) = \prod_{j=1}^N \left( 1 - \frac{1}{2}\beta^2 \text{Re}^2(\mathcal{U}_j) - \frac{1}{2}\beta^2 \text{Im}^2(\mathcal{U}_j) + \frac{1}{6} \Theta \rho^3 |\mathcal{U}_j|^3 \right)
\]

\[
= \prod_{j=1}^N \left( 1 - \frac{1}{2}\beta^2 |\mathcal{U}_j|^2 + \frac{1}{6} \Theta \rho^3 |\mathcal{U}_j|^3 \right). \quad (5.11)
\]

Letting \(B_N = \max\{|b_1|, \ldots, |b_N|\}\), one also knows by (5.4) that

\[
\frac{B_N}{s_N} \to 0 \quad (5.9')
\]

is equivalent to (5.9). With this point noted, we now keep \(N\) large enough to have, say,

\[
4\beta^2 \left( \frac{AB_N}{s_N} \right)^2 + 2 \sum_{j=1}^N \left( \rho \frac{A|b_j|}{s_N} \right)^3 < 1.
\]

Bearing in mind that \(|\mathcal{U}_j| \leq AB_N/s_N\), for some branch of the logarithm we then get:

\[
\log \varphi_N(\xi) = \sum_{j=1}^N \left( -\frac{1}{2}\beta^2 |\mathcal{U}_j|^2 + \frac{1}{6} \Theta \rho^3 |\mathcal{U}_j|^3 \right)
\]

\[
+ \sum_{j=1}^N \Theta \left( \frac{1}{2}\beta^2 |\mathcal{U}_j|^2 + \frac{1}{6} \rho^3 |\mathcal{U}_j|^3 \right)^2
\]

\[
= -\frac{1}{2}\beta^2 \sum_{j=1}^N |\mathcal{U}_j|^2 + \frac{1}{6} \Theta \sum_{j=1}^N \rho^3 |\mathcal{U}_j|^3
\]
\[
\frac{1}{2} \sum_{j=1}^{N} (\beta \mathcal{U}_j)^4 + \frac{1}{18} \Theta \sum_{j=1}^{N} (\rho^3 \mathcal{U}_j)^2
\]

\[
= -\frac{1}{2} \beta^2 \sum_{j=1}^{N} |\mathcal{U}_j|^2 + \frac{1}{6} \Theta \sum_{j=1}^{N} \rho^3 |\mathcal{U}_j|^3
\]

\[
+ \frac{1}{4} \Theta \sum_{j=1}^{N} (\beta^3 |\mathcal{U}_j|^3 + \rho^3 |\mathcal{U}_j|^3) \quad \{\beta \leq \rho\}
\]

\[
= -\frac{1}{2} \beta^2 \sum_{j=1}^{N} |\mathcal{U}_j|^2 + \frac{2}{3} \Theta \sum_{j=1}^{N} \rho^3 |\mathcal{U}_j|^3
\]

\[
= \frac{1}{2} \beta^2 |\xi|^2 + \frac{2}{3} \Theta \rho^3 A^3 \sum_{j=1}^{N} \frac{|b_j|^3}{s_N^3}.
\]

Upon exponentiating and letting \(N \to \infty\), the desired convergence of \(\varphi_N(\xi)\) follows at once. ■

Prior to continuing, it is convenient to pause for several comments concerning the \(\mathbb{R}\)–analog of Theorem 5.4. To this end, let \(E(x)\) be any probability distribution on \(\mathbb{R}\) satisfying an analog of (5.6)+(5.7) with respective values \(\{\rho^{3}; 0, \beta^{2}, \beta^{2}\}\). Let \(\{(X_j, Y_j)\}_{j=1}^{\infty}\) be any i.i.d. sequence of random variables on \(\mathbb{R}^2\) associated with the product measure \(dE \times dE\). The corresponding \(F(z)\) clearly satisfies (5.6)+(5.7). The following result is now immediate either by imitating the proof of Theorem 5.4 or simply putting \(y = \infty\) in (5.10).

**Corollary 5.5.** (The \(\mathbb{R}\)–analog of Theorem 5.4.) Given \(E\) and independent random variables \(\{X_j\}_{j=1}^{\infty}\) as above. Let \(\{b_j\}_{j=1}^{\infty}\) be any sequence of real numbers \((\neq 0)\) satisfying condition (5.9). Write

\[
T_N \equiv \frac{1}{s_N} \sum_{j=1}^{N} b_j X_j.
\]

In the limit of large \(N\), we then have

\[
\Pr\{T_N \leq x\} \to \frac{1}{\beta \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2\beta^2}\right) du
\]

for every \(x \in \mathbb{R}\). The convergence will be uniform w.r.t. \(x\).

Corollary 5.5 dates back to Liapounov (1901); cf. [Lia]. Either of the indicated proofs immediately adapts to encompass the nominally broader framework of \((2 + \delta)\)-moments and non-identically distributed, \(\beta^2\)-normalized \(X_j\) that one traditionally uses in stating this result. Cf. estimate (5.5) and [Usp, p.284], [Cr, §VI.4], [Bil, p.371(27.16)]. For effective versions of Corollary 5.5,

---

20The numerators in (5.9) are now modified to include appropriate \(\rho^{2+\delta}\) factors arising from the measures \(dE_j\). (Cf. the many \(\rho^3 |\mathcal{U}_j|^3\) terms in the proof of Theorem 5.4.)
see, e.g., [Ess, p.43] or [Fel2, p.544]. A quick look at [Fel3, theorem 1] is also illuminating.

