Topics in Probability and Random Processes

Michael Taylor

Contents

Introduction

These notes treat selected topics in probability theory and the theory of random processes, more or less related to variants of the theory of Brownian motion. They can serve as a follow-up to an introductory account of such material, taking off from such background in probability theory as is given in the following three chapters of [MEAS]:

Chapter 14. Ergodic theory,
Chapter 15. Probability spaces and random variables,
Chapter 16. Wiener measure and Brownian motion.

These notes also make use of basic material on Fourier analysis and functional analysis, such as can be found, for example, in the early sections of the following chapter and appendix of [PDE]:

Chapter 3. Fourier analysis, distributions, and constant-coefficient linear PDE,
Appendix A. Outline of functional analysis.

The following chapter from [PDE], though not needed as background for these notes, would nevertheless illuminate various subjects treated here:

Chapter 11. Brownian motion and potential theory.

Chapter 1 of this text treats the central limit theorem. The basic result is that if $\{f_j\}$ is a sequence of independent, identically distributed random variables on a probability space $(\Omega, \mathcal{F}, \mu)$, satisfying

$$(1) \qquad \|f_j\|^2_{L^2(\Omega)} = \sigma \in (0, \infty), \quad \int_\Omega f_j \, d\mu = 0,$$

then the sequence of sums, suitably rescaled, has probability distributions that converge to a Gaussian distribution on $\mathbb{R}$.

In more detail, a random variable g on Ω induces a probability measure ν_g on $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, by

$$(2) \qquad \nu_g(S) = \mu(g^{-1}(S)),$$

where $S \subset \widehat{\mathbb{R}}$ is a Borel set. If (1) holds, we set

$$(3) \qquad g_k = \frac{1}{\sqrt{k}} \sum_{j=1}^{k} f_j,$$

and conclude that

$$\nu_{g_k} \longrightarrow \gamma^\sigma, \quad \text{weak}^* \text{ in } \mathcal{M}(\widehat{\mathbb{R}}), \tag{4}$$

where γ^σ is the Gaussian probability measure, satisfying

$$\langle f, \gamma^\sigma \rangle = \frac{1}{\sqrt{2\pi\sigma}} \int\limits_{\mathbb{R}} f(x) e^{-x^2/2\sigma}\, dx, \quad f \in C(\widehat{\mathbb{R}}). \tag{5}$$

The rescaling factor $1/\sqrt{k}$ in (3) arises so that $\|g_k\|^2_{L^2(\Omega)} \equiv \sigma$.

Going further, we look for various quantitative improvements of (4), including associated estimates on the rate of convergence of the distribution functions

$$\Phi_k(y) = \nu_{g_k}((-\infty, y]), \quad G(y) = \gamma^\sigma((-\infty, y]), \tag{6}$$

of the form

$$\sup_y |\Phi_k(y) - G(y)| \le C k^{-r/2}, \tag{7}$$

under various hypotheses on f_j, typically involving the characteristic function

$$\chi_{f_j}(\xi) = \int\limits_\Omega e^{-i\xi f_j}\, d\mu. \tag{8}$$

One guide to this involves a rather detailed analysis of the coin toss, in §2 of this chapter. A particularly interesting case of (7) is the Berry-Esseen theorem, treated in §5.

In §8 we produce a variant of the CLT, applicable to sequences of IID random variables whose suitably scaled sums converge in distribution to certain non-Gaussian distributions, arising in the study of fractional diffusions, a topic that is taken up in detail in Chapters 3 and 7.

Appendix A to this chapter discusses an interesting functional analytic aspect of passing from (4), which says

$$\int f(x)\, d\nu_{g_k}(x) \longrightarrow \int f(x)\, d\gamma^\sigma(x), \tag{9}$$

for $f \in C(\widehat{\mathbb{R}})$, to $\Phi_k(y) \to G(y)$, for $y \in \mathbb{R}$, which says (9) holds when $f(x) = 1$ for $x \le y$, 0 for $x > y$. We establish a general result that if (4) holds, then (9) holds for each bounded Borel function $f : \widehat{\mathbb{R}} \to \mathbb{R}$ that is Riemann integrable on $\widehat{\mathbb{R}} \approx S^1$. Further aspects of this are pursued in Appendix B.

In Chapter 1 we have concentrated on sequences of real-valued random variables. Later on we will deal with $\mathbb{R}^n$-valued random variables. Regarding the CLT in this more general context, see Appendix A of Chapter 5.

Chapter 2 deals with stochastic operators and infinite dimensional versions of the Perron-Frobenius theorem. The basic setup concerns bounded linear operators

$$A : C(X) \longrightarrow C(X), \tag{10}$$

where X is a compact Hausdorff space. We say A is positive if

$$f \in C(X),\ f \geq 0 \Longrightarrow Af \geq 0. \tag{11}$$

A positive operator A is said to be a stochastic operator if

$$A1 = 1. \tag{12}$$

A bounded operator A has the adjoint

$$A^t : \mathcal{M}(X) \longrightarrow \mathcal{M}(X), \tag{13}$$

and if A is positive we have $A^t : \mathcal{M}_+(X) \to \mathcal{M}_+(X)$, the space of positive, finite, regular Borel measures on X. If A is stochastic, we have

$$A^t : \mathcal{P}(X) \longrightarrow \mathcal{P}(X), \tag{14}$$

the space of regular probability measures on X. We show that if A is a stochastic operator on $C(X)$, there exists $\mu \in \mathcal{P}(X)$ such that

$$A^t \mu = \mu. \tag{15}$$

One Perron-Frobenius type theorem is the following:

Proposition 2.A. *Let A be a compact, stochastic operator, and assume A is primitive, i.e., some power A^m is strictly positive. Then, as $k \to \infty$,*

$$A^k \longrightarrow P, \quad \textit{and} \ \ (A^t)^k \longrightarrow P^t, \tag{16}$$

in operator norm, where P is the projection of $C(X)$ onto Span(1) *that annihilates*

$$V = \{f \in C(X) : \langle f, \mu \rangle = 0\}. \tag{17}$$

Note that

$$Pf = \langle f, \mu \rangle 1, \quad P^t \lambda = \langle 1, \lambda \rangle \mu, \tag{18}$$

for $f \in C(X),\ \lambda \in \mathcal{M}(X)$.

This leads to the following result.

Proposition 2.B. *Let A be a compact stochastic operator on $C(X)$, and assume A is irreducible, i.e.,*

$$B = \sum_{k=1}^{\infty} 2^{-k} A^k \quad \textit{is strictly positive.} \tag{19}$$

Then the measure μ in (15) is unique.

The proof involves observing that, by Proposition 2.A, $(B^t)^k \to P^t$, given by (18).

In §3 we consider variants, such as crypto-stochastic operators and operators that are crypto-stochastic up to scaling.

In the classical Perron-Frobenius theorems, A is a real $n \times n$ matrix, so we are in the setting (10) with X a set with n elements. Section 4 addresses the limiting case of infinite matrices, yielding

$$A : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N}). \tag{20}$$

This can be converted to the setting of (1), with X being the Stone-Cech compactification of $\mathbb{N}$, or equivalently the maximal ideal space of the C^*-algebra $\ell^\infty(\mathbb{N})$.

Chapter 3 studies Lévy processes. These are variants of the Wiener process, which models Brownian motion. For Brownian motion on $\mathbb{R}^n$, we take

$$p(t,x) = (2\pi)^{-n} \int e^{-t|\xi|^2} e^{ix\cdot\xi}\, d\xi = (4\pi t)^{-n/2} e^{-|x|^2/4t}. \tag{21}$$

One desires to specify a probability measure on the space of paths in $\mathbb{R}^n$, having the following property. Given $0 < t_1 < t_2$ and given that $\omega(t_1) = x_1$, the probability density for the location of $\omega(t_2)$ is $p(t_2 - t_1, x - x_1)$. More generally, given $0 < t_1 < \cdots < t_k$, and given Borel sets $E_j \subset \mathbb{R}^n$, the probability that a path, starting at x_0 at time $t = 0$, lies in E_j at time t_j for each $j \in \{1, \dots, k\}$ is

$$\int_{E_1} \cdots \int_{E_k} p(t_k - t_{k-1}, x_k - x_{k-1}) \cdots p(t_1, x_1)\, dx_k \cdots dx_1. \tag{22}$$

One elegant approach to the construction of Wiener measure is given in [Nel].

P. Lévy initiated extensions of this theory to non-Gaussian distributions. In this chapter we will adapt the method of [Nel] to treat these Lévy processes. To start, we replace (21) by

$$p(t,x) = (2\pi)^{-n} \int e^{-t\psi(\xi)} e^{ix\cdot\xi}\, d\xi = e^{-t\psi(D)} \delta(x). \tag{23}$$

We take $\psi(\xi)$ to have the property that

$$p(t,x) \ge 0, \quad \forall\, t > 0,\ x \in \mathbb{R}^n, \tag{24}$$

and we require $\psi(0) = 0$, so $\int p(t,x)\,dx \equiv 1$. The example $\psi(\xi) = |\xi|^2$ gives the Gaussian case, in (21). Other examples include

$$\begin{aligned} \psi_\alpha(\xi) &= |\xi|^{2\alpha}, \\ \varphi_\alpha(\xi) &= (|\xi|^2+1)^\alpha - 1, \quad \alpha \in (0,1), \end{aligned} \tag{25}$$

as seen in §1 of Chapter 3. For $\alpha \in (0,1)$, the operators $-\psi_\alpha(D)$ generate fractional diffusions, mentioned in §8 of Chapter 1 as giving rise to variants of CLT. Another family of examples is

$$\psi(\xi) = 1 - e^{-iy\xi}, \tag{25A}$$

generating Poisson processes. In this case,

$$e^{-t\psi(D)}\delta(x) = \sum_{k=0}^{\infty} \frac{t^k}{k!} e^{-t} \delta_{ky} \tag{25B}$$

is a family of probability measures (not absolutely continuous with respect to Lebesgue measure). There is a general formula, the Lévy-Khinchin formula, for functions $\psi(\xi)$ such that (24) holds, discussed in Appendix A of Chapter 3, with complements in Appendix B.

In §2 of Chapter 3 we construct a probability measure on the "path space"

$$\mathfrak{P} = \prod_{t\in\mathbb{Q}^+} \dot{\mathbb{R}}^n, \tag{26}$$

where $\dot{\mathbb{R}}^n = \mathbb{R}^n \cup \{\infty\}$. This differs from [Nel], which took the Cartesian product over $t \in [0,\infty)$. Using the countable Cartesian product involves more elementary measure theory. We define

$$E : \mathcal{C}^\# \longrightarrow \mathbb{R}, \tag{27}$$

where $\mathcal{C}^\#$ consists of continuous functions on $\mathfrak{P}$ of the form

$$\varphi(\omega) = F(\omega(t_1),\dots,\omega(t_k)), \quad t_1 < \cdots < t_k, \tag{28}$$

with F continuous on $\prod_1^k \dot{\mathbb{R}}^n$ and $t_j \in \mathbb{Q}^+$, by a formula parallel to (22). We verify that E is a positive linear functional on the linear space $\mathcal{C}^\# \subset C(\mathfrak{P})$, satisfying $E(1) = 1$. Furthermore, $\mathcal{C}^\#$ satisfies the conditions of the Stone-Weierstrass theorem, so (27) has a unique extension to a continuous, positive linear function $E : C(\mathfrak{P}) \to \mathbb{R}$, giving rise to the desired probability measure.

We can define $X_t : \mathfrak{P} \to \dot{\mathbb{R}}^n$ for $t \in \mathbb{Q}^+$ by $X_t(\omega) = \omega(t)$, and a calculation gives, for $0 \le s < t,\ s,t \in \mathbb{Q}^+$,

$$E(|X_t - X_s|^q) = \int p(t-s,y)|y|^q\,dy. \tag{29}$$

If, for example, $\psi(\xi) = \psi_\alpha(\xi)$, as in (25), one obtains

$$E(|X_t - X_s|^q) = C_{n\alpha q}|t-s|^{q/2\alpha}, \quad -n < q < 2\alpha,\ 0 < \alpha < 1. \tag{30}$$

In particular, we have, for $t \in \mathbb{Q}^+$,

$$X_t \in L^q(\mathfrak{P}), \quad 0 < q < 2\alpha,\ 0 < \alpha < 1, \tag{31}$$

and by (30) we can extend this uniquely to $t \in \mathbb{R}^+$, depending continuously on $t \in \mathbb{R}^+$. (For $0 < q < 1$, the space $L^q(\mathfrak{P})$ is not a normed space, but it is a complete metric space.) We have $\|X_t\|_{L^q} = \infty$ for $q \ge 2\alpha$ when $\psi = \psi_\alpha$. By contrast, when $\psi(\xi) = \varphi_\alpha(\xi)$ in (25), we get

$$\begin{aligned} E(|X_t - X_s|^2) &= -\Delta e^{-|t-s|\varphi_\alpha(\xi)}\Big|_{\xi=0} \\ &= |t-s|\Delta\varphi_\alpha(0), \end{aligned} \tag{30A}$$

and, more generally, for $k \in \mathbb{N}$,

$$\begin{aligned} E(|X_t - X_s|^{2k}) &= (-\Delta)^k e^{-|t-s|\varphi_\alpha(\xi)}\Big|_{\xi=0} \\ &= -|t-s|(-\Delta)^k\varphi_\alpha(0) + O(|t-s|^2), \end{aligned} \tag{30B}$$

as $|t-s| \to 0$. Such results as (30), (30A), and (30B) express "stochastic continuity" of X_t, when $\psi = \psi_\alpha$ or φ_α as in (25). The study of stochastic continuity is extended to more general Lévy processes in §3.

For the Wiener process, almost all paths are continuous (more on which in Chapter 4), but other Lévy processes do not have this property. This is apparent for the Poisson processes, since by (25B) the only allowed moves are jumps, but also such processes as those generated by $-\psi_\alpha(D)$ and $-\varphi_\alpha(D)$, as in (25), do not have continuous paths. The probability distributions $e^{-t\psi_\alpha(D)}\delta(x)$ have "heavy tails," examined in Appendix C of Chapter 3. Further examination of the short and long time behavior of $e^{-t\psi(D)}\delta(x)$ is carried out in Appendix D, and other qualitative studies of these probability distributions are made in Appendices E and F.

Appendix M of Chapter 3 extends the scope of this chapter, beyond the study of random processes on Euclidean space $\mathbb{R}^n$ to other classes of manifolds. Appendix N goes further, exploring a class of Markov processes

$$e^{tA} : C(X) \longrightarrow C(X), \tag{32}$$

satisfying

$$e^{tA}1 = 1, \quad f \in C(X),\ f \ge 0 \Rightarrow e^{tA}f \ge 0, \tag{33}$$

which makes contact with material on stochastic operators in Chapter 2.

Chapter 4 deals with the modulus of continuity of paths for the 1D Wiener process. This has the basic form

$$|X_t(\omega) - X_s(\omega)| \le M_1(\omega) h(|t-s|), \tag{34}$$

for $s, t \in [0,1]$, with $M_1(\omega) < \infty$ for a.e. $\omega \in \mathfrak{P}$, and

$$h(\delta) = \Bigl(\delta \log \frac{1}{\delta}\Bigr)^{1/2}, \tag{35}$$

for $0 < \delta \le 1/e$. We present a proof of this due to M. Pinsky, which makes use of a representation of $X_t(\omega)$ as a Haar series, for $t \in [0,1]$. This proof also gives the classical estimate

$$P(M_1(\omega) > \lambda) \le C e^{-a\lambda^2}, \tag{36}$$

for some $C, a \in (0,\infty)$. Going further, we tweak the technique to establish the following.

Proposition 4.A. *Take $K \in (0,\infty)$, and set $\delta = e^{-K}$. Then there exist $a = a(K) > 0$ and $C_j = C_j(K)$ such that for $t, s \in [0,1]$, $|t-s| \le \delta$,*

$$|X_t(\omega) - X_s(\omega)| \le A_K(\omega) h(|t-s|) + B_K(\omega)|t-s|, \tag{37}$$

with

$$P(A_K(\omega) \ge \lambda) \le C_1 e^{-K\lambda^2}, \quad P(B_K(\omega) \ge \lambda) \le C_2 e^{-a\lambda^2}. \tag{38}$$

Chapter 5 treats stochastic integrals arising in the setting of square integrable, n-dimensional Lévy processes. In more detail, we assume that (29) is finite for $q = 2$ and that

$$E(X_t) = 0, \tag{39}$$

which leads to

$$E(|X_t - X_s|^2) = A|t-s|, \quad A = \Delta\psi(0), \tag{40}$$

with ψ as in (23). We define the Wiener stochastic integral

$$I_t(f) = \int_0^t f(x)\, dX_s, \tag{41}$$

for $f : \mathbb{R}^+ \to M(n,\mathbb{R})$, and show that

$$\|I_t f\|^2_{L^2(\Omega)} \le A \|f\|^2_{L^2([0,t])}, \tag{42}$$

with equality if f is scalar. Going further, we discuss more general Ito-type stochastic integrals, such as

$$\int_0^t f(s, X_s)\, dX_s, \tag{43}$$

and bring in Ito formulas, which for the Wiener process, defined by (21), take the form

$$df(X_s) = f'(X_s)\, dX_s + f''(X_s)\, ds. \tag{44}$$

An appendix to Chapter 5 relates the large time behavior of the probability distribution of

$$\frac{1}{\sqrt{t}} X_t \tag{44}$$

to the CLT, when X_t is a square integrable Lévy process satisfying (39).

Chapter 6 deals with multidimensional random fields, of the form

$$Z : \mathbb{F}^n \longrightarrow L^2(\Omega, \mu), \tag{45}$$

where (Ω, μ) is a probability space and $\mathbb{F} = \mathbb{R}$ or $\mathbb{Z}$. (These are also known as random processes indexed by $\mathbb{F}^n$). We assume the map Z in (45) is continuous. To such a field we associate a map

$$F : \Omega \longrightarrow \mathcal{O}, \quad F(\xi)(x) = Z(x)(\xi), \tag{46}$$

where $\mathcal{O}$ consists of functions $\mathbb{F}^n \to \mathbb{R}$, and we endow $\mathcal{O}$ with the probability measure ν, given by $\nu(S) = \mu(F^{-1}(S))$. For example,

$$\begin{aligned} &\varphi(\eta) = \psi(\eta(x_1), \dots, \eta(x_k)) \\ &\Rightarrow \int\limits_{\mathcal{O}} \varphi \, d\nu = \langle \psi(Z(x_1), \dots, Z(x_k)) \rangle, \end{aligned} \tag{47}$$

where we set

$$\langle X \rangle = E(X) = \int\limits_{\Omega} X(\xi) \, d\mu(\xi). \tag{48}$$

We have $\mathbb{F}^n$ acting on $\mathcal{O}$ by

$$\tau_y \eta(x) = \eta(x + y), \quad x, y \in \mathbb{F}^n, \ \eta : \mathbb{F}^n \to \mathbb{R}. \tag{49}$$

We say the random field Z is stationary if τ_y preserves the probability measure ν, for each y. Note that stationarity implies

$$\langle \psi(Z(x_1 + y), \dots, Z(x_k + y)) \rangle = \langle \psi(Z(x_1), \dots, Z(x_k)) \rangle, \tag{50}$$

for all $x_j, y \in \mathbb{F}^n$. In particular,

$$\langle Z(x_1 + y) \rangle = \langle Z(x_1) \rangle, \quad \langle Z(x_1 + y) Z(x_2 + y) \rangle = \langle Z(x_1) Z(x_2) \rangle. \tag{51}$$

Note that if Z is stationary, then (51) implies

$$\|Z(x + y) - Z(x)\|^2_{L^2(\Omega)} = 2\|Z(0)\|^2_{L^2} - 2\langle Z(0) Z(y) \rangle. \tag{52}$$

If Z is stationary, we say Z is ergodic if the action $\{\tau_y : y \in \mathbb{F}^n\}$ on $(\mathcal{O}, \nu)$ is ergodic. In §2 we consider implications of the ergodic theorem for ergodic random fields. One useful tool is Proposition 2.2, which says that if $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ is continuous, then the action of $\{\tau_y : y \in \mathbb{R}^n\}$ on $L^1(\mathcal{O}, \nu)$ is strongly continuous. This allows us to apply classical ergodic theorems. Notable consequences are

$$\begin{aligned} \langle Z(x_1)\rangle &= \lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{B_R} \eta(x_1+y)\, dy, \\ \langle Z(x_1)Z(x_2)\rangle &= \lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{B_R} \eta(x_1+y)\eta(x_2+y)\, dy, \end{aligned} \tag{53}$$

for ν-almost all $\eta \in \mathcal{O}$, provided Z is stationary and ergodic.

Sections 3 and 4 discuss Gaussian fields, for which all finite linear combinations $\sum a_j Z(x_j) \in L^2(\Omega, \mu)$ are Gaussian random variables. Proposition 3.1 shows that a Gaussian field is stationary if and only if (51) holds. A key object associated to a stationary Gaussian field is the covariance function $C : \mathbb{F}^n \to \mathbb{R}$, satisfying

$$C(x_1 - x_2) = \langle (Z(x_1) - M)(Z(x_2) - M)\rangle, \quad M = \langle Z(x)\rangle \equiv \langle Z(0)\rangle. \tag{54}$$

In Theorem 3.3 – Corollary 3.6, it is shown that if $C : \mathbb{R}^n \to \mathbb{R}$ is even, continuous, and integrable, then there exists a stationary Gaussian field $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ satisfying (54) provided the Fourier transform satisfies $\widehat{C}(\xi) \geq 0$, for all $\xi \in \mathbb{R}^n$. Examples are given in (3.24)–(3.27) and in (3.32)–(3.33). In §4 it is shown that all these examples yield ergodic Gaussian fields, except (3.32), which does not. The positive results here are a consequence of Proposition 4.2, which shows that if $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ is a stationary Gaussian field, with continuous covariance, and if

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{B_R} |C(y)|\, dy = 0, \tag{55}$$

then Z is ergodic.

Section 5 expands the scope of the study of random fields to

$$Z : G \longrightarrow L^2(\Omega, \mu), \tag{56}$$

where G is a Lie group. There is a natural extension of the notion of stationarity, in which (51) is replaced by

$$\langle Z(gx_1)\rangle = \langle Z(x_1)\rangle, \quad \langle Z(gx_1)Z(gx_2)\rangle = \langle Z(x_1)Z(x_2)\rangle, \tag{57}$$

for $g, x_1, x_2 \in G$. The covariance function is now $C : G \to \mathbb{R}$, given by

$$C(x^{-1}y) = \langle Z(x)Z(y)\rangle - M^2, \quad M = \langle Z(x)\rangle. \tag{58}$$

We concentrate on the case that G is compact. The Fourier transform appearing in §3 is replaced by

$$C^\pi = \int_G C(x)\pi(x)\,dx, \quad Z^\pi = \int_G Z(x)\pi(x)\,dx, \tag{59}$$

known as the spectral data of Z. Here π runs over the irreducible unitary representations of G. In §6 we consider random fields on a compact homogeneous space $X = G/K$,

$$Y : X \longrightarrow L^2(\Omega, \mu). \tag{60}$$

Use of the natural projection $\gamma : G \to X$ yields $Z = Y \circ \gamma$, and the considerations of §5 apply. Section 7 studies the inverse problem of producing $Z(x)$ from spectral data.

In §8 we take a finite-dimensional vector space V and discuss V-valued random fields, first on a general homogeneous space X, then specializing to $X = \mathbb{R}^n$, with special attention to $V = \mathbb{R}^n$, i.e., to random vector fields. In §9 we discuss random divergence-free vector fields on $\mathbb{R}^n$.

In §10 we discuss generalized random fields on $\mathbb{R}^n$, which are distributions on $\mathbb{R}^n$ with values in $L^2(\Omega, \mu)$. We define stationary generalized random fields and develop some of their properties.

Chapter 6 has three appendices. Appendix A gives background on ergodic theorems, and Appendix B relates the criterion on the covariance function given in §4 to the behavior of its Fourier transform. Appendix C discusses the Fourier transform of a continuous stationary field, first on $\mathbb{T}^n$ (obtaining a special case of results of §5) and then on $\mathbb{R}^n$, where we need to regard $\widehat{Z}$ as a vector-valued tempered distribution.

In Chapter 7 we consider fractional diffusion equations, such as

$${}^c\partial_t^\beta u = -(-\Delta)^\alpha u, \quad t \geq 0, \quad u(0,x) = f(x), \tag{61}$$

where $\alpha, \beta \in (0,1)$, and ${}^c\partial_t^\beta$ is the Caputo fractional derivative, given by

$${}^c\partial_t^\beta v(t) = \frac{1}{\Gamma(1-\beta)} \int_0^t (t-s)^{-\beta} \partial_s v(s)\,ds, \tag{62}$$

for $\beta \in (0,1)$. More generally, we consider

$${}^c\partial_t^\beta u = -Au, \quad u(0) = f, \tag{63}$$

where A is a self adjoint operator with the property that e^{-tA} is a semigroup of stochastic operators, so

$$f \geq 0 \Rightarrow e^{-tA} f \geq 0, \quad \int e^{-tA} f(x)\,dx = \int f(x)\,dx. \tag{64}$$

The Caputo operator (62) is a variant of the Riemann-Liouville fractional derivative, better suited for initial value problems. Indeed, as seen in §3, applying the Laplace transform to (63) gives

$$(s^\beta + A)\mathcal{L}u(s) = s^{\beta-1} f, \tag{65}$$

yielding for the solution to (63) the formula

$$u(t) = S_\beta^t f = E_\beta(-t^\beta A) f, \tag{66}$$

where $E_\beta(z)$ is the Mittag-Leffler function

$$E_\beta(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\beta n + 1)}, \tag{67}$$

and $E_\beta(-t^\beta A)$ is defined via the spectral representation of A.

Section 2 is devoted to a brief treatment of the $\beta = 1$ limit of (63),

$$\partial_t u = -Au, \tag{68}$$

with $A = L^\alpha,\ \alpha \in (0,1)$, and, e.g., $L = -\Delta$, via subordination identities.

In Section 3 we tackle (63), obtaining (66) with

$$E_\beta(-s) = \int_0^\infty M_\beta(r) e^{-rs}\, dr, \quad s \geq 0, \tag{69}$$

where

$$\beta r^{-\beta-1} M_\beta(r^{-\beta}) = \Phi_{1,\beta}(r) \geq 0, \tag{70}$$

and $\Phi_{t,\beta}$ given by (2.1)–(2.7). This analysis yields

$$M_\beta(r) \geq 0, \quad \text{hence} \quad E_\beta(-s) \geq 0, \tag{71}$$

leading via (69)–(70) and (64) to

$$f \geq 0 \Rightarrow E(-t^\beta A) f \geq 0 \Rightarrow S_\beta^t f \geq 0. \tag{72}$$

We also have

$$\int S_\beta^t f(x)\, dx = \int f(x)\, dx. \tag{73}$$

We make a brief comment on an analogue of (63) for $\beta \in (1,2]$ in §4. Section 5 discusses operators $A = \psi(D)$ that lead to stochastic semigroups e^{-tA}, making contact with material on Lévy processes in Chapter 3.

In section 6 we consider systems of fractional diffusion-reaction equations, of the form

$$\frac{\partial u}{\partial t} = -Lu + X(u), \quad u(0) = f, \tag{74}$$

where L is an $\ell \times \ell$ diagonal matrix whose diagonal entries are operators A_j yielding stochastic semigroups e^{-tA_j}. Section 7 discusses numerical attacks on such systems, based on a splitting method and the FFT, when $A_j = \psi_j(D)$.

In §§9–11 we tackle fractional diffusion-reaction equations of the form

$${}^c\partial_t^\beta u = -Au + F(u), \quad u(0) = f. \tag{75}$$

Chapter 7 ends with 3 appendices. Appendix A has basic material on Riemann-Liouville and Caputo fractional derivatives. Appendix B considers finite-dimensional linear systems of fractional differential equations. Appendix C discusses the derivation of the power series (67) for the Mittag-Leffler function $E_\beta(z)$.

References

[MEAS] M. Taylor, Measure Theory and Integration, Amer. Math. Soc., Providence RI, 2006.

[PDE] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer, New York, 1996 (2nd ed., 2011, 3rd ed., 2023).

1. Varieties of central limit theorems

0. Introduction

An eye-opening and side-splitting book review, [F], recently raised the interesting question of just what hypotheses on a sequence of IID random variables are needed for the sequence to satisfy a central limit theorem. One answer to this question is that one gets different central limit theorems depending on the specific hypotheses put forth. Our goal here is to describe explicitly some of the varieties of central limit theorems that arise.

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

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

In such a case, the independence implies

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

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

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

The proof is simple:

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

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

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

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

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

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

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

Similarly,

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

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

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

Given f as above, the function

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

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

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

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

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

One computes

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

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

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

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

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

We note that

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

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

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

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

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

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

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

Note that

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

In §1 we prove the following version of CLT:

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

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

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

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

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

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

REMARK 1. The weak* convergence (0.19) means

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

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

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

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

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

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

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

In §2 we study the coin toss, for which

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

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

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

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

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

This is proved as a consequence of the result that

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

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

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

and also that

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

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

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

refining (0.28), and

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

refining (0.27).

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

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

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

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

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

we have

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

where

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

In (0.35), we take

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

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

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

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

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

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

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

as $k \to \infty$, where

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

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

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

An elementary argument gives

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

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

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

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

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

Such estimates were established by Liapunov.

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

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

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

we have

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

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

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

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

in which case

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

In §6 we estabish the following.

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

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

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

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

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

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

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

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

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

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

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

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

Then

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

This result is complemented by the following.

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

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

Furthermore, for such ρ,

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

Here,

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

In §8 we expand the scope of CLT beyond results on approximating Gaussians. We look at probability measures on $\mathbb{R}$ arising from fractional diffusion equations (considered further in Chapters 3 and 6):

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

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

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

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

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

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

Then

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

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

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

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

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

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

In Appendix C we show that if f_j are IID random variables satisfying (0.1) and g_k are as in (0.16), and if

$$\chi_{g_j}(\xi) = e^{-\sigma\xi^2/2 + \xi^2\beta(\xi)}, \quad |\beta(\xi)| \le \frac{\sigma}{2}, \ \text{for } |\xi| \le a, \tag{0.65}$$

and

$$|\beta(\xi)| \le b|\xi|^r, \quad \text{for} \ \ |\xi| \le a, \ r \in (0,1], \tag{0.66}$$

then certain mollifications

$$\phi(k^{-s/2}D)\nu_{g_k} \longrightarrow \gamma^\sigma, \quad \text{uniformly, as } k \to \infty, \tag{0.67}$$

provided $s < r/(r+2)$, with an estimate on the rate of convergence. This result complements results brought to bear to establish the Berry-Esseen theorem.

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

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

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

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

and define γ^σ *as in (0.11). Then*

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

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

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

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

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

Hence

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

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

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

Hence

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

with

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

Therefore,

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

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

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

An equivalent statement is that

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

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

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

we also have

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

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

We can strengthen the conclusion of Theorem 1.1, by using

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

In particular,

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

and we have from (1.11) that

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

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

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

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

2. Coin toss

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

$$(2.1) \qquad \Omega = \prod_{j\in\mathbb{N}} \{1, -1\},$$

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

$$(2.2) \qquad f_j(\omega_1, \omega_2, \omega_3, \dots) = \omega_j,$$

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

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

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

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

as in (1.3).

To analyze this, we set

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

so

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

hence

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

where we have set

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

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

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

where

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

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

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

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

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

with

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

We also have

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

for some $a > 0$. It follows that

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

and furthermore

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

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

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

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

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

Hence

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

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

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

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

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

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

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

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

and

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

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

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

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

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

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

hence

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

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

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

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

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

as $k \to \infty$, where

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

We deduce from (2.16) that

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

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

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

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

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

we have

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

so, by (2.17),

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

Next,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

with

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

hence

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

We deduce that, with

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

we have

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

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

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

This refines (2.20) to

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

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

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

We can deduce from Jacobi's formula,

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

that

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

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

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

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

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

where

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

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

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

It is hence tempting to take

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

In light of (2.53)–(2.54), we have

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

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

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

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

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

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

for $k \ge 2$.

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

Here we derive some estimates on the rate at which

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

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

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

This implies that

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

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

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

where

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

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

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

we can expand the scope of (3.5) to

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

To start the estimate of (3.1), we have

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

Now

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

and (3.5) (or (3.7)) implies

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

It follows that

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

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

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

Shrinking a if necessary, we also arrange that

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

so

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

We will make do with the estimate

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

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

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

and obtain the following result.

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

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

we have

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

where

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

Note that

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

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

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

We toss in the conditions

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

Now we have

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

and estimates arising in the proof of Proposition 3.1 imply

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

where

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

We also have

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

This gives the following.

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

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

Other ways to present the last term arise via the identities

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

the latter via

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

We now have the task of estimating

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

Here is one straightforward result.

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

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

Then

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

Proof. Clearly

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

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

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

This gives (3.32).

The following result is a useful extension of Proposition 3.3.

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

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

Then

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

Proof. Set

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

and note that

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

We have

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

so

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

and (3.36) follows.

4. Convergence of distribution functions – Liapunov estimates

In this section we study the rate of convergence of

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

as $k \to \infty$, where

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

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

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

and

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

Recall that

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

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

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

where

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

Proposition 3.2 is applicable, and we have

$$|\Phi_k(y) - G(y)| \le Ck^{-r/2}\mathcal{A}(v_y) + \mathcal{B}_k(v_y), \tag{4.8}$$

where

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

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

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

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

$$\begin{aligned} w_{y,h}(x) = \quad & 0, && \text{if} \;\; x \ge y, \\ & \tfrac{h-(x-y)}{h}, && \text{if} \;\; y - h \le x \le y, \\ & 1, && \text{if} \;\; x \le y - h. \end{aligned} \tag{4.11}$$

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

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

and

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

hence

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

and

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

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

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

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

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

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

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

Hence (4.16) yields

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

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

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

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

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

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

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

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

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

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

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

In common with (4.11), we have

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

and

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

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

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

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

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

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

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

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

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

and we get the following extension of Proposition 4.1.

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

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

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

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

we have

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

One interesting corollary arises by writing

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

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

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

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

5. The Berry-Esseen theorem

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

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

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

Then there exists $C < \infty$ such that

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

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

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

and, recall,

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

with ψ as in (3.21).

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

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

We have

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

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

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

with

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

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

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

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

Once we have this, we get

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

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

Proof of Proposition 5.2. Pick ϕ satisfying

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

We desire to estimate

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

Note that its Fourier transform is

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

As in (3.13), we can assume

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

so

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

This gives an L^1 bound that implies

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

Note that

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

We can pick ϕ satisfying (5.11) and also

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

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

The proof of Theorem 5.1 is complete.

6. Faster convergence for more regular ν_{f_1}

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

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

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

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

We have

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

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

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

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

where, for ξ in this interval,

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

Add the following hypotheses:

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

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

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

Then, for $k \geq \ell$,

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

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

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

Proof. By Proposition 3.2, it remains to estimate

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

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

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

as desired.

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

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

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

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

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

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

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

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

7. Tail estimates

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

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

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

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

where, complementing (0.41), we set

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

We have

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

for $j \le 2\ell$, hence

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

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

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

and this gives (7.8).

We can now extend Proposition 1.2.

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

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

as $k \to \infty$.

Proof. We know from Theorem 1.1 that

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

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

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

This is equivalent to

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

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

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

which follows from (0.1).

Results of Appendix A then yield the following.

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

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

as $k \to 0$.

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

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

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

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

with

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

and ψ as in (0.37). Hence

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

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

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

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

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

We see that

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

and

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

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

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

We now have the following.

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

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

Furthermore, for such ρ,

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

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

8. CLT associated with a fractional diffusion

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

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

consist of positivity-preserving operators with the property that

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

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

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

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

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

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

For $\alpha = 2$, the operators $P_2^t = e^{t\partial_x^2}$ form the diffusion semigroup. For $\alpha < 2$, these are fractional diffusions. They give rise to stochastic processes belonging to the family of Levy processes. For material on this, see Chapter 3 of this text, which also treats the higher dimensional case.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

yielding

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

the last identity holding for

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

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

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

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

Proof. We see from (8.11)–(8.12) that

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

Arguing as in (1.10) yields

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

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

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

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

Then

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

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

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

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

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

then

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

Note that

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

A. Natural extension of weak* convergence of measures

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

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

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

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

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

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

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

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

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

so

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

Similarly

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

so we have (A.3).

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

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

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

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

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

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

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

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

are all equivalent. They say

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

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

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

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

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

we have

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

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

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

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

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

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

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

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

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

To start, pick N so large that

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

It follows that

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

hence, if (B.8) holds,

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

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

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

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

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

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

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

C. Convergence of mollifications of ν_{g_k}

As in §1, we assume f_j are IID random variables satisfying (0.1) and define g_k as in (0.16). Proposition 5.2, useful for our proof of the Berry-Esseen theorem, stated that in such a case there exists $C < \infty$ such that

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

In order to prove this, we took $\phi \in C^\infty(\mathbb{R})$, satisfying

$$\operatorname{supp} \phi \subset (-a, a), \quad \phi \ge 0, \quad \phi(0) = 1, \quad \phi(-\xi) = \phi(\xi), \quad \hat{\phi} \ge 0, \tag{C.2}$$

with a as in (3.4) and (3.13), so

$$\chi_{f_j}(\xi) = e^{-\sigma\xi^2/2 + \xi^2 \beta(\xi)}, \quad |\beta(\xi)| \le \frac{\sigma}{2}, \quad \text{for} \quad |\xi| \le a, \tag{C.3}$$

and established in (5.16) that

$$\sup_x |\phi(k^{-1/2} D) \nu_{g_k}(x)| \le C, \quad \forall\, k \in \mathbb{N}. \tag{C.4}$$

This in turn yields (C.1).

It follows from (C.4) and Theorem 1.1 that

$$\phi(k^{-1/2} D) \nu_{g_k} \longrightarrow \gamma^\sigma, \tag{C.5}$$

weak* in $L^\infty(\mathbb{R})$, as $k \to \infty$. A check on the behavior of ν_{g_k} for the coin toss in §2 shows that one cannot expect $\phi(k^{-1/2} D)\nu_{g_k}$ to converge to γ^σ uniformly as $k \to \infty$.

Here we show that if in (C.3)

$$|\beta(\xi)| \le b|\xi|^r, \tag{C.6}$$

for $|\xi| \le a$, and some $r \in (0, 1]$, then

$$\phi(k^{-s/2} D) \nu_{g_k} \longrightarrow \gamma^\sigma, \tag{C.7}$$

uniformly, as $k \to \infty$, if $s \in (0, 1)$ is taken sufficiently small.

To start, we note that $\phi(k^{-s/2} D)(\nu_{g_k} - \gamma^\sigma)$ is the Fourier transform of

$$\begin{aligned} &\phi(k^{-s/2}\xi)[\chi_{g_k}(\xi) - e^{-\sigma\xi^2/2}] \\ &= \phi(k^{-s/2}\xi)\big(e^{\xi^2 \beta(k^{-1/2}\xi)} - 1\big) e^{-\sigma\xi^2/2}. \end{aligned} \tag{C.8}$$

If (C.6) holds, then

$$|\beta(k^{-1/2}\xi)| \le bk^{-r/2}|\xi|^r, \tag{C.9}$$

for $|\xi| \le ak^{r/2}$. If $0 < s < r \le 1$, then

$$\begin{aligned} |\xi| \le ak^{s/2} &\Rightarrow |\xi|^r k^{-r/2} \le ak^{-(1-s)r/2} \\ &\Rightarrow |\beta(k^{-1/2}\xi)| \le ck^{-(1-s)r/2} \\ &\Rightarrow \xi^2|\beta(k^{-1/2}\xi)| \le ck^{s(1+r/2)}k^{-r/2}. \end{aligned} \tag{C.10}$$

This is a useful result provided

$$s\Bigl(1+\frac{r}{2}\Bigr) - \frac{r}{2} = -\tau < 0, \tag{C.11}$$

i.e., provided

$$s < \frac{r}{r+2}. \tag{C.12}$$

In such a case, we get

$$\bigl|\phi(k^{-s/2}\xi)\bigl(e^{\xi^2\beta(k^{-1/2}\xi)} - 1\bigr)\bigr| \le ck^{-\tau}, \tag{C.13}$$

and hence, taking into account the factor $e^{-\sigma\xi^2/2}$ in (C.8),

$$\|\phi(k^{-s/2}D)(\nu_{g_k} - \gamma^\sigma)\|_{L^\infty} \le ck^{-\tau}. \tag{C.14}$$

On the other hand, $\phi(k^{-s/2}D)\gamma^\sigma - \gamma^\sigma$ is the Fourier transform of

$$[\phi(k^{-s/2}\xi) - 1]e^{-\sigma\xi^2/2}, \tag{C.15}$$

whose absolute value is

$$\le Ck^{-s}\xi^2 e^{-\sigma\xi^2/2}, \tag{C.16}$$

so

$$\|\phi(k^{-s/2}D)\gamma^\sigma - \gamma^\sigma\|_{L^\infty} \le Ck^{-s}. \tag{C.17}$$

Putting (C.14) and (C.17) together, we have the following conclusion.

Proposition C.1. *Under the hypotheses given above, particularly with $r \in (0,1]$ satisfying (C.6), $s > 0$ satisfying (C.12), and $\tau > 0$ given by (C.11), we have*

$$\|\phi(k^{-s/2}D)\nu_{g_k} - \gamma^\sigma\|_{L^\infty} \le C(k^{-\tau} + k^{-s}). \tag{C.18}$$

The structure of $\phi(k^{-s/2}D)\nu_{g_k}$ and γ^σ as probability distributions allows us to deduce the following from Proposition C.1.

Corollary C.2. *In the setting of Proposition C.1,*

$$\|\phi(k^{-s/2}D)\nu_{g_k} - \gamma^\sigma\|_{L^1} \longrightarrow 0, \quad \textit{as} \;\; k \to \infty. \tag{C.19}$$

Proof. Set $\gamma_k = \phi(k^{-s/2}D)\nu_{g_k}$. Pick $\varepsilon > 0$. Then pick $M < \infty$ such that

$$\|\gamma^\sigma\|_{L^1(\mathbb{R}\setminus(-M,M))} < \varepsilon. \tag{C.20}$$

Now pick $k(\varepsilon)$ such that

$$k \geq k(\varepsilon) \Longrightarrow \|\gamma_k - \gamma^\sigma\|_{L^\infty} < \frac{\varepsilon}{M}, \tag{C.21}$$

hence

$$\|\gamma_k - \gamma^\sigma\|_{L^1((-M,M))} \leq \varepsilon, \tag{C.22}$$

hence

$$\Bigl|\int_{-M}^{M} \gamma_k(x)\,dx - 1\Bigr| \leq \Bigl|\int_{-M}^{M} \gamma^\sigma(x)\,dx - 1\Bigr| + \varepsilon < 2\varepsilon. \tag{C.23}$$

Since γ_k and γ^σ are both ≥ 0 and integrate over $\mathbb{R}$ to 1, it follows that, for $k \geq k(\varepsilon)$,

$$\|\gamma_k\|_{L^1(\mathbb{R}\setminus[-M,M])} < 2\varepsilon, \tag{C.24}$$

so, by (C.20) and (C.22),

$$\|\gamma_k - \gamma^\sigma\|_{L^1} < 4\varepsilon, \quad \text{for} \;\; k \geq k(\varepsilon). \tag{C.25}$$

This gives (C.19).

Finally, we can interpolate between (C.18) and (C.25).

Corollary C.3. *In the setting of Proposition C.1, for each $p \in [1,\infty]$,*

$$\|\phi(k^{-s/2}D)\nu_{g_k} - \gamma^\sigma\|_{L^p} \longrightarrow 0, \quad \textit{as} \;\; k \to \infty. \tag{C.26}$$

Proof. Writing $D_k = \phi(k^{-s/2}D)\nu_{g_k} - \gamma^\sigma$, we have

$$\|D_k\|_{L^p}^p = \int_{\mathbb{R}} |D_k|^{p-1}|D_k|\,dx \leq \|D_k\|_{L^\infty}^{p-1}\|D_k\|_{L^1}. \tag{C.27}$$

References

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

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

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

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

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

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

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

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

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

2. Stochastic operators and the Perron-Frobenius theorem in infinite dimensions

1. Introduction

Let X be a compact Hausdorff space. The space $C(X)$ of continuous, real-valued functions on X is a Banach space, with norm

$$\|f\|_{\sup} = \sup_{x \in X} |f(x)|. \tag{1.1}$$

Let

$$A : C(X) \longrightarrow C(X) \tag{1.2}$$

be a bounded linear map. We say A is *positive* if

$$f \in C(X),\ f \geq 0 \Longrightarrow Af \geq 0. \tag{1.3}$$

We say A is *strictly positive* if

$$f \in C(X),\ f \geq 0,\ f \neq 0 \Longrightarrow Af(x) > 0, \quad \forall\, x \in X. \tag{1.4}$$

We say A is *primitive* if A is positive and some power A^m is strictly positive. We say A is *irreducible* if A is positive and

$$f \in C(X),\ f \geq 0,\ f \neq 0 \Longrightarrow \sup_k A^k f(x) > 0, \quad \forall\, x \in X. \tag{1.5}$$

The dual of $C(X)$ is

$$\mathcal{M}(X) = C(X)', \tag{1.6}$$

where $\mathcal{M}(X)$ denotes the space of finite, signed, regular Borel measures on X. The norm on $\mathcal{M}(X)$ is the total variation, which satisfies

$$\|\mu\|_{\mathrm{TV}} = \sup\, \{\langle f, \mu \rangle : f \in C(X),\ \|f\|_{\sup} \leq 1\}, \tag{1.7}$$

where

$$\langle f, \mu \rangle = \int_X f\, d\mu. \tag{1.8}$$

The operator A in (1.2) has the adjoint

$$A^t : \mathcal{M}(X) \longrightarrow \mathcal{M}(X), \tag{1.9}$$

satisfying

$$\langle f, A^t\mu\rangle = \langle Af, \mu\rangle. \tag{1.10}$$

We have

$$\|A^t\|_1 = \|A\|_\infty, \tag{1.11}$$

where $\|A\|_\infty$ denotes the operator norm of A on $C(X)$ and $\|A^t\|_1$ that of A^t on $\mathcal{M}(X)$. Note that, if A is positive, then

$$A^t : \mathcal{M}_+(X) \longrightarrow \mathcal{M}_+(X), \tag{1.12}$$

where $\mathcal{M}_+(X)$ denotes the set of positive, finite, regular Borel measures on X.