**Remark 5.6.** When the variables $X_j$ have their support restricted to a finite interval $[-\Delta, \Delta]$, it is also feasible to obtain Corollary 5.5 by a consideration of $T_N$ - moments. Cf. [Bil, pp.407 (example 30.1) – 410]. Note that, by (5.4), Billingsley’s limit hypothesis (30.5) basically coincides with (5.9).

Markov showed in 1913 that, with a bit of ingenuity in forming truncated variables, the moment method could be successfully recast (i.e., “partitioned”) to yield Corollary 5.5 in full generality. Cf. [Usp, pp.388–395] and [Bil, p.600 (30.1)]. An analogous splitting will come up later in connection with logarithms of Euler products.

5.7. In our second C.L.T. variant, we employ a kind of vector “twist” to “lift” the Gaussian of Theorem 5.4 up to a related one on $C^J$. The twist is induced by replacing the earlier coefficients $b_n$ (note the change of index $j \mapsto n$) with a string of appropriately-behaved column vectors $\tilde{b}_n \in C^J$.

To keep the new Gaussian as unencumbered as possible, we normalize things by insisting that

$$\lim_{N \to \infty} \frac{1}{\sigma_N} \sum_{n=1}^{N} (b_{n,j})(\tilde{b}_{n,\ell}) = [\beta_j \delta_{j\ell}]$$

(5.12)

hold in the sense of matrix addition for certain $\beta_j > 0$ and certain monotonically increasing positive scaling factors $\sigma_N$. [Observe that each $J \times J$ summand in (5.12) is nonnegative-definite; the thought of transforming (5.12) under a unitary change of basis is thus present here virtually ab initio. Notice too that each summand is preserved under the “local” action $\tilde{b}_n \mapsto \exp(i\omega_n)\tilde{b}_n$.]

**Theorem 5.8.** Let $\{X_n\}_{n=1}^{\infty}$ be any sequence of complex-valued independent random variables having a common distribution function $F(z)$, where $F$ satisfies both (5.6) and (5.7). Let $\{b_{n,j}, \sigma_N, \beta_j\}$ be as in (5.12). Assume that

$$\lim_{N \to \infty} \sum_{n=1}^{N} \sum_{j=1}^{J} \left|\frac{b_{n,j}}{\sigma_N^3}\right|^3 = 0.$$  

(5.13)

In the limit of large $N$, the $C^J$ - valued random variable

$$T_N = \frac{1}{\sigma_N} \sum_{n=1}^{N} X_n(b_{n,j})$$

then becomes distributed like a multivariate Gaussian

$$\int_{E} \prod_{j=1}^{J} \frac{1}{2\pi \beta_j \beta^2} \exp \left( -\frac{|w_j|^2}{2\beta_j \beta^2} \right) \prod_{j=1}^{J} du_j dv_j$$

(5.14)

wherein $E \subseteq C^J$ and $w_j \equiv u_j + iv_j$ corresponds to the $j^{th}$ component of $T_N$. 

More on The Central Limit Theorem
Proof. One simply modifies the approach used over $\mathbb{C}^1$. We need to check [in an obvious notation] that

$$\varphi_N(\zeta_1, \kappa_1, \zeta_2, \kappa_2, \ldots, \zeta_J, \kappa_J) \equiv \prod_{j=1}^{J} \exp\left(-\frac{1}{2} \beta_j \beta^2 (\zeta_j^2 + \kappa_j^2)\right)$$

holds on every compact subset of $\mathbb{R}^{2J}$. To verify this, it is natural to write $\xi_j = \zeta_j + i\kappa_j$ and then look at

$$\varphi_N(\xi_1, \ldots, \xi_J) = \mathbb{E}\left[\exp(i\sum_{j=1}^{J} \Re(\xi_j T_{N, j}))\right]$$

$$\varphi_N(\xi_1, \ldots, \xi_J) = \mathbb{E}\left[\exp(i\frac{1}{\sigma_N} \sum_{j=1}^{J} \sum_{n=1}^{N} \Re(\xi_j b_{n,j} X_n))\right]$$

on the multidisk $\{||\xi_1|| \leq A, \ldots, ||\xi_J|| \leq A\}$. Putting

$$\varphi(\xi) = \int_{\mathcal{C}} \exp[i\Re(\overline{\xi} z)]dF(z)$$

as in (5.8) and using the independence of $X_n$, we first see that

$$\varphi_N(\xi_1, \ldots, \xi_J) = \prod_{n=1}^{N} \varphi(U_n), \quad U_n = \frac{1}{\sigma_N} \sum_{j=1}^{J} b_{n,j} \xi_j$$

From (5.11), we then get

$$\varphi_N(\xi_1, \ldots, \xi_J) = \prod_{n=1}^{N} \left(1 - \frac{1}{2} \beta^2 ||U_n||^2 + \frac{1}{6} \Theta \rho^3 ||U_n||^3\right)$$

By applying (5.3) with $x_j = |w_j|$, one immediately checks that

$$\frac{1}{J} \leq \frac{||\overline{w}||_{\infty}}{||\overline{w}||_1} \leq \frac{||\overline{w}||_{\lambda}}{||\overline{w}||_1} \leq 1$$

(5.15)

for any $\overline{w} \in \mathbb{C}^J$ and $1 \leq \lambda < \infty$. Temporarily fixing any $\mu \in [1, \infty)$, we now write $C_n = ||\overline{b}_n||_\mu$ and avail ourselves of (5.4) + (5.15) to see that

$$\frac{D_N}{\sigma_N} \to 0$$

(5.13')

with $D_N = \max\{C_1, \ldots, C_N\}$ is equivalent to (5.13). Cf. (5.9').

At this point, we set $\mu = 1$, so that $||U_n|| \leq AC_n/\sigma_N$, and then keep $N$ large enough to have

$$4\beta^2 \left(\frac{AD_N}{\sigma_N}\right)^2 + 2 \sum_{n=1}^{N} \left(\rho \frac{AC_n}{\sigma_N}\right)^3 < 1.$$

For some branch of the logarithm, we finally get
More on The Central Limit Theorem

\[
\log \varphi_N(\xi_1, \ldots, \xi_J) = \sum_{n=1}^{N} \left( -\frac{1}{2} \beta^2 |U_n|^2 + \frac{1}{6} \Theta \rho^3 |U_n|^3 \right) \\
+ \sum_{n=1}^{N} \Theta \left( \frac{1}{2} \beta^2 |U_n|^2 + \frac{1}{6} \rho^3 |U_n|^3 \right)^2 \\
= -\frac{1}{2} \beta^2 \sum_{n=1}^{N} |U_n|^2 + \frac{2}{3} \Theta \sum_{n=1}^{N} \rho^3 |U_n|^3 \\
= -\frac{1}{2} \beta^2 \sum_{n=1}^{N} |U_n|^2 + \frac{2}{3} \Theta \rho^3 A^3 \sum_{n=1}^{N} \frac{C^3}{\sigma_N^3} 
\]

exactly as in the proof of Theorem 5.4. Since \( C_n \leq J^{2/3} \| \vec{b}_n \|_3 \) by (5.3), and

\[
\frac{1}{2} \beta^2 \sum_{n=1}^{N} |U_n|^2 = \frac{1}{2} \beta^2 \sum_{j=1}^{J} \sum_{t=1}^{J} \xi_j \xi_t \left( \frac{1}{\sigma_N} \sum_{n=1}^{N} b_{n,j} b_{n,t} \right),
\]

the desired convergence of \( \varphi_N(\xi_1, \ldots, \xi_J) \) follows at once. ■

In situations where the coefficients \( b_{n,j} \) are real and the probability distribution \( E(x) \) is as in Corollary 5.5, a temporary passage to \( dF = dE \times dE \) in Theorem 5.8 immediately shows that

\[
S_N = \frac{1}{\sigma_N} \sum_{n=1}^{N} X_n(b_{n,j})
\]

becomes Gaussian distributed like

\[
\int_{E} \prod_{j=1}^{J} \frac{1}{\beta \sqrt{2\pi \beta_j}} \exp \left( -\frac{u_j^2}{2\beta_j \beta^2} \right) \prod_{j=1}^{J} du_j (5.16)
\]
in the limit of large \( N \) over \( \mathbb{R}^J \). The case \( J = 1 \) reduces to a minor extension of Corollary 5.5.

**Remark 5.9.** Theorem 5.8 is clearly a bit more “exotic” than Theorem 5.4. It is important to note, however, that both results are easy corollaries of the (now standard) Lindeberg formulation of the C.L.T. over \( \mathbb{R}^{2J} \); condition (5.7) enables the key link from \( \mathbb{C}^J \) to \( \mathbb{R}^{2J} \). Cf. [Cr, §§X.3, VI.3–4] and, for instance, [Fel2, pp.260–261, 263 (lines 5–6, 11–16)]. A quick look at the change-of-basis idea utilized in [Ber, pp.44–48 (top)] is also helpful. Seen from this perspective, Theorem 5.8 (with \( J > 1 \)) is simply a form of the C.L.T. in which the successive random variable summands have their supports restricted to “thin” sets.

One very slight advantage to our “pedestrian level” proofs of Theorems 5.4 and 5.8 is that, when viewed in conjunction with Theorems 4.8A–4.8C, matters are explicit enough therein to furnish what would seem to be a convenient springboard for the development of some parallel results having well-controlled error terms, at least for box-like \( E^{[21]} \). Cf. [Fel2, pp.544–545].

\[\text{21} \text{In our later work with } L\text{-functions, } dF \text{ will just be ordinary Haar measure on } \{|z| = 1\}.\]
Remark 5.10. Coming back to a point just touched on [concerning (5.7)], suppose for a few moments that, beyond having mean 0, distribution function \( F(z) \) merely satisfies \( E(|z|^q) < \infty \) with some \( q \geq 3 \). (Cf. (5.6).) Let \( \| \cdot \| \) denote the Euclidean norm on \( \mathbb{C}^J \). In formulating Theorem 5.8, we have been guided by a desire to remain in a structural setting where:

(i) matters are robust enough to produce a good limit distribution for \( T_N \) anytime (5.12) holds and \( \max\{\|b_1\|, \ldots, \|b_N\|\}/\sigma_N = o(1) \); and,

(ii) in the process of that, every \( \mathbb{R}^2J \)-style second moment of \( T_N \) is seen to converge to some limit as \( N \to \infty \).