A positive operator A on $C(X)$ is said to be a *stochastic* operator if, in addition,

$$A1 = 1. \tag{1.13}$$

For such operators, we have

$$A^t : \mathcal{P}(X) \longrightarrow \mathcal{P}(X), \tag{1.14}$$

where $\mathcal{P}(X)$ denotes the set of positive, regular Borel measures on X of total mass 1, i.e., probability measures on X.

The Perron-Frobenius theorem is a circle of results about the various sorts of operators defined above. The classical setting is the finite-dimensional case, i.e., where X is a finite point set. We establish such results here, in the infinite-dimensional setting. For a number of these results, we make the additional hypothesis that A in (1.2) is compact, which implies that A^t in (1.9) is compact.

In §2 we establish results in the Perron-Frobenius circle for stochastic operators. We first show that if A is a stochastic operator on $C(X)$, there exists $\mu \in \mathcal{P}(X)$ such that $A^t\mu = \mu$. (For this, compactness is not needed.) It follows that

$$A : V \longrightarrow V, \quad \text{where} \quad V = \{f \in C(X) : \langle f, \mu\rangle = 0\}. \tag{1.15}$$

We then show that if A is stochastic and strictly positive, then, for $f \in C(X)$,

$$f \notin \operatorname{Span}(1) \Longrightarrow \|Af\|_{\sup} < \|f\|_{\sup}. \tag{1.16}$$

This is used to show in Propositions 2.3–2.4 that if A is a compact, stochastic operator on $C(X)$, and A is strictly positive, or more generally if A is primitive, then $A_V = A|_V$ has spectral radius $\rho(A_V) < 1$. Using this result, we establish the following in Proposition 2.5.

Proposition 1.1. *Let A be a compact, stochastic operator, and assume A is primitive. Then*

$$A^k \longrightarrow P, \quad as \quad k \to \infty, \tag{1.17}$$

in operator norm, where P is the projection of $C(X)$ onto Span(1) *that annihilates V.*

It follows that

$$(A^t)^k \longrightarrow P^t \tag{1.18}$$

in operator norm, and P^t is the projection of $\mathcal{M}(X)$ onto Span(μ) that annihilates $W = \{\lambda \in \mathcal{M}(X) : \langle 1, \lambda \rangle = 0\}$. For P and P^t, we have the formulas

$$Pf = \langle f, \mu \rangle 1, \quad P^t\lambda = \langle 1, \lambda \rangle \mu, \tag{1.19}$$

given $f \in C(X)$, $\lambda \in \mathcal{M}(X)$. Making use of Proposition 1.1, we establish in Propositions 2.7–2.8 the following.

Proposition 1.2. *Let A be a compact, stochastic operator on $C(X)$, and assume A is irreducible. Then the measure $\mu \in \mathcal{P}(X)$ such that $A^t\mu = \mu$ is unique. Furthermore,* 1 *is an eigenvalue of A of algebraic multiplicity one, i.e., the generalized eigenspace $\mathcal{GE}(A, 1)$ is one-dimensional (equal to* Span(1)*).*

In §3 we turn to other classes of positive operators. We say a positive operator A on $C(X)$ is *crypto-stochastic* provided there exists

$$\psi \in C(X) \text{ such that } \psi(x) > 0, \ \forall\, x \in X, \text{ and } A\psi = \psi. \tag{1.20}$$

Then, with $M_\psi f = \psi f$, $\widetilde{A} = M_\psi^{-1} A M_\psi$ is stochastic, and results of §2 apply. More generally, we say A is crypto-stochastic up to scaling if there exists $\lambda \in (0, \infty)$ such that $\lambda^{-1}A$ is crypto-stochastic. Clearly a necessary condition for A to have this property is that

$$A1(x) = \varphi(x) > 0, \quad \forall\, x \in X. \tag{1.21}$$

We show in Proposition 3.4 that if A is a positive, irreducible, compact operator on $C(X)$ satisfying three hypotheses, given as (H1)–(H3), then A is crypto-stochastic, up to scaling.

Turning away from crypto-stochastic operators, we establish the following in Proposition 3.5.

Proposition 1.3. *Let A be a positive, compact operator on $C(X)$. Assume that A satisfies (1.21). Then there exists $\lambda > 0$ and $\mu \in \mathcal{P}(X)$ such that $A^t\mu = \lambda\mu$.*

In §4 we take a look at positive infinite matrices that define bounded linear maps

$$A : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N}). \tag{1.22}$$

Such maps are treated in [Sen]. We show that their analysis fits into the material developed in §§2–3, via the natural, positivity-preserving, isometric isomorphism

$$\ell^\infty(\mathbb{N}) \approx C(X_{\mathbb{N}}), \tag{1.23}$$

where $X_{\mathbb{N}}$ is the Stone-Cech compactification of $\mathbb{N}$, which can also be characterized as the maximal ideal space of $\ell^\infty(\mathbb{N})$, viewed as a commutative C^*-algebra.

2. Stochastic operators

Our first result in the circle of Perron-Frobenius theorems is the following. Actually, this result does not require A to be compact (nor does Proposition 2.2).

Proposition 2.1. *Assume A is a stochastic operator. Then there exists*

$$\mu \in \mathcal{P}(X) \quad \textit{such that} \quad A^t\mu = \mu. \tag{2.1}$$

Proof. The set $\mathcal{P}(X)$ is a compact, convex subset of $\mathcal{M}(X)$, endowed with the weak* topology, and A^t is continuous on $\mathcal{M}(X)$ in this topology. Also,

$$A^t : \mathcal{P}(X) \longrightarrow \mathcal{P}(X). \tag{2.2}$$

The existence of a fixed point $\mu \in \mathcal{P}(X)$ is then a consequence of the Markov-Kakutani fixed point theorem (cf. [DS], p. 456).

Given μ as in (2.1), we set

$$V = \{f \in C(X) : \langle f, \mu\rangle = 0\}, \tag{2.3}$$

a closed linear subspace of $C(X)$, of codimension 1. We have a direct sum decomposition

$$C(X) = V \oplus \mathrm{Span}(1). \tag{2.4}$$

Also, whenever (2.1) holds,

$$A : V \longrightarrow V, \tag{2.5}$$

since

$$\langle Af, \mu\rangle = \langle f, A^t\mu\rangle = \langle f, \mu\rangle. \tag{2.6}$$

Proposition 2.2. *Let A be a stochastic operator. Assume in addition that A is strictly positive, so (1.4) holds. Then, for $f \in C(X)$,*

$$f \notin \operatorname{Span}(1) \Longrightarrow \|Af\|_{\sup} < \|f\|_{\sup}. \tag{2.7}$$

Proof. It suffices to treat the case $\|f\|_{\sup} = 1$, so $-1 \le f \le 1$. If $f(x) < 1$ for some x, there exists $\varphi \in C(X)$ such that $\varphi \ge 0$, $\varphi(x) > 0$, and $f + \varphi \le 1$. Hence $Af + A\varphi \le 1$. The hypothesis (1.4) implies $A\varphi(x) > 0$ for all $x \in X$, so $\sup Af(x) < 1$. Similarly, if $f(x) > -1$ for some $x \in X$, we have $\inf Af(x) > -1$. If $-1 \le f \le 1$ and $f \notin \operatorname{Span}(1)$, both of these conditions hold, and we have (2.7).

Before stating the next result, we note that if A is a stochastic operator on $C(X)$, then

$$\|A\|_\infty = \|A^t\|_1 = 1. \tag{2.8}$$

Also, having (2.5), let us denote the restriction of A to V by A_V.

Proposition 2.3. *Let A be a compact stochastic operator, and assume A is strictly positive. Then*

$$\alpha \in \operatorname{Spec} A_V \Longrightarrow |\alpha| < 1. \tag{2.9}$$

Hence the spectral radius of A_V is < 1, i.e.,

$$\rho(A_V) < 1. \tag{2.10}$$

Proof. Note that $A_V : V \to V$ is compact, so each nonzero $\alpha \in \operatorname{Spec}(A_V)$ must be an eigenvalue. The conclusion (2.9) then follows directly from (2.7). Also, compactness of A_V implies that $\operatorname{Spec} A_V$ is a countable subset of $\mathbb{C}$, whose only possible accumulation point is 0. Hence (2.10) follows from (2.9).

Remark. We recall the following useful formula for the spectral radius:

$$\rho(A_V) = \limsup_{k\to\infty} \|A_V^k\|^{1/k}. \tag{2.11}$$

The following result extends the scope of Proposition 2.3 a bit.

Proposition 2.4. *Let A be a compact stochastic operator, and assume A is primitive, i.e.,*

$$A^m \text{ is strictly positive for some } m \in \mathbb{N}. \tag{2.12}$$

Then the conclusions (2.9)–(2.10) hold.

Proof. We still have (2.5), and we can define A_V as before. Also $(A_V)^m = (A^m)_V$. Now if $\alpha \in \operatorname{Spec} A_V$, and $\alpha \neq 0$, compactness implies α is an eigenvalue of A_V, hence α^m is an eigenvalue of $(A_V)^m = (A^m)_V$. But Proposition 2.3 applies to A^m, so $|\alpha^m| < 1$. This gives (2.9), and (2.10) follows.

REMARK. In case $X = \{1, 2\}$ so $C(X) = \mathbb{R}^2$, the following is an example of a stochastic matrix that is irreducible but not primitive:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.13}$$

Note that (2.9)–(2.10) fail for this matrix.

We can now prove the following key result.

Proposition 2.5. *Let A be a compact stochastic operator, and assume A is primitive. Then*

$$A^k \longrightarrow P, \quad as \ \ k \to \infty, \tag{2.14}$$

in operator norm on $C(X)$, where P is the projection of $C(X)$ onto Span(1) *that annihilates V.*

Proof. We have

$$\begin{aligned} A^k &= A^k P + A^k (I - P) \\ &= P + A_V^k (I - P), \end{aligned} \tag{2.15}$$

so

$$\begin{aligned} \|A^k - P\|_\infty &= \|A_V^k (I - P)\|_\infty \\ &\le \|A_V^k\|_\infty \cdot \|I - P\|_\infty, \end{aligned} \tag{2.16}$$

and the fact that this converges to 0 (at an exponential rate) follows from (2.10)–(2.11).

Corollary 2.6. *In the setting of Proposition 2.5,*

$$(A^t)^k \longrightarrow P^t, \quad as \ \ k \to \infty, \tag{2.17}$$

in operator norm on $\mathcal{M}(X)$. In this case, P^t is the projection of $\mathcal{M}(X)$ onto Span(μ) *that annihilates*

$$W = \{\lambda \in \mathcal{M}(X) : \langle 1, \lambda \rangle = 0\}. \tag{2.18}$$

Proof. To get (2.17), just apply the transpose to (2.15)–(2.16):

$$(A^t)^k = P^t + (A_V^k(I-P))^t, \tag{2.19}$$

and note that

$$\|(A_V^k(I-P))^t\|_1 = \|A_V^k(I-P)\|_\infty. \tag{2.20}$$

Let us note that P is given by the formula

$$Pf = \langle f, \mu\rangle 1, \tag{2.21}$$

and then the identity

$$\langle f, P^t\lambda\rangle = \langle Pf, \lambda\rangle = \langle f, \mu\rangle\, \langle 1, \lambda\rangle \tag{2.22}$$

yields the formula

$$P^t\lambda = \langle 1, \lambda\rangle \mu, \quad \text{for} \ \ \lambda \in \mathcal{M}(X). \tag{2.23}$$

From (2.17) and (2.23), we deduce that, when A is a compact stochastic operator that is primitive, the measure μ in (2.1) is unique. The following extends the scope of this result.

Proposition 2.7. *Let A be a compact stochastic operator on $C(X)$, and assume A is irreducible. Then the measure μ in (2.1) is unique.*

Proof. Form

$$B = \sum_{k=1}^{\infty} 2^{-k} A^k = \frac{1}{2} A \Bigl(I - \frac{1}{2}A\Bigr)^{-1}, \tag{2.24}$$

which is a convergent series by (2.8), and defines a compact stochastic operator on $C(X)$. If A is irreducible, then B is strictly positive. Hence Proposition 2.3 and Corollary 2.6 apply to B. On the other hand, Proposition 2.1 applies to A, and clearly

$$\mu \in \mathcal{P}(X),\ A^t\mu = \mu \Longrightarrow B^t\mu = \mu. \tag{2.25}$$

By Corollary 2.6, applied to B, $(B^t)^k \to P^t$, given by (2.23). This establishes uniqueness of μ in (2.25).

Proposition 2.8. *Let A be a compact stochastic operator on $C(X)$, and assume A is irreducible. Then 1 is an eigenvalue of A of algebraic multiplicity 1, i.e., the generalized eigenspace $\mathcal{GE}(A,1)$ is 1-dimensional.*

Proof. With B as in (2.24), we have $B - I = (A - I)(I - A/2)^{-1}$, and hence

$$f \in \mathcal{GE}(A,1) \Longleftrightarrow f \in \mathcal{GE}(B,1). \tag{2.26}$$

But Proposition 2.3 applies to B, and the conclusion (2.9) for B_V implies $\mathcal{GE}(B,1)$ has dimension 1.

3. Other classes of positive compact operators

We move on from compact stochastic operators to other classes of positive compact operators on $C(X)$. To begin, we say a positive operator A on $C(X)$ is *crypto-stochastic* if there exists

$$\psi \in C(X) \text{ such that } \psi(x) > 0,\ \forall\, x \in X, \text{ and } A\psi = \psi. \tag{3.1}$$

Then, with $M_\psi f = \psi f$, we have the positive operator

$$\widetilde{A} = M_\psi^{-1} A M_\psi, \quad \text{stochastic}, \tag{3.2}$$

and the results of §2 apply to $\widetilde{A}$. Note that if A is strictly positive, resp., primitive, or irreducible, so is $\widetilde{A}$. Note also that strict positivity of ψ is required in order that M_ψ^{-1} be a well defined, bounded operator on $C(X)$. In connection with this, we have the following.

Proposition 3.1. *Assume the positive operator A is irreducible. Then*

$$\psi \in C(X),\ \psi \ge 0,\ \psi \ne 0,\ A\psi = \psi \Longrightarrow \psi(x) > 0, \quad \forall\, x \in X. \tag{3.3}$$

Proof. Let

$$E = \sum_{k=1}^{\infty} \frac{1}{k!} A^k = e^A - I. \tag{3.4}$$

If A is irreducible, then E is strictly positive. Now

$$A\psi = \psi \Longrightarrow E\psi = (e-1)\psi. \tag{3.5}$$

But $\psi \ge 0,\ \psi \ne 0 \Rightarrow E\psi(x) > 0$ for all $x \in X$, so we have (3.5).

Clearly a necessary condition for a positive operator A on $C(X)$ to be crypto-stochastic is that

$$A1(x) = \varphi(x) > 0, \quad \forall\, x \in X. \tag{3.6}$$

However, this condition is not sufficient. For example, in one picks a positive $\lambda \neq 1$ and a strictly positive compact stochastic operator A_0 on $C(X)$, the operator $A = \lambda A_0$ is positive and satisfies (3.6), but (3.1) cannot hold. This motivates the definition of a more general class of operators. We say a positive operator A on $C(X)$ is crypto-stochastic up to scaling if there exist

$$\psi \in C(X),\ \lambda \in (0,\infty) \ \text{ such that } \psi(x) > 0,\ \forall\, x \in X \text{ and } A\psi = \lambda\psi. \tag{3.7}$$

In such a case, the operator $A^{\#} = \lambda^{-1}A$ is crypto-stochastic.

These considerations lead to the problem of determining when a positive, compact operator on $C(X)$ is crypto-stochastic, up to scaling. In connection with this, we mention the following weaker problem.

Problem PF. Given a positive, compact operator on $C(X)$, find

$$\psi \in C(X),\ \lambda > 0 \ \text{ such that } \ \psi \geq 0,\ \psi \neq 0, \text{ and } \ A\psi = \lambda\psi. \tag{3.8}$$

The weakening consists in not requiring ψ in (3.8) to be strictly positive. Part of the classical Perron-Frobenius theory is that this problem is always solvable when X is a finite point set, so, for some $n \in \mathbb{N}$, $C(X) \approx \mathbb{R}^n$. Here is that result. We phrase its formulation and proof in a way that lends itself to extension beyond the finite case.

Proposition 3.2. *Assume X has n points, $n \in \mathbb{N}$, and A is a positive operator on $C(X)$. Assume*

$$f \in C(X),\ f \geq 0,\ f \neq 0 \Longrightarrow Af \neq 0. \tag{3.9}$$

Then there exist $\lambda > 0$ and $\psi \in C(X)$ satisfying (3.8).

Proof. Let ν_0 be the probability measure on X that assigns the mass $1/n$ to each of its points. With the notation

$$C_+(X) = \{f \in C(X) : f \geq 0\}, \tag{3.10}$$

let

$$\Sigma = \{f \in C_+(X) : \langle f, \nu_0\rangle = 1\}. \tag{3.11}$$

Thus Σ is a compact, convex subset of $C(X)$. We define

$$\Phi : \Sigma \longrightarrow \Sigma \tag{3.12}$$

by

$$(3.13)\qquad \Phi(f) = \frac{1}{\langle Af, \nu_0\rangle} Af.$$

The hypothesis (3.11) implies $\langle Af, \nu_0\rangle > 0$ for all $f \in \Sigma$, and by compactness we have a positive lower bound. Now the Brouwer fixed point theorem applies to (3.13). (A proof of his result can be found in Chapter 1 of [T].) Hence there exists $f \in \Sigma$ such that

$$(3.14)\qquad Af = \langle Af, \nu_0\rangle f.$$

This proves Proposition 3.2.

Recalling Proposition 3.1, we see that if X is a finite point set, every positive, irreducible A on $C(X)$ is crypto-stochastic, up to scaling.

We return to cases where $C(X)$ is infinite dimensional, and investigate ways to extend the proof of Proposition 3.2 to cover positive, compact operators on $C(X)$, under some additional hypotheses. To start, we make the following three hypotheses:

(H1) There is a measure $\nu \in \mathcal{P}(X)$ such that $\nu(U) > 0$ for each nonempty open $U \subset X$. Equivalently,

$$(3.15)\qquad f \in C(X),\ f \ge 0,\ f \ne 0 \Longrightarrow \langle f, \nu\rangle > 0.$$

(H2) The positive operator A satisfies

$$(3.16)\qquad A : L^1(X, \nu) \longrightarrow C(X), \quad \text{compactly.}$$

(H3) With $C_+(X)$ as in (3.10), and

$$(3.17)\qquad \Sigma = \{f \in C_+(X) : \langle f, \nu\rangle = 1\},$$

there is a $\delta > 0$ such that

$$(3.18)\qquad f \in \Sigma \Longrightarrow \|Af\|_{\sup} \ge \delta.$$

These hypotheses imply that $A(\Sigma)$ is a relatively compact, convex subset of $C(X)$. The following is a useful improvement of (3.18).

Lemma 3.3. *Under hypotheses (H1)–(H3), there exists $\alpha > 0$ such that*

$$f \in \Sigma \Longrightarrow \langle Af, \nu\rangle \geq \alpha. \tag{3.19}$$

Proof. If (3.19) fails, there exist $f_k \in \Sigma$ such that $\langle Af_k, \nu\rangle \leq 2^{-k}$. Since $A(\Sigma)$ is relatively compact in $C(X)$, we have a subsequence f_{k_j} such that $Af_{k_j} \to g \in C_+(X)$, uniformly. Consequently, $\langle g, \nu\rangle = 0$, which by (H1), implies $g = 0$. This contradicts the condition (3.18) in (H3).

Now define

$$\Phi : \Sigma \longrightarrow \Sigma, \quad \Phi(f) = \frac{1}{\langle Af, \nu\rangle} Af. \tag{3.20}$$

By (3.19), this is a well defined, continuous map, and the relative compactness of $A(\Sigma)$ in $C(X)$ yields

$$\Phi : \Sigma \longrightarrow \mathcal{K}, \tag{3.21}$$

where $\mathcal{K}$ is a compact, convex subset of $\Sigma \subset C(X)$. The Schauder fixed point theorem (a proof of which can be found in Chapter 13 of [T]) applies, to yield $\psi \in \mathcal{K} \subset \Sigma$ satisfying $\Phi(\psi) = \psi$, hence $A\psi = \langle A\psi, \nu\rangle\psi$. We have proved the first part of the following.

Proposition 3.4. *Let $A : C(X) \to C(X)$ be a positive operator. Assume hypotheses (H1)–(H3). Then there exist $\lambda > 0$ and $\psi \in C(X)$ such that (3.8) holds.*

If also A is irreducible, then $A^{\#} = \lambda^{-1}A$ satisfies

$$A^{\#}\psi = \psi, \quad \textit{and} \quad \psi(x) > 0, \quad \forall\, x \in X, \tag{3.22}$$

and hence $A^{\#}$ is crypto-stochastic.

Proof. The first part was established above, and (3.22) follows from Proposition 3.1.

Suppose now that A_0 is a compact, positive operator on $C(X)$ and that (H1)–(H3) hold for $A = A_0^m$, for some $m \in \mathbb{N}$. If A_0 is irreducible, so is A, so, with λ as in Proposition 3.4, $A_1^m = A^{\#}$ is crypto-stochastic, where $A_1 = \lambda^{-1/m}A_0$, and we have (3.22). It follows that there exists $\mu \in \mathcal{P}(X)$ such that $(A^{\#})^t\mu = \mu$, and, with V as in (2.3), ψ as in (3.22), we have

$$C(X) = V \oplus \mathrm{Span}(\psi), \quad A^{\#} : V \to V. \tag{3.23}$$

If $A^{\#}$ is primitive, so is $A^{\#}$. One deduces via Proposition 2.4 that $A^{\#}_V = A^{\#}|_V$ has spectral radius $\rho < 1$, and

$$\mathcal{GE}(A^{\#}, 1) = \mathrm{Span}(\psi). \tag{3.24}$$

Note also that

$$A^{\#}(A_1\psi) = A_1^{m+1}\psi = A_1(A^{\#}\psi) = A_1\psi, \tag{3.25}$$

hence $A_1\psi \in \operatorname{Span}(\psi)$. If $A_1\psi = \beta\psi$, then $A^{\#}\psi = \beta^m\psi = \psi$, so $\beta^m = 1$. Since A_1 is positive, this implies $\beta = 1$, so

$$A_1\psi = \psi. \tag{3.26}$$

Consequently A_1 itself is crypto-stochastic.

We temporarily leave results related to (H1)–(H3), and look directly for positive measures on X that are eigenvectors of A^t.

Proposition 3.5. *Let A be a positive, compact operator on $C(X)$. Assume that $\varphi = A1$ satisfies (3.6). Then there exist $\lambda > 0$ and $\mu \in \mathcal{P}(X)$ such that*

$$A^t\mu = \lambda\mu. \tag{3.27}$$

Proof. First note that there exists $\delta > 0$ such that

$$\langle 1, A^t\mu\rangle = \langle\varphi, \mu\rangle \geq \delta, \quad \forall\, \mu \in \mathcal{P}(X), \tag{3.28}$$

given (3.6). Hence we can define

$$\Psi : \mathcal{P}(X) \longrightarrow \mathcal{P}(X), \quad \Psi(\mu) = \frac{1}{\langle 1, A^t\mu\rangle} A^t\mu, \tag{3.29}$$

and Ψ is continuous. Since $A^t(\mathcal{P}(X))$ is a relatively compact, convex subset of $\mathcal{M}_+(X)$, we have

$$\Psi : \mathcal{P}(X) \longrightarrow \mathcal{K}, \tag{3.30}$$

where $\mathcal{K}$ is a compact, convex subset of $\mathcal{P}(X)$. It follows from the Schauder fixed point theorem that Ψ has a fixed point, say μ, in $\mathcal{K}$, and then

$$A^t\mu = \langle 1, A^t\mu\rangle\mu, \tag{3.31}$$

giving (3.27).

Having Proposition 3.5, we again make contact with (H1):

Proposition 3.6. *Let A be a positive operator on $C(X)$. If A is irreducible and $\mu \in \mathcal{P}(X)$ satisfies (3.27), with $\lambda > 0$, then*

$$f \in C(X),\ f \ge 0,\ f \ne 0 \Longrightarrow \langle f, \mu \rangle > 0. \tag{3.32}$$

Proof. For each $k \in \mathbb{N}$,

$$\lambda^k \langle f, \mu \rangle = \langle f, (A^t)^k \mu \rangle = \langle A^k f, \mu \rangle, \tag{3.33}$$

hence, for $E = e^A - I$ as in (3.4) and f as in (3.32),

$$(e - 1)\langle f, \mu \rangle = \langle Ef, \mu \rangle > 0, \tag{3.34}$$

since irreducibility of A implies $Ef(x) > 0$ for all $x \in X$.

Remark. Compactness of A is not required for Proposition 3.6. This fact is particularly significant in light of Proposition 2.1.

It follows from Propositon 3.6 that, in the setting of Proposition 3.5, and with A irreducible, hypothesis (H1) holds with $\nu = \mu$. Furthermore, with

$$\Sigma = \{f \in C_+(X) : \langle f, \mu \rangle = 1\}, \tag{3.35}$$

we have

$$\lambda^{-1} A : \Sigma \longrightarrow \Sigma, \tag{3.36}$$

so also (H3) and (3.19) hold. Thus Proposition 3.4 implies the following.

Proposition 3.7. *Let A be a positive, compact, irreducible operator on $C(X)$, and assume $\varphi = A1$ satisfies (3.6). Take $\mu \in \mathcal{P}(X)$ such that (3.27) holds. Assume that (H2) holds with $\nu = \mu$, i.e.,*

$$A : L^1(X, \mu) \longrightarrow C(X), \quad \textit{compactly.} \tag{3.37}$$

Then $\lambda^{-1} A$ is crypto-stochastic.

4. Connections with infinite positive matrices

Here we look at infinite matrices $A = (a_{jk})$, defined for $j, k \in \mathbb{N}$, having a bound on the row sums:

$$\sum_{k=1}^{\infty} |a_{jk}| \le \alpha < \infty, \quad \forall\, j \in \mathbb{N}. \tag{4.1}$$

Then we have

$$A : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N}), \tag{4.2}$$

acting as a bounded operator, by

$$(Af)_j = \sum_{k=1}^{\infty} a_{jk} f_k, \quad \|A\|_\infty \le \alpha. \tag{4.3}$$

Here, $\ell^\infty(\mathbb{N})$ denotes the space of bounded real sequences, i.e., the space of bounded functions $f : \mathbb{N} \to \mathbb{R}$, a Banach space with norm $\|f\|_\infty = \sup_k |f_k|$. We identify $f_k = f(k)$. We say A is positive if $a_{jk} \ge 0$ for each $j, k \in \mathbb{N}$. We say a positive matrix is stochastic if each row sum is 1, i.e., $\sum_k a_{jk} = 1$ for each j, or equivalently

$$A1 = 1, \tag{4.4}$$

where here 1 denotes the function on $\mathbb{N}$ that is identically 1.

To relate the study of such matrices to material in §§1–3, we use the natural isometric isomorphism

$$\ell^\infty(\mathbb{N}) \approx C(X_{\mathbb{N}}), \tag{4.5}$$

where $X_{\mathbb{N}}$ denotes the Stone-Cech compactification of $\mathbb{N}$. This is a compact Hausdorff space. There is a natural inclusion

$$\mathbb{N} \subset X_{\mathbb{N}}, \tag{4.6}$$

as an open, dense subset. Since $\ell^\infty(\mathbb{N})$ is not separable, $X_{\mathbb{N}}$ is not metrizable. Using the isomorphism (4.5), we identify A in (4.2) with

$$A : C(X_{\mathbb{N}}) \longrightarrow C(X_{\mathbb{N}}). \tag{4.7}$$

Positivity in (4.2) turns into positivity in (4.7), since the isomorphism (4.5) is also positivity preserving. Also, if (4.4) holds on $\ell^\infty(\mathbb{N})$, it holds on $C(X_{\mathbb{N}})$.

As in (1.6), we have the duality

$$\ell^\infty(\mathbb{N})' \approx C(X_{\mathbb{N}})' = \mathcal{M}(X_{\mathbb{N}}), \tag{4.8}$$

where $\mathcal{M}(X_{\mathbb{N}})$ is the space of finite, signed, regular Borel measures on $X_{\mathbb{N}}$. There are a natural injection and a natural projection

$$J : \ell^1(\mathbb{N}) \longrightarrow \mathcal{M}(X_{\mathbb{N}}), \quad \Pi : \mathcal{M}(X_{\mathbb{N}}) \longrightarrow \ell^1(\mathbb{N}), \tag{4.9}$$

induced by (4.6).

As in (1.9), the map (4.7) has a transpose

$$A^t : \mathcal{M}(X_{\mathbb{N}}) \longrightarrow \mathcal{M}(X_{\mathbb{N}}). \tag{4.10}$$

We also have a map

$$A^\tau : \ell^1(\mathbb{N}) \longrightarrow \ell^1(\mathbb{N}), \tag{4.11}$$

given by

$$A^\tau = \Pi A^t J. \tag{4.12}$$

Results of §§2–3 yield conditions under which A^t has an invariant measure $\mu \in \mathcal{P}(X_{\mathbb{N}})$. Such μ is also an invariant element of A^τ if and only if $\operatorname{supp}\mu \subset \mathbb{N}$. We will see examples below for which $\ell^1(\mathbb{N})$ does not have a positive element that is invariant under A^τ.

While A^t has a richer structure that A^τ, the operator A^τ does not lose information about A. In fact, since $\ell^\infty(\mathbb{N}) = \ell^1(\mathbb{N})'$, a bounded operator on $\ell^1(\mathbb{N})$ has a transpose:

$$B : \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N}) \Longrightarrow B^t : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N}), \tag{4.13}$$

and we see that, for A in (4.2),

$$(A^\tau)^t = A. \tag{4.14}$$

Note also that $\ell^1(\mathbb{N})$ is weak* dense in $\mathcal{M}(X_{\mathbb{N}})$ (though not norm dense), and A^t is the unique extension of A^τ to a linear operator on $\mathcal{M}(X_{\mathbb{N}})$ that is continuous in the weak* topology of $\mathcal{M}(X_{\mathbb{N}})$.

At this point it is useful to note that, in place of the set of positive integers $\mathbb{N}$, we could use any countably infinite set S, and extend results derived above from the setting of (4.2) to

$$A : \ell^\infty(S) \to \ell^\infty(S), \quad \ell^\infty(S) \approx C(X_S), \quad C(X_S)' = \mathcal{M}(X_S), \tag{4.15}$$

where X_S is the Stone-Cech compactification of X_S. Particularly useful examples include

$$\mathbb{Z}, \quad \mathbb{Z}^n, \quad S\ell(n, \mathbb{Z}). \tag{4.16}$$

An example involving $S = \mathbb{Z}$ is

$$A : \ell^\infty(\mathbb{Z}) \longrightarrow \ell^\infty(\mathbb{Z}), \quad Af(k) = f(k+1). \tag{4.17}$$

An element $\mu \in \mathcal{P}(X_{\mathbb{Z}})$ invariant under A^t defines a linear functional on $\ell^\infty(\mathbb{Z})$ known as an *invariant mean.* The existence of such invariant means is a special case of Proposition 2.1. It is clear that such μ satisfies $\mu(\mathbb{Z})=0$, and there does not exist an A^τ-invariant element of $\ell^1(\mathbb{Z})$. Such results hold for a number of related operators, such as

$$A:\ell^\infty(\mathbb{Z})\longrightarrow\ell^\infty(\mathbb{Z}),\quad Af(k)=\frac{1}{2}f(k+1)+\frac{1}{2}f(k-1). \tag{4.18}$$

The Markov process associated to this operator is "the standard random walk" on $\mathbb{Z}$.

In many cases, the operator A might have special structure that allows one to replace $\ell^\infty(S)$ by a smaller subspace, such as

$$\begin{aligned}\ell^\infty_\#(\mathbb{N})&=\{f\in\ell^\infty(\mathbb{N}):\lim_{k\to\infty}f(k)\text{ exists}\},\quad\text{or}\\ \ell^\infty_\#(\mathbb{Z})&=\{f\in\ell^\infty(\mathbb{Z}):\lim_{k\to+\infty}f(k)\text{ and }\lim_{k\to-\infty}f(k)\text{ exist}\},\end{aligned} \tag{4.19}$$

having natural isomorphisms

$$\begin{aligned}\ell^\infty_\#(\mathbb{N})&\approx C(\widehat{\mathbb{N}}),\quad\widehat{\mathbb{N}}=\mathbb{N}\cup\{\infty\},\\ \ell^\infty_\#(\mathbb{Z})&\approx C(\widehat{\mathbb{Z}}),\quad\widehat{\mathbb{Z}}=\mathbb{Z}\cup\{+\infty,-\infty\}.\end{aligned} \tag{4.20}$$

For example, (4.18) has the variant

$$A:\ell^\infty_\#(\mathbb{Z})\longrightarrow\ell^\infty_\#(\mathbb{Z}),\quad Af(k)=\frac{1}{2}f(k+1)+\frac{1}{2}f(k-1). \tag{4.21}$$

Using the same letter A in (4.18) and (4.21) is perhaps an abuse of notation, a practice we continue by writing

$$A^t:\mathcal{M}(\widehat{\mathbb{Z}})\longrightarrow\mathcal{M}(\widehat{\mathbb{Z}}). \tag{4.22}$$

In the setting of (4.21), we can write the set of elements of $\mathcal{P}(\widehat{\mathbb{Z}})$ that are invariant under A^t as

$$\{a\delta_{+\infty}+(1-a)\delta_{-\infty}:0\le a\le 1\}. \tag{4.23}$$

While Proposition 2.1 applies to the operators in (4.17), (4.18), and (4.21), most of the rest of §2 does not. For one thing, the operators just mentioned are not compact. Also, the operators in (4.18) and (4.21) have the *weak* irreducibility property that

$$f\in\ell^\infty(\mathbb{Z}),\ f\ge 0,\ f\ne 0\Longrightarrow\sup_m A^mf(k)>0,\quad\forall\,k\in\mathbb{Z}. \tag{4.24}$$

However, whenever $f(k)\to 0$ as $|k|\to\infty$, so does $A^mf(k)$ for each m, so $\sup_m A^mf$ vanishes on $X_{\mathbb{Z}}\backslash\mathbb{Z}$ in case (4.18), and on $\widehat{\mathbb{Z}}\backslash\mathbb{Z}$ in case (4.21), for each such f, violating the condition (1.5) for irreducibility. This also explains why Proposition 3.6 does not apply to A in (4.18) and (4.21). For a positive operator $A : \ell^\infty(S) \to \ell^\infty(S)$, irreducibility in the sense of (1.5), for $A : C(X_S) \to C(X_S)$ (which, for emphasis, we will call *strong* irreducibility) is equivalent to the property

$$f \in \ell^\infty(S),\ f \geq 0,\ f \neq 0 \Longrightarrow \sup_m A^m f(k) \geq \delta(f) > 0, \quad \forall\, k \in S. \tag{4.25}$$

We next look at a family of operators on $\ell^\infty(\mathbb{N})$, examined in [Sen], given by the infinite matrices

$$A = \begin{pmatrix} q_1 & p_1 & & & \cdots \\ q_2 & 0 & p_2 & & \\ q_3 & 0 & 0 & p_3 & \\ \vdots & & & & \ddots \end{pmatrix}, \tag{4.26}$$

with

$$0 < q_j < 1, \quad q_j + p_j = 1. \tag{4.27}$$

Note that

$$\{q_j\} \text{ bounded away from } 0 \Longrightarrow A \text{ is strongly irreducible,} \tag{4.28}$$

as defined above, since then (4.25) holds. One also readily verifies that

$$p_j \to 0 \ \text{ as } \ j \to \infty \Longrightarrow A \ \text{ is compact on } \ \ell^\infty(\mathbb{N}). \tag{4.29}$$

One can further extend the scope of the approach to positive infinite matrices described above. For example, we can consider bounded linear operators

$$A : L^\infty(Y, \sigma) \longrightarrow L^\infty(Y, \sigma), \tag{4.30}$$

where (Y, σ) is a measure space. The positivity condition becomes

$$f \in L^\infty(Y, \sigma),\ f \geq 0 \Longrightarrow Af \geq 0. \tag{4.31}$$

This fits into the framework of §§2–3 as follows. The space $L^\infty(Y, \sigma)$ is a commutative C^*-algebra, and we have a natural, positivity-preserving, isometric isomorphism

$$L^\infty(Y, \sigma) \approx C(X), \tag{4.32}$$

where X is the maximal ideal space of the C^*-algebra $L^\infty(Y, \sigma)$. In (4.2), we have $Y = \mathbb{N}$ and $\sigma =$ counting measure.

References

[DS] N. Dunford and J. Schwartz, Linear Operators, Vol. 1, Wiley, New York, 1958.

[Nel] E. Nelson, The adjoint Markoff process, Duke Math. J. 25 (1958), 671–690.

[Sen] E. Seneta, Non-negative Matrices and Markov Chains, Springer-Verlag New York, 1981.

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

3. Lévy Processes

1. Introduction

Wiener measure is a measure on the space of paths in $\mathbb{R}^n$ having the following property. Consider the Gaussian probability distribution

$$p(t,x) = (2\pi)^{-n} \int e^{-t|\xi|^2} e^{ix\cdot\xi}\, d\xi = (4\pi t)^{-n/2} e^{-|x|^2/4t}. \tag{1.1}$$

Given $0 < t_1 < t_2$ and given that $\omega(t_1) = x_1$, the probability density for the location of $\omega(t_2)$ is $p(t_2 - t_1, x - x_1)$. More generally, given $0 < t_1 < t_2 < \cdots < t_k$ and given Borel sets $E_j \subset \mathbb{R}^n$, the probability that a path, starting at $x = 0$ at time $t = 0$, lies in E_j at time t_j for each $j \in \{1, \dots, k\}$ is

$$\int_{E_1} \cdots \int_{E_k} p(t_k - t_{k-1}, x_k - x_{k-1}) \cdots p(t_1, x_1)\, dx_k \cdots dx_1. \tag{1.2}$$

It takes some effort to prove that there is a countably additive measure characterized by these properties. This was first done by N. Wiener, who also proved that the associated measure, called Wiener measure, is supported on the space of continuous paths. One elegant approach to the construction of Wiener measure is due to [Nel]. Nelson's approach is taken in Chapter 16 of [T2] and also in Chapter 11 of [T1].

Extensions of this theory to non-Gaussian distributions have been pursued by many people, notably P. Lévy. Our main purpose here is to extend the method of [Nel] to treat these Lévy processes. To start, we replace (1.1) by

$$p(t,x) = (2\pi)^{-n} \int e^{-t\psi(\xi)} e^{ix\cdot\xi}\, d\xi = e^{-t\psi(D)}\delta(x). \tag{1.3}$$

Here $\psi(\xi)$ is a function with the property that

$$p(t,x) \ge 0, \quad \forall\, t > 0,\ x \in \mathbb{R}^n. \tag{1.4}$$

We require $\psi(0) = 0$, so

$$\int p(t,x)\, dx = 1, \quad \forall\, t > 0. \tag{1.5}$$

Note also that

$$\int p(t, x-y) p(s,y)\, dy = p(t+s, x). \tag{1.6}$$

The example $\psi(\xi) = |\xi|^2$ gives the Gaussian case, as in (1.1). Other examples include $\psi_\alpha(\xi) = |\xi|^{2\alpha}$, for $\alpha \in (0,1)$. The function $p_\alpha(t,x) = e^{-t\psi_\alpha(D)}\delta(x) = e^{-t(-\Delta)^\alpha}\delta(x)$ is related to (1.1) by the subordination identity

$$e^{-tL^\alpha} = \int_0^\infty \Phi_{t,\alpha}(s) e^{-sL}\, ds, \quad 0 < \alpha < 1, \tag{1.7}$$

valid for any positive self adjoint operator L, where $\Phi_{t,\alpha}$ has the property

$$e^{-t\lambda^\alpha} = \int_0^\infty \Phi_{t,\alpha}(s) e^{-s\lambda}\, ds, \quad \lambda > 0. \tag{1.8}$$

The fact that

$$(-\partial_\lambda)^k e^{-t\lambda^\alpha} \ge 0, \quad \text{for } \ \lambda, t > 0,\ k \in \mathbb{Z}^+, \tag{1.9}$$

implies

$$\Phi_{t,\alpha}(s) \ge 0, \quad \text{for } \ s \in [0,\infty), \tag{1.10}$$

given $t \in (0,\infty),\ \alpha \in (0,1)$. One also has

$$\int_0^\infty \Phi_{t,\alpha}(s)\, ds = 1. \tag{1.11}$$

This is discussed in a more general context in §IX.11 of [Y]. The most familiar case is the case $\alpha = 1/2$, where

$$\Phi_{t,1/2}(s) = \frac{t}{2\pi^{1/2}}\, e^{-t^2/4s}\, s^{-3/2}; \tag{1.12}$$

compare [T1], Chapter 3, (5.22)–(5.31).

The positivity (1.10) implies positivity in (1.4) when $\psi(\xi) = |\xi|^{2\alpha},\ \alpha \in (0,1)$. There is a good characterization of functions $\psi(\xi)$ for which (1.4) holds, the Lévy-Khinchin formula. We discuss this in Appendix A of this chapter, with further details on such homogeneous cases as $|\xi|^{2\alpha}$ in Appendix B. Here we give another example, arising from applying (1.7) to $e^{t(\Delta-1)}$. This leads to

$$\varphi_\alpha(\xi) = (|\xi|^2 + 1)^\alpha - 1, \quad 0 < \alpha < 1. \tag{1.13}$$

We contrast $e^{-t\varphi_\alpha(D)}\delta(x)$ with $e^{-t\psi_\alpha(D)}\delta(x)$. The latter has a "heavy tail":

$$e^{-t\psi_\alpha(D)}\delta(x) \sim C_{n\alpha t}|x|^{-n-2\alpha}, \quad |x| \to \infty. \tag{1.14}$$

This is contrasted with the exponential decay:

$$e^{-t\varphi_\alpha(D)}\delta(x) \sim C'_{n\alpha t} e^{-|x|}, \quad |x| \to \infty, \tag{1.15}$$

which is much more rapid decay than in (1.14), though not as rapid as the decay in (1.1). Further results on heavy tails are given in Appendix C, and a detailed analysis of the long-time and short-time behavior of (1.15) in Appendix D.

Appendix E discusses vanishing and superexponential decay on cones. Appendix F treats regularity properties of the semigroup $e^{-t\psi(D)}$.

Results mentioned above all deal with Levy processes on Euclidean space. Appendix M extends the notion to Riemannian manifolds. Here we emphasize constructions that extend elements arising in the Levy-Khinchin formula, but such variants quickly lead further afield, and it becomes natural to include Appendix N, discussing the production of more general Markov semigroups. Material here includes continuous-time finite Markov chains and denumerable Markov chains, as well as more general cases.

2. Construction of the probability measures

We will anticipate that the stochastic processes to be constructed here are determined by their values at positive *rational* t. Thus we consider the set of "paths"

$$\mathfrak{P} = \prod_{t\in\mathbb{Q}^+} \dot{\mathbb{R}}^n. \tag{2.1}$$

Here $\dot{\mathbb{R}}^n$ is the one point compactification of $\mathbb{R}^n$. Thus $\mathfrak{P}$ is a compact metrizable space. For each $\psi(\xi)$ such that (1.3)–(1.5) holds, we associate a probability measure on $\mathfrak{P}$.

In order to construct this measure, we will construct a certain positive linear functional $E : C(\mathfrak{P}) \to \mathbb{R}$, on the space $C(\mathfrak{P})$ of real valued continuous functions on $\mathfrak{P}$, satisfying $E(1) = 1$, and a condition motivated by (1.2), which we give below. We first define E on the space $\mathcal{C}^\#$ consisting of continuous functions on $\mathfrak{P}$ of the form

$$\varphi(\omega) = F\big(\omega(t_1), \dots, \omega(t_k)\big), \quad t_1 < \cdots < t_k, \tag{2.2}$$

where F is continuous on $\prod_1^k \dot{\mathbb{R}}^n$, and $t_j \in \mathbb{Q}^+$. Motivated by (1.2), we take

$$\begin{aligned} E(\varphi) = \int \cdots \int p(t_1, x_1) p(t_2 - t_1, x_2 - x_1) \cdots p(t_k - t_{k-1}, x_k - x_{k-1}) \\ F(x_1, \dots, x_k)\, dx_k \cdots dx_1. \end{aligned} \tag{2.3}$$

If $\varphi(\omega)$ in (2.2) actually depends on $\omega(t_\nu)$ for some proper subset $\{t_\nu\}$ of $\{t_1, \dots, t_k\}$, there arises a formula for $E(\varphi)$ with a different appearance from (2.3). The fact that these two expressions are equal follows from the identity (1.6). From this it follows that $E : \mathcal{C}^\# \to \mathbb{R}$ is well defined. It is also a positive functional, satisfying $E(1) = 1$.

Now, by the Stone-Weierstrass Theorem, $\mathcal{C}^{\#}$ is dense in $C(\mathfrak{P})$. Since $E : \mathcal{C}^{\#} \to \mathbb{R}$ is a positive linear functional and $E(1) = 1$, it follows that E has a unique continuous extension to $C(\mathfrak{P})$, possessing these properties. The Riesz representation theorem associates to E a probability measure W. Therefore we have:

Theorem 2.1. *Given $p(t,x) = e^{-t\psi(D)}\delta(x)$ satisfying (1.4)–(1.5), there is a unique probability measure W on $\mathfrak{P}$ such that (2.3) is given by*

$$E(\varphi) = \int_{\mathfrak{P}} \varphi(\omega)\, dW(\omega), \tag{2.4}$$

for each $\varphi(\omega)$ of the form (2.2) with F continuous on $\prod_1^k \dot{\mathbb{R}}^n$. In such a case, (2.3) then holds for any bounded Borel function F, and also for any positive Borel function F, on $\prod_1^k \dot{\mathbb{R}}^n$.

Let us do some basic examples of calculations of (2.4). Define functions X_t on $\mathfrak{P}$, taking values in $\dot{\mathbb{R}}^n$, by

$$X_t(\omega) = \omega(t). \tag{2.5}$$

We see that if $0 < s < t,\ q \in \mathbb{R}$,

$$\begin{aligned} E(|X_t - X_s|^q) &= \iint p(s,x_1)p(t-s,x_2-x_1)\,|x_2-x_1|^q\,dx_2\,dx_1 \\ &= \int p(t-s,y)|y|^q\,dy, \end{aligned} \tag{2.6}$$

making the change of variable $y = x_2 - x_1,\ z = x_1$ and using (1.5).

Let us specialize to $\psi(\xi) = \psi_\alpha(\xi) = |\xi|^{2\alpha}$, i.e.,

$$p(t,x) = e^{-t(-\Delta)^\alpha}\delta(x), \quad 0 < \alpha < 1. \tag{2.7}$$

Then $p(t,\cdot)$ is bounded and continuous on $\mathbb{R}^n$ for each $t > 0$ and we have the asymptotic behavior (1.14) as $|x| \to \infty$. We also have

$$p(t,x) = t^{-n/2\alpha}\,p(1,t^{-1/2\alpha}x), \tag{2.8}$$

and hence

$$\begin{aligned} \int p(t,y)|y|^q\,dy &= t^{-n/2\alpha}\int p(1,t^{-1/2\alpha}y)|y|^q\,dy \\ &= C_{n\alpha q}\,t^{q/2\alpha}, \end{aligned} \tag{2.9}$$

where

$$C_{n\alpha q} = \int p(1,y)|y|^q\,dy. \tag{2.10}$$

Since $p(1,y)$ is bounded and

$$p(1,y) \sim C_\alpha |y|^{-n-2\alpha}, \quad |y| \to \infty, \tag{2.11}$$

we have

$$C_{n\alpha q} < \infty \Longleftrightarrow -n < q < 2\alpha, \tag{2.12}$$

given $0 < \alpha < 1$. Of course, in the Gaussian case $\alpha = 1$ one has $C_{n\alpha q} < \infty$ for all $q \in (-n, \infty)$. In light of (2.6), we have

$$E(|X_t - X_s|^q) = C_{n\alpha q} |t-s|^{q/2\alpha}, \quad -n < q < 2\alpha, \ 0 < \alpha < 1. \tag{2.13}$$

If $\alpha = 1$ this extends to all $q \in (-n, \infty)$.

The identity (2.13) measures the distance from X_t to X_s in $L^q(\mathfrak{P}, W)$, provided $q > 0$ and the hypotheses hold to yield $C_{n\alpha q} < \infty$. Note that $L^q(\mathfrak{P}, W)$ is a Banach space for $q \in [1, \infty)$. For $q \in (0,1)$, it is not a Banach space, but it is still a metric space. We see that $t \mapsto X_t$ extends continuously from $\mathbb{Q}^+$ to $\mathbb{R}^+$, yielding a continuous function of t with values in $L^q(\mathfrak{P}, W)$, for $q \in (0, \alpha/2)$.

The following is a useful generalization of (2.6); if $G : \mathbb{R}^n \to \mathbb{R}$ is positive or bounded (and Borel measurable) and $0 < s < t$,

$$\begin{aligned} E(G(X_t - X_s)) &= \iint p(s, x_1) p(t-s, x_2 - x_1) G(x_2 - x_1)\, dx_2\, dx_1 \\ &= \int p(t-s, y) G(y)\, dy \\ &= E(G(X_{t-s})). \end{aligned} \tag{2.14}$$

In other words, $X_t - X_s$ has the same statistical behavior as X_{t-s}. The following result asserts that if $t > s \geq 0$ then $X_t - X_s$ is independent of X_σ for $\sigma \leq s$.

Proposition 2.2. *Assume* $0 < s_1 < \cdots < s_k < s < t$ *and consider functions on* $\mathfrak{P}$ *of the form*

$$\varphi(\omega) = F\big(\omega(s_1), \dots, \omega(s_k)\big), \quad \psi(\omega) = G\big(\omega(t) - \omega(s)\big). \tag{2.15}$$

Then

$$E(\varphi\psi) = E(\varphi) E(\psi). \tag{2.16}$$

Proof. Note that $E(\psi)$ is given by (2.14). Meanwhile, we have

$$\begin{aligned} E(\varphi\psi) = \int & p(s_1, x_1) p(s_2 - s_1, x_2 - x_1) \cdots p(s_k - s_{k-1}, x_k - x_{k-1}) \\ & p(s - s_k, y_1 - x_k) p(t-s, y_2 - y_1) F(x_1, \dots, x_k) \\ & G(y_2 - y_1)\, dx_1 \cdots dx_k\, dy_1\, dy_2. \end{aligned} \tag{2.17}$$

If we change variables to $x_1, \dots, x_k, y_1, z = y_2 - y_1$, then comparison with (2.14) shows that $E(\psi)$ factors out of (2.17). Then use of $\int p(s - s_k, y_1 - x_k)\, dy_1 = 1$ shows that the other factor is equal to $E(\varphi)$, so we have (2.16).

Note the characteristic function calculation

$$E(e^{i\xi\cdot X_t}) = \int p(t,y) e^{iy\cdot\xi}\, dy = e^{-t\psi(\xi)}. \tag{2.18}$$

Then, by (2.14), we have

$$E(e^{i\xi\cdot(X_t - X_s)}) = e^{-|t-s|\psi(\xi)}, \tag{2.19}$$

and an iterative use of (2.16) shows that if $0 < t_1 < \cdots < t_k$ and $\xi_j \in \mathbb{R}^n$, then

$$\begin{aligned} &E\big(e^{i\xi_1\cdot X_{t_1} + i\xi_2\cdot(X_{t_2} - X_{t_1}) + \cdots + i\xi_k\cdot(X_{t_k} - X_{t_{k-1}})}\big) \\ &= e^{-t_1\psi(\xi_1) - (t_2 - t_1)\psi(\xi_2) - \cdots - (t_k - t_{k-1})\psi(\xi_k)}. \end{aligned} \tag{2.20}$$

3. Stochastic continuity and regularity of paths

In §2 we constructed a probability space $(\mathfrak{P}, W)$ and a family X_t of random variables on $\mathfrak{P}$, given by (2.5), when $t \in \mathbb{Q}^+$. We indicated how to extend X_t to $t \in \mathbb{R}^+$ in case $\psi(\xi) = |\xi|^{2\alpha},\ 0 < \alpha \le 1$. We begin this section by making such an extension for general $\psi(\xi)$ treated in Theorem 2.1, obtaining a *stochastically continuous* family of random variables on $\mathfrak{P}$.

This is obtained in a fashion parallel to (2.7)–(2.13), with $|y|^q$ replaced by a different function $G(y)$, namely

$$G(y) = 1 - e^{-|y|} = 1 - g(y). \tag{3.1}$$

By (2.14) we have (at first for $s, t \in \mathbb{Q}^+$),

$$\begin{aligned} E(G(X_t - X_s)) &= \int p(|t-s|, y) G(y)\, dy \\ &= 1 - \int p(|t-s|, y) g(y)\, dy = \vartheta(|t-s|). \end{aligned} \tag{3.2}$$

Note that $\hat{g} \in L^1(\mathbb{R}^n)$ and

$$\int p(t,y) g(y)\, dy = (2\pi)^{-n} \int e^{-t\psi(\xi)} \hat{g}(\xi)\, d\xi. \tag{3.3}$$

Also, the function ψ satisfies

$$\operatorname{Re}\psi(\xi) \ge 0, \tag{3.4}$$

so the Lebesgue dominated convergence theorem applies, to give $\lim_{t\searrow 0} \int p(t,y)g(y)\,dy = 1$, and hence

$$E(G(X_t - X_s)) = \vartheta(|t-s|), \quad \lim_{t\to 0} \vartheta(t) = 0. \tag{3.5}$$

Observe that

$$\rho(X,Y) = E(G(X-Y)) \tag{3.6}$$

yields a *metric* on the space $M(\mathfrak{P}, W)$ of equivalence classes of measurable $\mathbb{R}^n$-valued functions on $\mathfrak{P}$, as a consequence of the monotonicity and concavity of $r \to 1 - e^{-r}$ on $[0,\infty)$. This metric defines the topology of convergence in measure on $\mathfrak{P}$.

In fact, $M(\mathfrak{P}, W)$ is a complete metric space with the metric (3.6). Given a Cauchy sequence, one can take a subsequence (Y_j) satisfying $\rho(Y_j, Y_{j+k}) \le 4^{-j}$, $\forall\, k \ge 1$. This sequence converges pointwise a.e. to a limit $Y \in M(\mathfrak{P}, W)$, by virtue of the estimate

$$W\Big(\{\omega \in \mathfrak{P} : |Y_j(\omega) - Y_{j+k}(\omega)| \ge 2^{-j}\}\Big) \le C\, 2^{-j},$$

and convergence also takes place in ρ-metric.

Given this completeness, the estimate (3.5) implies there is a unique continuous extension of $t \mapsto X_t$ from $\mathbb{Q}^+ \to M(\mathfrak{P}, W)$ to $\mathbb{R}^+ \to M(\mathfrak{P}, W)$. There results a stochastically continuous process $\{X_t : t \in \mathbb{R}^+\}$.

The result (3.5) is quite general, but it is much weaker than such results as (2.13). Here we mention some stronger results, valid for certain interesting classes of Lévy processes.

Stochastic continuity of Poisson processes

The Poisson process has transition probabilities given by

$$p(t,x)\,dx = \sum_{k=0}^{\infty} \frac{t^k}{k!} e^{-t} \delta_{ky}(x). \tag{3.7}$$

We can use the formula (2.6), and take $q = 1$, to get

$$\begin{aligned} E(|X_t - X_s|) &= \sum_{k=0}^{\infty} \frac{|t-s|^k}{k!} e^{-|t-s|} |yk| \\ &= |y| \sum_{k=1}^{\infty} \frac{|t-s|^k}{(k-1)!} e^{-|t-s|} \\ &= |y| \cdot |t-s|. \end{aligned} \tag{3.8}$$

Taking $q = 2$ in (2.6(, we have

$$
\begin{aligned}
E(|X_t - X_s|^2) &= |y|^2 \sum_{k=0}^{\infty} \frac{|t-s|^k}{k!} k^2 e^{-|t-s|} \\
&= |y|^2 \sum_{k=1}^{\infty} \frac{|t-s|^k}{(k-1)!} k e^{-|t-s|} \\
&= |y|^2 \Big\{ |t-s| e^{-|t-s|} + \sum_{k=2}^{\infty} \frac{|t-s|^k}{(k-2)!} \frac{k}{k-1} e^{-|t-s|} \Big\} \\
&\le |y|^2 \big\{ |t-s| + 2|t-s|^2 \big\}.
\end{aligned} \tag{3.9}
$$

Similarly one can estimate $E(|X_t - X_s|^k)$, for $k \ge 3$.

Stochastic continuity of modified fractional diffusion processes

In (2.7)–(2.13) we looked at the fractional diffusion process, for which $\psi(\xi) = |\xi|^{2\alpha},\ \alpha \in (0,1)$. Here we take $\psi(\xi) = \varphi_\alpha(\xi)$, where

$$
\varphi_\alpha(\xi) = (|\xi|^2 + 1)^\alpha - 1, \quad \alpha \in (0,1). \tag{3.10}
$$

In this case we get finite results for $E(|X_t - X_s|^q)$, for q larger than allowed in (2.13). In fact, use of the Parseval identity for the Fourier transform gives

$$
\begin{aligned}
E(|X_t - X_s|^2) &= -\Delta e^{-|t-s|\varphi_\alpha(\xi)}\Big|_{\xi=0} \\
&= |t-s| \Delta\varphi_\alpha(0),
\end{aligned} \tag{3.11}
$$

and more generally, for $k \in \mathbb{N}$,

$$
\begin{aligned}
E(|X_t - X_s|^{2k}) &= (-\Delta)^k e^{-|t-s|\varphi_\alpha(\xi)}\Big|_{\xi=0} \\
&= -|t-s|(-\Delta)^k \varphi_\alpha(0) + O(|t-s|^2),
\end{aligned} \tag{3.12}
$$

as $|t-s| \to 0$. This also works for $\alpha = 1$, where $\varphi_1(\xi) = |\xi|^2$ and we get the Wiener process. Note that

$$
k \ge 2 \Longrightarrow \Delta^k \varphi_1(0) = 0, \tag{3.13}
$$

which has a profound effect on the behavior of (3.12) as $|t-s| \to 0$, for $\alpha = 1$.

Appendix D of this chapter has further results on the behavior of $E(|X_t - X_s|^q)$, for q in the range $(-n, 2\alpha)$, for these processes.

The issue of path continuity

Regarding the behavior of individual paths $t \mapsto X_t(\omega)$, there is the following result of Kolmogorov. For a proof see [Kry], p. 20.

Proposition 3.1. *Suppose there exist* $q, \beta > 0,\ C < \infty$ *such that*

$$E(|X_t - X_s|^q) \le C|t-s|^{1+\beta}, \quad \forall\, s,t \ge 0. \tag{3.14}$$

Then the process $\{X_t\}$ *has a modification almost all of whose paths are continuous.*

Note that in (2.13) this estimate just barely fails, if one requires $q < 2\alpha$. As noted below (2.13), such an estimate works in the Gaussian case for all $q \in (-n, \infty)$, so (3.14) works there, which gives pathwise continuity for the Wiener process. For other Lévy processes, path continuity fails, but another result holds.

One says a path $t \mapsto \gamma(t)$ from $\mathbb{R}^+$ to $\mathbb{R}^n$ is cadlag provided that for each $t \in \mathbb{R}^+$,

$$\lim_{s \searrow t} \gamma(s) = \gamma(t), \quad \text{and} \quad \lim_{s \nearrow t} \gamma(s) \text{ exists,} \tag{3.15}$$

though the latter limit need not equal $\gamma(t)$. The following result is proven in [Kry], p. 136.

Proposition 3.2. *If* $\{X_t : t \in \mathbb{R}^+\}$ *is a stochastically continuous process with independent increments, then it admits a modification such that almost all paths are cadlag.*

4. Hausdorff dimension of Lévy paths and Lévy graphs

We restrict attention to Lévy processes on $\mathbb{R}^n$ generated by $(-\Delta)^\alpha$, with $\alpha \in (0,1]$. We estimate from below the Hausdorff dimension $\operatorname{Hdim} \omega(I)$ for a typical path $\omega(I) = \{\omega(t) : t \in I\}$, where $I = [0,T],\ T \in (0,\infty)$. We will show that for each such I,

$$\operatorname{Hdim} \omega(I) \ge \min(2\alpha, n), \quad \text{for a.e. } \ \omega. \tag{4.1}$$

Actually it is known that equality holds (see [Sat]), but we will not establish the reverse inequality. (For $\alpha = 1$ the reverse inequality is an immediate consequence of the modulus of continuity.) We will also estimate the Hausdorff dimension of a graph:

$$Z_\omega(t) = (t, \omega(t)). \tag{4.2}$$

With $Z_\omega(I) = \{Z_\omega(t) : t \in I\}$, we obtain the following estimates on $\operatorname{Hdim} Z_\omega(I)$. Namely,

$$n \ge 2,\ \frac{1}{2} \le \alpha \le 1 \Longrightarrow \operatorname{Hdim} Z_\omega(I) \ge 2\alpha, \tag{4.3}$$

for almost all ω, while

$$n = 1,\ \frac{1}{2} < \alpha \le 1 \Longrightarrow \operatorname{Hdim} Z_\omega(I) \ge 2 - \frac{1}{2\alpha}, \tag{4.4}$$

and for each $n \ge 1$,

$$0 < \alpha \le \frac{1}{2} \Longrightarrow \operatorname{Hdim} Z_\omega(I) \ge 1. \tag{4.5}$$

Perhaps one has equality in (4.3)–(4.5), for almost all ω, but we do not show this.

One tool we use to prove these estimates is the following (cf. [Fal], p. 78).

Lemma 4.1. *Let $K \subset \mathbb{R}^n$ be a compact set and take $b \in (0,\infty)$. Assume there is a positive Borel measure $\mu \neq 0$, supported on K, such that*

$$\iint \frac{d\mu(x)\, d\mu(y)}{|x-y|^b} < \infty. \tag{4.6}$$

Then the b-dimensional Hausdorff measure $\mathcal{H}^b(K) = \infty$, so Hdim $K \geq b$.

To prove (4.1), we use the following consequence of (2.13):

$$E\big(|X_s - X_t|^{-b}\big) = C_{n\alpha b}|t-s|^{-b/2\alpha}, \quad 0 < b < n, \tag{4.7}$$

with $C_{n\alpha b} < \infty$ in this range. Consequently

$$E\left(\int_0^T \int_0^T \frac{ds\, dt}{|\omega(t)-\omega(s)|^b}\right) = C \int_0^T \int_0^T \frac{ds\, dt}{|t-s|^{b/2\alpha}} < \infty, \tag{4.8}$$

provided $0 < b < 2\alpha$, so

$$0 < b < \min(2\alpha, n) \Longrightarrow \int_0^T \int_0^T \frac{ds\, dt}{|\omega(t)-\omega(s)|^b} < \infty, \quad \text{for } W\text{-a.e. } \omega. \tag{4.9}$$

Now we define the measure μ^ω on $\omega([0,T])$ by

$$\mu^\omega(S) = m\big(\{t \in [0,T] : \omega(t) \in S\}\big), \tag{4.10}$$

where m denotes Lebesgue measure on $[0,T]$. Thus (4.9) becomes

$$\int\limits_{\omega(I)} \int\limits_{\omega(I)} \frac{d\mu^\omega(x)\, d\mu^\omega(y)}{|x-y|^b} < \infty, \quad \text{for a.e. } \omega, \quad \text{if } \ 0 < b < \min(2\alpha, n). \tag{4.11}$$

While $\omega(I)$ is not compact (unless $\alpha = 1$), a modification of Lemma 4.1 should apply, to yield (4.1).

Moving on to graphs, we have

$$\begin{aligned} E\big(|Z(s) - Z(t)|^{-b}\big) &= \int p(|t-s|, y)\big(|t-s| + |y|\big)^{-b}\, dy \\ &= |t-s|^{-n/2\alpha} \int p(1, |t-s|^{-1/2\alpha} y)\big(|t-s| + |y|\big)^{-b}\, dy \\ &= \int p(1,z)\big(|t-s| + |t-s|^{1/2\alpha}|z|\big)^{-b}\, dz. \end{aligned} \tag{4.12}$$

Hence, since $p(1,\cdot)$ is integrable,

$$\begin{aligned} E\big(|Z(s) - Z(t)|^{-b}\big) &= |t-s|^{-b/2\alpha} \int p(1,z)\big(|t-s|^{1-1/2\alpha} + |z|\big)^{-b}\, dz \\ &\leq C|t-s|^{-b/2\alpha}\left[1 + \int \big(|t-s|^{1-1/2\alpha} + |z|\big)^{-b}\, dz\right]. \end{aligned} \tag{4.13}$$

If $0 < b < n$ and $1 - 1/2\alpha \geq 0$, the last integral is bounded independently of $s, t \in [0, T]$, and one has

(4.14)

$$0 < b < n,\ \frac{1}{2} \leq \alpha \leq 1 \Longrightarrow E\big(|Z(s) - Z(t)|^{-b}\big) \leq C|t - s|^{-b/2\alpha}$$
$$\Longrightarrow E\left(\int_0^T \int_0^T \frac{ds\,dt}{|Z(t) - Z(s)|^b}\right) \leq C \int_0^T \int_0^T \frac{ds\,dt}{|s - t|^{b/2\alpha}},$$

which is $< \infty$ provided $b < 2\alpha$. An argument parallel to that using (4.10)–(4.11) then yields

$$\text{Hdim}\, Z_\omega(I) \geq b, \quad \forall\, b < \min(2\alpha, n), \quad \text{if} \quad \alpha \geq \frac{1}{2}, \tag{4.15}$$

for almost all ω, which in turn gives (4.3).

Now assume that $1 = n < b$, while $\alpha \in (1/2, 1]$. To estimate the last integral in (4.3), write

$$\begin{aligned}\int_0^T \big(|t - s|^{1-1/2\alpha} + z\big)^{-b}\, dz &\leq \int_0^{|t-s|^\gamma} |t - s|^{-b(1-1/2\alpha)}\, dz + \int_{|t-s|^\gamma}^T z^{-b}\, dz \\ &\leq C + |t - s|^{\gamma - b(1-1/2\alpha)} + C|t - s|^{\gamma(1-b)}.\end{aligned} \tag{4.16}$$

Pick $\gamma = 1 - 1/2\alpha$ to make the exponents both equal to $(1 - b)(1 - 1/2\alpha)$. Hence

$$\begin{aligned}E\big(|Z(s) - Z(t)|^{-b}\big) &\leq C|t - s|^{-b/2\alpha}\big(1 + |t - s|^{(1-b)(1-1/2\alpha)}\big) \\ &\leq C|t - s|^{1-b-1/2\alpha} + C|t - s|^{-b/2\alpha}.\end{aligned} \tag{4.17}$$

Thus

$$E\left(\int_0^T \int_0^T \frac{ds\,dt}{|Z(s) - Z(t)|^b}\right) \leq C \int_0^T \int_0^T \left[\frac{ds\,dt}{|t - s|^{b+1/2\alpha-1}} + \frac{ds\,dt}{|t - s|^{b/2\alpha}}\right], \tag{4.18}$$

which is $< \infty$ provided $b < 2 - 1/2\alpha$. (Note that $\alpha \in (1/2, 1] \Rightarrow 2 - 1/2\alpha < 2\alpha$.) This plus another application of Lemma 4.1 (suitably modified) gives (4.4).

We now turn to the case $\alpha \in (0, 1/2]$. In that case, replace (4.13) by

$$\begin{aligned}E\big(|Z(s) - Z(t)|^{-b}\big) &= |t - s|^{-b} \int p(1, z)\big(1 + |t - s|^{1/2\alpha - 1}|z|\big)^{-b}\, dz \\ &\leq C|t - s|^{-b}.\end{aligned} \tag{4.19}$$

Thus

$$E\left(\int_0^T \int_0^T \frac{ds\,dt}{|Z(s) - Z(t)|^b}\right) \leq C \int_0^T \int_0^T \frac{ds\,dt}{|t - s|^b}, \tag{4.20}$$

which is $< \infty$ provided $b < 1$. Then a third application of Lemma 4.1 (suitably modified) gives (4.5).

A. Generators of Lévy processes

Given a function $\psi(\xi)$ on $\mathbb{R}^n$, we say $\psi(D)$ generates a Lévy process if $p(t,x) = e^{-t\psi(D)}\delta(x)$ satisfies

$$p(t,x) \ge 0, \quad \int p(t,x)\, dx = 1, \tag{A.1}$$

for all $t > 0$. (We allow $p(t,\cdot)$ to be a positive measure.) Examples include $\xi \cdot A\xi = \sum a_{jk}\xi_j\xi_k$, when $A = (a_{jk})$ is a positive semi-definite matrix, yielding Gaussians. Another family is $\psi(\xi) = ib \cdot \xi$, generating translations. Still another type is

$$\psi(\xi) = c(1 - e^{iy\cdot\xi}), \tag{A.2}$$

given $c \in (0,\infty)$ and $y \in \mathbb{R}^n$, generating a "Poisson process." In such a case we have

$$e^{-t\psi(\xi)} = \sum_{k=0}^{\infty} \frac{(ct)^k}{k!}\, e^{-ct}\, e^{iky\cdot\xi}. \tag{A.3}$$

Hence $e^{-t\psi(D)}\delta(x)$ is a countable sum of point masses, supported on $\{ky : k = 0,1,2,\dots\}$.

In light of the identity

$$e^{-t(\psi_1(D)+\psi_2(D))} = e^{-t\psi_1(D)}e^{-t\psi_2(D)}, \tag{A.4}$$

it is clear that positive superpositions of the various generators described above are also generators of Lévy processors. P. Lévy showed that this class, suitably completed, yields all such generators. (His proof was simplified by Khinchin.) The resulting formula

$$\psi(\xi) = \xi \cdot A\xi + ib \cdot \xi + \int_{\mathbb{R}^n} \Bigl(1 - e^{iy\cdot\xi} + iy \cdot \xi\chi_B(y)\Bigr)\, d\mu(y) \tag{A.5}$$

is called the Lévy-Khinchin formula. Here χ_B is one on the unit ball B and zero on the complement, and μ is a positive measure on $\mathbb{R}^n \setminus 0$ satisfying

$$\int (|y|^2 \wedge 1)\, d\mu(y) < \infty. \tag{A.6}$$

Often it is useful to modify the term $iy\cdot\xi\chi_B$; sometimes one will drop it altogether (i.e., absorb it into the term $ib\cdot\xi$). Examples of such modifications are given in (B.2)–(B.3) of the next appendix, where we discuss homogeneous generators.

We end this section with a brief discussion of radial generators. If $\psi(\xi)$ is a radial function of the form (A.5), we have

$$\psi(\xi) = a|\xi|^2 + \int_0^\infty (1-\psi_n(s|\xi|))\,d\rho(s), \tag{A.7}$$

where $a \geq 0$ and

$$\int_{S^{n-1}} e^{iy\cdot\xi}\,dS(y) = \psi_n(|\xi|) = (2\pi)^{n/2}|\xi|^{1-n/2}\,J_{n/2-1}(|\xi|). \tag{A.8}$$

Here ρ is a positive measure on $(0,\infty)$ satisfying $\int_0^\infty (s^2\wedge 1)\,d\rho(s) < \infty$. In case $n=1$, (A.7) takes the form

$$\psi(\xi) = a\xi^2 + \int_0^\infty (1-\cos s|\xi|)\,d\rho(s). \tag{A.9}$$

B. Homogeneous Lévy generators

Here we construct functions homogeneous of degree $\alpha\in(0,2)$ for which $p(t,x) = e^{-t\psi(D)}\delta(x)$ satisfies (A.1). Of course

$$\psi(\xi) = |\xi|^\alpha, \quad 0\leq\alpha\leq 2, \tag{B.1}$$

works, by the results of §1. We obtain further cases by specializing natural variants of the Lévy-Khinchin formula (A.5). In this way we obtain the following such homogeneous generators:

$$\Phi_{\alpha,g}(\xi) = -\int_{\mathbb{R}^n} (e^{iy\cdot\xi}-1)g(y)|y|^{-n-\alpha}\,dy, \qquad 0<\alpha<1, \tag{B.2}$$

$$\Psi_{\alpha,g}(\xi) = -\int_{\mathbb{R}^n} (e^{iy\cdot\xi}-1-iy\cdot\xi)g(y)|y|^{-n-\alpha}\,dy, \quad 1<\alpha<2. \tag{B.3}$$

The function g is assumed to be positive, bounded, and homogeneous of degree 0, i.e.,

$$g\geq 0, \quad g\in L^\infty(\mathbb{R}^n), \quad g(ry)=g(y), \quad \forall\, r>0. \tag{B.4}$$

It is easy to verify that both integrals in (B.2)–(B.3) are absolutely convergent, and, for $r > 0$,

$$\text{(B.5)} \qquad \begin{aligned} \Phi_{\alpha,g}(r\xi) &= r^{\alpha}\Phi_{\alpha,g}(\xi), \quad 0 < \alpha < 1, \\ \Psi_{\alpha,g}(r\xi) &= r^{\alpha}\Psi_{\alpha,g}(\xi), \quad 1 < \alpha < 2. \end{aligned}$$

When $g \equiv 1$ we obtain a positive multiple of (B.1).

We now specialize to $n = 1$ and $g = \chi_{\mathbb{R}^+}$, so we look at

$$\text{(B.6)} \qquad \begin{aligned} \varphi_\alpha(\xi) &= -\int_0^\infty (e^{iy\xi} - 1)y^{-1-\alpha}\, dy, \qquad\qquad 0 < \alpha < 1, \\ \psi_\alpha(\xi) &= -\int_0^\infty (e^{iy\xi} - 1 - iy\xi)y^{-1-\alpha}\, dy, \quad 1 < \alpha < 2. \end{aligned}$$

Clearly φ_α and ψ_α are holomorphic in $\{\xi \in \mathbb{C} : \operatorname{Im}\xi > 0\}$, and homogeneous of degree α in ξ. Also, for $\eta > 0$,

$$\text{(B.7)} \qquad \begin{aligned} \varphi_\alpha(i\eta) &= -\int_0^\infty (e^{-y\eta} - 1)y^{-1-\alpha}\, dy > 0, \qquad\qquad 0 < \alpha < 1, \\ \psi_\alpha(i\eta) &= -\int_0^\infty (e^{-y\eta} - 1 + y\eta)y^{-1-\alpha}\, dy < 0, \quad 1 < \alpha < 2, \end{aligned}$$

since, for $r > 0$, $1 - r < e^{-r} < 1$. It follows that $\varphi_\alpha(\xi)$ and $\psi_\alpha(\xi)$ are positive multiples of

$$\text{(B.8)} \qquad \begin{aligned} \varphi_\alpha^{\#}(\xi) &= (-i\xi)^\alpha, \qquad 0 < \alpha < 1, \\ \psi_\alpha^{\#}(\xi) &= -(-i\xi)^\alpha, \quad 1 < \alpha < 2, \end{aligned}$$

restrictions to $\mathbb{R}$ of functions holomorphic on $\{\xi \in \mathbb{C} : \operatorname{Im}\xi > 0\}$. Taking instead $g = \chi_{\mathbb{R}^-}$, we obtain positive multiples of

$$\text{(B.9)} \qquad \begin{aligned} \varphi_\alpha^{b}(\xi) &= (i\xi)^\alpha, \qquad 0 < \alpha < 1, \\ \psi_\alpha^{b}(\xi) &= -(i\xi)^\alpha, \quad 1 < \alpha < 2, \end{aligned}$$

restrictions to $\mathbb{R}$ of functions holomorphic on $\{\xi \in \mathbb{C} : \operatorname{Im}\xi < 0\}$, satisfying

$$\text{(B.10)} \qquad \varphi_\alpha^{b}(-i\eta) > 0, \quad \psi_\alpha^{b}(-i\eta) < 0, \quad \forall\, \eta > 0.$$

The functions in (B.8) and (B.9) are well known examples of homogeneous functions $\psi(\xi)$ for which $e^{-t\psi(D)}$ satisfies (A.1). The associated operators $\psi(D)$ are fractional derivatives.

It is also useful to observe the explicit formulas

(B.11)

$$e^{-t\varphi_\alpha^{\#}(\xi)} = e^{-t(\cos \pi\alpha/2)|\xi|^\alpha} \left[\cos\Bigl(t\bigl(\sin\frac{\pi\alpha}{2}\bigr)|\xi|^\alpha\Bigr) + i\sigma(\xi)\sin\Bigl(t\bigl(\sin\frac{\pi\alpha}{2}\bigr)|\xi|^\alpha\Bigr)\right].$$

for $t > 0,\ 0 < \alpha < 1$, where

$$\sigma(\xi) = \operatorname{sgn}\xi, \tag{B.12}$$

and

$$e^{-t\psi_\alpha^{\#}(\xi)} = e^{t(\cos\pi\alpha/2)|\xi|^\alpha}\left[\cos\Bigl(t\bigl(\sin\frac{\pi\alpha}{2}\bigr)|\xi|^\alpha\Bigr) - i\sigma(\xi)\sin\Bigl(t\bigl(\sin\frac{\pi\alpha}{2}\bigr)|\xi|^\alpha\Bigr)\right], \tag{B.13}$$

fior $t > 0,\ 1 < \alpha < 2$. Note that

$$0 < \alpha < 1 \Rightarrow \cos\frac{\pi\alpha}{2} > 0, \quad 1 < \alpha < 2 \Rightarrow \cos\frac{\pi\alpha}{2} < 0, \tag{B.14}$$

so of course we have decaying exponentials in both (B.11) and (B.13). We get similar formulas with # replaced by b, since in fact

$$\varphi_\alpha^b(\xi) = \varphi_\alpha^{\#}(-\xi), \quad \psi_\alpha^b(\xi) = \psi_\alpha^{\#}(-\xi). \tag{B.15}$$

Returning to the general formulas (B.2)–(B.3), we can switch to polar coordinates and write

$$\begin{aligned}\Phi_{\alpha,g}(\xi) &= -\int_{S^{n-1}}\int_0^\infty (e^{is\omega\cdot\xi} - 1)g(\omega)s^{-1-\alpha}\,ds\,dS(\omega),\\ \Psi_{\alpha,g}(\xi) &= -\int_{S^{n-1}}\int_0^\infty (e^{is\omega\cdot\xi} - 1 - is\omega\cdot\xi)g(\omega)s^{-1-\alpha}\,ds\,dS(\omega),\end{aligned} \tag{B.16}$$

and hence

$$\begin{aligned}\Phi_{\alpha,g}(\xi) &= \int_{S^{n-1}} \varphi_\alpha(\omega\cdot\xi)g(\omega)\,dS(\omega),\\ \Psi_{\alpha,g}(\xi) &= \int_{S^{n-1}} \psi_\alpha(\omega\cdot\xi)g(\omega)\,dS(\omega).\end{aligned} \tag{B.17}$$

We can extend the scope, replacing $g(\omega)\,dS(\omega)$ by a general positive, finite Borel measure on S^{n-1}. Taking into account the calculations yielding (B.8)–(B.9), we obtain homogeneous generators satisfying (A.1), of the form

$$\begin{aligned}\Phi_{\alpha,\nu}^b(\xi) &= \int_{S^{n-1}} (i\omega\cdot\xi)^\alpha\,d\nu(\omega), \qquad 0 < \alpha < 1,\\ \Psi_{\alpha,\nu}^b(\xi) &= -\int_{S^{n-1}} (i\omega\cdot\xi)^\alpha\,d\nu(\omega), \quad 1 < \alpha < 2,\end{aligned} \tag{B.18}$$

where ν is a positive, finite Borel measure on S^{n-1}.

It remains to discuss the case $\alpha = 1$. For $n = 1$ it is seen that positive multiples of

$$|\xi| + ia\xi, \quad a \in \mathbb{R}, \tag{B.19}$$

work. Hence the following functions on $\mathbb{R}^n$ work:

$$|\omega \cdot \xi| + ia\omega \cdot \xi, \quad \omega \in S^{n-1}, \ a \in \mathbb{R}.$$

We can take positive superpositions of such functions and, in analogy with (B.18), obtain generators of diffusion semigroups whose negatives are Fourier multiplication by

$$ib \cdot \xi + \Xi_\nu(\xi), \tag{B.20}$$

where $b \in \mathbb{R}^n$ and

$$\Xi_\nu(\xi) = \int\limits_{S^{n-1}} |\omega \cdot \xi| \, d\nu(\omega). \tag{B.21}$$

We now tie in results derived above with material given in Chapters 1–2 of [ST]. For such functions $\psi(\xi)$, homogeneous of degree $\alpha \in (0, 2]$, as constructed above, the probability distributions

$$p_t(x) = e^{-t\psi(D)} \delta(x) \tag{B.22}$$

are known as α-stable distributions. In the notation (1.1.6) of [ST], consider

$$\psi(\xi) = \sigma^\alpha \, |\xi|^\alpha \Big(1 - i\beta (\operatorname{sgn} \xi) \tan \frac{\pi\alpha}{2}\Big), \quad \xi \in \mathbb{R}. \tag{B.23}$$

Here

$$\sigma \in (0, \infty), \quad \beta \in [-1, 1], \tag{B.24}$$

and $\alpha \in (0, 2)$ but $\alpha \neq 1$. Also, take $\mu \in \mathbb{R}$. Then $e^{-\psi(D)+i\mu D}\delta(x)$ is a probability distribution on the line called an α-stable distribution with scale parameter σ, skewness parameter β, and shift parameter μ. It is clear from (B.11)–(B.13) that each function of the form (B.23) is a positive linear combination of $\varphi^\#_\alpha(\xi)$ and $\varphi^b_\alpha(\xi)$ if $\alpha \in (0, 1)$ and a positive linear combination of $\psi^\#_\alpha(\xi)$ and $\psi^b_\alpha(\xi)$ if $\alpha \in (1, 2)$.

In case $\alpha = 1$, one goes beyond $\psi(\xi)$ homogeneous of degree 1 in ξ, to consider

$$\psi(\xi) = \sigma \, |\xi| \Big(1 + i\frac{2\beta}{\pi} (\operatorname{sgn} \xi) \log |\xi|\Big) + i\mu\xi, \quad \xi \in \mathbb{R}, \tag{B.25}$$

again with $\beta \in [-1,1]$, $\mu \in \mathbb{R}$. Then $e^{-\psi(D)}\delta(x)$ is a probability distribution on $\mathbb{R}$ called a 1-stable distribution, with scale parameter σ, skewness β, and shift μ. The cases arising from (B.19) all have skewness $\beta = 0$.

Similarly, functions $\psi(\xi)$ of the form (B.18) and (B.20)–(B.21) produce probability distributions $e^{-\psi(D)}\delta(x)$ on $\mathbb{R}^n$ that are α-stable. These, plus analogues with a shift incorporated, comprise all of them except when $\alpha = 1$, in which case one generalizes (5.21) to

$$\widetilde{\Xi}_\nu(\xi) = \int\limits_{S^{n-1}} |\omega\cdot\xi|\Big(1 + \frac{2i}{\pi}(\operatorname{sgn}\omega\cdot\xi)\log|\omega\cdot\xi|\Big)\,d\nu(\omega). \tag{B.26}$$

Compare (2.3.1)–(2.3.2) in [ST].

We return to the case $n = 1$ and make some more comments on the probability distributions

$$\begin{aligned} p_t^\alpha(x) &= e^{-t\varphi_\alpha^\#(D)}\delta(x), \quad 0 < \alpha < 1, \\ p_t^\alpha(x) &= e^{-t\psi_\alpha^\#(D)}\delta(x), \quad 1 < \alpha < 2, \end{aligned} \tag{B.27}$$

and their variants with $\#$ replaced by b, which are simply $p_t^\alpha(-x)$. Explicitly, we have

$$p_t^\alpha(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ix\cdot\xi - t\varphi_\alpha^\#(\xi)}\,d\xi, \tag{B.28}$$

for $0 < \alpha < 1$, with $\varphi_\alpha^\#(\xi)$ replaced by $\psi_\alpha^\#(\xi)$ for $1 < \alpha < 2$. Recall that $\varphi_\alpha^\#$ and $\psi_\alpha^\#$ are holomorphic in $\{\xi \in \mathbb{C} : \operatorname{Im}\xi > 0\}$. It follows from the Paley-Wiener theorem that, for each $t > 0$,

$$p_t^\alpha(x) = 0, \quad \text{for } \ x \in [0,\infty), \ 0 < \alpha < 1. \tag{B.29}$$

This theorem does not apply when $\alpha \in (1,2)$, but a shift in the contour of integration to $\{\xi + ib : \xi \in \mathbb{R}\}$, with arbitrary $b > 0$ yields

$$p_t^\alpha(x) = e^{-bx} \times \text{ bounded function of } x, \tag{B.30}$$

for $x \in \mathbb{R}$, whenever $1 < \alpha < 2$, hence

$$p_t^\alpha(x) = o(e^{-bx}), \quad \forall\, b > 0, \ \text{ as } \ x \to +\infty, \ \text{ for } \ 1 < \alpha < 2. \tag{B.31}$$

A more precise asymptotic behavior is stated in (1.2.11) of [ST]. See also results in §E.

We also note that, for $\alpha \in (1,2)$, $p_t^\alpha(x)$ is real analytic in $x \in \mathbb{R}$, and in fact extends to an entire holomorphic function in $x \in \mathbb{C}$, for each $t > 0$, due to rapidity with which $\operatorname{Re}\psi_\alpha^\#(\xi) \to +\infty$ as $|\xi| \to \infty$, which of course forbids (B.29) in this case.

C. Asymptotic behavior of a class of stable distributions (heavy tails)

A fair number of Lévy generators $\psi(D)$, producing probability distributions $p(t,x) = e^{-t\psi(D)}\delta(x)$, have the following properties:

$$\psi \in C^\infty(\mathbb{R}^n \setminus 0), \tag{C.1}$$

$$\operatorname{Re}\psi(\xi) \geq C|\xi|^\beta, \quad \text{for some} \quad \beta \in (0,2), \ C > 0, \tag{C.2}$$

$$\psi(\xi) \sim \sum_{k\geq 0} a_k\Bigl(\frac{\xi}{|\xi|}\Bigr)\,|\xi|^{\gamma+k}, \quad |\xi| \to 0, \quad \text{for some} \quad \gamma \in (0,2), \tag{C.3}$$

with $a_k \in C^\infty(S^{n-1})$, (C.3) implying that $\psi(\xi) - \sum_{k=0}^m a_k(\xi/|\xi|)|\xi|^{\gamma+k} \in C^m(\mathbb{R}^n)$ for each m. Our goal is to derive the asymptotic behavior of $p(t,x)$ as $|x| \to \infty$, for fixed $t > 0$, under these hypotheses.

To start, we can write

$$e^{-t\psi(\xi)} = A_t(\xi) + B_t(\xi), \tag{C.4}$$

where, for each $t > 0$,

$$A_t \in \mathcal{S}(\mathbb{R}^n), \quad \operatorname{supp} B_t \subset \{\xi : |\xi| \leq 1\}, \tag{C.5}$$

and

$$B_t(\xi) \sim 1 + \sum_{j\geq 1, k\geq 0} b_{jkt}\Bigl(\frac{\xi}{|\xi|}\Bigr)|\xi|^{j\gamma+k}, \quad |\xi| \to 0. \tag{C.6}$$

In such a case,

$$p(t,x) = \widehat{A}_t(x) + \widehat{B}_t(x), \quad \widehat{A}_t \in \mathcal{S}(\mathbb{R}^n), \tag{C.7}$$

so the asymptotic behavior of $p(t,x)$ as $|x| \to \infty$, for fixed $t > 0$, is given by that of $\widehat{B}_t(x)$. Now if we set

$$B_{jkt}(\xi) = b_{jkt}\Bigl(\frac{\xi}{|\xi|}\Bigr)|\xi|^{j\gamma+k}, \quad \xi \in \mathbb{R}^n, \tag{C.8}$$

then $B_{jkt} \in \mathcal{S}'(\mathbb{R}^n)$, and if $\Phi \in C_0^\infty(\mathbb{R}^n)$, $\Phi(\xi) = 1$ for $|\xi| \leq 1$, then

$$B_t(\xi) - \Phi(\xi)\sum_{j=1}^N\sum_{k=0}^N B_{jkt}(\xi) \tag{C.9}$$

has a Fourier transform bounded by $C|x|^{-M}$ as $|x| \to \infty$, with $M = M(N) \to \infty$ as $N \to \infty$. Meanwhile,

$$\big(1 - \Phi(\xi)\big) B_{jkt}(\xi) \in S_1^{j\gamma+k}(\mathbb{R}^n), \tag{C.10}$$

so its Fourier transform is rapidly decreasing as $|x| \to \infty$. (Cf. [T1], Chapter 3, Proposition 8.2.) Hence, for each $t > 0$,

$$\widehat{B}_t(x) \sim \sum_{j\ge 1, k\ge 0} \widehat{B}_{jkt}(x), \quad |x| \to \infty. \tag{C.11}$$

As for $\widehat{B}_{jkt}(x)$, since $B_{jkt}(\xi)$ is a homogeneous element of $\mathcal{S}'(\mathbb{R}^n)$, of degree $j\gamma + k$, and smooth on $\mathbb{R}^n \setminus 0$, i.e.,

$$B_{jkt} \in \mathcal{H}^{\#}_{j\gamma+k}(\mathbb{R}^n), \tag{C.12}$$

in the notation (8.8) of [T1], Chapter 3, we have

$$\widehat{B}_{jkt} \in \mathcal{H}^{\#}_{-n-j\gamma-k}(\mathbb{R}^n), \tag{C.13}$$

by Proposition 8.1 in [T1], Chapter 3, the proof using Proposition 8.2, cited above. In other words,

$$\widehat{B}_{jkt}(x) = b^{\#}_{jkt}\Big(\frac{x}{|x|}\Big)\, |x|^{-n-j\gamma-k}, \tag{C.14}$$

with

$$b^{\#}_{jkt} \in C^\infty(S^{n-1}). \tag{C.15}$$

There are integral formulas for $b^{\#}_{jkt}$ in terms of b_{jkt}, $j\gamma + k$, and n, which we will not record here. (See, e.g., calculations in [Zai].) We have the following conclusion.

Proposition C.1. *If $\psi(\xi)$ satisfies (C.1)–(C.3), then for each $t > 0$, $p(t, x) = e^{-t\psi(D)}\delta(x)$ is C^∞ in x and satisfies*

$$p(t, x) \sim \sum_{j\ge 1, k\ge 0} b^{\#}_{jkt}\Big(\frac{x}{|x|}\Big)\, |x|^{-n-j\gamma-k}, \quad |x| \to \infty. \tag{C.16}$$

In particular, the leading term is

$$b^{\#}_{10t}\Big(\frac{x}{|x|}\Big)\, |x|^{-n-\gamma}. \tag{C.17}$$

Note that

$$\int\limits_{|x|\ge 1} |x|^{-n-\gamma+\ell}\, dx = A_n \int_1^\infty r^{-1-\gamma+\ell}\, dr, \tag{C.18}$$

which is $+\infty$ for $\ell = 1$ if $\gamma \in (0, 1]$ and is finite for $\ell = 1$ if $\gamma \in (1, 2)$, but $+\infty$ for $\ell = 2$ for all $\gamma \in (0, 2)$. Consequently, as long as $a_0(\xi/|\xi|)$ is not $\equiv 0$ in (C.3), we have, for each $t > 0$,

$$\begin{aligned} \int p(t,x)|x|^\ell\, dx = \infty \ \text{ if } \ \ell = 1 \ \text{ and } \ \gamma \in (0,1] \\ \infty \ \text{ if } \ \ell = 2 \ \text{ and } \ \gamma \in (0,2). \end{aligned} \tag{C.19}$$

Note that the homogeneous generators of degree α considered in §B that satisfy (C.1) and $\operatorname{Re}\psi(\xi) > 0$ for $\xi \neq 0$ also satisfy (C.2)–(C.3) with $\beta = \gamma = \alpha$.

D. Short time and long time behavior of $e^{-t\psi(D)}\delta(x)$: examples

Here we examine the asymptotic behavior of

$$p(t,x) = e^{-t\psi(D)}\delta(x), \tag{D.1}$$

both as $t \nearrow \infty$ and as $t \searrow 0$, for some specific examples of $\psi(\xi)$.