The essential point in (i) is that convergence should still occur even if the vectors \( b_n \) are multiplied by arbitrary phase factors \( \exp(i\omega_n) \).

In connection with (ii), thanks to the algebraic identity

\[
\left(\begin{array}{cc}
x_1 x_2 & x_1 y_2 \\
y_1 x_2 & y_1 y_2
\end{array}\right) = \frac{1}{2} \left(\begin{array}{cc}1 & 1 \\
-i & i
\end{array}\right) \left(\begin{array}{cc}z_1 \overline{z}_2 & \overline{z}_1 z_2 \\
z_1 \overline{z}_2 & \overline{z}_1 \overline{z}_2
\end{array}\right) \frac{1}{2} \left(\begin{array}{cc}1 & i \\
n & -i
\end{array}\right),
\]

we certainly want, e.g. for \( J = 1 \), the limiting covariance

\[
\mathcal{M} = \lim_{N \to \infty} \left(\begin{array}{cc}
E(T_N \overline{T}_N) & E(T_N^2) \\
E(\overline{T}_N^2) & E(T_N \overline{T}_N)
\end{array}\right)
\]
to exist. Taking \( |b_{n,1}| \equiv 1 \) and \( \sigma_N = \sqrt{N} \) produces

\[
\mathcal{M} = \lim_{N \to \infty} \left(\begin{array}{cc}
A & \frac{B}{N} C_N \\
\frac{B}{N} C_N & A
\end{array}\right),
\]

with

\[
A = \int_{\mathbb{C}} |z|^2 dF(z), \quad B = \int_{\mathbb{C}} z^2 dF(z), \quad C_N = \sum_{n=1}^N \|b_{n,1}\|^2.
\]

Putting \( A = 2\beta^2 > 0 \) w.l.o.g., and considering the effect of \( e^{i\omega_n} \)-twists in \( \mathcal{M} \), it is nearly self-evident that the coefficient \( B \) can only be 0. Similarly for \( J > 1 \). For this reason: the normalization (5.7) is basically forced on us in the present setting.

Further corroboration of this point ensues from the simple observation that, when \( q \geq 4 \) and assumption (5.12) holds, convergence of \( T_N \) in distribution automatically carries with it fulfillment of property (ii). One checks this by bounding the expectations of \( \text{Re}(T_{N,j})^4 \) and \( \text{Im}(T_{N,j})^4 \) for each \( j \), and then exploiting the comment made immediately after the proof of Corollary 4.12. Since \( \int z dF = 0 \), the only partitions of 4 that contribute nontrivially here are 4 and \( 2 + 2 \); the aforementioned expectations are thus majorized by

\[
\text{constant} \frac{\sum_{n=1}^N \|b_n\|^2}{\sigma_N^4} \left(\sum_{n=1}^N \|\overline{b}_n\|^2\right)^2.
\]
Compare: [Bil, pp.409(middle)–410(top)]. By making use of the more sophisticated Marcinkiewicz-Zygmund inequality (see [MZ1, §8] or [MZ2, §3]) and the convexity of $\Theta(u) = u^3$ for $\beta > 1$, the foregoing observation is readily seen to hold for $2 < q < 4$ as well.

5.11. To wrap things up, it is now helpful to “make a slight change of gears.” Once one has available a reasonable $\mathbb{R}^k$-analogue of Esseen’s lemma, it becomes natural to contemplate finding $k$-variable counterparts for not only the classical Berry-Esseen-type error estimates in the C.L.T., but also the related refinements of uniform asymptotic type traditionally associated with the names of Edgeworth and Cramér. See [Cr, §X.3 (para 1)] for some early remarks on both issues, albeit from the standpoint of a slightly different type of smoothing, viz., [Cr, §§VII.3, VII.5–6].

Though this report is clearly not the place for any sort of comprehensive discussion of these matters (the technical details being rather heavy even in the case of Theorems 5.4 and 5.8), it does seem reasonable to offer at least a few remarks on what can be achieved using Theorems 4.8A–4.8C in the classical C.L.T. setting of §4.14. [The ready availability of a pre-existing literature here facilitates preparation of an intelligible commentary within the span of just a few pages.]