The first example is

$$\psi(\xi) = (|\xi|^2 + 1)^\alpha - 1, \tag{D.2}$$

with $\alpha \in (0, 1)$. As noted in (1.15) we have, for each fixed $t > 0$,

$$p(t,x) \le C_{n\alpha t}\, e^{-|x|}, \quad |x| \to \infty.$$

We first treat the large t behavior. In light of the estimate above, the Central Limit Theorem applies. We have $p(t, x)$ behaving like

$$q(t,x) = e^{\alpha t\Delta}\delta(x) \tag{D.3}$$

as $t \nearrow \infty$, over the region $|x| \le Kt^{1/2}$, for each $K \in (0,\infty)$. More precisely, set

$$p^\#(t,x) = t^{n/2} p(t, t^{1/2}x), \tag{D.4}$$

and compare it with

$$q^\#(t,x) = t^{n/2} q(t, t^{1/2}x) = q(1,x). \tag{D.5}$$

We have

$$p^{\#}(t,x)=(2\pi)^{-n}\int e^{-t[(1+|\xi|^2/t)^\alpha-1]}e^{ix\cdot\xi}\,d\xi, \tag{D.6}$$

and standard arguments to establish versions of the Central Limit Theorem (cf. [T1], Chapter 3, §3, Exercises 7–13, or Chapter 1 of this text) yield

$$p^{\#}(t,x)\longrightarrow q(1,x),\quad \text{as}\ \ t\nearrow\infty, \tag{D.7}$$

both uniformly and in L^1-norm.

By contrast, we claim that as $t\searrow 0$, $p(t,x)$ behaves like

$$Q(t,x)=e^{-t(-\Delta)^\alpha}\delta(x). \tag{D.8}$$

To state this more precisely, in analogy with (D.4)–(D.7) we set

$$p^b(t,x)=t^{n/2\alpha}p(t,t^{1/2\alpha}x), \tag{D.9}$$

and compare it with

$$Q^b(t,x)=t^{n/2\alpha}Q(t,t^{1/2\alpha}x)=Q(1,x). \tag{D.10}$$

Then

$$\begin{aligned} p^b(t,x)&=(2\pi)^{-n}\int e^{-t[(1+|\xi|^2/t^{1/\alpha})^\alpha-1]}e^{ix\cdot\xi}\,d\xi\\ &=(2\pi)^{-n}\int e^{-[(t^{1/\alpha}+|\xi|^2)^\alpha-t]}e^{ix\cdot\xi}\,d\xi. \end{aligned} \tag{D.11}$$

Now

$$e^{-[(t^{1/\alpha}+|\xi|^2)^\alpha-t]}\longrightarrow e^{-|\xi|^{2\alpha}},\quad \text{as}\ \ t\searrow 0, \tag{D.12}$$

in $L^1(\mathbb{R}^n)$ and uniformly. The L^1-convergence implies

$$p^b(t,x)\longrightarrow Q(1,x),\quad \text{as}\ \ t\searrow 0, \tag{D.13}$$

uniformly, and the fact that $p^b(t,x)$ and $Q(1,x)$ are all positive functions of x integrating to 1 yields the convergence in L^1-norm.

We can say more. Note that

$$\begin{aligned} e^{-t}p^b(t,x)&=e^{-(-\Delta+t^{1/\alpha})^\alpha}\delta(x)\\ &=\int_0^\infty e^{-t^{1/\alpha}s}\,\Phi_{1,\alpha}(s)\,e^{s\Delta}\delta(x)\,ds, \end{aligned} \tag{D.14}$$

where we have used (1.7) with $L = -\Delta + t^{1/\alpha}$. Consequently, we also have

$$e^{-t} p^b(t,x) \nearrow Q(1,x), \quad t \searrow 0. \tag{D.15}$$

We can apply this to estimate the modulus of continuity of the stochastic process $\{X_t\}$ given in Theorem 2.1, with $\psi(D)$ given by (D.2). Recall (2.6):

$$E(|X_t - X_s|^q) = \int p(|t-s|, y)|y|^q \, dy. \tag{D.16}$$

Using (D.9) we have, for $-n < q < 2\alpha$,

$$\int p(t,y)|y|^q \, dy = t^{q/2\alpha} \int p^b(t,x)|x|^q \, dx, \tag{D.17}$$

and, by (D.15), as $t \searrow 0$,

$$e^{-t} \int p^b(t,x)|x|^q \, dx \nearrow \int Q(1,x)|x|^q \, dx, \tag{D.18}$$

which is a number $A \in (0,\infty)$, by arguments mentioned in (2.10)–(2.12). Thus, as $|t-s| \searrow 0$, $E(|X_t - X_s|^q)$ has the same asymptotic behavior for $\psi(\xi)$ given by (D.2) as it does for $\psi(\xi) = |\xi|^{2\alpha}$, given $\alpha \in (0,1)$, for this range of q.

Our second example is

$$\psi(\xi) = |\xi|^2 + |\xi|. \tag{D.19}$$

To treat the large time behavior, this time we examine

$$\begin{aligned} p^\#(t,x) &= t^n p(t,tx) \\ &= (2\pi)^{-n} \int e^{-t(t^{-2}|\xi|^2 + t^{-1}|\xi|)} e^{ix\cdot\xi} \, d\xi \\ &\to (2\pi)^{-n} \int e^{-|\xi|} e^{ix\cdot\xi} \, d\xi, \quad \text{as} \ \ t \nearrow \infty. \end{aligned} \tag{D.20}$$

In other words,

$$p^\#(t,x) \longrightarrow e^{-\sqrt{-\Delta}} \delta(x) = \frac{A_n}{(|x|^2+1)^{(n+1)/2}}, \quad \text{as} \ \ t \to \infty. \tag{D.21}$$

In this sense, $e^{t(\Delta - \sqrt{-\Delta})}\delta(x)$ behaves like $e^{-t\sqrt{-\Delta}}\delta(x)$ as $t \nearrow \infty$.

To treat small time behavior, we examine

$$\begin{aligned} p^b(t,x) &= t^{n/2} p(t, t^{1/2}x) \\ &= (2\pi)^{-n} \int e^{-t(t^{-1}|\xi|^2 + t^{-1/2}|\xi|)} e^{ix\cdot\xi} \, d\xi \\ &\to (2\pi)^{-n} \int e^{-|\xi|^2} e^{ix\cdot\xi} \, d\xi, \quad \text{as} \ \ t \searrow 0. \end{aligned} \tag{D.22}$$

In other words,

$$p^b(t,x) \longrightarrow e^{\Delta}\delta(x) = (4\pi)^{-n/2}\, e^{-|x|^2/4}, \quad \text{as } \ t \searrow 0. \tag{D.23}$$

In this sense, $e^{t(\Delta-\sqrt{-\Delta})}\delta(x)$ behaves like $e^{t\Delta}\delta(x)$ as $t \searrow 0$. However, one must be cautioned that the paths for the process associated to $e^{t(\Delta-\sqrt{-\Delta})}$ are not continuous, but rather cadlag, with jumps, so (D.23) does not tell the whole story about the short time behavior.

Let us now estimate $E(|X_t - X_s|^q)$ for the process $\{X_t\}$ generated by $\psi(D)$ with $\psi(\xi)$ given by (D.19). As in (D.16), we have

$$E(|X_t - X_s|^q) = \int p(|t-s|, y)|y|^q\, dy. \tag{D.24}$$

Using (D.22), we have the following analogue of (D.17):

$$\int p(t,y)|y|^q\, dy = t^{q/2} \int p^b(t,x)|x|^q\, dx. \tag{D.25}$$

However, from here the argument is different from (D.18). We have

$$p^b(t,x) = e^{\Delta - t^{1/2}\sqrt{-\Delta}}\delta(x) = h * q_t(x), \tag{D.26}$$

where

$$\begin{aligned} h(x) &= (4\pi)^{-n/2}\, e^{-|x|^2/4}, \quad q_t(x) = t^{-n/2}\, q_1(t^{-1/2}x), \\ q_1(x) &= \frac{C_n}{(1+|x|^2)^{(n+1)/2}}, \quad C_n = \pi^{-(n+1)/2}\, \Gamma\Big(\frac{n+1}{2}\Big). \end{aligned} \tag{D.27}$$

We deduce that

$$E(|X_t - X_s|^q) \le C_q |t-s|^{q/2}, \quad -n < q < 2. \tag{D.28}$$

As noted in (2.13) and the remark following it, for the process generated by Δ (i.e., Brownian motion) we have the estimate (D.28) over a larger range of q, namely $q \in (-n, \infty)$. Note that one can apply the Kolomogorov criterion (3.7) for sample path continuity as long as this estimate holds for some exponent $q/2 > 1$, but we do not get this in the product case, and this process is only cadlag.

On the other hand, using (D.28) with q close to $-n$, we obtain the following variant of (4.1), giving another respect in which the short time behavior of X_t in this case is like that of Brownian motion.

Proposition D.1. *For the process generated by $\Delta - \sqrt{-\Delta}$, if $n \geq 2$ we have for each interval $I = [0, T],\ T > 0$,*

$$\text{Hdim}\, \omega(I) \geq 2, \qquad \textit{for a.e. } \omega. \tag{D.29}$$

Here, as in (2.5), $\omega(t) = X_t(\omega)$. Presumably, one has equality in (D.29), but we do not have a proof of this.

E. Vanishing and super-exponential decay on cones

Let us set

$$\begin{aligned} \varphi_\alpha(x) &= (ix + i0)^\alpha, & 0 < \alpha < 1, \\ \psi_\alpha(x) &= -(ix + i0)^\alpha, & 1 < \alpha < 2, \end{aligned} \tag{E.1}$$

as in (B.9), i.e., φ_α is the boundary value on $\mathbb{R}$ of the function $(iz)^\alpha$ and ψ_α that of $-(iz)^\alpha$ on $\{z : \operatorname{Im} z < 0\}$, satisfying

$$\varphi_\alpha(-iy) > 0, \quad \psi_\alpha(-iy) < 0, \quad \forall\, y > 0. \tag{E.2}$$

As in (B.18), we consider

$$\begin{aligned} \Phi_{\alpha,\nu}(\xi) &= \int_{S^{n-1}} \varphi_\alpha(\xi \cdot \omega)\, d\nu(\omega), \quad 0 < \alpha < 1, \\ \Psi_{\alpha,\nu}(\xi) &= \int_{S^{n-1}} \psi_\alpha(\xi \cdot \omega)\, d\nu(\omega), \quad 1 < \alpha < 2, \end{aligned} \tag{E.3}$$

where ν is a positive, finite Borel measure on S^{n-1}, and we consider the associated probability distributions

$$\begin{aligned} P_{\alpha,\nu}(t,x) &= e^{-t\Phi_{\alpha,\nu}(D)}\delta(x), \quad 0 < \alpha < 1, \\ Q_{\alpha,\nu}(t,x) &= e^{-t\Psi_{\alpha,\nu}(D)}\delta(x), \quad 1 < \alpha < 2. \end{aligned} \tag{E.4}$$

Here we extend to n dimensions the vanishing result (B.29) (for $0 < \alpha < 1$) and the super-exponential decay result (B.31) (for $1 < \alpha < 2$), on a half-line, in case $n = 1$ and the measure ν on $S^0 = \{-1, 1\}$ has support in one point. We start with the extended vanishing result, when $0 < \alpha < 1$.

Proposition E.1. *Assume ν is a positive measure supported on $\Sigma \subset S^{n-1}$, and let $K \subset \mathbb{R}^n$ be the convex hull of the cone over Σ. Then, for $\alpha \in (0,1),\ t > 0$,*

$$\text{(E.5)} \qquad \mathit{supp}\, P_{\alpha,\nu}(t,\cdot) \subset K.$$

Proof. In this case we have (B.2), i.e.,

$$\text{(E.6)} \qquad \Phi_{\alpha,\nu}(\xi) = \int_\Sigma \int_0^\infty (1 - e^{iy\cdot\xi}) s^{-1-\alpha}\, ds\, d\nu(\omega), \quad y = s\omega.$$

This is a limit of finite, positive linear combinations of functions

$$\text{(E.7)} \qquad \psi_y(\xi) = 1 - e^{iy\cdot\xi}, \quad y = s\omega,\ \omega \in \Sigma,\ s > 0.$$

Hence $e^{-t\Phi_{\alpha,\nu}(D)}$ is a limit of compositions

$$\text{(E.8)} \qquad e^{-tc_1\psi_{y_1}(D)} \cdots e^{-tc_N\psi_{y_N}(D)},$$

and $P_{\alpha,\nu}(t,\cdot)$ is a limit in $\mathcal{S}'(\mathbb{R}^n)$ of a sequence of distributions of the form

$$\text{(E.9)} \qquad p_{tc_1,y_1} * \cdots * p_{tc_N,y_N}(x),$$

where

$$\text{(E.10)} \qquad p_{tc_j,y_j}(x) = \sum_{k=0}^\infty \frac{(tc_j)^k}{k!}\, e^{-c_j t}\, \delta(x - ky_j).$$

Note that p_{tc_j,y_j} is supported on the ray through y_j in $\mathbb{R}^n$. Hence the support of (E.9) is contained in the convex hull of the set of rays through $\{y_1,\dots,y_N\} \subset \Sigma$. In the limit we get (E.5).

Next we establish super-exponential decay of $Q_{\alpha,\nu}(t,x)$, not on the complement of K, but on the dual cone:

$$\text{(E.11)} \qquad L = \{x \in \mathbb{R}^n : x\cdot\omega < 0,\ \forall\, \omega \in \Sigma\}.$$

Proposition E.2. *Assume ν is a positive measure supported on $\Sigma \subset S^{n-1}$. Also assume that, for some $C > 0$,*

$$\text{(E.12)} \qquad \mathit{Re}\, \Psi_{\alpha,\nu}(\xi) \ge C|\xi|^\alpha, \quad \xi \in \mathbb{R}^n.$$

Then, for $\alpha \in (1,2),\ t > 0$,

$$\text{(E.13)} \qquad Q_{\alpha,\nu}(t,x) = o(e^{-b|x|}), \quad \forall\, b > 0,\ \mathit{as}\ \ |x| \to \infty,\ x \in L,$$

where L (which might be empty) is given by (E.11).

Proof. Under the stated hypotheses,

$$\Psi_{\alpha,\nu}(\xi + i\eta) = \int_{\Sigma} \psi_\alpha(\omega\cdot\xi + i\omega\cdot\eta)\, d\nu(\omega) \tag{E.14}$$

is well defined and holomorphic on

$$\{\xi + i\eta : \xi \in \mathbb{R}^n,\ \eta \in L\}. \tag{E.15}$$

Furthermore,

$$\operatorname{Re}\Psi_{\alpha,\nu}(\xi + i\eta) \geq C|\xi|^\alpha - C'|\eta|^\alpha, \tag{E.16}$$

with $C > 0$. Hence, for each $\eta \in L$,

$$\begin{aligned} Q_{\alpha,\nu}(t,x) &= (2\pi)^{-n} \int e^{ix\cdot\xi - t\Psi_{\alpha,\nu}(\xi)}\, d\xi \\ &= (2\pi)^{-n} e^{-x\cdot\eta} \int e^{ix\cdot\xi - t\Psi_{\alpha,\nu}(\xi+i\eta)}\, d\xi. \end{aligned} \tag{E.17}$$

Hence

$$|Q_{\alpha,\nu}(t,x)| \leq C_t(\eta) e^{-x\cdot\eta}. \tag{E.18}$$

If $x \in L$, we can pick $\eta = 2bx/|x|$ and deduce (E.13).

As for when (E.12) holds, note that, for $x \in \mathbb{R}$,

$$\operatorname{Re}\psi_\alpha(x) = \left|\cos\frac{\pi\alpha}{2}\right| \cdot |x|^\alpha, \tag{E.19}$$

(compare (B.13)–(B.14)), and hence, for $\xi \in \mathbb{R}^n,\ A_\alpha = |\cos \pi\alpha/2|$,

$$\operatorname{Re}\Psi_{\alpha,\nu}(\xi) = A_\alpha \int_{\Sigma} |\omega\cdot\xi|^\alpha\, d\nu(\omega). \tag{E.20}$$

Thus

$$\text{(E.12) holds} \Longleftrightarrow \int_{\Sigma} |\omega\cdot\xi|^\alpha\, d\nu(\omega) > 0, \quad \forall\, \xi \in \mathbb{R}^n \setminus 0. \tag{E.21}$$

Remark. In light of (E.16), we can restate (E.18) more precisely as

$$|Q_{\alpha,\nu}(t,x)| \leq C_t\Big(\frac{\eta}{|\eta|}\Big) e^{C'|\eta|^\alpha} e^{-x\cdot\eta}. \tag{E.22}$$

Hence, picking $\eta = bx/|x|$, we have

$$|Q_{\alpha,\nu}(t,x)| \le C_t\Big(\frac{x}{|x|}\Big) e^{-b|x|+C'tb^\alpha}, \quad b \in (0,\infty). \tag{E.23}$$

Optimizing over b, we then have

$$|Q_{\alpha,\nu}(t,x)| \le C_t\Big(\frac{x}{|x|}\Big) e^{-\kappa|x|^{\alpha/(\alpha-1)}/t^{1/(\alpha-1)}}, \quad x \in L. \tag{E.24}$$

F. Regularity properties of the semigroup $e^{-t\psi(D)}$

Let $e^{-t\psi(D)}$ be as in §1. In particular, we have (3.3)–(3.4). It is elementary that for each $t > 0$, $e^{-t\psi(D)}$ is a contraction on $L^p(\mathbb{R}^n)$ for each $p \in [1,\infty]$. It is also a contraction on $\mathrm{BC}(\mathbb{R}^n)$, the space of bounded continuous functions on $\mathbb{R}^n$, and on the closed linear subspace $\mathrm{UC}(\mathbb{R}^n)$ of bounded, uniformly continuous functions on $\mathbb{R}^n$, and on the space $C_*(\mathbb{R}^n)$ of continuous functions on $\mathbb{R}^n$ vanishing at infinity. It is positivity preserving on all these spaces.

The family $e^{-t\psi(D)}$ is a strongly continuous semigroup on most of these spaces, though not on $L^\infty(\mathbb{R}^n)$ or $\mathrm{BC}(\mathbb{R}^n)$. This continuity is quite elementary for $L^2(\mathbb{R}^n)$, by virtue of the identity

$$e^{-t\psi(D)}u(x) = (2\pi)^{-n}\int e^{-t\psi(\xi)}\hat{u}(\xi)e^{ix\cdot\xi}\,d\xi \tag{F.1}$$

and the Plancherel theorem. Also, if $u \in \mathcal{S}(\mathbb{R}^n)$, the integrand on the right side of (F.1) is a continuous function of $t \in [0,\infty)$ with values in $L^1(\mathbb{R}^n)$, so the Fourier integral is a continuous function of t with values in $C_*(\mathbb{R}^n)$. Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $C_*(\mathbb{R}^n)$, we have a strongly continuous semigroup on $C_*(\mathbb{R}^n)$. More generally, given $u \in \mathcal{S}(\mathbb{R}^n)$, $e^{-t\psi(\xi)}\hat{u}(\xi)$ is a continuous function of t with values in $L^p(\mathbb{R}^n)$ for each $p \in [1,2]$, so (F.1) is a continuous function of t with values in $L^{p'}(\mathbb{R}^n)$. The same density argument shows that we have a strongly continuous semigroup on $L^q(\mathbb{R}^n)$ for each $q \in [2,\infty)$.

We next argue that $e^{-t\psi(D)}$ is strongly continuous on $L^p(\mathbb{R}^n)$ for $p \in [1,2)$. To begin, take $u \in \mathcal{S}(\mathbb{R}^n)$ such that $u \ge 0$. Then $v(t) = e^{-t\psi(D)}u$ is ≥ 0 and $\int v(t,x)\,dx \equiv \int u(x)\,dx$. We already know $v(t) \to u$ uniformly as $t \searrow 0$. From these facts it follows that $v(t) \to u$ in L^1-norm. Hence for each $u \in \mathcal{S}(\mathbb{R}^n)$ (without the sign condition) we have $e^{-t\psi(D)}u \to u$ in L^1-norm as $t \searrow 0$. We also know this holds in L^2-norm, so it holds in L^p-norm for each $p \in [1,2]$. Again a density argument yields the asserted strong continuity on $L^p(\mathbb{R}^n)$, at $t = 0$. As is well known (cf. [Bob], p. 249) this suffices to establish strong continuity in $t \in [0,\infty)$.

Our next goal is to prove the following.

Proposition F.1. *The semigroup* $e^{-t\psi(D)}$ *is strongly continuous on* $\mathrm{UC}(\mathbb{R}^n)$.

Proof. As noted above, it suffices to prove strong continuity at $t=0$, so we need to show that if $u \in \mathrm{UC}(\mathbb{R}^n)$, then

$$e^{-t\psi(D)}u \to u \quad \text{uniformly, as} \;\; t \searrow 0. \tag{F.2}$$

It suffices to show that (F.2) holds for u in a dense subspace of $\mathrm{UC}(\mathbb{R}^n)$, and we take $\mathrm{Lip}(\mathbb{R}^n)$, which a mollifier argument shows to be dense. So suppose u is bounded and

$$|u(x+y)-u(x)| \le L(|y| \wedge 1), \tag{F.3}$$

for all $x, y \in \mathbb{R}^n$. As in (3.1), consider

$$G(y) = 1 - e^{-|y|} = 1 - g(y). \tag{F.4}$$

For each $y \in \mathbb{R}^n$,

$$u(y) - 2LG(x) \le u(x+y) \le u(y) + 2LG(x), \quad \forall\, x \in \mathbb{R}^n, \tag{F.5}$$

so for each $t > 0$,

$$u(y) - 2Le^{-t\psi(D)}G(x) \le e^{-t\psi(D)}u(x+y) \le u(y) + 2Le^{-t\psi(D)}G(x), \tag{F.6}$$

so

$$\left| e^{-t\psi(D)}u(y) - u(y) \right| \le 2Le^{-t\psi(D)}G(0). \tag{F.7}$$

As seen in (3.3)–(3.5), the right side of (F.7) tends to 0 as $t \searrow 0$, so the proof is complete.

M. Lévy processes on manifolds

Paralleling the study of translation invariant Lévy processes on Euclidean space $\mathbb{R}^n$, there is a theory of left (or right) invariant Lévy processes on Lie groups, initiated by G. Hunt. There are also studies of Lévy processes on more general Riemannian manifolds. Some articles in [BMR] discuss this, and give more references. Our goal here is to present various generalizations of the generators of Lévy processes given by the Lévy-Khinchin formula (A.5) that work in the manifold context.

The generators we seek are generators of semigroups $P^t = e^{tA}$ of positivity-preserving operators on $C_b(M)$ satisfying $P^t 1 = 1$. They have the form

$$P^t u(x) = \int_M p_t(x, dy)\, u(y), \tag{M.1}$$

where $p_t(x, \cdot)$ is a family of probability measures on M.

One general observation is that if A and B generate such semigroups and the Trotter product formula holds:

$$e^{t(A+B)} = \lim_{n\to\infty} \Bigl(e^{(t/n)A} e^{(t/n)B}\Bigr)^n, \tag{M.2}$$

then $A + B$ generates such a semigroup. Extending this, if $\{A(y) : y \in Y\}$ is a family of generators of such semigroups, then (frequently) so is

$$\int_Y A(y)\, d\mu(y), \tag{M.3}$$

given a positive measure μ on Y (perhaps with some sort of bound). Let us now get more specific.

The first two terms on the right side of (A.5) have well known generalizations to second order differential operators on M. In local coordinates,

$$L = \sum a^{jk}(x)\partial_j\partial_k + \sum b^j(x)\partial_j. \tag{M.4}$$

One makes various hypotheses, including $\sum a_{jk}(x)\xi_j\xi_k \ge 0$. There is a large literature on diffusion processes with such generators. See for example [Str].

We now point out various generators analogous to the last term in (A.5). To start, let $\varphi : M \to M$ be a continuous map, and consider

$$Tu(x) = u(\varphi(x)). \tag{M.5}$$

Now pick $c \in (0, \infty)$ and set

$$\begin{aligned} P^t_{T,c} u(x) &= e^{-tc(I-T)} u(x) \\ &= \sum_{k=0}^{\infty} \frac{(ct)^k}{k!} e^{-ct}\, T^k u(x). \end{aligned} \tag{M.6}$$

This has the form (M.1) with

$$p_t(x, \cdot) = \sum_{k=0}^{\infty} \frac{(ct)^k}{k!} e^{-ct}\, \delta_{\varphi^k(x)}. \tag{M.7}$$

One can impose further structure by requiring φ to be a diffeomorphism, or a volume preserving map, or an isometry (with respect to some Riemannian metric), etc. One can take a family $\varphi_y : M \to M$ of such maps and apply the process (M.3), obtaining generators of the form

$$-\int_Y (I - T(y))\, d\mu(y), \quad T(y)u(x) = u(\varphi_y(x)). \tag{M.8}$$

To specialize this construction, let $X_1, \dots, X_m$ be smooth vector fields on M, and assume that for each $y = (y_1, \dots, y_m) \in \mathbb{R}^m$, $y \cdot X = y_1 X_1 + \cdots + y_m X_m$ generates a global flow on M. Given a positive measure μ on $\mathbb{R}^n$, one has the generator

$$-\int_{\mathbb{R}^m} (I - e^{y \cdot X})\, d\mu(y), \tag{M.9}$$

given some bounds on μ. For example, one might require $\int (|y| \wedge 1)\, d\mu(y) < \infty$. Then (M.9) would be convergent if each $y \cdot X$ generated a volume preserving flow. Otherwise, further restrictions on μ might be needed. One can allow more singular behavior of μ near 0, i.e., $\int (|y|^2 \wedge 1)\, d\mu(y) < \infty$, upon replacing (M.9) by

$$-\int_{\mathbb{R}^m} (I - e^{y \cdot X} + \chi_B(y)\, y \cdot X)\, d\mu(y). \tag{M.10}$$

The operators (M.9) and (M.10) are often pseudodifferential operators of order $2\alpha \in (0, 2)$, for various measures $d\mu(y) = P(y)\, dy$, where $P(y)$ is smooth on $\mathbb{R}^m \setminus 0$, supported near 0 (this requirement can often be relaxed) and having a conormal singularity at 0, weaker than $|y|^{-m-1}$. One needs to require that $X_1^2 + \cdots + X_m^2$ be elliptic.

We mention the following problem. Suppose M has a Riemannian metric tensor, whose Laplace operator Δ generates a non-explosive diffusion, so $e^{t\Delta} 1 = 1$ for $t > 0$. For $\alpha \in (0, 1)$, one would like to write $-(-\Delta)^\alpha$ in the form (M.9) or (M.10) (perhaps with the term $\chi_B(y) y \cdot X$ suitably modified). Surely this is well known in some cases, but it would be nice to have a general result. It would also be interesting to find such representations for variants, such as

$$1 - (1 - \Delta)^\alpha, \quad \alpha \in (0, 1).$$

Leaving (M.9)–(M.10), we note the following more general context for (M.6)–(M.8). Namely $T : C_b(M) \to C_b(M)$ could be any positivity preserving operator satisfying $T1 = 1$. Then

$$T^k u(x) = \int_M \mu_k(x, dy)\, u(y), \tag{M.11}$$

where $\mu_k(x,\cdot)$ is a family of probability measures, and one replaces (M.7) by

$$p_t(x,\cdot) = \sum_{k=0}^{\infty} \frac{(ct)^k}{k!} e^{-ct}\, \mu_k(x,\cdot). \tag{M.12}$$

The processes associated to the semigroups $e^{-tc(I-T)}$ for such T are called Feller's pseudo-Poisson processes; cf. [Ap], pp. 160–162.

N. Other Markov processes

The transition beyond Lévy processes in the Euclidean setting to Riemannian manifolds, discussed in §M, motivates us to go a step further, and present some general results about Markov semigroups. We say only a little about this big area, referring to [D] for a more thorough introduction.

We start with continuous-time Markov processes on a finite set X, say with n points, also called a finite Markov chain. We have $C(X)$ isomorphic to $\mathbb{R}^n$, and Markov semigroups are given by $n \times n$ matrices,

$$e^{tA}, \quad A \in M(n,\mathbb{R}), \quad t \ge 0. \tag{N.1}$$

To say this is a Markov semigroup is to say

$$e^{tA}\mathbf{1} \equiv \mathbf{1}, \quad \text{and} \quad v \in \mathbb{R}^n,\ v \ge 0 \Rightarrow e^{tA}v \ge 0, \quad \text{for } t \ge 0. \tag{N.2}$$

Here $v \ge 0$ means each component is ≥ 0, and

$$\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{N.3}$$

We can restate the positivity condition as

$$e^{tA} = \Bigl(p_{jk}(t)\Bigr), \quad p_{jk}(t) \ge 0, \quad \text{for} \quad t \ge 0, \quad j,k \in \{1,\dots,n\}. \tag{N.4}$$

The set of probability measures on X is given by $\mathcal{P}$, where, for $w \in \mathbb{R}^n$,

$$w \in \mathcal{P} \Longleftrightarrow w \ge 0 \quad \text{and} \quad w \cdot \mathbf{1} = 1. \tag{N.5}$$

Then the action of the Markov semigroup on $\mathcal{P}$ is given by

$$(e^{tA})^* : \mathcal{P} \longrightarrow \mathcal{P}, \quad \text{for} \quad t \ge 0, \tag{N.6}$$

where, for $B \in M(n,\mathbb{R})$, B^* is the transpose of B.

The following result characterizes the generators A of all such semigroups.

Proposition N.1. *Given* $A = \big(a_{jk}\big) \in M(n,\mathbb{R})$, e^{tA} *satisfies (N.2) if and only if*

$$A\mathbf{1} = 0, \tag{N.7}$$

and

$$a_{jk} \ge 0 \quad \textit{whenever} \quad j \neq k. \tag{N.8}$$

Proof. Noting that

$$\frac{d}{dt} e^{tA}\Big|_{t=0} = A, \tag{N.9}$$

we see the relation $e^{tA}\mathbf{1} \equiv \mathbf{1}$ implies (N.7), and the positivity (N.8) follows from (N.4) plus $p_{jk}(0) = 0$ for $j \neq k$.

For the converse, if (N.8) is strengthened to $a_{jk} > 0$ whenever $j \neq k$, then, via

$$e^{tA} = I + tA + O(t^2), \tag{N.10}$$

one has $t_0 > 0$ such that $e^{tA} \ge 0$ for $0 \le t \le t_0$, and positivity for all $t \ge 0$ follows from $e^{ktA} = (e^{tA})^k$. Then the sufficiency of (N.7)–(N.8) in general can be established by a limiting argument. We leave the details to the reader. Or see [T3], p. 155, Proposition 4.4.10. An alternative approach to the converse, valid in a much more general setting, is described below, in Proposition N.2.

Remark. Clearly the conditions (N.7)–(N.8) imply for the diagonal elements of A that

$$a_{jj} \le 0, \quad \text{for} \quad j \in \{1, \dots, n\}. \tag{N.11}$$

Denumerable Markov chains are associated to processes on a countably infinite set, such as $\mathbb{N} = \{1, 2, 3, \dots\}$. One might have a semigroup

$$e^{tA} : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N}), \quad t \ge 0, \tag{N.12}$$

satisfying

$$f \in \ell^\infty(\mathbb{N}),\ f \ge 0 \Rightarrow e^{tA} f \ge 0 \ \text{ and } \ e^{tA} 1 \equiv 1. \tag{N.13}$$

Alternatively, one might consider sequences $f(n)$ that tend to a limit as $n \to \infty$, and

$$e^{tA} : C(\widehat{\mathbb{N}}) \longrightarrow C(\widehat{\mathbb{N}}), \quad t \ge 0, \tag{N.14}$$

where $\widehat{\mathbb{N}}$ is the one point compactification $\mathbb{N} \cup \{\infty\}$.

Extending the scope of (N.14), one can let X be a compact Hausdorff space, and consider semigroups

$$e^{tA} : C(X) \longrightarrow C(X), \quad t \geq 0, \tag{N.15}$$

satisfying

$$e^{tA}1 \equiv 1, \quad \text{and} \;\; f \in C(X),\; f \geq 0 \Rightarrow e^{tA}f \geq 0. \tag{N.16}$$

The class (N.15)–(N.16) actually contains (N.12)–(N.13) as a special case. In fact, we can regard $\ell^\infty(\mathbb{N})$ as a commutative C^* algebra and take X to be its maximal ideal space. Then the Gelfand transform provides a positivity-preserving isometric isomorphism $\ell^\infty(\mathbb{N}) \approx C(X)$.

The following result yields a large class of Markov semigroups. In particular, it provides a far-reaching generalization of the result of Proposition N.1 that any $A \in M(n, \mathbb{R})$ satisfying (N.7)–(N.8) generates a Markov semigroup on $\mathbb{R}^n$.

Proposition N.2. *Let X be a compact Hausdorff space. Let*

$$B : C(X) \longrightarrow C(X) \tag{N.17}$$

be continuous and positivity-perserving, i.e.,

$$f \in C(X), f \geq 0 \Longrightarrow Bf \geq 0. \tag{N.18}$$

Set $\varphi = B1 \in C(X)$, and define

$$A : C(X) \longrightarrow C(X), \quad Au = -\varphi u + Bu. \tag{N.19}$$

Then $\{e^{tA} : t \geq 0\}$ is a Markov semigroup on $C(X)$.

Proof. First, clearly

$$A1 = 0, \;\; \text{so} \;\; e^{tA}1 \equiv 1. \tag{N.20}$$

It remains to show that, for $t \geq 0$, e^{tA} has the positivity property given in (N.16). This follows from the Trotter product formula,

$$e^{tA}f = \lim_{n \to \infty} \Bigl(e^{-t\varphi/n} e^{(t/n)B}\Bigr)^n f, \tag{N.21}$$

plus the fact that, for $s \geq 0$, $f \in C(X)$,

$$f \geq 0 \Longrightarrow e^{-s\varphi} f \geq 0 \;\; \text{and} \;\; e^{sB} f \geq 0, \tag{N.22}$$

the latter via the power series expansion

$$e^{sB} = \sum_{k=0}^{\infty} \frac{s^k}{k!} B^k. \tag{N.23}$$

References

[Ap] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge Univ. Press, Cambridge, 2004.

[BMR] O. Barndorff-Nielsen, T. Mikosch, and S. Resnick (eds.), Lévy Processes, Theory and Applications, Birkhäuser, Boston, 2001.

[Ber] J. Bertoin, Lévy Processes, Cambridge Univ. Press, 1996.

[Bob] A. Bobrowski, Functional Analysis for Probability and Stochastic Processes, Cambridge Univ. Press, 2005.

[D] J. Doob, Stochastic Processes, J. Wiley, New York, 1953.

[Fal] K. Falconer, The Geometry of Fractal Sets, Cambridge Univ. Press, 1985.

[Jac] N. Jacob, Pseudo-Differential Operators and Markov Processes, World Scientific, Vol. 1, 2001; Vol. 2, 2002.

[Kry] N. Krylov, Introduction to the Theory of Random Processes, Graduate Studies in Math. #43, AMS, Providence, RI, 2002.

[Nel] E. Nelson, Feynman integrals and the Schrödinger equation, J. Math. Phys. 5 (1964), 332–343.

[ST] G. Samorodnitsky and M. Taqqu, Stable Non-Gaussian Random Processes, Chapman and Hall, New York, 1994.

[Sat] K. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge Univ. Press, 1999.

[Str] D. Stroock, An Introduction to the Analysis of Paths on a Riemannian Manifold, AMS, Providence, RI, 2000.

[T1] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer-Verlag, New York, 1996.

[T2] M. Taylor, Measure Theory and Integration, Graduate Studies in Math. #76, AMS, Providence, RI, 2006.

[T3] M. Taylor, Linear Algebra, AMS, Providence RI, 2020.

[Y] K. Yosida, Functional Analysis, Springer-Verlag, Berlin, 1965.

[Zai] A. Zaigraev, On asymptotic properties of multidimensional α-stable distributions, Math. Nachr. 16 (2006), 1835–1854.

4. Remarks on the modulus of continuity for Brownian paths

Let $\mathfrak{P}$ be path space for Brownian motion on the line, with Wiener measure P. It is classical (cf. [T], Proposition 16.5) that there is an estimate

$$|X_t(\omega) - X_s(\omega)| \le M_1(\omega) h(|t-s|), \tag{1}$$

valid for $s, t \in [0,1]$, with $M_1(\omega) < \infty$ for almost all $\omega \in \mathfrak{P}$. Here

$$h(\delta) = \Bigl(\delta \log \frac{1}{\delta}\Bigr)^{1/2} \tag{2}$$

for $0 < \delta \le 1/e$, and we set $h(\delta) = h(1/e) = 1/\sqrt{e}$ for $\delta \ge 1/e$. M. Pinsky [P] produced a pleasant proof of this, using Ciesielski's representation of $X_t(\omega)$ for $0 \le t \le 1$ as a Haar series:

$$X_t(\omega) = \sum_{N=1}^{\infty} Z_N(\omega) \int_0^t \varphi_N(s)\, ds, \tag{3}$$

where $\{\varphi_N : N \ge 1\}$ is the Haar orthonormal basis of $L^2([0,1])$. Path space over $0 \le t \le 1$ is parametrized by

$$\Omega = \prod_{N \ge 1} F_N, \quad F_N = \bigl(\mathbb{R}, (2\pi)^{-1/2} e^{-x^2/2}\, dx\bigr), \tag{4}$$

and $Z_N : \Omega \to \mathbb{R}$ is projection onto the Nth factor, identified with $\mathbb{R}$.

The estimate (1) can be compared and contrasted with the Lévy estimate

$$\limsup_{0 \le s,t \le 1, |s-t| \searrow 0} \frac{|X_t(\omega) - X_s(\omega)|}{h(|t-s|)} = \sqrt{2}, \tag{5}$$

valid for almost all $\omega \in \mathfrak{P}$. For one, (5) implies (1), but with no effective bound on $M_1(\omega)$. In fact, $M_1(\omega)$ cannot be essentially bounded on $\mathfrak{P}$; if it were, one would have for some $K < \infty$ an estimate

$$|X_t(\omega) - X_s(\omega)| \le K h(|t-s|), \tag{6}$$

valid for almost every $\omega \in \mathfrak{P}$, for all $s, t \in [0,1]$. In fact, for *fixed* $s < t \in (0,1]$ the Gaussian statistics for $X_t(\omega) - X_s(\omega)$ guarantee that (6) is violated for a set of ωs of positive measure.

Maximal estimates on Gaussian processes, such as given in Theorem 1.3.3 of [F], imply that once one has $M_1(\omega) < \infty$ almost everywhere in (1), then there is a bound

$$P\bigl(M_1(\omega) > \lambda\bigr) \le C e^{-a\lambda^2}, \tag{7}$$

for some $C < \infty$, $a > 0$. There is even a sharp result on the optimal value of a in (7).

Here we show that the method of proof of (1) in [P] can be pushed a little further to establish (7) directly (though without a sharp estimate for a). On the other hand, the estimate we establish in Proposition 1 below is in some ways more precise than (7).

To get started, we recall the ingredients of the proof of (1) in [P]. One ingredient is the following set of estimates on the Haar functions:

$$\|\varphi_N\|_{L^\infty(I)} \le C N^{1/2}, \tag{8}$$

and

$$\|\varphi_N\|_{L^1(I)} \le C N^{-1/2}, \tag{9}$$

(where $I = [0, 1]$) plus the fact that, over each range $2^{\nu-1} < N \le 2^\nu$, the functions

$$\psi_N(t) = \int_0^t \varphi_N(s)\, ds \tag{10}$$

have disjoint supports. Another ingredient is a study of the function

$$A(\omega) = \sup_{N \ge 2} \frac{|Z_N(\omega)|}{\sqrt{\log N}}. \tag{11}$$

It is shown in Lemma 1 of [P] that $A(\omega) < \infty$ for almost all $\omega \in \Omega$. Then the sum

$$X_t(\omega) - X_s(\omega) = \sum_{N=1}^{\infty} Z_N(\omega)[\psi_N(t) - \psi_N(s)] \tag{12}$$

is broken into two pieces and the estimate (1) is obtained, with $M_1(\omega) \le 1 + 2A(\omega)$. Hence (7) will follow from an associated estimate on $A(\omega)$.

We find it of interest to consider more generally

$$A_\mu(\omega) = \sup_{N \ge \mu} \frac{|Z_N(\omega)|}{\sqrt{\log N}}, \tag{13}$$

for $\mu \ge 2$. Now the nature of Z_N as a Gaussian random variable gives

$$P\big(|Z_N(\omega)| \ge x\big) \le e^{-x^2/2}, \tag{14}$$

hence

$$\mathcal{S}_{N,\lambda} = \{\omega \in \Omega : |Z_N(\omega)| \ge \lambda \sqrt{\log N}\} \Longrightarrow P(\mathcal{S}_{N,\lambda}) \le N^{-\lambda^2/2}. \tag{15}$$

Now

$$\{\omega \in \Omega : A_\mu(\omega) \ge \lambda\} = \bigcup_{N \ge \mu} \mathcal{S}_{N,\lambda}, \tag{16}$$

so

$$P\big(A_\mu(\omega) \ge \lambda\big) \le \sum_{N \ge \mu} N^{-\lambda^2/2}. \tag{17}$$

The convexity of the function $f(y) = y^{-s}$ implies that, for $s \ge 2$,

$$\sum_{N \ge \mu} N^{-s} \le \mu^{-s} + \int_{\mu+1/2}^{\infty} y^{-s}\, dy \le \Big(\mu + \frac{3}{2}\Big)\mu^{-s}, \tag{18}$$

so we have, for $\lambda \ge 2,\ \mu \ge 2$,

$$P\big(A_\mu(\omega) \ge \lambda\big) \le \Big(\mu + \frac{3}{2}\Big)\mu^{-\lambda^2/2} \le C_\mu e^{-K(\mu)\lambda^2}, \tag{19}$$

with $C_\mu = \mu + 3/2,\ K(\mu) = (1/2)\log \mu$.

The $\mu = 2$ case of this estimate is already enough to establish (7), in view of the estimate $M_1(\omega) \le 1 + 2A_2(\omega)$ established in [P]. However, we will go further (in a parallel fashion). Suppose $\mu = 2^\alpha + 1$ and $\alpha \ge 1$. We will estimate

$$\sum_{2^\nu < N \le 2^{\nu+1}} Z_N(\omega)[\psi_N(t) - \psi_N(s)] = D_\nu(t,s,\omega) \tag{20}$$

for $\nu \ge \alpha$. The observations in (8)–(10) imply the following two estimates:

$$\begin{aligned} |D_\nu(t,s,\omega)| &\le C A_\mu(\omega)\sqrt{\nu}\, 2^{\nu/2}|t-s|, \\ |D_\nu(t,s,\omega)| &\le C A_\mu(\omega)\sqrt{\nu}\, 2^{-\nu/2}. \end{aligned} \tag{21}$$

Hence

$$\begin{aligned} &\Big|\sum_{M \ge \mu} Z_N(\omega)[\psi_N(t) - \psi_N(s)]\Big| \\ &\qquad \le \sum_{\nu \ge \alpha} |D_\nu(t,s,\omega)| \\ &\qquad \le C A_\mu(\omega)\Big[\sum_{\alpha \le \nu \le \beta} 2^{\nu/2}\nu^{1/2}|t-s| + \sum_{\nu > \beta} 2^{-\nu/2}\nu^{1/2}\Big] \\ &\qquad \le C_2 A_\mu(\omega)\beta^{1/2}\big[2^{\beta/2}|t-s| + 2^{-\beta/2}\big]. \end{aligned} \tag{22}$$

We can optimize this by picking β such that $2^{-\beta/2} \approx |t-s|$, as long as $|t-s| \le 2^{-\alpha/2}$, say $|t-s| \le \mu^{-1/2}$. This gives

$$\Big|\sum_{N \ge \mu} Z_N(\omega)[\psi_N(t) - \psi_N(s)]\Big| \le C_3 A_\mu(\omega) h(|t-s|), \tag{23}$$

with $h(\delta)$ as in (2). As for the rest of (12), we crudely have

$$\Big|\sum_{N < \mu} Z_N(\omega)[\psi_N(t) - \psi_N(s)]\Big| \le C B_\mu(\omega)|t-s|, \quad B_\mu(\omega) = \mu^{1/2}\sum_{N<\mu}|Z_N(\omega)|. \tag{24}$$

This establishes the following (with slight change in notation):

Proposition 1. *Fix $K \in (0,\infty)$ and set $\delta = e^{-K}$. There exists $a = a(K) > 0$ and $C_j = C_j(K)$ such that, for $t, s \in [0,1]$, $|t-s| \le \delta$,*

$$|X_t(\omega) - X_s(\omega)| \le A_K(\omega) h(|t-s|) + B_K(\omega)|t-s|, \tag{25}$$

with

$$P\big(A_K(\omega) \ge \lambda\big) \le C_1 e^{-K\lambda^2}, \quad P\big(B_K(\omega) \ge \lambda\big) \le C_2 e^{-a\lambda^2}. \tag{26}$$

Returning to the context of the estimate (1), we make a concluding comment. It similarly follows that there is for each $k \in \mathbb{Z}^+$ an estimate

$$|X_t(\omega) - X_s(\omega)| \le M_k(\omega) h(|t-s|), \quad s, t \in [k-1, k]. \tag{27}$$

The functions M_k on $\mathfrak{P}$ can be taken to be independent random variables that are identically distributed.

References

[F] X. Fernique, Regularité des trajectoires des fonctions aléatoires Gaussiennes, pp. 1–96 in LNM #480, Springer-Verlag, New York, 1975.

[P] M. Pinsky, Brownian continuity modulus via series expansions, Jour. of Theoretical Probability 14 (2001), 261–266.

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

5. Stochastic integrals for square integrable Lévy processes

1. Introduction

Lévy processes were introduced in Chapter 3. We saw that if $\psi : \mathbb{R}^n \to \mathbb{C}$ is a continuous function satisfying $\psi(0) = 0$, such that, for $t \ge 0$, $e^{-t\psi(\xi)}$ is a tempered distribution, and

$$e^{-t\psi(D)}\delta(x) = p(t,x) \ge 0, \tag{1.1}$$

for $t \ge 0$ (or more generally is a positive probability measure), then there is a probability space (Ω, μ) and a continuous map

$$X : [0,\infty) \longrightarrow \mathcal{M}(\Omega, \mu, \mathbb{R}^n), \tag{1.2}$$

the space of measurable functions on Ω (modulo identity μ-a.e.) with the property that if $E_j \subset \mathbb{R}^n$ are Borel sets, and $0 < t_1 < \cdots < t_k$, the probability that $X_{t_j} \in E_j$ for each $j \in \{1, \dots, k\}$ is

$$\int\limits_{E_1} \cdots \int\limits_{E_k} p(t_k - t_{k-1}, x_k - x_{k-1}) \cdots p(t_1, x_1)\, dx_k \cdots dx_1. \tag{1.3}$$

Notable examples of such functions $\psi(\xi)$ include

$$\psi(\xi) = |\xi|^2, \quad e^{-t\psi(D)}\delta(x) = (4\pi t)^{-n/2} e^{-|x|^2/4t}, \tag{1.4}$$

yielding the Wiener process, and

$$\psi(\xi) = c(1 - e^{iy\cdot\xi}), \quad e^{-t\psi(D)}\delta(x) = \sum_{k=0}^{\infty} \frac{(ct)^k}{k!} e^{-ct} \delta_{ky}(x), \tag{1.5}$$

yielding Poisson processes, given $y \in \mathbb{R}^n,\ c > 0$.

As seen in (2.6) of Chapter 3, for $q \in \mathbb{R},\ 0 < s < t$,

$$E(|X_t - X_s|^q) = E(|X_{t-s}|^q) = \int p(t-s, x)|x|^q\, dx. \tag{1.6}$$

One condition we impose on the Lévy processes treated here is square integrability:

$$\int p(t,x)|x|^2\, dx < \infty. \tag{1.7}$$

This implies the Fourier transform of $p(t,x)x^\alpha$ is bounded and continuous for $|\alpha| \le 2$, hence

$$\partial_\xi^\alpha e^{-t\psi(\xi)} \tag{1.8}$$

bounded and continuous for $|\alpha| \le 2$, and

$$E(|X_t|^2) = -\Delta e^{-t\psi(\xi)}\Big|_{\xi=0}. \tag{1.9}$$

Note that

$$\begin{aligned} \partial_k e^{-t\psi(\xi)} &= -t(\partial_k\psi)e^{-t\psi(\xi)}, \\ \partial_k^2 e^{-t\psi(\xi)} &= -t(\partial_k^2\psi)e^{-t\psi(\xi)} + t^2(\partial_k\psi)^2 e^{-t\psi(\xi)}, \end{aligned} \tag{1.10}$$

hence

$$E(|X_t|^2) = t\Delta\psi(0) - t^2\sum_k(\partial_k\psi(0))^2. \tag{1.11}$$

It is convenient to impose the additional condition

$$E(X_t) = 0, \quad \forall\, t \ge 0. \tag{1.12}$$

Since $X_{t_2} - X_{t_1}$ is independent of X_s for $s < t_1 < t_2$, this implies

$$s < t_1 < t_2 \Longrightarrow X_s \perp X_{t_2} - X_{t_1}, \tag{1.13}$$

with respect to the inner product on $L^2(\Omega,\mu)$. A calculation parallel to (1.9) gives

$$E(X_t) = i\nabla_\xi e^{-t\psi(\xi)}\Big|_{\xi=0} = -it\nabla\psi(0), \tag{1.14}$$

so we require

$$\nabla\psi(0) = 0. \tag{1.15}$$

In such a case, (1.11) yields

$$E(|X_t|^2) = t\Delta\psi(0). \tag{1.16}$$

Thus we have

$$X : \mathbb{R}^+ \longrightarrow L^2(\Omega,\mu,\mathbb{R}^n) \tag{1.17}$$

continuous, satisfying, for $0 \le s < t$,

$$\|X_t - X_s\|^2_{L^2(\Omega)} = A(t-s), \quad A = \Delta\psi(0). \tag{1.18}$$

Regarding the examples of $\psi(\xi)$ given in (1.4) and (1.5), we note that both yield square integrable Lévy processes. In (1.4), we have $\nabla\psi(0) = 0$, but in (1.5) we have $\nabla\psi(0) = -icy$. Hence (1.15) fails for the Poisson process. We modify it by adding a drift term, obtaining

$$\psi(\xi) = c(1 - e^{iy\cdot\xi} + iy\cdot\xi). \tag{1.19}$$

This yields a Lévy process satisfying (1.18), with $A = |y|^2$.

We recall some other examples of Lévy processes mentioned in Chapter 3. For example,

$$\varphi_\alpha(\xi) = (|\xi|^2 + 1)^\alpha - 1, \quad \alpha \in (0,1) \tag{1.20}$$

arises in (3.10) of Chapter 3, and (3.11) there verifies (1.18) of this chapter, with $A = \Delta\varphi_\alpha(0)$. Clearly $\nabla\varphi_\alpha(0) = 0$. By contrast,

$$\psi_\alpha(\xi) = |\xi|^{2\alpha}, \quad \alpha \in (0,1), \tag{1.21}$$

introduced above (1.7) of Chapter 3, yields $p(t,x)$ with the behavior described in (1.14) of Chapter 3, which implies that $\int p(t,x)|x|^2\,dx = \infty$ for such cases. One says these probability distributions have heavy tails. Our treatment of stochastic integrals in this chapter does not apply to such cases as arise from (1.21).

Before describing subsequent sections, we provide some useful complements to the identities (1.9)–(1.18). Here we make the standing assumption that (1.7) and (1.15) hold, hence (1.12) and (1.18) hold. Complementing this, we have

$$E(X_t^* X_t) = \int p(t,x)x^*x\,dx = C(t) \in M(n,\mathbb{R}), \tag{1.22}$$

where $C(t) = (c_{jk}(t))$ is symmetric and positive semidefinite, and, parallel to (1.9)–(1.16),

$$\begin{aligned} c_{jk}(t) &= \int p(t,x)x_j x_k\,dx \\ &= -\partial_j\partial_k e^{-t\psi(\xi)}\Big|_{\xi=0} \\ &= t\partial_j\partial_k\psi(0), \end{aligned} \tag{1.23}$$

the latter identity using (1.15). Hence

$$C(t) = Ct, \quad C = (c_{jk}) = \big(\partial_j\partial_k\psi(0)\big), \tag{1.24}$$

Choose an orthonormal basis of $\mathbb{R}^n$ with respect to which C is diagonal,

$$C = \begin{pmatrix} c_{11} & & \\ & \ddots & \\ & & c_{nn} \end{pmatrix}, \quad c_{jj} = \partial_j^2 \psi(0) = \int p(t,x) x_j^2 \, dx. \tag{1.25}$$

We see that

$$c_{jj} = 0 \Longrightarrow \operatorname{supp} p(t,\cdot) \subset \{x : x_j = 0\}, \tag{1.26}$$

so $p(t,x)$ is a singular measure. Effectively, we have a Lévy process taking values in a linear subspace of $\mathbb{R}^n$. We can eliminate the redundant variables, and re-notate, so we are left with a positive definite matrix C. For example, if $\psi(\xi)$ is given by (1.19), we retain only one variable, thus arriving at the situation

$$n = 1, \quad \psi(\xi) = c(1 - e^{iy\xi} + iy\xi), \quad y \in \mathbb{R}. \tag{1.27}$$

We proceed to describe the rest of this chapter. In §2 we define the Wiener-type stochastic integral

$$I_t(f) = \int_0^t f(s)\, dX_s, \tag{1.28}$$

when X is a Lévy process satisfying (1.12) and (1.18). If f is real valued, then we have

$$\|I_t f\|^2_{L^2(\Omega)} = A \|f\|^2_{L^2([0,t])}. \tag{1.29}$$

More generally, if f takes values in $M(n,\mathbb{R})$, we have

$$\|I_t f\|^2_{L^2(\Omega)} \le A \|f\|^2_{L^2([0,t])}. \tag{1.30}$$

We also devote some attention to Lévy-Langevin equations, of the form

$$dY_t = -LY_t\, dt + dX_t, \tag{1.31}$$

with $L \in M(n,\mathbb{R})$, obtaining explicit formulas for the solution.

In §2 we treat more general Ito-type stochastic integrals, such as

$$I_t(f) = \int_0^t f(s, X_s)\, dX_s, \tag{1.32}$$

and extend the scope of (1.30) to

$$\|I_t f\|^2_{L^2(\Omega)} \le C \int_0^t E\big(|f(s,X_s)|^2\big)\, ds. \tag{1.33}$$

We also present Ito formulas, such as

$$f(X_t) - f(0) = \int_0^t f'(X_s)\,dX_s + \int_0^t f''(X_s)\,ds, \tag{1.34}$$

valid in the special case of the Wiener process.

Appendix A discusses the central limit theorem for an n-dimensional square-integrable Lévy process, and presents it as a description of the long time behavior of the probability distribution of the normalized process

$$\frac{1}{\sqrt{t}} X_t, \tag{1.35}$$

converging in $\mathcal{S}'(\mathbb{R}^n)$, hence weak* in $\mathcal{M}(\widehat{\mathbb{R}}^n)$, to a Gaussian distribution, thus making contact with material of Chapter 1.

2. Wiener-Lévy stochastic integrals

As stated in §1, we are working with n-dimensional Lévy processes, of the form

$$X : \mathbb{R}^+ \longrightarrow L^2(\Omega, \mu, \mathbb{R}^n) \tag{2.1}$$

(where (Ω, μ) is a probability space), satisfying, for $0 \le s < t$,

$$E(|X_t - X_s|^2) = A|t - s|, \tag{2.2}$$

and, for $t \ge 0$,

$$E(X_t) = 0. \tag{2.3}$$

Since, for $s \le t_1 < t_2$,

$$X_s \text{ and } X_{t_2} - X_{t_1} \text{ are independent random variables,} \tag{2.4}$$

we see that (2.3) implies

$$X_{t_2} - X_{t_1} \perp X_{s_2} - X_{s_1}, \text{ for } 0 \le s_1 < s_2 \le t_1 < t_2, \tag{2.5}$$

and we use the $L^2(\Omega, \mu)$ inner product. We aim to define the Wiener-type stochastic integral

$$\int_0^t f(s)\, dX_s, \tag{2.6}$$

for certain classes of functions f. To start, we take $f(s)$ to be real valued.

To proceed, we fix $t \in (0, \infty)$ and define (2.6) f in the space $\mathrm{PK}([0, t))$ of piecewise constant functions on $[0, t)$, where

$$f \in \mathrm{PK}([0, t)) \tag{2.7}$$

provided there exist t_j, satisfying $0 = t_0 < t_1 < \cdots < t_k = t$, such that f is constant on each interval $[t_j, t_{j+1})$, $0 \le j \le k - 1$. We define

$$I_t : \mathrm{PK}([0, t)) \longrightarrow L^2(\Omega, \mu, \mathbb{R}^n) \tag{2.8}$$

by

$$\begin{aligned} I_t(f) &= f(0)(X_{t_1} - X_{t_0}) + f(t_1)(X_{t_2} - X_{t_1}) + \cdots + f(t_{k-1})(X_{t_k} - X_{t_{k-1}}) \\ &= \sum_{j=0}^{k-1} f(t_j)(X_{t_{j+1}} - X_{t_j}). \end{aligned} \tag{2.9}$$

One can check that this is stable under further refinement of the partition of $[0,t)$, that I_t in (2.8) is linear, and that

$$\begin{aligned}
&f \in \mathrm{PK}([0,t)) \Rightarrow \\
&\|I_t f\|^2_{L^2(\Omega)} = \sum_{j=1}^{k-1} |f(t_j)|^2 \|X_{t_{j+1}} - X_{t_j}\|^2_{L^2(\Omega)} \\
&\qquad = A \sum_{j=0}^{k-1} |f(t_j)|^2 (t_{j+1} - t_j) \\
&\qquad = A \int_0^t |f(s)|^2 \, ds.
\end{aligned} \tag{2.10}$$

This leads to the following.

Proposition 2.1. *Let X be a Lévy process that satisfies (2.1)–(2.3). For each $t > 0$, the map I_t in (2.8) has a unique continuous linear extension to*

$$I_t : L^2([0,t), dt) \longrightarrow L^2(\Omega, \mu, \mathbb{R}^n), \tag{2.11}$$

satisfying

$$\|I_t f\|^2_{L^2(\Omega)} = A\|f\|^2_{L^2([0,t))}. \tag{2.12}$$

For such f, we define

$$\int_0^t f(s)\, dX_s = I_t(f). \tag{2.13}$$

As a corollary, we have a well defined linear map

$$I_t : C([0,t]) \longrightarrow L^2(\Omega, \mu, \mathbb{R}^n), \tag{2.14}$$

satisfying (2.12). It is routine to verify that, for $f \in C([0,t])$,

$$I_t(f) = \lim_{k\to\infty} \sum_{j=0}^{k-1} f(t_j)(X_{t_{j+1}} - X_{t_j}), \quad t_j = \frac{jt}{k}, \tag{2.15}$$

the limit taken in $L^2(\Omega, \mu)$-norm.

Note that applying summation by parts to (2.15) and making a limiting argument yields the following.

Proposition 2.2. *Given* $f \in C^1([0,t])$,

$$\int_0^t f(s)\,dX_s = f(t)X_t - \int_0^t f'(s)X_s\,ds. \tag{2.16}$$

More generally, (2.16) holds whenever $f \in C([0,t])$ has integrable weak derivative

$$f' \in L^1([0,t]). \tag{2.17}$$

More generally still, an approximation argument yields

Corollary 2.3. *Given* $f \in \mathrm{BV}([0,t])$,

$$\int_0^t f(x)\,dX_s = f(t-)X_t - \int\limits_{(0,t)} X_s\,df(s). \tag{2.18}$$

If we allow $f(s)$ to take values in $M(n,\mathbb{R})$,

$$f : [0,t) \longrightarrow M(n,\mathbb{R}), \tag{2.19}$$

then most computations above extend. We simply replace (2.10) by

$$\begin{aligned} f \in \mathrm{PK}([0,t), M(n,\mathbb{R})) \Rightarrow \\ \|I_t f\|^2_{L^2(\Omega)} &= \sum_{j=0}^{k-1} \|f(t_j)(X_{t_{j+1}} - X_{t_j})\|^2_{L^2(\Omega)} \\ &\le \sum_{j=0}^{k-1} \|f(t_j)\|^2 \|X_{t_{j+1}} - X_{t_j}\|^2_{L^2(\Omega)} \\ &= A\int_0^t \|f(s)\|^2\,ds, \end{aligned} \tag{2.20}$$

where $\|f(s)\|$ denotes the matrix operator norm of $f(s) \in M(n,\mathbb{R})$. From this, Proposition 2.1 extends, with (2.12) replaed by

$$\|I_t f\|^2_{L^2(\Omega)} \le A\|f\|^2_{L^2([0,t),M(n,\mathbb{R}))}. \tag{2.21}$$

We continue to have Proposition 2.2 and Corollary 2.3, except that noncommutativity of matrix multiplication forces us to change (2.18) to

$$\int_0^t f(s)\,dX_s = f(t-)X_t - \int\limits_{(0,t)} df(s)\,X_s. \tag{2.22}$$

Lévy-Langevin equations

Let X be an n-dimensional Lévy process satisfying (2.1)–(2.3), and take $L \in M(n, \mathbb{R})$. We consider Langevin equations, of the form

$$dY_t = -LY_t\,dt + dX_t, \tag{2.23}$$

for a stochastic process $Y : \mathbb{R}^+ \to L^2(\Omega, \mu, \mathbb{R}^n)$, to be determined. Equivalently,

$$Y_t = Y_0 - \int_0^t LY_s\,ds + X_t. \tag{2.24}$$

Anticipating that arguments parallel to (2.15)–(2.16) apply to Y_t, we write

$$d(f(t)Y_t) = f(t)\,dY_t + f'(t)Y_t\,dt, \tag{2.25}$$

when $f \in C^1(\mathbb{R}^+, M(n, \mathbb{R}))$. We apply this with $f(t) = e^{tL}$, so

$$d(e^{tL}Y_t) = e^{tL}(dY_t + LY_t\,dt). \tag{2.26}$$

Then (2.23) yields

$$d(e^{tL}Y_t) = e^{tL}dX_t, \tag{2.27}$$

which integrates to

$$e^{tL}Y_t = Y_0 + \int_0^t e^{sL}\,dX_s. \tag{2.28}$$

Hence the solution to (2.24) is

$$\begin{aligned} Y_t &= e^{-tL}Y_0 + e^{-tL}\int_0^t e^{sL}\,dX_s \\ &= e^{-tL}Y_0 + X_t - L\int_0^t e^{-(t-s)L}X_s\,ds, \end{aligned} \tag{2.29}$$

the latter identity by (2.16).

In the Langevin theory it is also useful to consider the integrated process

$$Z_t = \int_0^t Y_\tau\,d\tau. \tag{2.30}$$

In case $Y_0 = 0$, we obtain from (2.29) that

$$Z_t = \int_0^t e^{-(t-s)L}X_s\,ds. \tag{2.31}$$

In such a case,

$$LZ_t = X_t - \int_0^t e^{-(t-s)L}\, dX_s = X_t - R_t, \tag{2.32}$$

and

$$\|R_t\|^2_{L^2(\Omega)} \le A \int_0^t \|e^{-(t-s)L}\|^2_{\mathcal{L}(\mathbb{R}^n)}\, ds. \tag{2.33}$$

In particular, if, for some $\lambda > 0$,

$$\|e^{tL}\|^2_{\mathcal{L}(\mathbb{R}^n)} \le e^{-t\lambda}, \tag{2.34}$$

then

$$\|R_t\|^2_{L^2(\Omega)} \le A\frac{1 - e^{-t\lambda}}{\lambda}. \tag{2.35}$$

3. Ito-type stochastic integrals, basic case

A. Square integrable Lévy processes vs. the CLT

Let $X : \mathbb{R}^+ \to L^2(\Omega, \mu, \mathbb{R}^n)$ be an n-dimensional Lévy process, satisfying (1.12) and (1.18), hence (1.7), (1.15), and (1.16). If we fix $t > 0$ and let $\{Y_{t,j} : j \in \mathbb{N}\}$ be a family of IID random variables with the same probability distribution as X_t. the central limit theorem applies to this family. An example would be

$$Y_{t,j} = X_{jt} - X_{(j-1)t}. \tag{A.0}$$

This yields conclusions that at first glance might seem to be at odds with the Lévy-process behavior of $\{X_t : t \in \mathbb{R}^+\}$, in the non-Gaussian case. Here we provide some formulas that clarify the situation.

As explained in the introduction to this chapter, we can if necessary replace $\mathbb{R}^n$ by a linear subspace and arrange that the semigroup $e^{-t\psi(D)}$ has the property that

$$C = \bigl(\partial_j \partial_k \psi(0)\bigr) \quad \text{is positive definite.} \tag{A.1}$$

We assume that done.

Now the probability distributions that arise for the sequence

$$\frac{1}{\sqrt{k}}\bigl(Y_{t,1} + \cdots + Y_{t,k}\bigr), \tag{A.2}$$

e.g., the probability distributions of

$$\frac{1}{\sqrt{k}} X_{kt}, \tag{A.2A}$$

in the case (A.0), have the form (cf. (1.3) of Chapter 1)

$$e^{-kt\psi(D/\sqrt{k})}\delta(x) = q_k(t, \cdot), \tag{A.3}$$

while those associated to the Lévy process X_t have the form

$$e^{-t\psi(D)}\delta(x) = p(t, \cdot). \tag{A.4}$$

We see that (A.3) and (A.4) are distinct objects. Let us examine the behavior of (A.3) as $k \to \infty$.

We have for the Fourier transform

$$\hat{q}_k(t, \xi) = e^{-kt\psi(\xi/\sqrt{k})} = e^{-t\psi_k(\xi)}. \tag{A.5}$$

Note that

$$\operatorname{Re}\psi(\xi) \ge 0, \quad \forall\, \xi \in \mathbb{R}^n, \tag{A.6}$$

and, since $\psi(0) = 0$ and $\nabla\psi(0) = 0$,

$$\psi(\xi) = \tfrac{1}{2}\xi \cdot C\xi + o(|\xi|^2), \quad \text{for} \ \ |\xi| \le 1, \tag{A.7}$$

hence

$$\begin{aligned} &\psi_k(\xi) = \tfrac{1}{2}\xi \cdot C\xi + R_k(\xi), \quad \text{for} \ \ |\xi| \le k, \\ &|R_k(\xi)| \le \rho_k |\xi|^2, \quad \rho_k \to 0 \text{ as } k \to \infty. \end{aligned} \tag{A.8}$$

Hence

$$|e^{-t\psi_k(\xi)}| \le 1, \quad e^{-t\psi_k(\xi)} \to e^{-(t/2)\xi\cdot C\xi}, \tag{A.9}$$

as $k \to \infty$, for each $\xi \in \mathbb{R}^n$. It follows that

$$e^{-t\psi_k(\xi)} \longrightarrow e^{-(t/2)\xi\cdot C\xi} \text{ in } \mathcal{S}'(\mathbb{R}^n), \tag{A.10}$$

so

$$q_k(t,\cdot) \longrightarrow h(t,\cdot) \text{ in } \mathcal{S}'(\mathbb{R}^n), \tag{A.11}$$

where

$$h(t,x) = \det(4\pi t C)^{-1} e^{-x\cdot C^{-1}x/4t}. \tag{A.12}$$

Since $q_k(t,\cdot)$ and $h(t,\cdot)$ are all probability measures on $\mathbb{R}^n$, we also have

$$q_k(t,\cdot) \longrightarrow h(t,\cdot), \ \text{weak}^* \text{ in } \mathcal{M}(\widehat{\mathbb{R}}^n). \tag{A.13}$$

We have thus reaffirmed the CLT for the random variables $\{Y_{t,j} : j \in \mathbb{N}\}$, and seen in (A.3)–(A.4) how this fits in with the Lévy-process behavior of $\{X_t : t \in \mathbb{R}^+\}$.

Example 1. Take

$$\varphi_\alpha(\xi) = (|\xi|^2 + 1) - 1, \quad \alpha \in (0,1), \tag{A.14}$$

as in (1.20). Then

$$e^{-kt\varphi_\alpha(\xi/\sqrt{k})} \longrightarrow e^{-\alpha t|\xi|^2}, \tag{A.15}$$

pointwise and boundedly, hence in $\mathcal{S}'(\mathbb{R}^n)$, as $k \to \infty$.

Example 2. Take $n = 1$, $y \in \mathbb{R}$, and consider the Poisson process, associated to

$$\psi(\xi) = 1 - e^{iy\xi} + iy\xi. \tag{A.16}$$

Then

$$e^{-kt\psi(\xi/\sqrt{k})} \longrightarrow e^{-(y^2/2)t\xi^2}, \tag{A.17}$$

pointwise and boundedly, hence in $\mathcal{S}'(\mathbb{R})$, as $k \to \infty$.

References

[Ap] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge Univ. Press, 2004.

[CW] K. Chung and R. Williams, Introduction to Stochastic Integration, Birkhauser, Boston, 1990.

[I] K. Ito, On Stochastic Differential Equations, Memoirs AMS #4, 1951.

[K] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge Univ. Press, 1990.

[McK] Stochastic Integrals, Academic Press NY, 1977.

[T] M. Taylor, Partial Differential Equations, Vol. 2, Springer NY, 1996 (2nd ed., 2011).

6. Multidimensional random fields: stationarity, ergodicity, and spectral behavior

1. Introduction and definitions

A random field Z on n-dimensional Euclidean space $\mathbb{R}^n$ or the lattice $\mathbb{Z}^n$ (also called a random function, or a stochastic process indexed by $\mathbb{R}^n$ or $\mathbb{Z}^n$) assigns to each $x \in \mathbb{R}^n$ (resp., $\mathbb{Z}^n$) a random variable $Z(x)$ on some probability space (Ω, μ) (where μ is a probability measure on the set Ω). For definiteness, let us say

$$Z : \mathbb{F}^n \longrightarrow L^2(\Omega, \mu), \tag{1.1}$$

where $L^2(\Omega, \mu)$ denotes the space of square-integrable functions (random variables with finite first and second moments) on Ω. Here and below, $\mathbb{F}$ denotes either $\mathbb{R}$ or $\mathbb{Z}$. We assume the random variables $Z(x)$ are real valued, until §8.

We use Z to assign a probability measure ν on the set $\mathcal{O} = \mathbb{R}^{\mathbb{F}^n}$ of all functions from $\mathbb{F}^n$ to $\mathbb{R}$, as follows. First, Z gives rise to a map

$$F : \Omega \longrightarrow \mathcal{O}, \tag{1.2}$$

defined as follows. If $\xi \in \Omega$, $F(\xi) \in \mathcal{O}$ is a function on $\mathbb{F}^n$ whose value at $x \in \mathbb{F}^n$ is $Z(x)(\xi)$, i.e.,

$$F(\xi)(x) = Z(x)(\xi), \quad \xi \in \Omega,\ x \in \mathbb{F}^n. \tag{1.3}$$

(Recall that $Z(x)$ is a function on Ω, for each $x \in \mathbb{F}^n$.) Then ν is defined by

$$\nu(S) = \mu(F^{-1}(S)), \tag{1.4}$$

when $S \subset \mathcal{O}$ is a measurable set. The probability measure ν incorporates the joint probability distributions of the random variables $Z(x)$, as x runs over $\mathbb{F}^n$, as we indicate below. Another way to write (1.4) is as

$$\int_{\mathcal{O}} \varphi(\eta)\, d\nu(\eta) = \int_{\Omega} \varphi(F(\xi))\, d\mu(\xi). \tag{1.5}$$

Let us consider some special cases. Pick $x_1, x_2 \in \mathbb{F}^n$ and set

$$\varphi_1(\eta) = \eta(x_1), \quad \varphi_2(\eta) = \eta(x_1)\eta(x_2). \tag{1.6}$$

Then

$$\varphi_1(F(\xi)) = Z(x_1)(\xi), \quad \varphi_2(F(\xi)) = Z(x_1)(\xi)\, Z(x_2)(\xi), \tag{1.7}$$

so

$$\int_{\mathcal{O}} \varphi_1 \, d\nu = \int_{\Omega} Z(x_1)(\xi) \, d\mu(\xi) = \langle Z(x_1) \rangle, \tag{1.8}$$

and

$$\int_{\mathcal{O}} \varphi_2 \, d\nu = \int_{\Omega} Z(x_1)(\xi) \, Z(x_2)(\xi) \, d\mu(\xi) = \langle Z(x_1) Z(x_2) \rangle. \tag{1.9}$$

We see that (1.8) is the mean of the random variable $Z(x_1)$. The quantity (1.9) together with the means of $Z(x_1)$ and of $Z(x_2)$ are ingredients in the formula for the covariance of $Z(x_1)$ and $Z(x_2)$.

In further preparation for defining the concepts of stationarity and ergodicity, we bring in the action of $\mathbb{F}^n$ on $\mathcal{O}$,

$$\tau_y : \mathcal{O} \longrightarrow \mathcal{O}, \quad y \in \mathbb{F}^n, \tag{1.10}$$

defined as follows. If $y \in \mathbb{F}^n$ and $\eta \in \mathcal{O}$ (so η is a function, $\eta : \mathbb{F}^n \to \mathbb{R}$), $\tau_y \eta \in \mathcal{O}$ is given by

$$\tau_y \eta(x) = \eta(x + y), \quad x, y \in \mathbb{F}^n. \tag{1.11}$$

Definition 1.1. The random field Z is *stationary* provided τ_y preserves the probability measure ν, for each $y \in \mathbb{F}^n$. Equivalently, if $\varphi \in L^1(\mathcal{O}, \nu)$,

$$\int_{\mathcal{O}} \varphi(\tau_y \eta) \, d\nu(\eta) = \int_{\mathcal{O}} \varphi(\eta) \, d\nu(\eta), \quad \forall \, y \in \mathbb{F}^n. \tag{1.12}$$

An alternative label for such a field Z is *homogeneous.*

If φ_1 and φ_2 are defined as in (1.6), then

$$\varphi_1(\tau_y \eta) = \eta(x_1 + y), \quad \text{and} \quad \varphi_2(\tau_y \eta) = \eta(x_1 + y) \eta(x_2 + y), \tag{1.13}$$

so, parallel to (1.8)–(1.9), we have

$$\begin{aligned} \int_{\mathcal{O}} \varphi_1(\tau_y \eta) \, d\nu(\eta) &= \int_{\Omega} Z(x_1 + y)(\xi) \, d\mu(\xi) \\ &= \langle Z(x_1 + y) \rangle, \end{aligned} \tag{1.14}$$

and

$$\begin{aligned} \int_{\mathcal{O}} \varphi_2(\tau_y \eta) \, d\nu(\eta) &= \int_{\Omega} Z(x_1 + y)(\xi) \, Z(x_2 + y)(\xi) \, d\mu(\xi) \\ &= \langle Z(x_1 + y) Z(x_2 + y) \rangle, \end{aligned} \tag{1.15}$$

so stationarity implies

$$\langle Z(x_1)\rangle = \langle Z(x_1+y)\rangle, \quad \langle Z(x_1)Z(x_2)\rangle = \langle Z(x_1+y)Z(x_2+y)\rangle, \tag{1.16}$$

for each $x_1, x_2, y \in \mathbb{F}^n$.

Definition 1.2. The action $\{\tau_y : y \in \mathbb{F}^n\}$ on $(\mathcal{O}, \nu)$ is *ergodic* provided it preserves the measure ν and the following holds. If $\varphi \in L^1(\mathcal{O}, \nu)$ and

$$\varphi \circ \tau_y = \varphi \ \text{ in } \ L^1(\mathcal{O}, \nu), \quad \forall\, y \in \mathbb{F}^n, \tag{1.17}$$

then φ must be constant (ν-a.e.).

Definition 1.3. Assume Z is a stationary random field. Then Z is ergodic if and only if the action $\{\tau_y : y \in \mathbb{F}^n\}$ on $(\mathcal{O}, \nu)$ is ergodic.

It is useful to introduce the following auxiliary random field, namely

$$\mathcal{Z} : \mathbb{F}^n \longrightarrow L^2(\mathcal{O}, \nu), \tag{1.18}$$

given by

$$\mathcal{Z}(x)(\eta) = \eta(x), \quad x \in \mathbb{F}^n, \quad \eta \in \mathcal{O}. \tag{1.19}$$

By (1.3),

$$\mathcal{Z}(x)(F(\xi)) = Z(x)(\xi), \quad x \in \mathbb{F}^n, \ \xi \in \Omega. \tag{1.20}$$

The process $\mathcal{Z}$ has the same joint distributions as Z. In fact, given $x_1, \dots, x_k \in \mathbb{F}^n$ and suitable $\psi : \mathbb{R}^k \to \mathbb{R}$, we have

$$\begin{aligned} &\int_{\mathcal{O}} \psi(\mathcal{Z}(x_1), \dots, \mathcal{Z}(x_k))\, d\nu \\ &\quad = \int_{\Omega} \psi\big(\mathcal{Z}(x_1)(F(\xi)), \dots, \mathcal{Z}(x_k)(F(\xi))\big)\, d\mu(\xi) \\ &\quad = \int_{\Omega} \psi(Z(x_1), \dots, Z(x_k))\, d\mu, \end{aligned} \tag{1.21}$$

the first identity by (1.5) and the second by (1.20). It follows that the construction described in the first paragraph yields again the same space $(\mathcal{O}, \nu)$. In particular, if Z is stationary and ergodic, so is $\mathcal{Z}$.

The following sequence of identities will prove to be valuable:

$$\mathcal{Z}(x)(\tau_y\eta) = (\tau_y\eta)(x) = \eta(x+y) = \mathcal{Z}(x+y)(\eta), \tag{1.22}$$

valid for $x, y \in \mathbb{F}^n,\ \eta \in \mathcal{O}$.

The rest of this chapter is structured as follows. In §2 we relate spatial averages and ensemble averages of quantities associated to a random field, particularly means and covariances, when the field is ergodic. In §3 we discuss stationary Gaussian fields, and in §4 we give a criterion, involving the behavior of the covariance, that such fields are ergodic. We note that stationary Gaussian fields with covariances given by (3.24), (3.25), (3.26), (3.27), or (when $n > 1$) by (3.33) are ergodic, while those with covariance given by (3.32) are not ergodic.

In §5 we consider stationary random fields on Lie groups, and in §6 we consider stationary random fields on homogeneous spaces. In §§5–6, we focus not on ergodicity but on spectra. In §7 we consider the inverse problem of constructing a random field on a compact homogeneous space, given spectral data. These sections bring in basic concepts from the representation theory of Lie groups, which can be found in [T3].

In §8 we take a finite-dimensional vector space V and discuss V-valued random fields, defined first on a homogeneous space X, though we specialize to $X = \mathbb{R}^n$, with special attention to $V = \mathbb{R}^n$, i.e., to random vector fields. In §9 we discuss random divergence-free vector fields on $\mathbb{R}^n$.

In §10 we discuss generalized random fields on $\mathbb{R}^n$, which are distributions on $\mathbb{R}^n$ with values in $L^2(\Omega, \mu)$. We define stationary generalized random fields and develop some of their properties.

We have three appendices. Appendix A gives background on ergodic theorems, and Appendix B relates the criterion on the covariance function given in §4 to the behavior of its Fourier transform. Appendix C discusses the Fourier transform of a continuous, stationary field, first on $\mathbb{T}^n$ (obtaining a special case of results of §5) and then on $\mathbb{R}^n$, where we need to regard $\widehat{Z}$ as a vector-valued tempered distribution.

2. Implications of the ergodic theorem

The significance of the property of ergodicity, defined in §1, arises from the following result, known as the ergodic theorem. As before, $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{Z}$.

Theorem 2.1. *Let $\{\tau_y : y \in \mathbb{F}^n\}$ consist of measure preserving maps on the probability space $(\mathcal{O}, \nu)$, satisfying $\tau_{y_1+y_2} = \tau_{y_1}\tau_{y_2}$, for $y_1, y_2 \in \mathbb{F}^n$. Assume the action is ergodic. Take $\varphi \in L^1(\mathcal{O}, \nu)$.*

(A) If $\mathbb{F} = \mathbb{Z}$, then

$$\lim_{R\to\infty} \frac{1}{V(R)} \sum_{y\in\mathbb{Z}^n\cap B_R} \varphi(\tau_y\eta) = \int_{\mathcal{O}} \varphi\, d\nu, \tag{2.1}$$

for ν-almost every $\eta \in \mathcal{O}$.

(B) If $\mathbb{F} = \mathbb{R}$, and if the action of τ_y on $L^1(\mathcal{O}, \nu)$ is strongly continuous in y, then

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{B_R} \varphi(\tau_y \eta)\, dy = \int\limits_{\mathcal{O}} \varphi\, d\nu, \tag{2.2}$$

for ν-almost every $\eta \in \mathcal{O}$.

Here, $B_R = \{y \in \mathbb{R}^n : |y| \le R\}$ is a ball and $V(R)$ is its volume (which is a good approximation to the number of points in $\mathbb{Z}^n \cap B_R$). The left sides of (2.1) and (2.2) are spatial averages, and the right sides are ensemble averages.

We apply Theorem 2.1 to results discussed in §1. First, take $\mathbb{F} = \mathbb{Z}$, so $Z : \mathbb{Z}^n \to L^2(\Omega, \mu)$ is a random field on the discrete lattice $\mathbb{Z}^n$. We construct $(\mathcal{O}, \nu)$ and $\tau_y : \mathcal{O} \to \mathcal{O}$ as in §1. If Z is stationary and ergodic, then (2.1) holds, for ν-almost every $\eta \in \mathcal{O}$. If φ_1 and $\varphi_2 \in L^1(\mathcal{O}, \nu)$ are defined as in (1.6), then (1.8)–(1.9) and (1.13), in concert with (2.1), give, for each $x_1, x_2 \in \mathbb{Z}^n$,

$$\langle Z(x_1) \rangle = \lim_{R\to\infty} \frac{1}{V(R)} \sum_{y \in \mathbb{Z}^n \cap B_R} \eta(x_1 + y), \tag{2.3}$$

and

$$\langle Z(x_1) Z(x_2) \rangle = \lim_{R\to\infty} \frac{1}{V(R)} \sum_{y \in \mathbb{Z}^n \cap B_R} \eta(x_1 + y)\eta(x_2 + y), \tag{2.4}$$

for ν-almost every $\eta \in \mathcal{O}$.

In case $\mathbb{F} = \mathbb{R}$, matters are not so simple, because the requirement that $\varphi \circ \tau_y \in L^1(\mathcal{O}, \nu)$ be continuous in y for $\varphi \in L^1(\mathcal{O}, \nu)$ can fail to hold. If this continuity did hold it would apply to φ_1 and φ_2, given by (1.6). In such a case, (2.2) would yield, for each $x_1, x_2 \in \mathbb{R}^n$,

$$\langle Z(x_1) \rangle = \lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{B_R} \eta(x_1 + y)\, dy, \tag{2.5}$$

and

$$\langle Z(x_1) Z(x_2) \rangle = \lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{B_R} \eta(x_1 + y)\eta(x_2 + y)\, dy, \tag{2.6}$$

for ν-almost every $\eta \in \mathcal{O}$, provided the random field Z is stationary and ergodic.

Suppose for example that the random variables $Z(x)$ are identically distributed and *independent*, as x runs over $\mathbb{R}^n$, and that the distribution of $Z(0)$ is not concentrated at a single point. Then ν is a product measure on $\mathcal{O}$, an *uncountable* product measure. With φ_1 as above, we have

$$\|\varphi_1 \circ \tau_y - \varphi_1\|_{L^1(\mathcal{O},\nu)} = \int\limits_{\mathcal{O}} |\eta(x_1 + y) - \eta(x_1)|\, d\nu(\eta) \tag{2.7}$$

equal to 0 for $y = 0$, and to a nonzero constant *independent of* y if $y \neq 0$. It follows that $\varphi_1 \circ \tau_y$ is an everywhere discontinuous function of y, with values in $L^1(\mathcal{O}, \nu)$. Furthermore, we expect that, for ν-almost every $\eta \in \mathcal{O}$, the function $y \mapsto \eta(y)$ is not Lebesgue measurable, so the right sides of (2.5) and (2.6) are not well defined.

On the other hand, for many important random fields on $\mathbb{R}^n$, matters are more tractable.

Proposition 2.2. *If $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ is stationary and continuous, then the action of $\{\tau_y : y \in \mathbb{R}^n\}$ on $L^1(\mathcal{O}, \nu)$ is strongly continuous.*

Proof. Since $\{\tau_y\}$ is a group of isometries of $L^1(\mathcal{O}, \nu)$, it suffices to show that $y \mapsto \varphi \circ \tau_y$ is continuous from $\mathbb{R}^n$ to $L^1(\mathcal{O}, \nu)$, for φ in a dense subspace of $L^1(\mathcal{O}, \nu)$. We consider functions φ of the form

$$\varphi(\eta) = \psi(\eta(x_1), \dots, \eta(x_k)), \tag{2.8}$$

where $x_1, \dots, x_k \in \mathbb{R}^n$ and $\psi : \mathbb{R}^k \to \mathbb{R}$ is globally Lipschitz.

We have

$$\begin{aligned} &|\varphi \circ \tau_y(\eta) - \varphi(\eta)| \\ &= |\psi(\eta(x_1 + y), \dots, \eta(x_k + y)) - \psi(\eta(x_1), \dots, \eta(x_k))| \\ &\le C \sum_{j=1}^{k} |\eta(x_j + y) - \eta(x_j)|. \end{aligned} \tag{2.9}$$

Hence

$$\begin{aligned} \|\varphi \circ \tau_y - \varphi\|_{L^1(\mathcal{O},\nu)} &\le C \sum_j \int_{\mathcal{O}} |\eta(x_j + y) - \eta(x_j)| \, d\nu(\eta) \\ &= C \sum_j \int_{\mathcal{O}} |\mathcal{Z}(x_j + y) - \mathcal{Z}(x_j)| \, d\nu \\ &= C \sum_j \int_{\Omega} |Z(x_j + y) - Z(x_j)| \, d\mu, \end{aligned} \tag{2.10}$$

the first identity by (1.22) and the second by (1.21). The last line is bounded by $C \sum_j \|Z(x_j + y) - Z(x_j)\|_{L^2(\Omega,\mu)}$, so Proposition 2.2 is proven.

Note that if $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ is stationary, then

$$\begin{aligned} \|Z(x + y) - Z(x)\|_{L^2}^2 &= \|Z(x + y)\|_{L^2}^2 + \|Z(x)\|_{L^2}^2 - 2\langle Z(x)Z(x + y)\rangle \\ &= 2\|Z(0)\|_{L^2}^2 - 2\langle Z(0)Z(y)\rangle, \end{aligned} \tag{2.11}$$

so Z is continuous if and only if

$$\lim_{y \to 0} \langle Z(0)Z(y)\rangle = \|Z(0)\|_{L^2}^2. \tag{2.12}$$

When dealing with $\mathbb{F} = \mathbb{R}$, we will henceforth assume Z is continuous. However, we note that [R] emphasizes the importance of such discontinuous examples as described above to stochastic hydrogeology. This might point to some mathematical problems that need further study.

3. Stationary Gaussian fields

A random field $Z : \mathbb{F}^n \to L^2(\Omega, \mu)$ is said to be a Gaussian field if the following holds. For each $x_j \in \mathbb{F}^n,\ a_j \in \mathbb{R},\ k \in \mathbb{N}$,

$$\sum_{j=1}^{k} a_j Z(x_j) \ \text{ is a Gaussian random variable.} \tag{3.1}$$

The following important property is special to Gaussian fields.

Proposition 3.1. *If Z is a Gaussian field, then Z is stationary provided*

$$\langle Z(x) \rangle = \langle Z(0) \rangle, \quad \text{and} \quad \langle Z(x) Z(x+y) \rangle = \langle Z(0), Z(y) \rangle, \quad \forall\, x, y, \in \mathbb{F}^n. \tag{3.2}$$

The proof uses the Gaussian property to obtain that, for each $k \ge 1$,

$$\langle Z(x_1 + y) \cdots Z(x_k + y) \rangle \ \text{ is independent of } \ y \in \mathbb{F}^n, \ \forall\, x_1, \dots, x_k \in \mathbb{F}^n. \tag{3.3}$$

This follows from the fact that the data

$$\{ \langle Z(x_1) \rangle, \langle Z(x_1) Z(x_2) \rangle : x_1, x_2 \in \mathbb{F}^n \} \tag{3.4}$$

uniquely determine the data

$$\{ \langle Z(x_1) \cdots Z(x_k) \rangle : x_j \in \mathbb{F}^n,\ k \in \mathbb{N} \}, \tag{3.5}$$

under the hypothesis (3.1). In fact, the data (3.4) determine the data

$$\{ \langle e^{i \sum \lambda_j Z(x_j)} \rangle : x_j \in \mathbb{F}^n,\ \lambda_j \in \mathbb{R} \}, \tag{3.6}$$

which in turn determine (3.5). See (3.11A) below for more on (3.6).

The fact that (3.3) implies the stationarity asserted in Proposition 3.1 is a special case of the following.

Lemma 3.2. *Let $Z : \mathbb{F}^n \to \cap_{p<\infty} L^p(\Omega, \mu)$. If (3.3) holds for each $k \in \mathbb{N}$, then Z is stationary.*

Sketch of proof. For a k-tuple $\overline{x} = (x_1, \dots, x_k)$, define $\varphi_{\overline{x}} \in L^1(\mathcal{O}, \nu)$ by

$$\varphi_{\overline{x}}(\eta) = \eta(x_1) \dots \eta(x_k). \tag{3.7}$$

Then, via (1.21), (3.3) implies

$$\int_{\mathcal{O}} \varphi_{\overline{x}} \circ \tau_y \, d\nu = \int_{\mathcal{O}} \varphi_{\overline{x}} \, d\nu, \quad \forall \, y \in \mathbb{F}^n. \tag{3.8}$$

Now one can show that, in this situation, the set of functions of the form (3.7) has dense linear span in $L^1(\mathcal{O}, \nu)$. This implies the desired stationarity.

We next consider existence of Gaussian fields with given first and second moments. The following is proven in [D], p. 72.

Theorem 3.3. *Let $M : \mathbb{F}^n \to \mathbb{R}$ and $R : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{R}$. Assume*

$$R(x, y) = R(y, x), \tag{3.9}$$

and, for all $k \geq 1$, $x_1, \dots, x_k \in \mathbb{F}^n$, and $a_1, \dots, a_k \in \mathbb{C}$,

$$\sum_{i,j} R(x_i, x_j) a_i \overline{a}_j \geq 0. \tag{3.10}$$

Then there exists a Gaussian field $Z : \mathbb{F}^n \to L^2(\Omega, \mu)$ such that, for all $x_1, x_2 \in \mathbb{F}^n$,

$$\langle Z(x_1) \rangle = M(x_1), \quad \langle (Z(x_1) - M(x_1))(Z(x_2) - M(x_2)) \rangle = R(x_1, x_2). \tag{3.11}$$

We refer to [D] for the proof, but remark that one ingredient is the formula

$$\Big\langle e^{i \sum \lambda_j Z(x_j)} \Big\rangle = \mathrm{Exp}\Big\{ -\frac{1}{2} \sum_{j,k} R(x_j, x_k) \lambda_j \lambda_k + i \sum_j M(x_j) \lambda_j \Big\}, \tag{3.11A}$$

given $x_1, \dots, x_\ell \in \mathbb{F}^n$, $\lambda_1, \dots, \lambda_\ell \in \mathbb{R}$, and $\ell \geq 1$.

Remark. The conditions (3.9)–(3.10) are necessary, as well as sufficient, for the existence of such a field Z.

In concert with Proposition 3.1, Theorem 3.3 yields the following.

Corollary 3.4. *Let $M \in \mathbb{R}$ and let $C : \mathbb{F}^n \to \mathbb{R}$ satisfy*

$$C(x) = C(-x), \tag{3.12}$$

and, for each $k \geq 1$, $x_1, \dots, x_k \in \mathbb{F}^n$, and $a_1, \dots, a_k \in \mathbb{C}$,

$$\sum_{i,j} C(x_i - x_j) a_i \overline{a}_j \geq 0. \tag{3.13}$$

Then there exists a stationary Gaussian field $Z : \mathbb{F}^n \to L^2(\Omega, \mu)$ such that, for each $x_1, x_2 \in \mathbb{F}^n$,

$$\langle Z(x_1) \rangle = M, \tag{3.14}$$

and

$$\langle (Z(x_1) - M)(Z(x_2) - M)) \rangle = C(x_1 - x_2). \tag{3.15}$$

Remark. Given (3.14), the condition (3.15) is equivalent to

$$\langle Z(x_1) Z(x_2) \rangle - M^2 = C(x_1 - x_2).$$

Also, (2.11) implies

$$\frac{1}{2} \| Z(x+y) - Z(x) \|_{L^2}^2 = C(0) - C(y). \tag{3.16}$$

One example of a function $C : \mathbb{R}^n \to \mathbb{R}$ satisfying (3.12)–(3.13) is

$$\begin{aligned} C(x) = 1 \quad &\text{if} \quad x = 0, \\ 0 \quad &\text{if} \quad x \neq 0, \end{aligned} \tag{3.17}$$

which yields a special case of the class of discontinuous random fields discussed in the paragraph following (2.6). If $C : \mathbb{R}^n \to \mathbb{R}$ is continuous and satisfies (3.12)–(3.13), then the stationary Gaussian field Z arising in Corollary 3.4 is continuous, by (3.16).