5.12. In the i.i.d. situation of §4.14, one has

$$\bar{X}_n = \frac{1}{\sqrt{n}}(\bar{X}_1 + \cdots + \bar{X}_n) \quad \text{and} \quad \psi(\bar{v}) = \exp\left(-\frac{1}{2} \sum v_j a_{j\ell} v_{\ell}\right),$$

wherein $a_{j\ell} = \mathbb{E}(x_j x_\ell)$; cf. (4.44). The characteristic function $\psi(\bar{v})$ corresponds to that of a $k$-variate Gaussian ([Cr, p.109(144)]) having covariance matrix $[a_{j\ell}]$. Anytime the initial probability distribution $F(\bar{x})$ of §4.14 satisfies

$$\int_{\mathbb{R}^k} \left(\|\bar{x}\|_\infty\right)^q dF < \infty \quad (5.17)$$

with an appropriately big $q \geq 3$, there is a concomitant expectation that, after a modicum of calculation and parameter optimization, Theorems 4.8A–4.8C will each be found capable of producing quantitative refinements in the C.L.T. analogous to those known to hold in the case of one-variable.

In particular, insofar as the characteristic function $\varphi$ of $dF$ satisfies

$$\limsup_{\|\bar{v}\|\to\infty} |\varphi(\bar{v})| < 1 \quad (5.18)$$

[i.e., Cramér’s condition (C)], it should emerge rather quickly that there exist asymptotic developments of Edgeworth-type over $\mathbb{R}^k$ comparable to those articulated in, say, [Cr, §VII.5], [Ess, pp.48–52], and [Fel2, pp.539, 541(bottom)] for $\mathbb{R}^1$. To help one appreciate the more formal side of things, a quick review of [Cr, pp.26(lemma 1), 71(lemma 2), 74(lemma 3), 75(102), 81(C),

[22] In this connection, see also [Po2, pp.172 (lines 4–18), 177 (para 2)].
Gaussians, Zeros, and Linear Combinations [Part A]

85, 86(101a), [Ess, p.44(lemma 2)], and [vBa, pp.72–74(top)] is very useful. The paper by von Bahr deals specifically with $\mathbb{R}^k$ and has the added advantage of addressing derivatives of $\varphi(v/\sqrt{n})^n$ as well; cf. his Lemmas 2 and 3.

Just as in the case of $\mathbb{R}^1$, there are no real difficulties implementing these technical formalities in the overall calculation until it comes time to bound the final error term. The essential question then shifts to how small a size can be achieved in the latter through an appropriate choice of parameters. The outcome in this will naturally depend on the particular type of smoothing inequality that is available.

Theorem 2 in [vBa] shows that, under hypothesis (5.18), matters behave precisely as one would expect given the analogous results over $\mathbb{R}^1$. von Bahr’s theorem serves to refine an earlier estimate (of Berry-Esseen type) obtained around twenty years earlier by H. Bergström; see [Berg1, pp.109, 121(top)] and [Berg2, p.40(D)].

In perusing this work, what becomes evident nearly immediately is that our Theorem 4.8 (with its $D_C$-terms and related sum over $\mathcal{P}$) manifests a certain structural similarity with [vBa, pp.77(8), 75(lemma 4)]. Likewise for Theorem 4.8C. In proving Theorem 4.8C, we have clearly been influenced by several aspects of this parallelism. (Cf., in particular, [vBa, pp.76(7), 77(line 14)]. The decomposition of $I$ towards the middle of our proof can be seen as a natural counterpart of p.76(7).)

With these remarks as a backdrop, it should come as no big surprise that applying Theorem 4.8C in conjunction with [vBa, Lemma 3] is readily checked to lead to a result essentially identical to von Bahr’s Theorem 2. One simply takes

$$\Omega_j = \frac{1}{\sqrt{k}}(\sqrt{n})^{q-1}$$

as in [vBa, p.77(lines 17–18)] with $q \geq 3$, observes that

$$0 \leq |t|^n \exp(-\frac{1}{2}\beta|t|^{\frac{2}{\psi-1}}) \leq c_1(n, \beta, q)$$

$$0 < c_2(k, \psi) \leq \frac{\left(\sum_{j=1}^{k}|t_j|^n\right)^\psi}{\sum_{j=1}^{k}|t_j|^\psi} \leq c_3(k, \psi) < \infty$$

for every $n \geq 0$, $\beta > 0$, $\psi > 0$, and then focuses on estimating the $d\tilde{v}$-integrals in (4.40) one partition at a time, aided by von Bahr’s Lemma 3. Writing

$$\mathcal{C} = \{1, 2, ..., \delta\}, \quad \mathcal{D} = \{\delta + 1, ..., r\}, \quad \mathcal{B} = \{r + 1, ..., k\},$$

w.l.o.g., it is helpful to collapse the region of integration down to

$$\prod_{j=1}^{r-1}[\Omega_j, \Omega_j] \times \prod_{j=r+1}^{k}(v_j = 0)$$

Note that the $\alpha$ in Lemma 3 depends on the specifics of how (5.18) is fulfilled.

It is illuminating at this juncture to also have a look at [Rao, p.360, theorem 2] and the less polished, partial refinement obtained by [Dun, Sad] around 1966–68 utilizing Sadi-kova’s slightly nonlinear, but highly suggestive, two-dimensional variant of (3.1).