The search for continuous functions $C : \mathbb{R}^n \to \mathbb{R}$ satisfying (3.12)–(3.13) is aided by the Fourier transform, as we now discuss. (Note that, for such C, the restriction to $\mathbb{Z}^n$ also satisfies (3.12)–(3.13).) The Fourier transform of a function $F \in L^1(\mathbb{R}^n)$ is given by

$$\hat{F}(x) = (2\pi)^{-n/2} \int\limits_{\mathbb{R}^n} F(\xi) e^{-ix\cdot\xi} \, d\xi. \tag{3.18}$$

In such a case, $C = \hat{F}$ is continuous on $\mathbb{R}^n$. If F is even (i.e., $F(\xi) = F(-\xi)$) and real valued, so is $C = \hat{F}$, so (3.12) holds. Also,

$$\sum_{j,k} \hat{F}(x_j - x_k) a_j \overline{a}_k = (2\pi)^{-n/2} \int_{\mathbb{R}^n} F(\xi) B(\xi)\, d\xi, \tag{3.19}$$

where

$$\begin{aligned} B(\xi) &= \sum_{j,k} a_j \overline{a}_k e^{-i(x_j - x_k)\cdot\xi} \\ &= \Big| \sum_j a_j e^{-ix_j\cdot\xi} \Big|^2 \\ &\ge 0, \end{aligned} \tag{3.20}$$

so we have the following.

Proposition 3.5. *Let $F \in L^1(\mathbb{R}^n)$ be even and real valued. If*

$$F(\xi) \ge 0, \quad \forall\, \xi \in \mathbb{R}^n, \tag{3.21}$$

then $C(x) = \hat{F}(x)$ is a continuous function satisfying (3.12)–(3.13).

If C is also integrable, the Fourier inversion formula gives $\tilde{C}(\xi) = F(\xi)$, where

$$\tilde{C}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} C(x) e^{ix\cdot\xi}\, dx. \tag{3.22}$$

If C is even, then $\tilde{C} = \hat{C}$, so we have the following.

Corollary 3.6. *Assume $C : \mathbb{R}^n \to \mathbb{R}$ is even, continuous, and integrable. Then (3.13) holds provided*

$$\hat{C}(\xi) \ge 0, \quad \forall\, \xi \in \mathbb{R}^n. \tag{3.23}$$

Remark. If C is real, even, continuous, and integrable, (3.23) is known to be necessary, as well as sufficient, for the validity of (3.13). When C satisfies all these conditions, including (3.23), it can be shown that $\hat{C} \in L^1(\mathbb{R}^n)$. In fact, $\|\hat{C}\|_{L^1} = 2^{n/2} C(0)$.

Here are some examples to which Corollary 3.6 applies.

$$C(x) = e^{-|x|^2/2} \Longrightarrow \hat{C}(\xi) = e^{-|\xi|^2/2}, \tag{3.24}$$

$$C(x) = e^{-|x|} \Longrightarrow \hat{C}(\xi) = c_n(|\xi|^2 + 1)^{-(n+1)/2}, \tag{3.25}$$

where $c_n = 2^{n/2}\pi^{-1/2}\Gamma((n+1)/2)$. These calculations can be found in many places, e.g., Chapter 3 of [T]. Note that applying the Fourier inversion formula to (3.25) gives

$$C(x) = (|x|^2 + 1)^{-(n+1)/2} \Longrightarrow \hat{C}(\xi) = c_n^{-1} e^{-|\xi|}. \tag{3.26}$$

Here is an example where $C(x)$ is not ≥ 0 everywhere. Let $\chi_B(\xi) = 1$ for $|\xi| \leq 1,\ 0$ for $|\xi| > 1$. Then

$$F(\xi) = \chi_B(\xi) \Longrightarrow C(x) = \hat{F}(x) = c_n \, \frac{J_{n/2}(|x|)}{|x|^{n/2}}, \tag{3.27}$$

where c_n are positive constants, and J_ν is the Bessel function of order ν.

Variations on these examples can be obtained by linear changes of variables. If $b_j > 0$, then

$$C_b(x) = C(b_1 x_1, \dots, b_n x_n) \Longrightarrow \hat{C}_b(\xi) = (b_1 \cdots b_n)^{-1} \hat{C}(b_1^{-1}\xi_1, \dots, b_n^{-1}\xi_n). \tag{3.28}$$

More generally, if T is an $n \times n$ real matrix and $\det T \neq 0$,

$$C_T(x) = C(Tx) \Longrightarrow \hat{C}_T(\xi) = |\det T|^{-1} \hat{C}(T^{-1}\xi). \tag{3.29}$$

The following extension of Proposition 3.5 yields more general covariance functions for stationary Gaussian fields.

Proposition 3.7. *If σ is a finite, positive measure on $\mathbb{R}^n$, invariant under $x \mapsto -x$, then*

$$C(x) = \hat{\sigma}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \, d\sigma(\xi) \tag{3.30}$$

is a continuous function satisfying (3.12)–(3.13).

The proof is a slight variant of that of Proposition 3.5. In place of (3.19), we have

$$\sum_{j,k} C(x_j - x_k) a_j \overline{a}_k = (2\pi)^{-n/2} \int_{\mathbb{R}^n} B(\xi) \, d\sigma(\xi), \tag{3.31}$$

with $B(\xi)$ as in (3.20).

The Bochner-Herglotz theorem implies that, conversely, if C is a continuous function satisfying (3.12)–(3.13), then there exists a finite, positive measure σ such that $\hat{\sigma} = C$.

If $p \in \mathbb{R}^n \setminus 0$ and δ_p is the point mass concentrated at p, then

$$\sigma = \delta_p + \delta_{-p} \Longrightarrow C(x) = \hat{\sigma}(x) = \Bigl(\frac{2}{\pi}\Bigr)^{1/2} \cos(p \cdot x). \tag{3.32}$$

If $\delta(|\xi|-1)$ denotes the surface measure of the unit sphere $S^{n-1} = \{\xi \in \mathbb{R}^n : |\xi| = 1\}$, then

$$\sigma = \delta(|\xi|-1) \Longrightarrow C(x) = \hat{\sigma}(x) = |x|^{1-n/2} J_{n/2-1}(|x|), \tag{3.33}$$

where, as in (3.27), J_ν is the Bessel function of order ν. The cases $n = 1$ and $n = 3$ yield

$$|x|^{1/2} J_{-1/2}(|x|) = \Bigl(\frac{2}{\pi}\Bigr)^{1/2} \cos|x|, \quad |x|^{-1/2} J_{1/2}(|x|) = \Bigl(\frac{2}{\pi}\Bigr)^{1/2} \frac{\sin|x|}{|x|}. \tag{3.34}$$

Of course, the $n = 1$ case agrees with (3.32), with $p = 1$.

4. Ergodic Gaussian fields

Here we discuss conditions under which a stationary Gaussian field $Z : \mathbb{F}^n \to L^2(\Omega, \mu)$ is ergodic, i.e., the action $\{\tau_y\}$ on $(\mathcal{O}, \nu)$ is ergodic. The first result gives a condition that implies this action is *mixing*, i.e.,

$$\lim_{|y|\to\infty} \langle \varphi \circ \tau_y \, \psi \rangle = \langle \varphi \rangle \, \langle \psi \rangle, \quad \forall \, \varphi, \psi \in L^2(\mathcal{O}, \nu). \tag{4.1}$$

This condition implies ergodicity. Compare (4.10) below, and see Appendix A for further discussion.

Proposition 4.1. *Let $Z : \mathbb{F}^n \to L^2(\mathcal{O}, \nu)$ be a stationary Gaussian field, with mean $\langle Z(x) \rangle = M$ and covariance*

$$C(y) = \langle Z(x)Z(x+y) \rangle - M^2. \tag{4.2}$$

If $\mathbb{F} = \mathbb{R}$, assume $C : \mathbb{F}^n \to \mathbb{R}$ is continuous. If

$$\lim_{|y|\to\infty} C(y) = 0, \tag{4.3}$$

then this field is mixing, i.e., (4.1) holds.

For the proof, there is no loss of generality to assume $M = 0$ (and it simplifies some formulas). Also, it suffices to show that (4.1) holds for φ, ψ in some dense subspace of $L^2(\mathcal{O}, \nu)$. We prove it for φ and ψ of the form

$$\varphi(\eta) = f(\eta(x_1), \dots, \eta(x_\ell)), \quad \psi(\eta) = g(\eta(x_1), \dots, \eta(x_\ell)), \tag{4.3A}$$

where $\ell \in \mathbb{N}$, $x_1, \dots, x_\ell \in \mathbb{F}^n$, and $f, g \in \mathcal{S}(\mathbb{R}^\ell)$. Applying the Fourier inversion formula to f and g, we get

$$\varphi \circ \tau_y(\eta) = \int \hat{f}(v) e^{i \sum \eta(x_j+y)v_j} \, dv, \tag{4.4}$$

and a similar formula for $\psi(\eta)$, hence

$$\langle \varphi \circ \tau_y \, \psi \rangle = \iint \hat{f}(v)\hat{g}(w) \Big\langle e^{i\sum \eta(x_j+y)v_j} e^{i\sum \eta(x_j)w_j} \Big\rangle \, dv \, dw. \tag{4.5}$$

Now (3.11A), with $R(x_j, x_k) = C(x_j - x_k)$ amd $M(x_j) \equiv 0$, yields

$$\Big\langle e^{i\sum \eta(x_j)\lambda_j} \Big\rangle = \operatorname{Exp}\Big\{ -\frac{1}{2} \sum_{j,k} C(x_j - x_k)\lambda_j \lambda_k \Big\}, \tag{4.6}$$

hence

$$\begin{aligned} &\Big\langle e^{i\sum \eta(x_j+y)v_j} e^{i\sum \eta(x_j)w_j} \Big\rangle \\ &= e^{-\sum C(x_j - x_k)v_j w_k/2} e^{-\sum C(x_j-x_k)w_j w_k/2} \\ &\quad \times e^{-\sum C(x_j+y-x_k)v_j w_k}. \end{aligned} \tag{4.7}$$

In this setting, $x_1, \dots, x_\ell \in \mathbb{F}^n$ are fixed. The hypothesis (4.3) implies that, as $|y| \to \infty$, the last factor on the right side of (4.7) tends to 1, so

$$\begin{aligned} &\Big\langle e^{i\sum \eta(x_j+y)v_j} e^{i\sum \eta(x_j)w_j} \Big\rangle \\ &\quad \longrightarrow \Big\langle e^{i\sum \eta(x_j)v_j} \Big\rangle \Big\langle e^{i\sum \eta(x_j)w_j} \Big\rangle, \end{aligned} \tag{4.8}$$

pointwise in $v, w \in \mathbb{R}^\ell$. These quantities are ≤ 1 in absolute value, so the dominated convergence theorem applies to (4.5), giving

$$\begin{aligned} &\lim_{|y|\to\infty} \langle \varphi \circ \tau_y \, \psi \rangle \\ &= \iint \hat{f}(v)\hat{g}(w) \Big\langle e^{i\sum \eta(x_j)v_j} \Big\rangle \Big\langle e^{i\sum \eta(x_j)w_j} \Big\rangle \, dv \, dw \\ &= \langle \varphi \rangle \, \langle \psi \rangle, \end{aligned} \tag{4.9}$$

completing the proof of Proposition 4.1.

We move on to more general conditions on C that imply ergodicity. We will work with $\mathbb{F} = \mathbb{R}$. The action of $\{\tau_y\}$ on $(\mathcal{O}, \nu)$ is ergodic provided

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} \langle \varphi \circ \tau_y \, \psi \rangle \, dy = \langle \varphi \rangle \, \langle \psi \rangle, \quad \forall \, \varphi, \psi \in L^2(\mathcal{O}, \nu). \tag{4.10}$$

See Appendix A. Note that (4.1) implies (4.10). To establish (4.10), it suffices to check it for φ, ψ in a dense subspace of $L^2(\mathcal{O}, \nu)$, such as functions of the form

(4.3A). Via (4.5) and the dominated convergence theorem, we see that (4.10) holds for such functions provided

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} \Bigl\langle e^{i\sum \eta(x_j+y)v_j} e^{i\sum \eta(x_j)w_j} \Bigr\rangle \, dy = \Bigl\langle e^{i\sum \eta(x_j)v_j} \Bigr\rangle \Bigl\langle e^{i\sum \eta(x_j)w_j} \Bigr\rangle, \tag{4.11}$$

for each $x_1, \dots, x_\ell \in \mathbb{R}^n$, $v, w \in \mathbb{R}^\ell$, $\ell \in \mathbb{N}$. In turn, by (4.6)–(4.7), we see that (4.11) holds provided

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} e^{-\sum C(x_j+y-x_k)v_jw_k} \, dy = 1, \tag{4.12}$$

for all such x_j, v, w. This holds provided

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} \operatorname{Exp}\Bigl\{-\sum_{j=1}^{\ell} \lambda_j C(x_j+y)\Bigr\} \, dy = 1, \tag{4.13}$$

for each $\ell \in \mathbb{N}$, $x_1, \dots, x_\ell \in \mathbb{R}^n$, $\lambda_1, \dots, \lambda_\ell \in \mathbb{R}$. Note that

$$|C(x_j+y)| \le C(0), \quad \forall\, x_j, y \in \mathbb{R}^n. \tag{4.14}$$

so the integrand in (4.13) is bounded by $e^{C(0)\sum|\lambda_j|}$. Now, for $s \in \mathbb{R}$,

$$e^s = 1 + \rho(s), \tag{4.15}$$

with $\rho(0) = 0$, $\rho'(s) = e^s$, hence

$$|s| \le C(0)L \Longrightarrow |\rho(s)| \le |s| e^{C(0)L}. \tag{4.16}$$

We have the following.

Proposition 4.2. *Let $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ be a stationary Gaussian field, with continuous covariance. If*

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} |C(y)| \, dy = 0, \tag{4.17}$$

then (4.10) holds, and Z is ergodic.

We remark that (4.17) is equivalent to

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} |C(y)|^2 \, dy = 0, \tag{4.18}$$

one implication by Cauchy's inequality and the other by the bound (4.14). In turn, (4.18) is equivalent to the assertion that the measure $\sigma = \hat{C}$ has no atoms. See Appendix B. In such a case, we say Z has continuous spectrum. Thus we have that Z is ergodic provided it has continuous spectrum. The converse is also true. A stationary Gaussian field on $\mathbb{R}^n$ with continuous covariance is ergodic if and only if it has continuous spectrum. This was proved in [M] and [G] when $n = 1$ and in [BE] when $n > 1$.

In light of these results, we see that stationary Gaussian fields with covariances given by (3.24), (3.25), (3.26), (3.27), or (when $n > 1$) by (3.33) are ergodic (in fact, mixing), while those with covariance given by (3.32) are not ergodic.

There is a straightforward analogue of Proposition 4.2 for $\mathbb{F} = \mathbb{Z}$.

5. Stationary random fields on Lie groups

We consider a random field on a Lie group G,

$$Z : G \longrightarrow L^2(\Omega, \mu), \tag{5.1}$$

where (Ω, μ) is a probability space. We assume that the random variables $Z(x)$ are real-valued, and that Z is continuous. Parallel to (1.2), we have

$$F : \Omega \longrightarrow \mathcal{O} = \mathbb{R}^G, \tag{5.2}$$

given by

$$F(\xi)(x) = Z(x)(\xi), \quad \xi \in \Omega,\ x \in G. \tag{5.3}$$

Then we get a probability measure ν on $\mathcal{O}$:

$$\nu(S) = \mu(F^{-1}(S)), \tag{5.4}$$

so

$$\int_{\mathcal{O}} \varphi(\eta)\, d\nu(\eta) = \int_{\Omega} \varphi(F(\xi))\, d\mu(\xi). \tag{5.5}$$

Formulas parallel to (1.6)–(1.9) hold. Parallel to (1.10)–(1.11), we have a g-action on $\mathcal{O}$:

$$\tau_g : \mathcal{O} \to \mathcal{O}, \quad \tau_g \eta(x) = \eta(gx), \quad x, g \in G. \tag{5.6}$$

We say Z is stationary if this G-action preserves ν, i.e.,

$$\int_{\mathcal{O}} \varphi(\tau_g \eta)\, d\nu(\eta) = \int_{\mathcal{O}} \varphi(\eta)\, d\nu(\eta), \quad \forall\, \varphi \in L^1(\mathcal{O}, \nu),\ g \in G. \tag{5.7}$$

Parallel to (1.16), we see that stationarity implies

$$\langle Z(gx_1)\rangle = \langle Z(x_1)\rangle, \quad \langle Z(gx_1)Z(gx_2)\rangle = \langle Z(x_1)Z(x_2)\rangle, \tag{5.8}$$

for all $g, x_1, x_2 \in G$. Consequently

$$\langle Z(x)\rangle = M \tag{5.9}$$

is independent of $x \in G$ and the covariance, given by

$$R(x,y) = \langle Z(x)Z(y)\rangle - M^2, \tag{5.10}$$

satisfies

$$R(x,y) = R(y,x), \quad R(gx,gy) = R(x,y), \tag{5.11}$$

hence there exists $C : G \to \mathbb{R}$ such that

$$R(x,y) = C(x^{-1}y), \quad C(x) = C(x^{-1}), \quad x, y \in G. \tag{5.12}$$

There is a positivity condition parallel to (3.10), which translates, for stationary fields, to

$$\sum_{i,j} C(x_i^{-1}x_j)a_i\overline{a}_j \geq 0. \tag{5.13}$$

Note that if $e \in G$ is the identity element,

$$\begin{aligned} \|Z(x) - Z(y)\|_{L^2}^2 &= \|Z(e) - Z(x^{-1}y)\|_{L^2}^2 \\ &= 2C(e) - 2C(x^{-1}y), \end{aligned} \tag{5.14}$$

so the continuity of a stationary field $Z : G \to L^2(\Omega, \mu)$ is equivalent to the continuity of $C : G \to \mathbb{R}$ at e (and implies the continuity of C on G). As mentioned above, we work under this continuity hypothesis.

Given such continuity, the condition (5.13) is equivalent to

$$\int_G \int_G C(x^{-1}y)f(x)\overline{f(y)}\,dx\,dy \geq 0, \tag{5.15}$$

for all $f \in C_0^\infty(G)$, where dx denotes Haar measure on G. We henceforth assume

$$G \text{ is unimodular,} \tag{5.16}$$

so left Haar measure and right Haar measure coincide. Note that, by (5.12), we can replace $C(x^{-1}y)$ by $C(y^{-1}x)$ in (5.15). Now the *convolution* is defined by

$$f * C(x) = \int_G f(y)C(y^{-1}x)\,dy, \tag{5.17}$$

so the condition (5.15) is equivalent to

$$(f, f * C)_{L^2} \geq 0, \tag{5.18}$$

for all $f \in C_0^\infty(G)$.

From here on in this section we will assume G is *compact*, which of course implies unimodularity. The Peter-Weyl theorem yields a unitary isomorphism

$$\mathcal{F} : L^2(G) \longrightarrow \bigoplus_{\pi \in \widehat{G}} \operatorname{End}(V^\pi), \tag{5.19}$$

where $\widehat{G}$ consists of (equivalence classes of) the irreducible unitary representations of G. This is given by

$$\mathcal{F}(f)(\pi) = \pi(f) = \int_G f(x)\pi(x)\,dx, \tag{5.20}$$

with Plancherel formula

$$(f, g)_{L^2} = \sum_{\pi \in \widehat{G}} d_\pi \operatorname{Tr}(\pi(f)\pi(g)^*), \tag{5.21}$$

and inversion formula

$$f(x) = \sum_{\pi \in \widehat{G}} d_\pi \operatorname{Tr}(\pi(f)\pi(x)^*). \tag{5.22}$$

Since $\pi(f * g) = \pi(f)\pi(g)$, we get

$$(f, f * C)_{L^2} = \sum_{\pi \in \widehat{G}} d_\pi \operatorname{Tr}(\pi(f)\pi(C)^*\pi(f)^*), \tag{5.23}$$

and the condition (5.18) on $C \in C(G)$ becomes

$$\operatorname{Tr}(A\pi(C)^*A^*) \geq 0, \quad \forall\, \pi \in \widehat{G},\ A \in \operatorname{End}(V^\pi). \tag{5.24}$$

Note that if A is an orthogonal projection, $Av = (v, w)w,\ \|w\| = 1$, then $A\pi(C)^*A^*v = (v, w)A\pi(C)^*w = (v, w)(\pi(C)^*w, w)w$, so $\operatorname{Tr}(A\pi(X)^*A^*) = (\pi(C)^*w, w)$. Thus (5.24) implies

$$\pi(C) \geq 0, \quad \forall\, \pi \in \widehat{G}. \tag{5.25}$$

The reverse implication is also readily established.

For compact G, a stationary random field $Z : G \to L^2(\Omega, \mu)$ yields random variables

$$Z_{ij}^{\pi} = \int_G Z(x)\pi_{ij}(x)\, dx \in L^2(\Omega, \mu), \tag{5.26}$$

where $\pi_{ij}(x)$ are the matrix entries of $\pi(x)$, with respect to some given orthonormal basis of V^{π}. These entries fit together to produce

$$Z^{\pi} = \int_G Z(x)\pi(x)\, dx \in L^2(\Omega, \mu, \mathrm{End}(V^{\pi})). \tag{5.27}$$

Let us assume

$$\langle Z(x)\rangle \equiv M = 0. \tag{5.28}$$

Then

$$\begin{aligned} \langle Z_{ij}^{\pi}\overline{Z}_{k\ell}^{\pi}\rangle &= \int_G\int_G C(y^{-1}x)\pi_{ij}(x)\overline{\pi}_{k\ell}(y)\, dx\, dy \\ &= \int_G\int_G C(z)\pi_{ij}(yz)\overline{\pi}_{k\ell}(y)\, dz\, dy \\ &= \sum_m \int_G\int_G C(z)\pi_{im}(y)\pi_{mj}(z)\overline{\pi}_{k\ell}(y)\, dz\, dy \\ &= \frac{1}{d_\pi}\sum_m \delta_{ik}\delta_{m\ell}\int_G C(z)\pi_{mj}(z)\, dz \\ &= \frac{1}{d_\pi}\delta_{ik}\int_G C(z)\pi_{\ell j}(z)\, dz \\ &= \frac{1}{d_\pi}\delta_{ik}\big(\pi(C)\big)_{\ell j}. \end{aligned} \tag{5.29}$$

If π and λ are distinct elements of $\widehat{G}$,

$$\langle Z_{ij}^{\pi}\overline{Z}_{k\ell}^{\lambda}\rangle \equiv 0. \tag{5.30}$$

The Peter-Weyl theorem gives

$$Z(x) = \sum_{\pi\in\widehat{G}} Z^{\pi}(x), \tag{5.31}$$

with

$$Z^\pi(x) = d_\pi \sum_{i,j} Z^\pi_{ij}\overline{\pi}_{ij}(x) = d_\pi \operatorname{Tr}(Z^\pi \pi(x)^*). \tag{5.32}$$

We have

$$\begin{aligned}
\langle Z^\pi(x)\overline{Z}^\pi(y)\rangle &= d_\pi^2 \sum_{i,j}\sum_{k,\ell} \langle Z^\pi_{ij}\overline{Z}^\pi_{k\ell}\rangle \overline{\pi}_{ij}(x)\pi_{k\ell}(y) \\
&= d_\pi \sum_{i,j}\sum_{k,\ell} \delta_{ik}\big(\pi(C)\big)_{\ell j}\overline{\pi}_{ij}(x)\pi_{k\ell}(y) \\
&= d_\pi \sum_{j,k,\ell} \big(\pi(C)\big)_{\ell j}\overline{\pi}_{kj}(x)\pi_{k\ell}(y) \\
&= d_\pi \sum_{j,k} \big(\pi(y)\pi(C)\big)_{kj}\overline{\pi}_{kj}(x) \\
&= d_\pi \operatorname{Tr}\big(\pi(x)^*\pi(y)\pi(C)\big) \\
&= d_\pi \operatorname{Tr}\big(\pi(x^{-1}y)\pi(C)\big),
\end{aligned} \tag{5.33}$$

and $\langle Z^\pi(x)\overline{Z}^\lambda(y)\rangle \equiv 0$ if π and λ are distinct elements of $\widehat{G}$.

Note that, since $Z(x)$ is real valued, $\overline{Z}^\pi = Z^{\overline{\pi}}$, so

$$\begin{aligned}
\langle Z^\pi_{ij} Z^\pi_{k\ell}\rangle = \langle Z^\pi_{ij}\overline{Z}^\pi_{k\ell}\rangle \quad &\text{if} \quad \pi = \overline{\pi}, \\
0 \qquad\qquad &\text{if} \quad \pi \neq \overline{\pi}.
\end{aligned} \tag{5.34}$$

Hence

$$\begin{aligned}
\langle Z^\pi(x) Z^\pi(y)\rangle = \langle Z^\pi(x)\overline{Z}^\pi(y)\rangle \quad &\text{if} \quad \pi = \overline{\pi}, \\
0 \qquad\qquad &\text{if} \quad \pi \neq \overline{\pi}.
\end{aligned} \tag{5.35}$$

Furthermore, $\langle Z^\pi(x) Z^\lambda(y)\rangle \equiv 0$ if π and $\overline{\lambda}$ are distinct elements of $\widehat{G}$.

Remark. The condition (5.8) is the condition that the field Z is "2-weakly stationary." This condition is weaker than stationarity, but it suffices for all the results in (5.9)–(5.35).

We now give a result that follows from stationarity but not from 2-weak stationarity. First, some notation. Let Y^σ and $\widetilde{Y}^\sigma$ denote two families of elements of $L^1(\Omega, \mu)$, indexed by $\sigma \in \Sigma$. We write

$$Y^\sigma \leftrightarrow_\sigma \widetilde{Y}^\sigma \tag{5.36}$$

provided that, for arbitrary $\sigma_1, \dots, \sigma_N \in \Sigma,\ N \in \mathbb{N}$, the random variables

$$\{Y^{\sigma_1}, \dots, Y^{\sigma_N}\} \ \text{ and } \ \{\widetilde{Y}^{\sigma_1}, \dots, \widetilde{Y}^{\sigma_N}\} \tag{5.37}$$

have the same joint distribution. Note that a field $Z : G \to L^2(\Omega, \mu)$ is stationary if and only if

$$Z(gx) \leftrightarrow_g Z(x), \quad \forall\, x \in X. \tag{5.38}$$

Now if Z^π is defined by (5.27), then

$$\begin{aligned} \pi(g) Z^\pi &= \int\limits_G Z(x) \pi(gx)\, dx \\ &= \int\limits_G Z(g^{-1}y) \pi(y)\, dy, \end{aligned} \tag{5.39}$$

so (5.38) gives

$$\pi(g) Z^\pi \leftrightarrow_\pi Z^\pi, \quad \forall\, g \in G, \tag{5.40}$$

provided Z is stationary. This result does not follow from 2-weak stationarity.

6. Stationary random fields on homogeneous spaces

Let X be a Riemannian manifold with a transitive group G of isometries. If $K \subset G$ is the subgroup fixing a point $p_0 \in X$, then K is compact and $X \approx G/K$. We have

$$\gamma : G \longrightarrow X, \quad \gamma(g) = g \cdot p_0. \tag{6.1}$$

Given a continuous random field

$$Y : X \longrightarrow L^2(\Omega, \mu), \tag{6.2}$$

we have

$$Z = Y \circ \gamma : G \longrightarrow L^2(X, \mu). \tag{6.3}$$

We say Y is stationary if Z is stationary. Note that a field $Z : G \to L^2(\Omega, \mu)$ has the form (6.3) if and only if

$$Z(xk) = Z(x), \quad \forall\, x \in G,\ k \in K. \tag{6.4}$$

In such a case, $R(x, y) = \langle Z(x) Z(y) \rangle - M^2$ satisfies

$$R(xk_1, yk_2) = R(x, y), \quad \forall\, k_j \in K,\ x, y \in G, \tag{6.5}$$

so, given stationarity, with $R(x,y) = C(x^{-1}y)$, we have

$$C((xk_1)^{-1}yk_2) = C(x^{-1}y), \quad \forall\, x, y, \in G,\ k_j \in K, \tag{6.6}$$

or equivalently

$$C(k_2 x k_1) = C(x), \quad \forall\, x \in G,\ k_j \in K. \tag{6.7}$$

In particular, we have a function

$$\mathcal{C} : X \longrightarrow \mathbb{R}, \quad C = \mathcal{C} \circ \gamma, \tag{6.8}$$

and

$$\mathcal{C}(kp) = \mathcal{C}(p), \quad \forall\, k \in K,\ p \in X. \tag{6.9}$$

Conversely, given a continuous $\mathcal{C} : X \to \mathbb{R}$ satisfying (6.9), $C = \mathcal{C} \circ \gamma$ satisfies (6.7).

A Riemannian manifold X might have more than one group of isometries acting transitively, so one might use the phrase "G-stationary" to be more precise (though we will not). For example, if $X = \mathbb{R}^n$, one has $G = \mathbb{R}^n$ acting by translations, and also the larger group $G = E(n)$ of rigid motions, a semidirect product of $\mathbb{R}^n$ and $SO(n)$. For an $\mathbb{R}^n$-stationary field on $X = \mathbb{R}^n$ to be $E(n)$-stationary ([MP] prefers the term "isotropic"), one needs $\mathcal{C} : \mathbb{R}^n \to \mathbb{R}$ to be radial. (And this condition would suffice for Gaussian fields.) For another example, if $X = S^{n-1}$ is the unit sphere in $\mathbb{R}^n$, one has $SO(n)$ acting transitively as a group of isometries, and if $n = 2k$, one has the subgroups $U(k)$ and $SU(k)$ also acting transitively.

For the rest of this section, we assume X is *compact*, hence G is compact. As shown in [Z], p. 80, the regular representation R of G on $L^2(X)$,

$$R(g)f(p) = f(g^{-1}p), \tag{6.10}$$

decomposes into a family of finite-dimensional representations,

$$L^2(X) = \bigoplus_{\pi \in \widehat{G}_0} L^2_\pi(X). \tag{6.11}$$

Here $\widehat{G}_0 \subset \widehat{G}$ is defined by

$$\pi \in \widehat{G}_0 \Leftrightarrow V_0^\pi = \{\varphi \in V^\pi : \pi(k)\varphi = \varphi,\ \forall\, k \in K\} \neq 0, \tag{6.12}$$

and we have isomorphisms

$$\begin{aligned} &\Psi_\pi : V_0^\pi \otimes V^\pi \longrightarrow L^2_\pi(X), \\ &\Psi_\pi(\varphi \otimes \psi)(g \cdot p_0) = (\pi(g)\varphi, \psi). \end{aligned} \tag{6.13}$$

Note that, given $h \in G,\ \varphi \in V_0^\pi,\ \psi \in V^\pi$,

$$
\begin{aligned}
R(h)\Psi_\pi(\varphi \otimes \psi)(g \cdot p_0) &= (\pi(h^{-1}g)\varphi, \psi) \\
&= (\pi(g)\varphi, \pi(h)\psi) \\
&= \Psi_\pi(\varphi \otimes \pi(h)\psi)(g \cdot p_0).
\end{aligned}
\tag{6.14}
$$

In other words, for each $\varphi \in V_0^\pi$, we have

$$
\Phi_\varphi : V^\pi \to L^2(X), \quad \Phi_\varphi(\psi)(g \cdot p_0) = (\pi(g)\varphi, \psi), \tag{6.15}
$$

and

$$
R(h)\Phi_\varphi(\psi) = \Phi_\varphi(\pi(h)\psi). \tag{6.16}
$$

The case $X = S^2,\ G = SO(3)$ was emphasized in [MP]. Then (6.11) becomes

$$
L^2(S^2) = \bigoplus_{j \ge 0} V_j, \tag{6.17}
$$

where V_j is an eigenspace of the Laplace-Beltrami operator on S^2, of dimension $2j+1$, and $SO(3)$ acts on V_j by the representation denoted D_j. Elements of V_j are called spherical harmonics. In this case, $V_{j,0}$ is one dimensional, spanned by the zonal harmonic in V_j. It is desired to understand the behavior of the spherical harmonic expansion of the continuous function $\mathcal{C} : S^2 \to \mathbb{R}$ arising from a stationary field $Y : S^2 \to L^2(\Omega, \mu)$, via (6.2)–(6.8).

More generally, for a compact homogeneous space $X = G/K$, we want to understand the behavior of $\pi(C)$, as π runs over $\widehat{G}_0$. Results of §5 apply, of course, particularly (5.25). Further structure arises from (6.7), which implies

$$
\pi(C) = \pi(k_1)\pi(C)\pi(k_2), \quad \forall\, k_j \in K,\ \pi \in \widehat{G}. \tag{6.18}
$$

Note that

$$
P_0 = \int_K \pi(k)\, dk \Longrightarrow P_0 : V^\pi \to V_0^\pi, \ \text{ orthogonal projection.} \tag{6.19}
$$

Integrating (6.18) yields

$$
\pi(C) = P_0 \pi(C) P_0. \tag{6.20}
$$

Conversely, (6.20) $\Rightarrow$ (6.18). Note that if (6.18) holds, then

$$
C(x) = \sum_{\pi \in \widehat{G}} d_\pi \ \mathrm{Tr}\big(\pi(C)\pi(x)^*\big) \tag{6.21}
$$

satisfies

$$C(k_2xk_1) = \sum_{\pi\in\widehat{G}} d_\pi \ \mathrm{Tr}\big(\pi(C)\pi(k_1)^*\pi(x)^*\pi(k_2)^*\big) = C(x), \quad \forall\, k_j \in K. \tag{6.22}$$

We also have

$$\begin{aligned} C(x^{-1}) &= \sum_{\pi\in\widehat{G}} d_\pi \ \mathrm{Tr}\big(\pi(C)\pi(x)\big) \\ &= \sum_{\pi} d_\pi \ \mathrm{Tr}\big(\pi(x)^*\pi(C)^*\big) \\ &= C(x), \end{aligned} \tag{6.23}$$

given that C is real valued and that $\pi(C)$, as a consequence of (5.25), is self-adjoint.

As we have mentioned, if $X = S^2,\ G = SO(3)$, then

$$\dim V_0^\pi = 1, \tag{6.24}$$

for all $\pi \in \widehat{G}_0$. This holds more generally for $X = S^{n-1},\ G = SO(n)$, but it does not hold for $X = S^3,\ G = SU(2)$. When (6.24) holds, we can write (6.20) as

$$\pi(C) = \tau_\pi(C)P_0, \tag{6.25}$$

with

$$\tau_\pi(C) = \mathrm{Tr}\,\pi(C) = \int_G C(x)\chi_\pi(x)\,dx, \tag{6.26}$$

where $\chi_\pi(x) = \mathrm{Tr}\,\pi(x)$ is the character of the representation π. The positivity condition (5.25) becomes

$$\tau_\pi(C) \geq 0, \quad \forall\, \pi \in \widehat{G}_0. \tag{6.26A}$$

Recall from (5.27) the construction of

$$Z^\pi = \int_G Z(x)\pi(x)\,dx \in L^2(\Omega, \mu, \mathrm{End}(V^\pi)). \tag{6.27}$$

If (6.4) holds, then

$$Z^\pi = Z^\pi\pi(k), \quad \forall\, k \in K, \tag{6.28}$$

and integration over $k \in K$ gives

$$Z^\pi = Z^\pi P_0, \tag{6.29}$$

with P_0 as in (6.19). Conversely, (6.29) $\Rightarrow$ (6.28), which in turn implies that

$$Z(x) = \sum_{\pi \in \widehat{G}} d_\pi \ \mathrm{Tr}\big(Z^\pi \pi(x)^*\big) \tag{6.30}$$

(note that the sum is actually over $\pi \in \widehat{G}_0$) satisfies

$$Z(xk) = \sum_\pi d_\pi \ \mathrm{Tr}\big(Z^\pi \pi(k)^* \pi(x)^*\big) = Z(x), \quad \forall\, k \in K. \tag{6.31}$$

If (6.24) holds and $M = 0$, we deduce from (5.29) and (6.25) that

$$\langle Z^\pi_{ij} \overline{Z}^\pi_{k\ell} \rangle = \frac{\tau_\pi(C)}{d_\pi} \delta_{ik} (P_0)_{\ell j} \quad (\text{if } \dim V_0^\pi = 1). \tag{6.32}$$

It would be typical to pick an orthonormal basis of V^π such that $(P_0)_{\ell j} = \delta_{\ell 0} \delta_{j0}$, so

$$\langle Z^\pi_{ij} \overline{Z}^\pi_{k\ell} \rangle = \frac{\tau_\pi(C)}{d_\pi} \delta_{\ell 0} \delta_{j0} \delta_{ik}. \tag{6.33}$$

We also have from (5.33) and (6.25) that

$$\begin{aligned} \langle Z^\pi(x) \overline{Z}^\pi(y) \rangle &= d_\pi \tau_\pi(C) \ \mathrm{Tr}(\pi(x^{-1}y) P_0) \\ &= d_\pi \tau_\pi(C) \ \mathrm{Tr}(P_0 \pi(x^{-1}y) P_0). \end{aligned} \tag{6.34}$$

Note that (5.34)–(5.35) apply to $\langle Z^\pi_{ij} Z^\pi_{k\ell} \rangle$ and $\langle Z^\pi(x) Z^\pi(y) \rangle$. We mention that, for $X = S^2$, $G = SO(3)$ (more generally, for $X = S^{n-1}$, $G = SO(n)$) all the representations $\pi \in \widehat{G}_0$ are real, and $Z^\pi(x) = \overline{Z}^\pi(x)$, for all x.

Remark. Parallel to the remark following (5.35), we mention that the results (6.5)–(6.34) hold for Z as in (6.3), whenever Z is 2-weakly stationary (we then say Y is 2-weakly stationary). This condition is weaker than the assumption that Y is stationary. See the remarks at the end of §7 for more on this.

7. The inverse problem: constructing $Z(x)$ from spectral data

As in the latter part of §6, X will be a compact Riemannian manifold, G a transitive group of isometries of X, $K \subset G$ the subgroup fixing a given point $p_0 \in X$. We are given data

$$C^\pi \in \mathrm{End}(V^\pi), \quad Z^\pi \in L^2(\Omega, \mu, \mathrm{End}(V^\pi)), \tag{7.1}$$

for $\pi \in \widehat{G}_0$, defined in (6.12). We want to specify the conditions on this data that guarantee the existence of a continuous, real valued, $Y : X \to L^2(\Omega, \mu)$ such that, with $Z = Y \circ \gamma$, as in (6.3), we have

$$\begin{aligned} Z^\pi &= \int_G Z(x)\pi(x)\,dx, \quad \pi(C) = C^\pi, \\ \langle Z(x)\rangle &= 0, \quad \langle Z(x)Z(y)\rangle = C(x^{-1}y). \end{aligned} \tag{7.2}$$

Necessary conditions follow from results of §§5–6. Here we want to show they are sufficient. We start out in the general setting described above, but later on we will make some simplifying assumptions, which are satisfied when $X = S^{n-1},\ G = SO(n),\ n \geq 3$.

We begin by seeing what condition on $\{C^\pi\}$ gives rise to a continuous, positive-definite function C on G satisfying

$$C(k_2 x k_1) = C(x), \quad \forall\, x \in G,\ k_j \in K. \tag{7.3}$$

As seen in §§5–6, a necessary condition is

$$C^\pi = P_0 C^\pi P_0 \geq 0, \quad \forall\, \pi \in \widehat{G}_0, \tag{7.4}$$

where $P_0 : V^\pi \to V_0^\pi$ is the orthogonal projection. (For notational simplicity, we do not record the dependence of P_0 on π.) We now show that (7.4) is sufficient for the existence of the desired function C (given appropriate decay of C^π as $\pi \to \infty$). In fact, taking a cue from (5.22), we set

$$C(x) = \sum_{\pi \in \widehat{G}_0} d_\pi \ \mathrm{Tr}(C^\pi \pi(x)^*). \tag{7.5}$$

Given sufficient decay (cf. (7.34) below), this converges to $C \in C(G)$, and

$$\pi(C) = C^\pi, \tag{7.6}$$

for all π. Furthermore, given $k_j \in K$,

$$\begin{aligned} C(k_2 x k_1) &= \sum_\pi d_\pi \ \mathrm{Tr}(C^\pi \pi(k_2)^* \pi(x)^* \pi(k_1)^*) \\ &= \sum_\pi d_\pi \ \mathrm{Tr}(\pi(k_1^{-1}) C^\pi \pi(k_2^{-1}) \pi(x)^*), \end{aligned} \tag{7.7}$$

and (7.4) implies

$$\pi(k_1^{-1}) C^\pi \pi(k_2^{-1}) = C^\pi, \quad \forall\, k_j \in K, \tag{7.8}$$

so (7.3) holds. By (7.6) and the argument around (5.25), the function C in (7.5) is positive definite. Let us also note that if $C(x)$ is given by (7.5), then

$$\begin{aligned} C(x^{-1}) &= \sum_\pi d_\pi \operatorname{Tr}(C^\pi \pi(x)) \\ &= \sum_\pi d_\pi \overline{\operatorname{Tr}(\pi(x)^* C^\pi)} \\ &= \overline{C(x)}, \end{aligned} \tag{7.9}$$

where we have used self-adjointness of C^π. We want $C(x)$ to be real valued, so we will impose the following restriction:

$$\text{Each representation } \pi \in \widehat{G}_0 \text{ is real,} \tag{7.10}$$

with respect to some orthonormal basis of V^π. As mentioned in §6, this holds when $X = S^{n-1},\ G = SO(n), n \geq 3$. We also complement (7.4) with the condition that

$$C^\pi \quad \text{is real,} \tag{7.11}$$

with respect to such a basis of V^π.

We move on to Z^π, with matrix entries $Z^\pi_{ij} \in L^2(\Omega, \mu)$. We set

$$Z(x) = \sum_\pi d_\pi \operatorname{Tr}(Z^\pi \pi(x)^*), \tag{7.12}$$

which yields a continuous function $Z : G \to L^2(\Omega, \mu)$, given appropriate decay of $\{Z^\pi\}$ as $\pi \to \infty$, and we have

$$Z^\pi = \int_G Z(x)\pi(x)\, dx. \tag{7.13}$$

Now

$$\langle Z(x) \rangle = \sum_\pi d_\pi \operatorname{Tr}(\langle Z^\pi \rangle \pi(x)^*) = 0, \tag{7.14}$$

provided

$$\langle Z^\pi_{ij} \rangle = 0, \quad \forall\, i, j, \pi. \tag{7.15}$$

As seen in §6, a necessary condition on Z^π is

$$Z^\pi = Z^\pi P_0. \tag{7.16}$$

This implies $Z^\pi = Z^\pi \pi(k)$ for all $k \in K$, hence

$$Z(xk) = \sum_\pi d_\pi \ \mathrm{Tr}(Z^\pi \pi(k)^* \pi(x)^*) = Z(x), \tag{7.17}$$

for all $x \in G,\ k \in K$.

At this point we bring in the following simplifying assumption.

$$\dim V_0^\pi = 1, \quad \forall\, \pi \in \widehat{G}_0. \tag{7.18}$$

As mentioned in in §6, this holds for $X = S^{n-1},\ G = SO(n),\ n \geq 3$. Given (7.18), (7.4) becomes

$$C^\pi = \tau_\pi P_0, \quad \tau_\pi \in [0, \infty). \tag{7.19}$$

It is common to take an orthonormal basis $\{v_j\}$ of V^π for which (7.10) holds and $V_0^\pi = \mathrm{Span}(v_0)$, so $(P_0)_{ij} = \delta_{i0}\delta_{j0}$. Then $Z_{ij}^\pi = 0$ unless $j = 0$, so

$$Z_{ij}^\pi = \zeta_i^\pi \delta_{j0}, \quad \zeta_i^\pi \in L^2(\Omega, \mu) \ \ \text{(real valued)}. \tag{7.20}$$

Finally, we need to make an appropriate hypothesis on ζ_i^π. The condition (7.15) gives

$$\langle \zeta_i^\pi \rangle = 0, \quad \forall\, i, \pi, \tag{7.21}$$

and the formula (6.33) (plus (5.30)) yields the necessary condition

$$\langle \zeta_i^\pi \zeta_k^\lambda \rangle = \frac{\tau_\pi}{d_\pi} \delta_{ik} \delta_{\pi\lambda}, \quad \pi, \lambda \in \widehat{G}_0. \tag{7.22}$$

It remains to check the covariance identity in (7.2). To break this down, we write (7.12) as

$$Z(x) = \sum_\pi Z^\pi(x), \tag{7.23}$$

with

$$\begin{aligned} Z^\pi(z) &= d_\pi \ \mathrm{Tr}(Z^\pi \pi(x)^*) \\ &= d_\pi \sum_{i,j} Z_{ij}^\pi \overline{\pi}_{ij}(x) \\ &= d_\pi \sum_i \zeta_i^\pi \overline{\pi}_{i0}(x), \end{aligned} \tag{7.24}$$

the last identity by (7.20). From here, we get the following. (We keep the bar, but recall that in this setting $\overline{Z}^\pi(y) = Z^\pi(y)$.)

$$\begin{aligned} \langle Z^\pi(x)\overline{Z}^\pi(y)\rangle &= d_\pi^2 \sum_{i,k} \langle \zeta_i^\pi \overline{\zeta}_k^\pi\rangle \overline{\pi}_{i0}(x)\pi_{k0}(y) \\ &= d_\pi \tau_\pi \sum_i \overline{\pi}_{i0}(x)\pi_{i0}(y) \\ &= d_\pi \tau_\pi \pi(x^{-1}y)_{00}. \end{aligned} \tag{7.25}$$

Hence (celebrating reality and dropping the bars),

$$\begin{aligned} \langle Z(x)Z(y)\rangle &= \sum_{\pi,\lambda} \langle Z^\pi(x)Z^\lambda(y)\rangle \\ &= \sum_\pi \langle Z^\pi(x)Z^\pi(y)\rangle \\ &= \sum_\pi d_\pi \tau_\pi \pi(x^{-1}y)_{00}, \end{aligned} \tag{7.26}$$

the second identity by (7.22). Meanwhile,

$$\begin{aligned} C(x) &= \sum_\pi d_\pi \operatorname{Tr}(C^\pi \pi(x)^*) \\ &= \sum_\pi d_\pi \tau_\pi \operatorname{Tr}(P_0 \pi(x)^*) \\ &= \sum_\pi d_\pi \tau_\pi \operatorname{Tr}(P_0 \pi(x) P_0) \\ &= \sum_\pi d_\pi \tau_\pi \pi(x)_{00}, \end{aligned} \tag{7.27}$$

so

$$\langle Z(x)Z(y)\rangle = C(x^{-1}y). \tag{7.28}$$

We formulate our result.

Proposition 7.1. *Assume on $\widehat{G}_0$ that (7.10) and (7.18) hold. Take $\tau_\pi \in [0,\infty)$, decreasing sufficiently rapidly as $\pi \to \infty$, and define C^π by (7.19). Let $\zeta_j^\pi \in L^2(\Omega,\mu)$ satisfy (7.21)–(7.22), and define $Z(x)$ by (7.23)–(7.24) and $C(x)$ by (7.5). Then*

$$C : G \to \mathbb{R}, \quad Z : G \to L^2(\Omega,\mu) \tag{7.29}$$

are continuous, $Z(xk) = Z(k)$ for all $x \in G$, $k \in K$, and the identities in (7.2) hold.

Let us record just what decay is required on $\{\tau_\pi : \pi \in \widehat{G}_0\}$. We have

$$Z(x) = \sum_{\pi \in \widehat{G}_0} d_\pi \sum_i \zeta_i^\pi \overline{\pi}_{i0}(x), \tag{7.30}$$

and $\{\zeta_i^\pi\}$ consists of mutually orthogonal elements of $L^2(\Omega, \mu)$, with square norm τ_π / d_π. Hence

$$\begin{aligned} \|Z(x)\|^2_{L^2(\Omega)} &= \sum_\pi d_\pi^2 \sum_i \frac{\tau_\pi}{d_\pi} |\pi_{i0}(x)|^2 \\ &= \sum_\pi d_\pi \tau_\pi, \end{aligned} \tag{7.31}$$

since, by unitarity,

$$\sum_i |\pi_{i0}(x)|^2 \equiv 1. \tag{7.32}$$

Hence, as long as

$$\sum_{\pi \in \widehat{G}_0} d_\pi \tau_\pi < \infty, \tag{7.33}$$

the infinite series (7.30) converges uniformly on G to a continuous function with values in $L^2(\Omega, \mu)$.

Remark. The random field $Z : G \to L^2(\Omega, \mu)$ constructed in Proposition 7.1 satisfies

$$\langle Z(x) \rangle = 0, \quad \langle Z(gx) Z(gy) \rangle = \langle Z(x) Z(y) \rangle, \quad \forall\, x, y, g \in G. \tag{7.34}$$

As mentioned in §5, one says such a random field is "2-weakly stationary." If $\{\zeta_j^\pi\}$ are mutually independent Gaussian random variables satisfying (7.22), then Z is a Gaussian field, and arguments mentioned in §3 show that (7.34) implies stationarity. In the non-Gaussian case, 2-weak stationarity does not imply stationarity.

Here is a result that follows from stationarity but not from 2-weak stationarity. Namely, with respect to the orthonormal basis of V^π mentioned below (7.19), the elements $\zeta_i^\pi \in L^2(\Omega, \mu)$ introduced in (7.20) are the components of

$$\zeta^\pi \in L^2(\Omega, \mu, V^\pi). \tag{7.35}$$

Then (5.40) implies

$$\pi(g)\zeta^\pi \leftrightarrow_\pi \zeta^\pi, \quad \forall\, g \in G, \tag{7.36}$$

provided G is stationary. (See (5.36)–(5.37) for the notation used in (7.36).) It follows from (7.36) that

$$S_{\pi_1\cdots\pi_N} = \langle \zeta^{\pi_1} \otimes \cdots \otimes \zeta^{\pi_N} \rangle \in V^{\pi_1} \otimes \cdots \otimes V^{\pi_N} \tag{7.37}$$

satisfies

$$\pi_1(g) \otimes \cdots \otimes \pi_N(g) S_{\pi_1\cdots\pi_N} = S_{\pi_1\cdots\pi_N}, \quad \forall\, g \in G. \tag{7.38}$$

Let us note, parenthetically, that (7.20) is equivalent to

$$Z^\pi = \zeta^\pi \otimes v_0^\pi, \quad V_0^\pi = \text{Span}\,(v_0^\pi),\ \|v_0^\pi\| = 1, \tag{7.39}$$

that is,

$$Z^\pi v = (v, v_0^\pi)\zeta^\pi, \quad v \in V^\pi. \tag{7.40}$$

In Chapter 6 of [MP] the following result is established, in the case $X = S^2,\ G = SO(3)$. Assume $Y : S^2 \to L^2(X,\mu)$ is stationary. Take $\pi \in \widehat{G}$ and assume the elements $\zeta_i^\pi \in L^2(\Omega,\mu)$ (known to be mutually orthogonal, as i varies, by (7.22)) are actually independent. Then these random variables must be Gaussian. The proof makes use of (7.36). This analysis is extended to general compact homogeneous spaces $X = G/K$ in [BMV].

Now we can find non-Gaussian ζ_i^π, that are mutually independent and satisfy the hypotheses of Proposition 7.1. Then this proposition yields a continuous field $Z : G \to L^2(\Omega,\mu)$ that is 2-weakly stationary, but (by the result of [MP] stated above) not stationary.

Non-Gaussian stationary fields $Z : G \to L^2(\Omega,\mu)$ can be obtained from a Gaussian stationary field Z_G by taking

$$Z(x) = F(Z_G(x)), \tag{7.41}$$

where $F : \mathbb{R} \to \mathbb{R}$ is continuous and satisfies moderate bounds. Such a class of stationary fields, called Gaussian-subordinated stationary fields, are studied in [MP] (with $X = S^2,\ G = SO(3)$).

8. V-valued random fields

To start, let X be a Riemannian manifold with a transitive group G of isometries. Let V be a finite dimensional inner product space (over $\mathbb{R}$) and π an orthogonal representation of G on V. We take a continuous function

$$Z : X \longrightarrow L^2(\Omega,\mu), \tag{8.1}$$

where (Ω, μ) is a probability space. We induce a measure ν on $\mathcal{O} = V^X$ as follows. We have a map

$$F : \Omega \longrightarrow \mathcal{O}, \quad F(\xi)(x) = Z(x)(\xi), \quad \xi \in \Omega,\ x \in X, \tag{8.2}$$

giving rise to

$$\nu(S) = \mu(F^{-1}(S)). \tag{8.3}$$

Parallel to (1.10)–(1.11) and to (5.6), we have a G-action on $\mathcal{O}$:

$$\tau_g : \mathcal{O} \longrightarrow \mathcal{O}, \quad (\tau_g \eta)(x) = \pi(g)^{-1}\eta(gx), \quad x \in X,\ g \in G,\ \eta \in \mathcal{O}. \tag{8.4}$$

With this convention, $\tau_{gh} = \tau_h \tau_g$. We say that Z is stationary (G-stationary, for clarity, when needed) provided the action $\{\tau_g\}$ preserves ν. We say Z is ergodic if in addition this action is ergodic on $(\mathcal{O}, \nu)$.

We now specialize to $X = \mathbb{R}^n$, and consider two cases of G:

$$\mathbb{R}^n, \quad E(n) = SO(n) \times_\varphi \mathbb{R}^n. \tag{8.5}$$

The group operation on $E(n)$ is given by

$$(g, x) \cdot (h, y) = (x + \varphi(g)y, gh), \quad x, y \in \mathbb{R}^n,\ g, h \in SO(n), \tag{8.6}$$

where φ is the standard action of $SO(n)$ on $\mathbb{R}^n$, i.e, $\varphi(g)y = gy$. In case $G = \mathbb{R}^n$, we take the trivial representation on V. In case $G = E(n)$, we consider representations of the form

$$\lambda(g, x)v = \pi(g)v, \tag{8.7}$$

where π is a unitary representation of $SO(n)$ on V. The resulting actions on $\mathcal{O} = V^{\mathbb{R}^n}$ are

$$\tau_y \eta(x) = \eta(x + y), \quad \tau_{(g,y)}\eta(x) = \pi(g)^{-1}\eta(\varphi(g)x + y). \tag{8.8}$$

Given a continuous $Z : \mathbb{R}^n \to L^2(\Omega, \mu, V)$, in case $G = \mathbb{R}^n$ and Z is G-stationary, we say Z is a homogeneous random field. In case $G = E(n)$ and Z is G-stationary, we say Z is an isotropic random field. A case of central importance is

$$V = \mathbb{R}^n, \quad \pi = \varphi, \tag{8.9}$$

the standard action of $SO(n)$ on $\mathbb{R}^n$. Then we say Z is a random vector field. For $G = \mathbb{R}^n$ or $E(n)$, respectively, we say a G-stationary Z is a homogeneous random vector field or, respectively, an isotropic random vector field.

Let us return to the general setting (8.1) and note that we have expectations and correlations,

$$\langle Z(x)\rangle \in V, \quad \langle Z(x)\otimes Z(y)\rangle = \mathcal{R}(x,y)\in V\otimes V, \quad x,y\in X. \tag{8.10}$$

G-stationarity implies, in the language of (5.36)–(5.37),

$$Z(gx) \leftrightarrow_g \pi(g)Z(x), \quad \forall\, x\in X. \tag{8.11}$$

Hence G-stationarity implies

$$\begin{gathered}\langle Z(gx)\rangle = \pi(g)\langle Z(x)\rangle,\\ \mathcal{R}(gx,gy) = (\pi(g)\otimes\pi(g))\mathcal{R}(x,y),\end{gathered} \tag{8.12}$$

for all $x,y\in X,\ g\in G$.

The inner product on V gives rise to an isomorphism,

$$j : V\otimes V \xrightarrow{\approx} \operatorname{End}(V), \quad j(v\otimes w)u = (u,w)v. \tag{8.13}$$

We can also write $j(v\otimes w) = vw^t$, with $w^t(u) = (u,w)$. We then have

$$R(x,y) = j\mathcal{R}(x,y)\in \operatorname{End}(V), \tag{8.14}$$

for $x,y\in X$. A useful alternative notation is

$$R(x,y) = \langle Z(x)Z(y)^t\rangle. \tag{8.14A}$$

Note that, for $A\in\operatorname{End}(V)$,

$$j(\pi(g)\otimes\pi(h))j^{-1}A = \pi(g)A\pi(h)^{-1}. \tag{8.15}$$

Hence (8.12) implies

$$R(gx,gy) = \pi(g)R(x,y)\pi(g)^{-1}, \quad \forall\, x,y\in X,\ g\in G. \tag{8.16}$$

Specializing to $X=\mathbb{R}^n,\ G=\mathbb{R}^n$, we have, for G-stationary Z,

$$\begin{gathered}\langle Z(x)\rangle = \langle Z(y)\rangle, \quad \forall\, x,y\in\mathbb{R}^n,\\ R(x,y) = C(x-y), \quad C:\mathbb{R}^n\to\operatorname{End}(V).\end{gathered} \tag{8.17}$$

If $G=E(n)$, we also have (8.17), and in addition

$$\begin{gathered}\langle Z(gx)\rangle = \pi(g)\langle Z(x)\rangle, \quad \text{hence } \ \pi(g)\langle Z(x)\rangle = \langle Z(x)\rangle,\\ C(gx) = \pi(g)C(x)\pi(g)^{-1}, \quad \forall\, x\in\mathbb{R}^n,\ g\in SO(n).\end{gathered} \tag{8.18}$$

NOTE. Our definition of $C(x-y)$ as $\langle Z(x)Z(y)\rangle$ differs slightly from that in (4.2) and (5.10), but the definitions coincide when $\langle Z(x)\rangle \equiv 0$, which is the typical situation.

Another symmetry property is the following. By (8.14A), $R(y,x) = R(x,y)^t$ (the adjoint in $\mathrm{End}(V)$), hence

$$C(-x) = C(x)^t, \quad \forall\, x \in \mathbb{R}^n. \tag{8.19}$$

We have a positivity property parallel to (3.10). Let $k \geq 1$, $x_1, \dots, x_k \in \mathbb{R}^n$, and $a_1, \dots, a_k \in \mathbb{C}$. Then

$$\begin{gathered} \sum_{i,j} R(x_i, x_j) a_i \overline{a}_j = \langle WW^* \rangle \geq 0 \ \text{ in } \ \mathrm{End}(V), \\ W = \sum_i a_i Z(x_i) \in L^2(\Omega, \mu, V). \end{gathered} \tag{8.20}$$

Hence, in the setting of (8.17),

$$\sum_{i,j} C(x_i - x_j) a_i \overline{a}_j \geq 0 \ \text{ in } \ \mathrm{End}(V). \tag{8.21}$$

Given that $C : \mathbb{R}^n \to \mathrm{End}(V)$ is continuous, (8.21) is equivalent to

$$\iint C(x-y) f(x) \overline{f(y)}\, dx\, dy \geq 0, \quad \forall\, f \in \mathcal{S}(\mathbb{R}^n), \tag{8.22}$$

and also, via Bochner-Herglotz, to

$$\widetilde{C} \text{ is a (finite) positive } \mathrm{End}(V_{\mathbb{C}})\text{-valued measure on } \mathbb{R}^n, \tag{8.23}$$

given (8.19), which implies

$$\widetilde{C}^* = \widetilde{C} \ \text{ in } \ \mathcal{S}'(\mathbb{R}^n, \mathrm{End}(V_{\mathbb{C}})). \tag{8.24}$$

Also (8.19) implies

$$\widetilde{C}(-\xi) = \widetilde{C}(\xi)^t. \tag{8.25}$$

(Given $A \in \mathrm{End}(V_{\mathbb{C}})$, $A^* = \overline{A}^t$.) Note that the Fourier transform of $C_g(x) = C(gx)$ is

$$\widetilde{C}_g(\xi) = (2\pi)^{-n/2} \int C(gx) e^{-ix\cdot\xi}\, dx = \widetilde{C}(g\xi), \tag{8.26}$$

so (8.18) implies

$$\widetilde{C}(g\xi) = \pi(g)\widetilde{C}(\xi)\pi(g)^{-1}. \tag{8.27}$$

It follows from (8.18) that if $Z : \mathbb{R}^n \to L^2(\Omega, \mu, V)$ is isotropic, then the hypothesis

$$\pi \text{ does not contain a trivial representation of } SO(n) \tag{8.28}$$

implies

$$\langle Z(x)\rangle \equiv 0, \tag{8.29}$$

and the hypothesis

$$\pi \text{ acts irreducibly on } V_{\mathbb{C}} \tag{8.30}$$

implies

$$C(0) = \alpha I, \quad \alpha \in \mathbb{R}^+. \tag{8.31}$$

Also, if $\widetilde{C}$ is continuous in a neighborhood of $0 \in \mathbb{R}^n$, (8.30) implies

$$\widetilde{C}(0) = \beta I, \quad \beta \in \mathbb{R}^+ \tag{8.32}$$

(inclusion in $\mathbb{R}^+$ by (8.23)).

If π is the standard representation of $SO(n)$ on $\mathbb{R}^n$, then (8.28) holds whenever $n \geq 2$, and (8.30) holds whenever $n \geq 3$. The hypothesis (8.30) fails for $n = 2$, but nevertheless (8.31) continues to hold. In fact, if $n = 2$, (8.18) implies $C(0)$ must be a scalar multiple of a rotation on $\mathbb{R}^2$. Since also $C(0) = \langle Z(0)Z(0)^t\rangle \geq 0$ in $\mathrm{End}(\mathbb{R}^2)$, (8.31) follows. A similar argument applies to (8.32).

We continue to take $X = \mathbb{R}^n$, $G = E(n)$, $V = \mathbb{R}^n$, and π the standard representation of $SO(n)$ on $\mathbb{R}^n$. The result (8.18) on C implies that it is uniquely specified by $C(re_n)$, $r \in [0, \infty)$, where $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{R}^n$. If $SO(n-1)$ acts on $\mathbb{R}^n$, fixing e_n and taking the standard $SO(n-1)$ action on $\mathrm{Span}(e_1, \dots, e_{n-1}) = \mathbb{R}^{n-1}$, then C is well defined on $\mathbb{R}^n \setminus 0$ if and only if

$$\pi(g)C(re_n)\pi(g)^{-1} = C(re_n), \quad \forall\, g \in SO(n-1), \tag{8.33}$$

the case of $C(0)$ having been discussed above. Now $\mathbb{C}^n$ splits into two factors, $\mathbb{C}e_n$ and $\mathbb{C}$-$\mathrm{Span}(e_1, \dots, e_{n-1})$, on each of which $SO(n-1)$ acts irreducibly. Hence (8.33) is equivalent to

$$C(re_n) = A(r)P_{e_n} + B(r)(I - P_{e_n}), \tag{8.34}$$

where, for $x \in \mathbb{R}^n$,

$$P_x = \text{ orthogonal projection of } \mathbb{R}^n \text{ onto } \operatorname{Span}(x), \tag{8.34A}$$

and A and B are scalar. Now,

$$g \in SO(n) \Longrightarrow \pi(g)P_{e_n}\pi(g)^{-1} = P_{ge_n}, \tag{8.35}$$

so we get

$$C(x) = A(|x|)P_x + B(|x|)(I - P_x). \tag{8.36}$$

From (8.31), we have $A(0) = B(0) = \alpha$. In view of (8.27), a similar analysis holds for $\widetilde{C}$. Assuming $\widetilde{C}$ is continuous on $\mathbb{R}^n \setminus 0$, we have

$$\widetilde{C}(\xi) = A^{\#}(|\xi|)P_\xi + B^{\#}(|\xi|)(I - P_\xi), \tag{8.37}$$

with $A^{\#}$ and $B^{\#}$ scalar. If in addition $\widetilde{C}$ is continuous in a neighborhood of 0, we have $A^{\#}(0) = B^{\#}(0) = \beta$. To celebrate the positivity result (8.23), we also write

$$\widetilde{C} = A^{\#}P_\xi + B^{\#}(I - P_\xi), \tag{8.38}$$

where

$$A^{\#} \text{ and } B^{\#} \text{ are finite, positive (scalar) radial measures on } \mathbb{R}^n. \tag{8.39}$$

Since P_ξ is not continuous at $\xi = 0$, we elaborate on (8.38). We have

$$\widetilde{C} = A^b P_\xi + B^b(I - P_\xi) + \gamma I\delta, \tag{8.40}$$

where A^b and B^b are finite, positive, scalar, rotationally invariant measures on $\mathbb{R}^n$ with no atom at 0, and $\gamma \geq 0$.

Note that (8.38)–(8.40) imply a reality condition, sharpening (8.24) to

$$\widetilde{C}^* = \widetilde{C}^t = \widetilde{C} \text{ in } \mathcal{S}'(\mathbb{R}^n, \operatorname{End}(\mathbb{R}^n)), \tag{8.41}$$

in the isotropic case, and converting (8.25) to

$$\widetilde{C}(-\xi) = \widetilde{C}(\xi). \tag{8.42}$$

9. Random divergence-free vector fields

We take the setting of §8, with $X = \mathbb{R}^n$, $G = \mathbb{R}^n$ or $E(n)$, and $V = \mathbb{R}^n$. As usual, (Ω, μ) is a probability space. A continuous function

$$Z : \mathbb{R}^n \longrightarrow L^2(\Omega, \mu) \tag{9.1}$$

is divergence-free provided $\operatorname{div} Z = 0$, i.e., with $Z = (Z_1, \dots, Z_n)^t$,

$$\sum_j \partial_j Z_j = 0, \tag{9.2}$$

considered as an element of $\mathcal{S}'(\mathbb{R}^n, L^2(\Omega, \mu))$. Equivalently,

$$\int_{\mathbb{R}^n} Z(x) \cdot \nabla f(x) = 0, \quad \forall\, f \in C_0^\infty(\mathbb{R}^n). \tag{9.3}$$

Recall that if Z is homogeneous (i.e., $\mathbb{R}^n$-stationary) and $\langle Z(x) \rangle \equiv 0$, we have

$$C(x) = \langle Z(x) Z(0)^t \rangle \in \operatorname{End}(\mathbb{R}^n), \quad \text{i.e.,} \;\; C_{ij}(x) = \langle Z_i(x) Z_j(0) \rangle. \tag{9.4}$$

Then (9.2) implies

$$\sum_i \partial_i C_{ij} = 0, \quad \text{hence} \;\; \sum_j \partial_j C_{ij} = 0, \tag{9.5}$$

the latter identity following since $C_{ji}(x) = C_{ij}(-x)$.

Applying the Fourier transform (cf. Appendix C) to (9.2) gives

$$\sum_j \xi_j \widehat{Z}_j = 0 \;\text{ in }\; \mathcal{S}'(\mathbb{R}^n, L^2(\Omega, \mu)), \tag{9.6}$$

and, in case Z is homogeneous, applying the Fourier transform to (9.5) gives

$$\sum_i \xi_i \widetilde{C}_{ij}(\xi) = 0, \quad \sum_j \xi_j \widetilde{C}_{ij}(\xi) = 0. \tag{9.7}$$

If Z is isotropic, then (8.40), i.e.,

$$\widetilde{C} = A^b P_\xi + B^b (I - P_\xi) + \gamma I \delta, \tag{9.8}$$

plus (9.7) gives

$$\widetilde{C} = B^b (I - P_\xi) + \gamma I \delta, \tag{9.9}$$

where B^b is a finite, positive (scalar), rotationally invariant measure on $\mathbb{R}^n$, with no atom at 0 (hence no atoms at all), and $\gamma \geq 0$. As shown in Appendix C, mild decay conditions on $C(x)$ as $|x| \to \infty$ imply no atoms for $\widetilde{C}$, hence $\gamma = 0$.

We next discuss the existence of nontrivial homogeneous (or isotropic) divergence-free vector fields. Take a continuous $C : \mathbb{R}^n \to \mathrm{End}(\mathbb{R}^n)$ satisfying $C(x) = C(-x)^t$ and $\widetilde{C}$ positive, e.g., $\widetilde{C}$ as in (9.8). Parallel to Theorem 3.3 and Corollary 3.4, there is (no doubt) an existence result for a Gaussian random field $Y : \mathbb{R}^n \to L^2(\Omega, \mu, \mathbb{R}^n)$ such that

$$\langle Y(x)\rangle \equiv 0, \quad \langle Y(x)Y(y)^t\rangle = C_Y(x-y), \tag{9.10}$$

and such a Gaussian field will be homogeneous ($\mathbb{R}^n$-stationary). If $C_Y(x)$ satisfies (8.18), e.g., if $\widetilde{C}_Y$ is given by (9.8), then (no doubt) such a Gaussian field Y will be isotropic ($E(n)$-stationary). (Justifying this would involve establishing a generalization of Proposition 3.1.)

To get a random field satisfying (9.2), we might need to alter Y. Consider

$$Z(x) = f * Y(x) = \int f(x-y)Y(y)\,dy, \tag{9.11}$$

with

$$f \in L^1(\mathbb{R}^n, \mathrm{End}(\mathbb{R}^n)). \tag{9.12}$$

Note that

$$\widehat{Z} = \hat{f}\,\hat{Y}. \tag{9.12A}$$

It easily follows from (9.11) that Z is homogeneous if Y is; compare similar results in Appendix C. We next investigate when Z can be said to be isotropic, given that Y is isotropic, i.e., Y is homogeneous, and, in the terminology of (5.36)–(5.37), with g running over $SO(n)$ and π the standard action of $SO(n)$ on $\mathbb{R}^n$,

$$Y(gx) \leftrightarrow_g \pi(g)Y(x), \quad \forall\, x \in \mathbb{R}^n. \tag{9.13}$$

Note that

$$\begin{aligned} Z(gx) &= \int f(gx-y)Y(y)\,dy \\ &= \int f(g(x-y))Y(gy)\,dy \\ &\leftrightarrow_g \int f(g(x-y))\pi(g)Y(y)\,dy. \end{aligned} \tag{9.14}$$

Thus, to achieve

$$Z(gx) \leftrightarrow_g \pi(g)Z(x), \quad \forall\, x \in \mathbb{R}^n, \tag{9.19}$$

we need f to satisfy

$$f(gx) = \pi(g)f(x)\pi(g)^{-1}, \tag{9.16}$$

or equivalently

$$\hat{f}(g\xi) = \pi(g)\hat{f}(\xi)\pi(g)^{-1}. \tag{9.17}$$

Note that

$$P_{g\xi} = \pi(g)P_\xi \pi(g)^{-1}, \quad \forall\, g \in SO(n),\ \xi \in \mathbb{R}^n \setminus 0. \tag{9.18}$$

Hence, we take $f \in L^1(\mathbb{R}^n, \mathrm{End}(\mathbb{R}^n))$ such that

$$\hat{f}(\xi) = a(\xi)(I - P_\xi), \tag{9.19}$$

with $a : \mathbb{R}^n \to \mathbb{R}$ radial, and sufficiently regular, and vanishing sufficiently as $\xi \to 0$ and as $|\xi| \to \infty$ to ensure that f is integrable. With such a choice of f, Z, defined by (9.11), will be divergence free. It will be homogeneous if Y is, and it will be isotropic if Y is.

We remark that if Y and Z are related by (9.11), then $\langle Y(x)\rangle \equiv 0 \Rightarrow \langle Z(x)\rangle \equiv 0$ and, with $C_Y(x-y)$ as in (9.10) and

$$C_Z(x-y) = \langle Z(x)Z(y)^t\rangle, \tag{9.20}$$

a calculation gives

$$C_Z = f * C_Y * f^\#, \quad f^\#(x) = f(-x)^t, \tag{9.21}$$

hence

$$\begin{aligned} \widetilde{C}_Z &= (2\pi)^n \hat{f}(\xi)\widetilde{C}_Y \hat{f}(-\xi)^t \\ &= (2\pi)^n \hat{f}(\xi)\widetilde{C}_Y \hat{f}(\xi)^*. \end{aligned} \tag{9.22}$$

If $\hat{f}$ is given by (9.19), we obtain

$$\widetilde{C}_Z = (2\pi)^n a(\xi)(I - P_\xi)\widetilde{C}_Y(I - P_\xi)a(\xi). \tag{9.23}$$

Recall from (8.31)–(8.32) that if Z is an isotropic random vector field on $\mathbb{R}^n$ and $\langle Z(x)\rangle \equiv 0$ and $C = C_Z$ is given by (9.20), then $C(0)$ is a scalar multiple of the identity. If $\widetilde{C}$ is continuous on a neighborhood of 0, $\widetilde{C}(0)$ is also a scalar multiple of the identity. We note that if Z is also divergence free, then

$$\widetilde{C}(\xi)\xi = 0 \Longrightarrow \widetilde{C}(0) = 0, \tag{9.24}$$

given such continuity. Indeed, fixing $\omega \in S^{n-1} \subset \mathbb{R}^n$, we have $\widetilde{C}(r\omega)\omega = 0$, for all $r > 0$, and letting $r \to 0$ yields $\widetilde{C}(0)\omega = 0$, for all $\omega \in S^{n-1}$. (In fact, this argument applies more generally to homogeneous, divergence-free random vector fields.)

On the other hand, there exist isotropic, divergence-free random fields on $\mathbb{R}^n$ for which $\widetilde{C}$ is continuous on $\mathbb{R}^n \setminus 0$ and does not tend to 0 at the origin. Examples can be obtained as follows. Set

$$\widetilde{C}_Y(\xi) = |\xi|^{-a} e^{-|\xi|} I, \quad a \in (0, n), \tag{9.25}$$

which is positive and integrable. Then there exists a homogeneous Gaussian random vector field Y such that $\langle Y(x) \rangle \equiv 0$ and (9.10) holds. Then form Z as in (9.11), with

$$\hat{f}(\xi) = |\xi|^{a/2} e^{-|\xi|} (I - P_\xi). \tag{9.26}$$

Then $f \in C^\infty(\mathbb{R}^n)$ and $|f(x)| \le C(1 + |x|)^{-n-\alpha/2}$, so $f \in L^1(\mathbb{R}^n, \mathrm{End}(\mathbb{R}^n))$. We have an isotropic, divergence-free random vector field Z, and, by (9.23),

$$\widetilde{C}_Z(\xi) = (2\pi)^n e^{-3|\xi|} (I - P_\xi). \tag{9.27}$$

This is bounded, continuous on $\mathbb{R}^n \setminus 0$, and has no limit as $\xi \to 0$. If instead of (9.26) we took $\hat{f}(\xi) = |\xi|^{b/2} e^{-|\xi|}(I - P_\xi)$, with $b \in (0, a)$, we would get an isotropic, divergence-free random vector field Z for which $\widetilde{C}_Z$ blows up at the origin.

10. Generalized random fields

As before, we fix a probability space (Ω, μ). Let us explicitly assume that $L^2(\Omega, \mu)$ is separable. A generalized random field on $\mathbb{R}^n$ is an $L^2(\Omega, \mu)$-valued distribution, $Z \in \mathcal{D}'(\mathbb{R}^n, L^2(\Omega, \mu))$, i.e., a continuous linear map

$$Z : C_0^\infty(\mathbb{R}^n) \longrightarrow L^2(\Omega, \mu). \tag{10.1}$$

More generally, we can take $L^2(\Omega, \mu, V)$, as in §8, but for now we drop the V. Given $f \in C_0^\infty(\mathbb{R}^n)$, we have the convolution

$$\begin{aligned} &f * Z \in C^\infty(\mathbb{R}^n, L^2(\Omega, \mu)), \\ &f * Z(x) = Z(\check{f}_x), \quad \check{f}_x = f(x - y). \end{aligned} \tag{10.2}$$

Definition. A generalized random field Z is stationary if and only if $f * Z$ is stationary (as a continuous random field) for all $f \in C_0^\infty(\mathbb{R}^n)$.

If $Z \in \mathcal{D}'(\mathbb{R}^n, L^2(\Omega,\mu))$ is stationary, the continuous linear map

$$K_Z : C_0^\infty(B) \longrightarrow C^\infty(\mathbb{R}^n, L^2(\Omega,\mu)),\qquad K_Z f(x) = f * Z(x) = Z(\check{f}_x), \tag{10.3}$$

where $B = \{x \in \mathbb{R}^n : |x| \le 1\}$, has the property that

$$K_Z : C_0^\infty(B) \longrightarrow L^\infty(\mathbb{R}^n, L^2(\Omega,\mu)). \tag{10.4}$$

It follows that K_Z in (10.4) is a closed linear map from a Frechet space to a Banach space, hence continuous. Thus there exist $k \in \mathbb{N}$ and $C \in (0,\infty)$ (depending on Z) such that

$$\sup_x \|Z(\check{f}_x)\|_{L^2(\Omega,\mu)} \le C\|f\|_{C^k}, \quad \forall\, f \in C_0^\infty(B). \tag{10.5}$$

This estimate leads to an extension of the action of such Z, as follows. If $\mathcal{Q} = \{Q_\alpha\}$ denotes the tiling of $\mathbb{R}^n$ by n-dimensional cubes with vertices in $\mathbb{Z}^n$, we can define

$$f \in L^1C^k(\mathbb{R}^n) \Leftrightarrow f \in C^k(\mathbb{R}^n) \text{ and } \|f\|_{L^1C^k} = \sum_{Q_\alpha \in \mathcal{Q}} \|f\|_{C^k(Q_\alpha)} < \infty. \tag{10.6}$$

A partition of unity argument leads from (10.5) to

$$\|Z(f)\|_{L^2(\Omega,\mu)} \le C\|f\|_{L^1C^k}, \tag{10.7}$$

for all $f \in C_0^\infty(\mathbb{R}^n)$, and from there to a continuous extension

$$Z : L^1C^k(\mathbb{R}^n) \longrightarrow L^2(\Omega,\mu), \tag{10.8}$$

whenever $Z \in \mathcal{D}'(\mathbb{R}^n, L^2(\Omega,\mu))$ is stationary. In particular,

$$Z : \mathcal{S}(\mathbb{R}^n) \longrightarrow L^2(\Omega,\mu), \tag{10.9}$$

i.e.,

$$Z \in \mathcal{S}'(\mathbb{R}^n, L^2(\Omega,\mu)). \tag{10.9A}$$

Furthermore, K_z in (10.3) extends to a continuous linear map

$$K_Z : L^1C^k(\mathbb{R}^n) \longrightarrow L^\infty \cap C(\mathbb{R}^n, L^2(\Omega,\mu)). \tag{10.10}$$

The following "Tauberian theorem" provides a useful characterization of stationary generalized random fields.

Proposition 10.1. *Let $Z \in \mathcal{S}'(\mathbb{R}^n, L^2(\Omega,\mu))$, and assume there exists a single $f \in \mathcal{S}(\mathbb{R}^n)$ such that $\hat{f}(\xi)$ is nowhere vanishing and*

$$f * Z \quad \textit{is stationary,} \tag{10.11}$$

as a continuous random field. Then (10.11) holds for all $f \in \mathcal{S}(\mathbb{R}^n)$, so Z is stationary. Furthermore, given $k \in \mathbb{N}$ such that (10.7) holds, the result (10.11) holds for all $f \in L^1C^k(\mathbb{R}^n)$.

Proof. Given that $Z \in \mathcal{S}'(\mathbb{R}^n, L^2(\Omega,\mu))$, it suffices to show that (10.11) holds for a set of functions with dense linear span in $\mathcal{S}(\mathbb{R}^n)$. Thus it suffices to note that if $f \in \mathcal{S}(\mathbb{R}^n)$ and $\hat{f}(\xi)$ is nowhere vanishing, then $\{\check{f}_x : x \in \mathbb{R}^n\}$ has dense linear span in $\mathcal{S}(\mathbb{R}^n)$, where $\check{f}_x(y) = f(x-y)$. The well known proof goes as follows. If $\omega \in \mathcal{S}'(\mathbb{R}^n)$ annihilates this span, then $f * \omega = 0$. This implies $\hat{f}\hat{\omega} = 0$, which implies $\omega = 0$, given that $\hat{f}(\xi)$ never vanishes. The asserted density then follows from the Hahn-Banach theorem.

We consider the following class of generalized random fields. Assume

$$Z : \mathbb{R}^n \longrightarrow L^2(\Omega,\mu) \quad \text{is weakly continuous.} \tag{10.12}$$

Assume Z is stationary, as an element of $\mathcal{D}'(\mathbb{R}^n, L^2(\Omega,\mu))$. Take $f_1 \in C_0^\infty(\mathbb{R}^n)$, satisfying $f_1 \ge 0$ and $\int f\,dx = 1$, let $f_k(x) = k^n f(kx)$, and set

$$Z_k = f_k * Z \in C^\infty(\mathbb{R}^n, L^2(\Omega,\mu)). \tag{10.13}$$

which are stationary as continuous random fields. We have $\langle Z(x)\rangle = \langle Z(x)\,1\rangle$, continuous in x, and

$$\langle Z_k\rangle = f_k * \langle Z\rangle \longrightarrow \langle Z\rangle, \quad \text{locally uniformly on } \mathbb{R}^n. \tag{10.14}$$

Since each $\langle Z_k(x)\rangle = M_k$ is constant, so is $\langle Z(x)\rangle \equiv M = \lim M_k$. Subtracting M, we assume $\langle Z(x)\rangle \equiv 0$, and then $M_k \equiv 0$.

Our aim is to prove the following.

Proposition 10.2. *If Z satisfies (10.12) and is statinary, as a generalized random field, then*

$$Z : \mathbb{R}^n \longrightarrow L^2(\Omega,\mu) \quad \textit{is norm-continuous,} \tag{10.15}$$

and Z is stationary, as a continuous random field.

To start the proof, using the constructions above, we define Z_k as in (10.13) and reduce to the case $\langle Z(x)\rangle \equiv 0$, so $\langle Z_k(x)\rangle \equiv 0$. We have

$$\langle Z_k(x)\,\varphi\rangle \longrightarrow \langle Z(x)\,\varphi\rangle, \quad \text{locally uniformly in } x, \quad \forall\,\varphi \in L^2(\Omega,\mu). \tag{10.16}$$

Stationarity of Z_k implies

$$\|Z_k(x)\|_{L^2} \equiv E_k \text{ (independent of } x). \tag{10.17}$$

Hence

$$\|Z(x)\|_{L^2} \le \liminf_{k\to\infty} E_k, \quad \forall\, x \in \mathbb{R}^n. \tag{10.18}$$

The Dunford-Pettis theorem implies

$$Z : \mathbb{R}^n \longrightarrow L^2(\Omega, \mu) \ \text{ is strongly measurable.} \tag{10.19}$$

Hence, for each $x \in \mathbb{R}^n$,

$$Z_k(x) = \int f_k(x-y) Z(y)\, dy \tag{10.20}$$

exists as a Bochner integral, and

$$E_k \equiv \|Z_k(x)\|_{L^2} \le \int f_k(x-y) \|Z(y)\|_{L^2}\, dy. \tag{10.21}$$

Also, since $y \mapsto \|Z(y)\|_{L^2}$ is bounded (by (10.18)) and measurable,

$$\int f_k(x-y)\|Z(y)\|_{L^2}\, dy \longrightarrow \|Z(x)\|_{L^2}, \quad \text{for a.e. } x, \tag{10.22}$$

and hence

$$\limsup_{k\to\infty} E_k \le \|Z(x)\|_{L^2}, \quad \text{for a.e. } x. \tag{10.23}$$

We are in a posiiton to establish the following.

Lemma 10.3. *There exists $S \subset \mathbb{R}^n$ such that $m(\mathbb{R}^n \setminus S) = 0$ and*

$$\|Z(x)\|_{L^2} = E = \lim_{k\to\infty} E_k, \quad \forall\, x \in S, \tag{10.24}$$

$$Z_k(x) \longrightarrow Z(x) \text{ in } L^2(\Omega,\mu)\text{-norm}, \quad \forall\, x \in S, \tag{10.25}$$

$$\langle Z_k(x) Z_k(y)\rangle \longrightarrow \langle Z(x)Z(y)\rangle, \quad \forall\, x \in S,\ y \in \mathbb{R}^n. \tag{10.26}$$

$$\langle Z_k(x-y) Z_k(0)\rangle \longrightarrow \langle Z(x-y)Z(0)\rangle, \quad \forall\, x-y \in S. \tag{10.27}$$

Proof. We get (10.24) from (10.18) and (10.23). Then (10.27) follows from (10.11) and (10.24). Next, (10.26) follows from (10.16) and (10.25), and then (10.27) follows from (10.26).

To proceed, we know that

$$\langle Z_k(x)Z_k(y)\rangle = C_k(x-y), \tag{10.28}$$

and $C_k : \mathbb{R}^n \to \mathbb{R}$ is continuous, for each k. Let us define

$$C(x) = \langle Z(x)Z(0)\rangle, \tag{10.29}$$

so $C : \mathbb{R}^n \to \mathbb{R}$ is continuous, by (10.12). We want to show that

$$\langle Z(x)Z(y)\rangle = C(x-y), \quad \forall\, x, y \in \mathbb{R}^n. \tag{10.30}$$

Note that

$$C_k(x) = \langle Z_k(x)Z_k(0)\rangle \longrightarrow C(x), \quad \forall\, x \in S, \tag{10.31}$$

by (10.26). As noted in (10.27), it follows that

$$\langle Z_k(x)Z_k(y)\rangle = \langle Z_k(x-y)Z_k(0)\rangle \to C(x-y), \quad \forall\, x-y \in S. \tag{10.36}$$

Comparison with (10.26) yields

$$\langle Z(x)Z(y)\rangle = C(x-y), \quad \text{provided} \;\; x,\; x-y \in S, \tag{10.33}$$

which is a special case of (10.30). Fixing $x \in S$ and using (10.12) and the continuity of C, we have

$$\langle Z(x)Z(y)\rangle = C(x-y), \quad \forall\, x \in S,\; y \in \mathbb{R}^n. \tag{10.34}$$

Then taking $y \in \mathbb{R}^n$ and applying a similar argument, we have (10.30).

From (10.30), the norm continuity (10.15) follows readily. We have

$$\begin{aligned} \|Z(x+y) - Z(y)\|_{L^2}^2 &= \langle (Z(x+y) - Z(y))(Z(x+y) - Z(y))\rangle \\ &= 2C(0) - 2C(y), \end{aligned} \tag{10.35}$$

which tends to 0 as $|y| \to 0$.

A. Multiparameter ergodic theory

We assume $\{\tau_y : y \in \mathbb{F}^n\}$ is a family of measure preserving transformations on the probability space $(\mathcal{O}, \nu)$, satisfying $\tau_{y_1+y_2} = \tau_{y_1} \circ \tau_{y_2}$. To be definite, we take $\mathbb{F} = \mathbb{R}$, and we assume the induced action on $L^p(\mathcal{O}, \nu)$,

$$U(y)\varphi(\eta) = \varphi(\tau_y(\eta)), \tag{A.1}$$

is strongly continuous in y, for each $p \in [1, \infty)$. Note that $U(y)$ is an invertible isometry on $L^p(\mathcal{O}, \nu)$ (unitary on $L^2(\mathcal{O}, \nu)$) and $U(y_1 + y_2) = U(y_1)U(y_2)$. We aim to discuss ergodic theorems, dealing with averages of the form

$$A_R\varphi = \frac{1}{V(R)} \int\limits_{|y|\le R} U(y)\varphi \, dy. \tag{A.2}$$

First, there is an abstract mean ergodic theorem, valid when $\{U(y) : y \in \mathbb{R}^n\}$ is a strongly continuous unitary group on a Hilbert space H. It starts as follows.

Lemma A.1. *We have the orthogonal direct sum $H = K \oplus \overline{R}$, where*

$$\begin{aligned} K &= \{\varphi \in H : U(y)\varphi = \varphi, \ \forall\, y\}, \\ R &= \bigoplus_y \mathit{Range}\,(I - U(y)). \end{aligned} \tag{A.3}$$

Proof. This follows from the observation that

$$R^\perp = \bigcap_y \operatorname{Ker}\,(I - U(y)^*) = \bigcap_y \operatorname{Ker}\,(I - U(y)). \tag{A.4}$$

Here is the resulting abstract mean ergodic theorem.

Proposition A.2. *For all $\varphi \in H$, $A_R\varphi \to P\varphi$ in H-norm, where P is the orthogonal projection of H onto K.*

Proof. Note that $\varphi \in K \Rightarrow A_R\varphi \equiv \varphi$. Next, if $\varphi = (I - U(y_0))\psi, \ \psi \in H$, then

$$\frac{1}{V(R)} \int\limits_{|y|\le R} U(y)(I - U(y_0))\psi \, dy \to 0, \quad \text{as} \ \ R \to \infty. \tag{A.5}$$

In view of Lemma A.1, this yields the asserted result.

Proposition A.2 applies to the case $H = L^2(\mathcal{O}, \nu)$, with $U(y)$ given by (A.1). We record the following extension.

Proposition A.3. *Let $\{\tau_y : y \in \mathbb{R}^n\}$ satisfy the hypotheses given above, and take $U(y)$, A_R as in (A.1)–(A.2). Then, for $p \in [1,2]$, P extends to a continuous projection $P : L^p(\mathcal{O},\nu) \to L^p(\mathcal{O},\nu)$, and*

$$A_R\varphi \longrightarrow P\varphi \ \text{ in } \ L^p\text{-norm}, \quad \forall\, \varphi \in L^p(\mathcal{O},\nu). \tag{A.6}$$

Proof. Note that $\|A_R\|_{\mathcal{L}(L^p)} \le 1$ and use denseness of $L^2(\mathcal{O},\nu)$ in such $L^p(\mathcal{O},\nu)$.

Using other arguments, one can extend the scope of Proposition A.3 to all $p \in [1,\infty)$, but we omit details here.

The action of $\{\tau_y\}$ on $(\mathcal{O},\nu)$ is ergodic precisely when only constant functions on $\mathcal{O}$ belong to K, defined in (A.3). Then, and only then,

$$P\varphi = \Bigl(\int\limits_{\mathcal{O}} \varphi\, d\nu\Bigr)\, 1, \quad \forall\, \varphi \in L^2(\mathcal{O},\nu). \tag{A.7}$$

Note that this implies the criterion (4.10) for ergodicity.

It is of interest to extend Proposition A.2, replacing $\mathbb{R}^n$ by a broader class of Lie groups. Let G be a Lie group, endowed with a right-invariant Haar measure. Let $U : G \to \mathcal{L}(H)$ be a strongly continuous unitary representation of G on a Hilbert space H. Take, for $R \in \mathbb{R}^+$,

$$f_R \in L^1(G), \quad f_R \ge 0, \quad \int\limits_G f_R(y)\, dy \equiv 1, \tag{A.8}$$

and set

$$A_R\varphi = \int\limits_G f_R(y)\, U(y)\varphi\, dy. \tag{A.9}$$

We seek conditions that lead to a result of the form $A_R\varphi \to P\varphi$ as $R \to \infty$.

To start, we note that Lemma A.1 holds in this more general setting, with y running over G to define K and R as in (A.3). Again (A.4) provides the proof.

To proceed, clearly

$$\varphi \in K \Longrightarrow A_R\varphi \equiv \varphi. \tag{A.10}$$

Next, if $\varphi = (I - U(y_0))\psi,\ \psi \in H$, then

$$\begin{aligned} A_R\varphi &= \int\limits_G f_R(y)\bigl(U(y) - U(y)U(y_0)\bigr)\psi\, dy \\ &= \int\limits_G [f_R(y) - f_R(yy_0^{-1})]U(y)\psi\, dy. \end{aligned} \tag{A.11}$$

We can deduce that

$$A_R\varphi \longrightarrow 0 \quad \text{as} \quad R \to \infty, \tag{A.12}$$

for all $\varphi \in R$, hence for all $\varphi \in \overline{R}$, provided $\{f_R\}$ satisfies (A.8) and also

$$\lim_{R\to\infty} \int_G |f_R(y) - f_R(yy_0^{-1})|\, dy = 0, \quad \forall\, y_0 \in G. \tag{A.13}$$

We record the conclusion.

Proposition A.4. *Let $\{f_R : R \in \mathbb{R}^+\}$ satisfy (A.8) and define $A_R : H \to H$ by (A.9). For all $\varphi \in H$, $A_R\varphi \to P\varphi$ in H-norm, where P is the orthogonal projection of H on K, provided $\{f_R\}$ also satisfies (A.13).*

A Lie group G for which a family $\{f_R\}$ satisfying (A.8) and (A.13) exists is said to be *amenable*. For $G = \mathbb{R}^n$, one can pick $f_1 \geq 0$ such that $\int_{\mathbb{R}^n} f_1(y)\, dy = 1$ and set $f_R(y) = R^{-n} f(R^{-1}y)$. Many non-abelian Lie groups are amenable, but not all of them are.

So far, we have discussed mean ergodic theorems. The demonstrations given above are straightforward variants of the classical case of the real line. (Compare [T2], Chapter 14.) There are also pointwise a.e. results, known as Birkhoff ergodic theorems, that are classical for $\mathbb{R}^n$, having been extended from $n = 1$ to general n in [W]. See [L] for treatments of some other groups.