Readers for whom reference [vBa] is new may find it helpful to view this latter assertion in the more specific setting of [vBa, pp.74(line 5)(hT)(5), 75(top), 76(bottom), 77(lines 1-10, 13-15, 18-19)]. Compare: [Berr, p.127(28)], [Fel2, pp.537(3.3)(3.6), 538(3.12)].
and then introduce slabs for \( j \in [1, \delta] \). In view of (4.39), there is no need to look at anything worse than a first derivative in Lemma 3. Since \( |v_j^*| \geq 1 \) and
\[
\int_{-\Omega_j}^{\Omega_j} \frac{1}{\Omega_j} dv_j = 2, \quad \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\beta|t|^{\frac{q}{q-1}}\right) dt < \infty,
\]
one immediately recovers Theorem 2 of [vBa].

Utilization of Theorem 4.8B in place of 4.8C clearly produces a counterpart of Theorem 2 having an extra factor of size at most \( c_{10}(D+1)^k \) in the final error term. (With a little more effort, \( (D+1)^k \) can be replaced by \( \log^{k-1}(D + e) \); cf. (5.21)+(5.22) below.)

In the case of Theorem 4.8A, matters are slightly more involved due to the need to select appropriate values for both \( \alpha \) and \( \Delta \) in (4.29). As will soon become apparent, however, the choice of \( \alpha \in (0, q] \) is largely immaterial – in that any modifications therein lead to effects felt only at the level of the implied constant in the final error term. (That \( E(\|T_n\|^\alpha) = O(1) \) for every \( \alpha \) in \( [2, q] \) is an easy consequence of the Marcinkiewicz-Zygmund inequality; cf. the last four lines of Remark 5.10. Setting \( \alpha = 2 \) is, of course, simplest.)

To bound (4.29), one starts by taking \( \Omega_j \) as above and letting \( \gamma = 2q - 1 \).

We then note (using (5.20)) that
\[
\int_{-\Omega_1}^{\Omega_1} \cdots \int_{-\Omega_k}^{\Omega_k} \frac{\|\vec{v}\|^q \exp(-\beta \|\vec{v}\|^{\gamma})}{|v_1^*| \cdots |v_k^*|} d\vec{v} \leq c_4(k, q) I(q) I(0)^{k-1}, \tag{5.21}
\]
wherein \( \|\vec{v}\| \) is the standard Euclidean norm, \( \Delta > 1 \),
\[
I(m) = \int_{0}^{\infty} v^m \exp\left(-\beta v^{\gamma}\right) dv \quad \text{for} \quad m \geq 0,
\]
and \( \hat{\beta} \) is some elementary fraction of \( \beta \). Since \( 1/v^* = \min(\Delta, 1/v) \) for \( v > 0 \), it is natural to set \( v = w/\Delta \) in \( I(m) \). By elementary calculus, this promptly gives
\[
I(m) = O(1)[1 + \delta m_0(\log \Delta)], \tag{5.22}
\]
with an implied constant depending solely on \( (m, \beta, \gamma, k) \). Taking account of Lemma 3 and page 77 (lines 17–22) in [vBa], things are now optimized (modulo unimportant constants) by declaring \( \Delta = 10n^{(q-1)/2\alpha} \). The resulting error term matches that of von Bahr’s Theorem 2 apart from an extra factor of \( (\log n)^{k-1} \). That is to say, with Theorem 4.8A, one ultimately obtains
\[
|\text{Remainder}| \leq C d(n) \left(\frac{1}{\sqrt{n}}\right) q^{-2} (\log n)^{k-1},
\]
with some \( d(n) = o(1) \). Exploitation of the final sentence in Theorem 4.8A produces the slightly weaker bound \( C d(n)(1/\sqrt{n}) q^{-2-\nu} \) with
\[
\nu = k \frac{q-2}{q+k} \in (0, k].
\]
Both error estimates seem eminently reasonable given the relative crudity of Theorem 4.8A. (The second bound is of interest mainly for large $q$. In both cases, note that $C$ and $d(n)$ typically depend on $\{k, q, F\}$; cf. (5.22).)

As a quick review of the foregoing considerations makes clear, there are undoubtedly very similar expansions that hold in the setting of Theorems 5.4 and 5.8 – at least under hypothesis (5.18) and some suitable magnitude restrictions on both $\|b_n\|$ and the remainder term in (5.12). Our subsequent work with $L$-functions will implicitly entail working out the details of at least one such case; see footnote 21. (Notice incidentally that the measure in footnote 21 has $\varphi(\xi) = J_0(|\xi|)$ for $\xi \in \mathbb{C}$; cf., e.g., [Fel2, p.523 (7.8)]. In (5.18), $|\varphi(\xi)|$ will thus decay like $1/\sqrt{|\xi|}$.)

5.13. Modern treatments of multivariate Edgeworth-type expansions have largely come to focus on regions and set-ups substantially more general than rectangles. See, for instance, [BhR, pp.208, 214(20.44)–215; 52–54, 210–212]; also [vBa, pp.85(line 5)–87(middle)].

In connection with such results, it is helpful to keep two key facts in mind. First: that in contexts where the relevant shapes $E$ are smoothly bounded and overshoots in $q$ are a non-issue, the uniformity aspect of §5.11–12 will invariably permit one to obtain some results of Edgeworth-type simply by making inner/outer approximations based on rectangular grids and then optimizing the choice of grid size(s) only at the very end. Second: though Theorems 4.8–4.8C have intentionally been kept quite close to (3.1) format-wise, it has been found increasingly expedient over the years for work involving multivariate normal approximation to make use of smoothing inequalities having less rigid, convolution-based formats. See [Ess, pp.101, 104(53)(56), 110(lines 5–9)], [Berg2, p.47(lemma 8)], [vBa, pp.79(9), 81(11)], [Saz, pp.184(lemma 2), 195(lemma 6)], and [BhR, pp.94(11.13), 98(11.26), 102(top), 210(top), 84–88, 274–278] for a few examples, in chronological order. Pages 274–278 in [BhR] relate to the much-discussed Stein method of establishing Berry-Esseen type bounds.

A concomitant look at [CGS, pp.336(12.66)–337(top), 4–6, 16(2.13)–18(top), 20(2.30), 46–48] and [Bar, pp.293–294] serves to provide some valuable additional perspective on these more recent matters.

§6. Acknowledgments

I thank Yumi Karlsson for her assistance in LaTEXing several earlier drafts of this report. The first of these dates back to 2010, while I was still active at Uppsala University in Sweden. More recently, I am grateful to the Institute for Advanced Study in Princeton for providing me with excellent working conditions and a supplemental research grant (from the Bell Companies) during the Spring Term of 2012, when work on the current, improved, version was being done.

---

26e.g., when $dF$ is compactly supported
27(lines 5–9 are nicely supplemented by [vBa2, pp.63(8), 64(lines 12–19), 67(C)–68])
28(Compare 266 lines 12–19, and note the missing $*K$ on 276 line 7.)
References

[Ach] N. I. Achieser, *Theory of Approximation*, F. Ungar Publishing Co., New York, 1956, especially pp.179(top), 201(1), 203(top), 204(4), 205(top). Cf. also 7(top), 177, 183(top half), 188–189(3), 190(bottom), and 191(middle)–194.

[Bar] A. D. Barbour, Asymptotic expansions based on smooth functions in the central limit theorem, *Prob. Theor. Rel. Fields* **72** (1986), 289–303.

[Berg1] H. Bergström, On the central limit theorem in the space $R_k$, $k > 1$, *Skand. Aktuarietidskr.* **28** (1945), 106–127.

[Berg2] H. Bergström, On the central limit theorem in the case of not equally distributed random variables, *Skand. Aktuarietidskr.* **32** (1949), 37–62.

[Ber] S. Bernstein, Sur l’extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes, *Math. Ann.* **97** (1927), 1–59.

[Berr] A. Berry, The accuracy of the Gaussian approximation to the sum of independent variates, *Trans. Amer. Math. Soc.* **49** (1941), 122–136.

[Beu] A. Beurling, On functions with a spectral gap (Uppsala University Seminar, March 24, 1942), in *The Collected Works of Arne Beurling*, vol. 2, Birkäuser, Boston, 1989, pp.370–372.

[BhR] R. N. Bhattacharya and R. Ranga Rao, *Normal Approximation and Asymptotic Expansions*, Updated edition, Classics in Applied Mathematics No. 64, S.I.A.M., Philadelphia, 2010.

[Bil] P. Billingsley, *Probability and Measure*, 2nd edition, John Wiley, New York, 1986.

[Boa] R. P. Boas, *Entire Functions*, Academic Press, New York, 1954.

[BH] E. Bombieri and D. A. Hejhal, On the distribution of zeros of linear combinations of Euler products, *Duke Math. J.* **80** (1995), 821–862.

[CGS] L. H. Y. Chen, L. Goldstein, and Q. Shao, *Normal Approximation by Stein’s Method*, Springer-Verlag, Berlin/Heidelberg, 2011.

[Cr] H. Cramér, *Random Variables and Probability Distributions*, Cambridge University Press, 1937.

[DL] D. L. Donoho and B. F. Logan, Signal recovery and the large sieve, *SIAM J. Appl. Math.* **52** (1992), 577–591, especially 582(middle)–587 (top).

[Dun] J. E. A. Dunnage, On Sadikova’s method in the central limit theorem, *J. London Math. Soc.* (2) **2** (1970), 49–59. (See also: *Proc. London Math. Soc.* (3) **20** (1970), 33–59.)
REFERENCES

[Erd] A. Erdélyi, et al., *Higher Transcendental Functions*, vol. 1, McGraw-Hill Book Co., New York, 1953.

[Ess] C.-G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law; *Acta Math.* 77 (1945), 1–125.

[Fel1] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 2nd edition, John Wiley, New York, 1957.

[Fel2] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, 2nd edition, John Wiley, New York, 1971.

[Fel3] W. Feller, On the Berry-Esseen theorem, *Z. Wahrscheinlichkeitstheorie verw. Geb.* 10 (1968), 261–268.

[Gau] C. F. Gauss, *Werke*, Band 8, Königliche Gesellschaft der Wissenschaften zu Göttingen, 1900. (Note that the assertions on p.88 date back to 1799~1813.)

[Har] G. H. Hardy, Notes on special systems of orthogonal functions (IV): the orthogonal functions of Whittaker’s cardinal series, *Proc. Camb. Philo. Soc.* 37 (1941), 331–348.

[Hig1] J. R. Higgins, Five short stories about the cardinal series, *Bull. Amer. Math. Soc.* (N.S.) 12 (1985), 45–89.

[Hig2] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis: Foundations*, Oxford Univ. Press, 1996.

[Jor] C. Jordan, *Calculus of Finite Differences*, 3rd edition, Chelsea Publishing Co., New York, 1965.

[KL] J.-P. Kahane and P.-G. Lemarié-Rieusset, Remarques sur la formule sommatoire de Poisson, *Studia Math.* 109 (1994), 303–316.

[Kat] Y. Katznelson, *An Introduction to Harmonic Analysis*, John Wiley, New York, 1968.

[KeS] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, vol.1 (Distribution Theory), 4th edition, Charles Griffin & Company, London, 1977.

[Lev] B. Ya. Levin, *Lectures on Entire Functions*, Transl. Math. Monographs, Vol. 150, American Math. Society, Providence, 1996.

[Lia] A. M. Liapounov, Nouvelle forme du théorème sur la limite de probabilité, *Mém. Acad. Sci. St. Pétersbourg* (8) 12 (1901), No. 5, 1–24. (See also: *Bull. Acad. Sci. St. Pétersbourg* (5) 13 (1900), No. 4, 359–386.)

[Log] B. F. Logan, Bandlimited functions bounded below over an interval, *Notices Amer. Math. Soc.* 24 (1977), A-331.

[Malm] C. J. Malmstén, Sur la formule $hu' = \Delta u - (h/2)\Delta u' + (1/2!)B_1h^2\Delta u'' - (1/4!)B_2h^4\Delta u^{(iv)} + \text{etc.}$, *J. Reine Angew. Math.* 35 (1847), 55–82. (See also the improved version in *Acta Math.* 5 (1884), 1–46.)

[MZ1] J. Marcinkiewicz and A. Zygmund, Sur les fonctions indépendantes, *Fund. Math.* 29 (1937), 60-90.
REFERENCES

[MZ2] J. Marcinkiewicz and A. Zygmund, Quelques théorèmes sur les fonctions indépendantes, *Studia Math.* 7 (1938), 104–120.

[PaW] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Amer. Math. Soc. Colloquium Publications Vol. 19, New York, 1934.

[PIP] M. Plancherel and G. Pólya, Fonctions entières et intégrales de Fourier multiples (II), *Comment. Math. Helv.* 10 (1937/38), 110–163.

[Po1] G. Pólya, Über die Nullstellen gewisser ganzer Funktionen, *Math. Zeit.* 2 (1918), 352–383.

[Po2] G. Pólya, Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem, *Math. Zeit.* 8 (1920), 171–181.

[PSz] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 1, Julius Springer, Berlin, 1925.

[Rao] R. Ranga Rao, On the central limit theorem in $R_k$, *Bull. Amer. Math. Soc.* 67 (1961), 359–361.

[Sad] S. M. Sadikova, Two-dimensional analogues of an inequality of Esseen with applications to the central limit theorem, *Theor. Prob. Appl.* 11 (1966), 325–335. (See also: 13 (1968) pp.165(2.3), 166–167 and 12 (1967), 343.)

[Saz] V. V. Sazonov, On the multi-dimensional central limit theorem, *Sankhyā*, Ser. A 30 (1968), 181–204.

[Sel] A. Selberg, Old and new conjectures and results about a class of Dirichlet series, in *Proceedings of the Amalfi Conference on Analytic Number Theory*, Sept. 1989 (E. Bombieri, et al., eds.), Università di Salerno, 1992, pp.367–385. Cf. also [Sel2, pp.47–63].

[Sel2] A. Selberg, *Collected Papers*, vol.2, Springer-Verlag, Berlin/Heidelberg, 1991.

[Ste] J. F. Steffensen, *Interpolation*, 2nd edition, Chelsea Publishing Co., New York, 1950.

[Tim] A. F. Timan, *Theory of Approximation of Functions of a Real Variable*, Pergamon Press/MacMillan, New York, 1963, especially §§5.1.3–5.1.4, 5.6.35, 5.11 (problems 4, 11, 12), 8.7(problems 23–25); cf. also §§9.21, 2.6.3, 2.11.3(10), 2.12.6, 3.3(para 2), and 3.5.3.

[Usp] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill Book Co., New York, 1937.

[Vaa] J. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc.* 12 (1985), 183–216.

[Val] G. Valiron, Sur la formule d’interpolation de Lagrange, *Bull. Sci. Mathematiques* (2) 49 (1925), 181–192, 203–224.

[vBa] B. von Bahr, Multi-dimensional integral limit theorems, *Arkiv för Matematik* 7 (1967), 71–88.

[vBa2] B. von Bahr, On the central limit theorem in $R_k$, *Arkiv för Matematik* 7 (1967), 61–69.
[Z1] A. Zygmund, *Trigonometric Series*, 2nd edition, vol. 1, Cambridge Univ. Press, 1959.

[Z2] A. Zygmund, *Trigonometric Series*, 2nd edition, vol. 2, Cambridge Univ. Press, 1959.

School of Mathematics
University of Minnesota
Minneapolis, MN. 55455 USA

E-mail address: hejhal@math.umn.edu