B. Atoms of $\widehat{C}$

Let σ be a finite (possibly complex) measure on $\mathbb{R}^n$. Its Fourier transform

$$C(x) = \hat{\sigma}(x) = (2\pi)^{-n/2} \int e^{-ix\cdot\xi}\, d\sigma(\xi) \tag{B.1}$$

is a bounded, continuous function on $\mathbb{R}^n$. We say σ has an atom at $p \in \mathbb{R}^n$ if $\sigma(\{p\}) \neq 0$. The set $\mathcal{A}(\sigma)$ of such points is countable, and we can write

$$\sigma = \sigma_0 + \sum_{p_j \in \mathcal{A}(\sigma)} a_j \delta_{p_j}, \tag{B.2}$$

where σ_0 has no atoms (we say σ_0 is a continuous measure). Here we prove the following result (due to N. Wiener), of interest in §4.

Proposition B.1. *With σ and C as above,*

$$\lim_{R\to\infty} \frac{1}{V(R)} \int\limits_{|y|\le R} |C(y)|^2\, dy = (2\pi)^{-n} \sum |a_j|^2. \tag{B.3}$$

Proof. We show that, more generally, if $f \in L^1(\mathbb{R}^n)$, $\int f(y)\, dy = 1$, and $f_R(y) = f(y/R)$, then

$$R^{-n} \int f_R(y) |\hat{\sigma}(y)|^2\, dy \longrightarrow (2\pi)^{-n} \sum |a_j|^2, \quad \text{as} \quad R \to \infty. \tag{B.4}$$

In fact, the left side of (B.4) is equal to

$$\begin{aligned} &(2\pi R)^{-n} \int f_R(y) \int e^{iy\cdot\xi}\, d\sigma(\xi) \int e^{-iy\cdot\eta}\, \overline{d\sigma(\eta)}\, dy \\ &= (2\pi)^{-n} \iint \Big\{ \int f(y) e^{iR(\xi-\eta)\cdot y}\, dy \Big\}\, d\sigma(\xi)\, \overline{d\sigma(\eta)}. \end{aligned} \tag{B.5}$$

Since the expression in brackets is bounded by $\|f\|_{L^1}$, we can pass to the limit under the integral sign. Now $\int f(y) e^{iR(\xi-\eta)\cdot y}\, dy$ tends to 0 as $R \to \infty$ if $\xi \neq \eta$, by the Riemann-Lebesgue lemma, while the expression is 1 at $\xi = \eta$. Thus

$$\lim_{R\to\infty} R^{-n} \int f_R(y) |\hat{\sigma}(y)|^2\, dy = (2\pi)^{-n} \iint\limits_{\xi=\eta} d\sigma(\xi)\, \overline{d\sigma(\eta)} = (2\pi)^{-n} \sum |a_j|^2. \tag{B.6}$$

This completes the proof.

C. Fourier transform of a stationary field

If we have a real-valued, continuous, stationary field on the n-torus,

$$Z : \mathbb{T}^n \longrightarrow L^2(\Omega, \mu), \tag{C.1}$$

it has a representation

$$Z(x) = \sum_{k\in\mathbb{Z}^n} \widehat{Z}(k) e^{ik\cdot x}, \tag{C.2}$$

with

$$\widehat{Z} : \mathbb{Z}^n \longrightarrow L^2(\Omega, \mu), \quad \widehat{Z}(k) = (2\pi)^{-n} \int\limits_{\mathbb{T}^n} Z(x) e^{-ik\cdot x}\, dx. \tag{C.3}$$

Note that $\widehat{Z}$ is not stationary. However, the various random variables $\widehat{Z}(k)$ are uncorrelated. In fact, in such a case,

$$\begin{aligned} \langle \widehat{Z}(k)\overline{\widehat{Z}(\ell)}\rangle &= (2\pi)^{-2n} \iint \langle Z(x)Z(y)\rangle e^{-ik\cdot x} e^{i\ell\cdot y}\,dx\,dy \\ &= (2\pi)^{-2n} \iint C(x-y) e^{-ik\cdot x} e^{i\ell\cdot y}\,dx\,dy \\ &= (2\pi)^{-n}\widehat{C}(k)\int e^{i(\ell-k)\cdot y}\,dy \\ &= \widehat{C}(k)\delta_{k\ell}. \end{aligned} \tag{C.4}$$

This is a special case of (5.29). (Note also that $\overline{\widehat{Z}(\ell)} = \widehat{Z}(-\ell)$.)

NOTE. The characterization of $C(x-y)$ as $\langle Z(x)Z(y)\rangle$ is equivalent to that in (4.2) and (5.10) if and only if $\langle Z(x)\rangle \equiv 0$. The same applies to (C.9).

Treating the Fourier transform of a real-valued, continuous, stationary field

$$Z : \mathbb{R}^n \longrightarrow L^2(\Omega,\mu) \tag{C.5}$$

will require the use of vector-valued tempered distributions, since Z is bounded but not integrable. We have

$$\widehat{Z} \in \mathcal{S}'(\mathbb{R}^n, L^2(\Omega,\mu)), \tag{C.6}$$

that is, $\widehat{Z} : \mathcal{S}(\mathbb{R}^n) \to L^2(\Omega,\mu)$, defined by

$$\widehat{Z}(f) = Z(\hat{f}) = \int Z(x)\hat{f}(x)\,dx, \quad f \in \mathcal{S}(\mathbb{R}^n), \tag{C.7}$$

where $\mathcal{S}(\mathbb{R}^n)$ denotes the Schwartz space of rapidly decreasing functions (cf. [T], Chapter 3). Formally (i.e., informally),

$$\widehat{Z}(f) = \int \widehat{Z}(\xi) f(\xi)\,d\xi. \tag{C.8}$$

Now, parallel to (C.4), we have

$$\begin{aligned} \langle \widehat{Z}(f), \overline{\widehat{Z}(g)}\rangle &= \langle Z(\hat{f})\overline{Z(\hat{g})}\rangle \\ &= \iint \langle Z(x)Z(y)\rangle \hat{f}(x)\overline{\hat{g}(y)}\,dx\,dy \\ &= \iint C(x-y)\hat{f}(x)\overline{\hat{g}(y)}\,dx\,dy. \end{aligned} \tag{C.9}$$

Note that

$$\text{(C.10)}\qquad \begin{aligned} \int C(x-y)\hat{f}(x)\,dx &= (2\pi)^{-n/2}\iint f(\xi)e^{-ix\cdot\xi}C(x-y)\,d\xi\,dx \\ &= \int e^{-iy\cdot\xi}\widehat{C}(\xi)f(\xi)\,d\xi, \end{aligned}$$

and

$$\text{(C.11)}\qquad \begin{aligned} \int e^{-iy\cdot\xi}\overline{\hat{g}(y)}\,dy &= \Big(\int e^{iy\cdot\xi}\hat{g}(y)\,dy\Big)^{*} \\ &= (2\pi)^{n/2}\overline{g(\xi)}, \end{aligned}$$

so

$$\text{(C.12)}\qquad \langle \widehat{Z}(f)\overline{\widehat{Z}(g)}\rangle = (2\pi)^{n/2}\int \widetilde{C}(\xi)f(\xi)\overline{g(\xi)}\,d\xi.$$

If we formally take $f=\delta_{\xi_1},\ g=\delta_{\xi_2}$ in (C.9), we get, formally,

$$\text{(C.13)}\qquad \begin{aligned} \langle \widehat{Z}(\xi_1)\overline{\widehat{Z}(\xi_2)}\rangle &= (2\pi)^{-n}\iint C(x-y)e^{-ix\cdot\xi_1}e^{iy\cdot\xi_2}\,dx\,dy \\ &= (2\pi)^{-n/2}\int \widehat{C}(\xi_1)e^{iy\cdot(\xi_2-\xi_1)}\,dy \\ &= \widehat{C}(\xi_1)\delta(\xi_1-\xi_2), \end{aligned}$$

at least if $\widehat{C}$ is continuous (which holds if $C\in L^1(\mathbb{R}^n)$). We have stated in §3 that $\widehat{C}$ is always a finite positive measure; call it σ. In this setting, we would write (C.12) as

$$\text{(C.14)}\qquad \langle \widehat{Z}(f),\overline{\widehat{Z}(g)}\rangle = (2\pi)^{n/2}\int f(\xi)\overline{g(\xi)}\,d\sigma(\xi).$$

In such a case, the bottom line of (C.13) can be interpreted as a finite positive measure on $\mathbb{R}^n\times\mathbb{R}^n$.

We also note that, in case $n=1$, we can write

$$\text{(C.15)}\qquad \widehat{Z}=\frac{d}{d\xi}F,\quad (1+|\xi|)^{-2}F\in L^2(\mathbb{R},L^2(\Omega,\mu)),$$

and hence

$$\text{(C.16)}\qquad Z(x)=(2\pi)^{-1/2}\int e^{ix\xi}\,dF(\xi).$$

To get this, we first note that (for general n) since $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ is bounded and continuous,

$$g \in L^2(\mathbb{R}^n) \Longrightarrow \widehat{Z} * g \in L^2(\mathbb{R}^n, L^2(\Omega, \mu)). \tag{C.17}$$

In case $n = 1$, we can take

$$\begin{aligned} g(\xi) = e^{-\xi}&, \quad \xi > 0, \\ 0&, \quad \xi < 0, \end{aligned} \tag{C.18}$$

which satisfies $g' = \delta - g$, and then

$$W = \widehat{Z} * g \Longrightarrow \widehat{Z} = W' + W, \quad W \in L^2(\mathbb{R}, L^2(\Omega, \mu)), \tag{C.19}$$

so (C.15) holds with

$$F(\xi) = W(\xi) + \int_0^\xi W(\eta)\, d\eta. \tag{C.20}$$

It follows readily from (C.14) and a limiting argument that the increments $F(\xi') - F(\xi)$ are uncorrelated over non-overlapping intervals.

A representation alternative to (C.16) is

$$Z(x) = (2\pi)^{-1/2}(1 - ix) \int e^{ix\xi} W(\xi)\, d\xi, \quad W \in L^2(\mathbb{R}, L^2(\Omega, \mu)), \tag{C.21}$$

with W as in (C.19). This has n-dimensional variants. We can take

$$g(\xi) = (1 - \Delta)^{-k} \delta(\xi), \quad k > \frac{n}{4}, \tag{C.22}$$

so $g \in L^2(\mathbb{R}^n)$ and $(1 - \Delta)^k g = \delta$. Then

$$W = \widehat{Z} * g \Longrightarrow \widehat{Z} = (1 - \Delta)^k W, \quad W \in L^2(\mathbb{R}^n, L^2(\Omega, \mu)), \tag{C.23}$$

and

$$Z(x) = (1 + |x|^2)^k (2\pi)^{-n/2} \int e^{ix\cdot\xi} W(\xi)\, d\xi. \tag{C.24}$$

Returning to (C.6)–(C.7), we note that $\widehat{Z}$ has a more precise description than being in $\mathcal{S}'(\mathbb{R}^n, L^2(\Omega, \mu))$. It is useful to introduce some notation. We fix the probability space (Ω, μ) and associated Hilbert space $L^2(\Omega, \mu)$. We say

$$Z \in \Sigma(\mathbb{R}^n) \tag{C.25}$$

provided $Z : \mathbb{R}^n \to L^2(\Omega, \mu)$ is a continuous, stationary field. We then set

$$\mathcal{F}\Sigma(\mathbb{R}^n) = \{\widehat{Z} : Z \in \Sigma(\mathbb{R}^n)\} \subset \mathcal{S}'(\mathbb{R}^n, L^2(\Omega, \mu)). \tag{C.26}$$

One observation is that if $Z \in \Sigma(\mathbb{R}^n)$, then (C.7) extends to $\hat{f} \in L^1(\mathbb{R}^n)$, i.e., to $f \in \mathcal{F}L^1(\mathbb{R}^n)$. More generally, given

$$\nu \in \mathcal{M}(\mathbb{R}^n), \quad \tilde{\nu} \in \mathcal{F}\mathcal{M}(\mathbb{R}^n), \tag{C.27}$$

where $\mathcal{M}(\mathbb{R}^n)$ is the space of finite Borel measures on $\mathbb{R}^n$, and $\tilde{\nu}(\xi) = (2\pi)^{-n/2} \int e^{ix\cdot\xi}\, d\nu(x)$ is the inverse Fourier transform, we have

$$\widehat{Z}(\tilde{\nu}) = Z(\nu) = \int Z(x)\, d\nu(x) \in L^2(\Omega, \mu), \tag{C.28}$$

so, extending (C.6), we have

$$\widehat{Z} : \mathcal{F}\mathcal{M}(\mathbb{R}^n) \longrightarrow L^2(\Omega, \mu), \tag{C.29}$$

if $Z \in \Sigma(\mathbb{R}^n)$.

Some related results arise as follows. First, given $f \in L^1(\mathbb{R}^n)$,

$$Z * f(x) = \int Z(x-y) f(y)\, dy \tag{C.30}$$

is well defined, and

$$f \in L^1(\mathbb{R}^n),\ Z \in \Sigma(\mathbb{R}^n) \Longrightarrow Z * f \in \Sigma(\mathbb{R}^n). \tag{C.31}$$

More generally, given $\nu \in \mathcal{M}(\mathbb{R}^n)$, we can set

$$Z * \nu(x) = \int Z(x-y)\, d\nu(y), \tag{C.32}$$

and then

$$\nu \in \mathcal{M}(\mathbb{R}^n),\ Z \in \Sigma(\mathbb{R}^n) \Longrightarrow Z * \nu \in \Sigma(\mathbb{R}^n). \tag{C.33}$$

Furthermore,

$$\mathcal{F}\Sigma(\mathbb{R}^n) \ \text{ is a module over } \ \mathcal{F}L^1(\mathbb{R}^n), \tag{C.34}$$

and, more generally,

$$\mathcal{F}\Sigma(\mathbb{R}^n) \ \text{ is a module over } \ \mathcal{F}\mathcal{M}(\mathbb{R}^n), \tag{C.35}$$

under pointwise multiplication, and

(C.36)
$$\widehat{Z * f} = \hat{f}\,\widehat{Z}, \quad \widehat{Z * \nu} = \hat{\nu}\,\widehat{Z}.$$

References

[AT] R. Adler and J. Taylor, Random Fields and Geometry, Springer NY, 2007.

[BMV] P. Baldi, D. Marinucci, and V. Varadarajan, On the characterization of isotropic random fields on homogeneous spaces of compact groups, Electronic Commun. in Probability 12 (2007), 291–302.

[BE] J. Blum and B. Eisenberg, Conditions for metric transitivity for stationary Gaussian processes on groups, Ann. Math. Stat. 43 (1972), 1737–1741.

[D] J. Doob, Stochastic Processes, J. Wiley, New York, 1953.

[G] U. Grenander, Stochastic processes and statistical inference, Ark. Mat. 1 (1950), 195–277.

[L] E. Lindenstrauss, Pointwise theorems for amenable groups, Invent. Math. 146 (2001), 259–295.

[Lum] J. Lumley, Stochastic Tools in Turbulence, Dover, NY, 2007.

[MP] D. Marinucci and G. Peccati, Random Fields on the Sphere, LMS Lecture Note #389, Cambridge Univ. Press, 2011.

[M] G. Maruyama, The harmonic analysis of stationary stochastic processes, Mem. Fac. Sci. Kyushu Univ. A 4 (1949), 45–106.

[MY] A. Monin and A. Yaglom, Statistical Fluid Mechanics, MIT Press, 1970 (Dover, NY, 2007).

[R] Y. Rubin, Applied Stochastic Hydrogeology, Oxford Univ. Press, 2003.

[T] M. Taylor, Partial Differential Equations, Vol. 1, Springer-Verlag, New York, 1996 (2nd ed., 2011).

[T2] M. Taylor, Measure Theory and Integration, Amer. Math. Soc., Providence RI, 2006.

[T3] M. Taylor, Introduction to Lie Groups, Open Math Notes, AMS, 2021.

[W] N. Wiener, The ergodic theorem, Duke Math. J. 5 (1939), 1–18.

[Y] A. Yaglom, Stationary Random Functions, Prentice-Hall, Englewood Cliffs NJ, 1962.

[Y2] A. Yaglom, Positive-definite functions and homogeneous random fields on homogeneous spaces, Soviet Math. Dokl. 1 (1960), 1402–1405.

[Z] D. Zelobenko, Compact Lie groups and their Representations, Transl. Math. Monogr. #40, AMS, Providence RI, 1973.

7. Fractional diffusion equations

1. Introduction

Work on non-Gaussian probability distributions has led people to consider "fractional diffusion equations" of the following sort:

$$ {}^c\partial_t^\beta u = -(-\Delta)^\alpha u, \quad t \ge 0; \qquad u(0,x) = f(x), \tag{1.1} $$

with $\alpha, \beta \in (0,1]$, the case $\alpha = \beta = 1$ being the standard diffusion equation. Here, Δ is the Laplace operator, the fractional power $(-\Delta)^\alpha$ is a positive self-adjoint operator, defined by the spectral theorem, and ${}^c\partial_t^\beta$ is a Caputo fractional derivative (a variant of the Riemann-Liouville fractional derivative, better suited for initial-value problems):

$$ {}^c\partial_t^\beta v(t) = \frac{1}{\Gamma(1-\beta)} \int_0^t (t-s)^{-\beta}\, \partial_s v(s)\, ds, \tag{1.2} $$

if $\beta \in (0,1)$. There have been a number of recent papers on this topic, with emphasis on the case $\Delta = \partial_x^2$, acting on functions on the line $\mathbb{R}$. See, for example, [CCL], [CCL2], [MPG], and references therein.

Here we point out that in a more general context the solution operator $S^t_{\beta,\alpha}$ to (1.1) yields a family of probability distributions, by virtue of being positivity-preserving:

$$ f \ge 0 \Longrightarrow S^t_{\beta,\alpha} f \ge 0, \tag{1.3} $$

and having the property

$$ \int S^t_{\beta,\alpha} f(x)\, dx = \int f(x)\, dx, \tag{1.4} $$

under appropriate hypotheses. This will hold, e.g., when Δ is the Laplace operator on $\mathbb{R}^n$, or on a bounded domain $\Omega \subset \mathbb{R}^n$, with the Neumann boundary condition. (With the Dirichlet boundary condition, (1.3) will hold, but not (1.4). In such a case one would have a diffusion with absorption.) The key behind this is the demonstration that

$$ S^t_{\beta,\alpha} = \int_0^\infty \Psi^t_{\beta,\alpha}(s) e^{s\Delta}\, ds, \tag{1.5} $$

where

$$ \Psi^t_{\beta,\alpha}(s) \ge 0 \quad \text{for} \ \ s, t > 0, \quad \alpha, \beta \in (0,1] \tag{1.6} $$

(and $(\alpha, \beta) \neq (1,1)$), and

$$\int_0^\infty \Psi^t_{\beta,\alpha}(s)\, ds = 1. \tag{1.7}$$

It will be convenient to work in the more general setting of symmetric diffusion semigroups. We also break up the analysis of positivity into two pieces. In §2 we analyze the case $\beta = 1$ of (1.1), generalized to

$$\partial_t u = -L^\alpha u, \quad u(0) = f, \tag{1.8}$$

where L is a positive self-adjoint operator and e^{-tL} a symmetric diffusion semigroup. This analysis is classical and we merely sketch the results, described in more detail in Chapter IX of [Y]. The basic conclusion is that e^{-tL^α} is also a symmetric diffusion semigroup, for $\alpha \in (0,1)$. It will be useful to have this analysis for the next step, tackled in §3:

$$^c\partial_t^\beta u = -Au, \quad u(0) = f, \tag{1.9}$$

where e^{-tA} is a symmetric diffusion semigroup and $\beta \in (0,1)$. A familiar Laplace transform analysis writes the solution operator S^t_β to (1.9) as

$$S^t_\beta = E_\beta(-t^\beta A), \tag{1.10}$$

where $E_\beta(z)$ is a special function (the Mittag-Leffler function) and the right side of (1.10) is defined by the functional calculus for self-adjoint operators. Known Laplace transform identities involving $E_\beta(z)$ (cf. (3.6), (3.11)) serendipitously allow us to deduce (1.5)–(1.7) (in a more general context, with $-\Delta$ replaced by L) from the results of §2.

In §4 we consider an extension of (1.1) to $\beta \in (1,2]$. In such a case (1.5)–(1.7) fails. One still has (1.3) for $\alpha = 1$ and $\Delta = \partial_x^2$ on functions on $\mathbb{R}$ (as shown in [MPG]), but we note that such positivity fails in higher dimension.

In §5 we construct functions $\psi(\xi)$, homogeneous of degree $\alpha \in (0,2)$, such that $e^{-t\psi(D)}$, acting on functions on $\mathbb{R}^n$, satisfies (1.3), including as special cases (with $n = 1$) various fractional derivatives. The probability distributions so obtained are known as α-stable distributions. We mention connections with material in [ST], and also Chapter 3 of this text.

In §6 we briefly discuss a class of fractional diffusion-reaction equations. In §7 we present the results of some numerical calculations of solutions to some linear diffusion and fractional diffusion equations and fractional diffusion-reaction equations of Fisher-Kolmogorov type, for functions $u(t,x)$ defined on $[0,\infty) \times S^1$.

In §8 we discuss formulas and estimates for the solution to inhomogeneous fractional diffusion equations, of the form

$$^c\partial_t^\beta u = -Au + q(t), \quad u(0) = f. \tag{1.11}$$

In §9 we apply results of §8 to establish the short time existence to fractional diffusion-reaction equations of the form

$$(1.12)\qquad {}^c\partial_t^\beta u = -Au + F(u), \quad u(0) = f, \quad A = (-\Delta)^{m/2}, \quad 0 < m \le 2.$$

when $\beta \in (0,1)$, the case $\beta = 1$ having been discussed in §6. We consider the cases $f \in C(M)$ and $f \in L^6(M)$, when M is a compact n-dimensional Riemannian manifold. The latter case requires the restriction $n/2 < m \le 2$. Also, for this result, and for the results of §§10-11, we essentially require $F(u)$ to be a cubic polynomial in u, a situation that is popular in the study of reaction-diffusion equations.

In §10 we consider (1.12) for $f \in L^{3q}(M)$, when

$$(1.13)\qquad q > 1 \ \text{ and } \ \frac{3n}{3q} < m \le 2.$$

In §11 we push this a bit, in the case $n = 2$, $m = 2$, and obtain local existence given $f \in L^p(M)$, $p > 2$.

In Appendix A we recall some basic material on Riemann-Liouville fractional integrals and the Caputo fractional derivative, used in the main body of this chapter. In Appendix B we briefly discuss results on finite linear systems, of the form

$$(1.14)\qquad {}^c\partial_t^\beta u = Lu, \quad u(0) = f,$$

where

$$(1.15)\qquad f \in V, \quad L \in \operatorname{End}(V), \quad \dim V < \infty.$$

In Appendix C we provide several approaches to deriving the formula (1.10), with the power series (3.5) for $E_\beta(z)$.

2. Subordination identities

Let L be a positive self-adjoint operator. By the spectral theorem, one has

$$(2.1)\qquad e^{-tL^\alpha} = \int_0^\infty \Phi_{t,\alpha}(s) e^{-sL}\, ds, \quad 0 < \alpha < 1,$$

for $t > 0$, where $\Phi_{t,\alpha}$ has the property

$$(2.2)\qquad e^{-t\lambda^\alpha} = \int_0^\infty \Phi_{t,\alpha}(s) e^{-s\lambda}\, ds, \quad \lambda > 0.$$

The fact that

$$(2.3)\qquad (-1)^k \partial_\lambda^k e^{-t\lambda^\alpha} \ge 0 \quad \text{for} \quad \lambda, t > 0, \quad k \in \mathbb{Z}^+$$

implies

$$\Phi_{t,\alpha}(s) \geq 0, \quad \text{for } \ s \in [0,\infty), \tag{2.4}$$

given $t \in (0,\infty)$, $\alpha \in (0,1)$. One also has

$$\int_0^\infty \Phi_{t,\alpha}(s)\, ds = 1. \tag{2.5}$$

This is discussed in a more general context in §IX.11 of [Y].

We recall that the most familiar case is the case $\alpha = 1/2$, where

$$\Phi_{t,1/2}(s) = \frac{t}{2\pi^{1/2}} e^{-t^2/4s}\, s^{-3/2}. \tag{2.6}$$

This particular subordination identity has numerous applications to analysis; cf. [T], Chapter 3, (5.22)–(5.31), and Chapter 11, (2.24), for some examples.

The positivity in (2.4) has the implication that if e^{-sL} is a diffusion semigroup, so is e^{-tL^α}, for each $\alpha \in (0,1)$.

We record some further useful properties of $\Phi_{t,\alpha}$. First, a change of variable gives

$$\Phi_{t,\alpha}(s) = t^{-1/\alpha}\Phi_{1,\alpha}(t^{-1/\alpha}s). \tag{2.7}$$

Next, up to a constant factor,

$$f_\alpha(\xi) = e^{-(i\xi)^\alpha} \tag{2.8}$$

is the Fourier transform of $\Phi_{1,\alpha}$, extended by 0 on $(-\infty,0]$. For $\alpha \in (0,1)$, f_α is rapidly decreasing, with all derivatives, as $|\xi| \to \infty$. It follows that $\Phi_{1,\alpha}(s)$, so extended, is C^∞ on $\mathbb{R}$, in particular, vanishing to all orders as $s \to 0$, as illustrated in case $\alpha = 1/2$ by

$$\Phi_{1,1/2}(s) = \frac{1}{2\pi^{1/2}} e^{-1/4s} s^{-3/2}, \quad s > 0. \tag{2.9}$$

On the other hand, the nature of the singularity of f_α at $\xi = 0$ implies that $\Phi_{1,\alpha}(s)$ has the following asymptotic behavior as $s \to +\infty$:

$$\Phi_{1,\alpha}(s) \sim \sum_{k\geq 1} \gamma_{\alpha k} s^{-k\alpha-1}, \quad s \to +\infty, \tag{2.10}$$

also illustrated by (2.9) in case $\alpha = 1/2$.

3. Fractional diffusion equations

Let A be a positive self-adjoint operator. We analyze the solution to

$$ {}^c\partial_t^\beta u = -Au, \quad t > 0; \qquad u(0) = f, \tag{3.1} $$

given $\beta \in (0,1)$, and show that if e^{-sA} is a diffusion semigroup the solution to (3.1) is also given by a diffusion, i.e., a family of positivity-preserving operators. As is standard, we use the fact that, with

$$ \mathcal{L}u(s) = \int_0^\infty e^{-st} u(t)\,dt, \tag{3.2} $$

the equation (3.1) becomes

$$ (s^\beta + A)\mathcal{L}u(s) = s^{\beta-1} f. \tag{3.3} $$

Application of Laplace inversion (cf. [MPG], Appendix A) gives

$$ u(t) = E_\beta(-t^\beta A) f, \tag{3.4} $$

where $E_\beta(z)$ is the Mittag-Leffler function

$$ E_\beta(z) = \sum_{n=0}^\infty \frac{z^n}{\Gamma(\beta n + 1)}, \tag{3.5} $$

and the linear operator $E_\beta(-t^\beta A)$ in (3.4) is given by the standard operator calculus for self-adjoint operators. As derived in (A.37) of [MPG], one has

$$ E_\beta(-s) = \int_0^\infty M_\beta(r) e^{-rs}\,dr, \quad s > 0, \tag{3.6} $$

given $\beta \in (0,1)$, where

$$ M_\beta(r) = \frac{1}{2\pi i} \int_\gamma e^{\zeta - r\zeta^\beta} \frac{d\zeta}{\zeta^{1-\beta}}, \tag{3.7} $$

and γ can be taken as a vertical line $\{i\sigma + \varepsilon : \sigma \in \mathbb{R}\}$, with small $\varepsilon > 0$. It follows that

$$ E_\beta(-t^\beta A) = \int_0^\infty M_\beta(r) e^{-rt^\beta A}\,dr, \quad t > 0, \quad \beta \in (0,1). \tag{3.8} $$

Some particular cases of $M_\beta(r)$, mentioned in (A.34)–(A.35) of [MPG], are

$$ M_{1/2}(r) = \pi^{-1/2} e^{-r^2/4}, \quad M_{1/3}(r) = 3^{2/3}\mathrm{Ai}(3^{-1/3} r). \tag{3.9} $$

These examples illustrate the following important result.

Proposition 3.1. *Given* $0 < \beta < 1,\ r \geq 0$*, we have*

$$M_\beta(r) \geq 0. \tag{3.10}$$

Proof. This can be deduced from the following identity, due to [P], and noted in (A.41) of [MPG]:

$$\beta \int_0^\infty r^{-\beta-1} M_\beta(r^{-\beta}) e^{-rs}\, dr = e^{-s^\beta}, \tag{3.11}$$

given $\beta \in (0,1)$. Comparison with (2.2) gives

$$\beta r^{-\beta-1} M_\beta(r^{-\beta}) = \Phi_{1,\beta}(r). \tag{3.12}$$

Thus the positivity (3.10) follows from (2.4)

We are now able to prove the positivity assertion made in the introduction. We merely plug (2.1) into (3.8) to obtain (1.5)–(1.7).

4. The case $\beta \in (1,2]$

Work in [MPG] also considered (1.1) for $\beta \in (1,2]$. Here the Caputo fractional derivative ${}^c\partial_t^\beta$ is given by

$${}^c\partial_t^\beta v(t) = \frac{1}{\Gamma(2-\beta)} \int_0^t (t-s)^{-\beta+1}\, \partial_s^2 v(s)\, ds, \quad 1 < \beta < 2.$$

One continues to get (3.4), i.e.,

$$u(t) = S^t_{\beta,\alpha} f = E_\beta(-t^\beta A) f, \quad A = (-\Delta)^\alpha. \tag{4.1}$$

One has in particular

$$E_2(-s) = \cos s^{1/2}, \tag{4.2}$$

and hence

$$S^t_{2,\alpha} = \cos t(-\Delta)^{\alpha/2}, \tag{4.3}$$

the solution operator to the Cauchy problem

$$\big(\partial_t^2 + (-\Delta)^\alpha\big) u = 0, \quad u(0) = f,\ \partial_t u(0) = 0. \tag{4.4}$$

For $\alpha = 1$ one gets the wave equation:

$$(\partial_t^2 - \Delta)u = 0, \quad u(0) = f,\ \partial_t u(0) = 0. \tag{4.5}$$

If $A = -\partial_x^2$, acting on functions on the line, then, as shown in [MPG], one has a diffusion. In fact, by (4.6) of [MPG], for $\beta < 2$,

$$E_\beta(t^\beta \partial_x^2)\delta(x) = \frac{1}{2t^{1/2}} M_{\beta/2}(t^{-\beta/2}|x|), \quad x \in \mathbb{R}, \tag{4.6}$$

for $t > 0$. For $\beta \in (1, 2)$ we have $\beta/2 \in (1/2, 1)$, and Proposition 3.1 yields positivity of (4.6). As for the endpoint case, $\beta = 2$, one has

$$\left(\cos t\sqrt{-\partial_x^2}\right)\delta(x) = \frac{1}{2}\big[\delta(x+t) + \delta(x-t)\big], \quad x \in \mathbb{R}. \tag{4.7}$$

Well known formulas for $\cos t\sqrt{-\Delta}\ \delta(x)$ with $x \in \mathbb{R}^n$ (cf. [T], Chapter 3, §5) involve distributions that are not positive measures. Hence positivity fails for $S^t_{2,1}$ on functions on $\mathbb{R}^n$ with $n \geq 2$. It follows by continuity that positivity fails for $S^t_{\beta,1}$ for β close to 2. One might investigate in more detail just how $S^t_{\beta,\alpha}$ behaves on functions on $\mathbb{R}^n$ for $n \geq 2$, $\beta \in (1, 2)$.

5. Diffusion semigroups with homogeneous generators

Here we consider semigroups of the form $e^{-t\psi(D)}$, where $\psi(D)$ acts on functions on $\mathbb{R}^n$ via Fourier multiplication by $\psi(\xi)$. We construct functions homogeneous of degree $\alpha \in (0, 2)$ for which $e^{-t\psi(D)}$ is positivity preserving and furthermore satisfies

$$0 \leq f \leq 1 \Longrightarrow 0 \leq e^{-t\psi(D)} f \leq 1, \quad \forall\, t > 0. \tag{5.1}$$

Of course

$$\psi(\xi) = |\xi|^\alpha, \quad 0 \leq \alpha \leq 2, \tag{5.2}$$

works, by the results of §2. We obtain further cases by specializing the Levy-Khinchin formula (cf. [J], §3.7). In this way we obtain the following such homogeneous generators:

$$\begin{aligned} \Phi_{\alpha,g}(\xi) &= -\int_{\mathbb{R}^n} (e^{iy\cdot\xi} - 1) g(y)|y|^{-n-\alpha}\, dy, && 0 < \alpha < 1, \\ \Psi_{\alpha,g}(\xi) &= -\int_{\mathbb{R}^n} (e^{iy\cdot\xi} - 1 - iy\cdot\xi) g(y)|y|^{-n-\alpha}\, dy, && 1 < \alpha < 2. \end{aligned} \tag{5.3}$$

The function g is assumed to be positive, bounded, and homogeneous of degree 0, i.e.,

$$g \geq 0, \quad g \in L^\infty(\mathbb{R}^n), \quad g(ry) = g(y), \quad \forall\, r > 0. \tag{5.4}$$

It is easy to verify that both integrals in (5.3) are absolutely convergent, and, for $r > 0$,

$$\begin{aligned} \Phi_{\alpha,g}(r\xi) &= r^\alpha \Phi_{\alpha,g}(\xi), \quad 0 < \alpha < 1, \\ \Psi_{\alpha,g}(r\xi) &= r^\alpha \Psi_{\alpha,g}(\xi), \quad 1 < \alpha < 2. \end{aligned} \tag{5.5}$$

When $g \equiv 1$ we obtain a positive multiple of (5.2).

We now specialize to $n = 1$ and $g = \chi_{\mathbb{R}^+}$, so we look at

$$\begin{aligned} \varphi_\alpha(\xi) &= -\int_0^\infty (e^{iy\xi} - 1) y^{-1-\alpha}\, dy, && 0 < \alpha < 1, \\ \psi_\alpha(\xi) &= -\int_0^\infty (e^{iy\xi} - 1 - iy\xi) y^{-1-\alpha}\, dy, && 1 < \alpha < 2. \end{aligned} \tag{5.6}$$

Clearly φ_α and ψ_α are holomorphic in $\{\xi \in \mathbb{C} : \operatorname{Im} \xi > 0\}$, and homogeneous of degree α in ξ. Also, for $\eta > 0$,

$$\begin{aligned} \varphi_\alpha(i\eta) &= -\int_0^\infty (e^{-y\eta} - 1) y^{-1-\alpha}\, dy > 0, && 0 < \alpha < 1, \\ \psi_\alpha(i\eta) &= -\int_0^\infty (e^{-y\eta} - 1 + y\eta) y^{-1-\alpha}\, dy < 0, && 1 < \alpha < 2, \end{aligned} \tag{5.7}$$

since, for $r > 0$, $1 - r < e^{-r} < 1$. It follows that $\varphi_\alpha(\xi)$ and $\psi_\alpha(\xi)$ are positive multiples of

$$\begin{aligned} \varphi_\alpha^\#(\xi) &= (-i\xi)^\alpha, && 0 < \alpha < 1, \\ \psi_\alpha^\#(\xi) &= -(-i\xi)^\alpha, && 1 < \alpha < 2, \end{aligned} \tag{5.8}$$

restrictions to $\mathbb{R}$ of functions holomorphic on $\{\xi \in \mathbb{C} : \operatorname{Im} \xi > 0\}$. Taking instead $g = \chi_{\mathbb{R}^-}$, we obtain positive multiples of

$$\begin{aligned} \varphi_\alpha^b(\xi) &= (i\xi)^\alpha, && 0 < \alpha < 1, \\ \psi_\alpha^b(\xi) &= -(i\xi)^\alpha, && 1 < \alpha < 2, \end{aligned} \tag{5.9}$$

restrictions to $\mathbb{R}$ of functions holomorphic on $\{\xi \in \mathbb{C} : \operatorname{Im} \xi < 0\}$, satisfying

$$\varphi_\alpha^b(-i\eta) > 0, \quad \psi_\alpha^b(-i\eta) < 0, \quad \forall\, \eta > 0. \tag{5.10}$$

The functions in (5.8) and (5.9) are well known examples of homogeneous functions $\psi(\xi)$ for which $e^{-t\psi(D)}$ satisfies (5.1). The associated operators $\psi(D)$ are fractional derivatives.

It is also useful to observe the explicit formulas

(5.11)
$$e^{-t\varphi^{\#}_\alpha(\xi)} = e^{-t(\cos \pi\alpha/2)|\xi|^\alpha} \Big[\cos\Big(t\big(\sin\frac{\pi\alpha}{2}\big)|\xi|^\alpha\Big) + i\sigma(\xi)\sin\Big(t\big(\sin\frac{\pi\alpha}{2}\big)|\xi|^\alpha\Big)\Big].$$

for $t > 0,\ 0 < \alpha < 1$, where

$$\sigma(\xi) = \operatorname{sgn}\xi, \tag{5.12}$$

and

$$e^{-t\psi^{\#}_\alpha(\xi)} = e^{t(\cos \pi\alpha/2)|\xi|^\alpha} \Big[\cos\Big(t\big(\sin\frac{\pi\alpha}{2}\big)|\xi|^\alpha\Big) - i\sigma(\xi)\sin\Big(t\big(\sin\frac{\pi\alpha}{2}\big)|\xi|^\alpha\Big)\Big], \tag{5.13}$$

fior $t > 0,\ 1 < \alpha < 2$. Note that

$$0 < \alpha < 1 \Rightarrow \cos\frac{\pi\alpha}{2} > 0, \quad 1 < \alpha < 2 \Rightarrow \cos\frac{\pi\alpha}{2} < 0, \tag{5.14}$$

so of course we have decaying exponentials in both (5.11) and (5.13). We get similar formulas with # replaced by b, since in fact

$$\varphi^b_\alpha(\xi) = \varphi^{\#}_\alpha(-\xi), \quad \psi^b_\alpha(\xi) = \psi^{\#}_\alpha(-\xi). \tag{5.15}$$

Returning to the general formulas (5.3), we can switch to polar coordinates and write

$$\begin{aligned}\Phi_{\alpha,g}(\xi) &= -\int_{S^{n-1}}\int_0^\infty (e^{is\omega\cdot\xi} - 1)g(\omega)s^{-1-\alpha}\,ds\,dS(\omega),\\ \Psi_{\alpha,g}(\xi) &= -\int_{S^{n-1}}\int_0^\infty (e^{is\omega\cdot\xi} - 1 - is\omega\cdot\xi)g(\omega)s^{-1-\alpha}\,ds\,dS(\omega),\end{aligned} \tag{5.16}$$

and hence

$$\begin{aligned}\Phi_{\alpha,g}(\xi) &= \int_{S^{n-1}} \varphi_\alpha(\omega\cdot\xi)g(\omega)\,dS(\omega),\\ \Psi_{\alpha,g}(\xi) &= \int_{S^{n-1}} \psi_\alpha(\omega\cdot\xi)g(\omega)\,dS(\omega).\end{aligned} \tag{5.17}$$

We can extend the scope, replacing $g(\omega)\,dS(\omega)$ by a general positive, finite Borel measure on S^{n-1}. Taking into account the calculations yielding (5.8)–(5.9), we obtain homogeneous generators satisfying (5.1), of the form

$$\begin{aligned}\Phi^b_{\alpha,\nu}(\xi) &= \int_{S^{n-1}} (i\omega\cdot\xi)^\alpha\,d\nu(\omega), \qquad 0 < \alpha < 1,\\ \Psi^b_{\alpha,\nu}(\xi) &= -\int_{S^{n-1}} (i\omega\cdot\xi)^\alpha\,d\nu(\omega), \quad 1 < \alpha < 2,\end{aligned} \tag{5.18}$$

where ν is a positive, finite Borel measure on S^{n-1}.

It remains to discuss the case $\alpha = 1$. For $n = 1$ it is seen that positive multiples of

$$|\xi| + ia\xi, \quad a \in \mathbb{R}, \tag{5.19}$$

work. Hence the following functions on $\mathbb{R}^n$ work:

$$|\omega \cdot \xi| + ia\omega \cdot \xi, \quad \omega \in S^{n-1}, \ a \in \mathbb{R}.$$

We can take positive superpositions of such functions and, in analogy with (5.18), obtain generators of diffusion semigroups whose negatives are Fourier multiplication by

$$ib \cdot \xi + \Xi_\nu(\xi), \tag{5.20}$$

where $b \in \mathbb{R}^n$ and

$$\Xi_\nu(\xi) = \int\limits_{S^{n-1}} |\omega \cdot \xi| \, d\nu(\omega). \tag{5.21}$$

We now tie in results derived above with material given in Chapters 1–2 of [ST]. For such functions $\psi(\xi)$, homogeneous of degree $\alpha \in (0, 2]$, as constructed above, the probability distributions

$$p_t(x) = e^{-t\psi(D)}\delta(x) \tag{5.22}$$

are known as α-stable distributions. In the notation (1.1.6) of [ST], consider

$$\psi(\xi) = \sigma^\alpha \, |\xi|^\alpha \Bigl(1 - i\beta(\operatorname{sgn} \xi) \tan \frac{\pi\alpha}{2}\Bigr), \quad \xi \in \mathbb{R}. \tag{5.23}$$

Here

$$\sigma \in (0, \infty), \quad \beta \in [-1, 1], \tag{5.24}$$

and $\alpha \in (0, 2)$ but $\alpha \neq 1$. Also, take $\mu \in \mathbb{R}$. Then $e^{-\psi(D)+i\mu D}\delta(x)$ is a probability distribution on the line called an α-stable distribution with scale parameter σ, skewness parameter β, and shift parameter μ. It is clear from (5.11)–(5.13) that each function of the form (5.23) is a positive linear combination of $\varphi^\#_\alpha(\xi)$ and $\varphi^b_\alpha(\xi)$ if $\alpha \in (0, 1)$ and a positive linear combination of $\psi^\#_\alpha(\xi)$ and $\psi^b_\alpha(\xi)$ if $\alpha \in (1, 2)$.

In case $\alpha = 1$, one goes beyond $\psi(\xi)$ homogeneous of degree 1 in ξ, to consider

$$\psi(\xi) = \sigma \, |\xi| \Bigl(1 + i\frac{2\beta}{\pi}(\operatorname{sgn} \xi) \log |\xi|\Bigr) + i\mu\xi, \quad \xi \in \mathbb{R}, \tag{5.25}$$

again with $\beta \in [-1,1]$, $\mu \in \mathbb{R}$. Then $e^{-\psi(D)}\delta(x)$ is a probability distribution on $\mathbb{R}$ called a 1-stable distribution, with scale parameter σ, skewness β, and shift μ. The cases arising from (5.19) all have skewness $\beta = 0$.

Similarly, functions $\psi(\xi)$ of the form (5.18) and (5.20)–(5.21) produce probability distributions $e^{-\psi(D)}\delta(x)$ on $\mathbb{R}^n$ that are α-stable. These, plus analogues with a shift incorporated, comprise all of them except when $\alpha = 1$, in which case one generalizes (5.21) to

$$\widetilde{\Xi}_\nu(\xi) = \int\limits_{S^{n-1}} |\omega\cdot\xi|\Big(1 + \frac{2i}{\pi}(\operatorname{sgn}\omega\cdot\xi)\log|\omega\cdot\xi|\Big)\,d\nu(\omega). \tag{5.26}$$

Compare (2.3.1)–(2.3.2) in [ST].

We return to the case $n = 1$ and make some more comments on the probability distributions

$$\begin{aligned} p_t^\alpha(x) &= e^{-t\varphi_\alpha^\#(D)}\delta(x), \quad 0 < \alpha < 1, \\ p_t^\alpha(x) &= e^{-t\psi_\alpha^\#(D)}\delta(x), \quad 1 < \alpha < 2, \end{aligned} \tag{5.27}$$

and their variants with # replaced by b, which are simply $p_t^\alpha(-x)$. Explicitly, we have

$$p_t^\alpha(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ix\cdot\xi - t\varphi_\alpha^\#(\xi)}\,d\xi, \tag{5.28}$$

for $0 < \alpha < 1$, with $\varphi_\alpha^\#(\xi)$ replaced by $\psi_\alpha^\#(\xi)$ for $1 < \alpha < 2$. Recall that $\varphi_\alpha^\#$ and $\psi_\alpha^\#$ are holomorphic in $\{\xi \in \mathbb{C} : \operatorname{Im}\xi > 0\}$. It follows from the Paley-Wiener theorem that, for each $t > 0$,

$$p_t^\alpha(x) = 0, \quad \text{for } \ x \in [0,\infty),\ 0 < \alpha < 1. \tag{5.29}$$

This theorem does not apply when $\alpha \in (1,2)$, but a shift in the contour of integration to $\{\xi + ib : \xi \in \mathbb{R}\}$, with arbitrary $b > 0$ yields

$$p_t^\alpha(x) = e^{-bx} \times \text{ bounded function of } x, \tag{5.30}$$

for $x \in \mathbb{R}$, whenever $1 < \alpha < 2$, hence

$$p_t^\alpha(x) = o(e^{-bx}), \quad \forall\, b > 0, \quad \text{as } \ x \to +\infty, \quad \text{for } \ 1 < \alpha < 2. \tag{5.31}$$

A more precise asymptotic behavior is stated in (1.2.11) of [ST].

We also note that, for $\alpha \in (1.2)$, $p_t^\alpha(x)$ is real analytic in $x \in \mathbb{R}$, and in fact extends to an entire holomorphic function in $x \in \mathbb{C}$, for each $t > 0$, due to rapidity with which $\operatorname{Re}\psi_\alpha^\#(\xi) \to +\infty$ as $|\xi| \to \infty$, which of course forbids (5.29) in this case.

6. Fractional diffusion-reaction equations

We consider $\ell \times \ell$ systems of equations

$$\frac{\partial u}{\partial t} = -Lu + X(u), \quad u(0) = f, \tag{6.1}$$

where $u = u(t,x)$ takes values in $\mathbb{R}^\ell$, X is a real vector field on $\mathbb{R}^\ell$, and L is a diagonal operator,

$$L = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_\ell \end{pmatrix}, \tag{6.2}$$

where each operator $-A_j$ generates a diffusion semigroup, satisfying

$$a \le f \le b \Longrightarrow a \le e^{-tA_j} f \le b, \quad \forall\, t > 0. \tag{6.3}$$

In case the operators A_j are second order differential operators satisfying (6.3), the system (6.1) is a reaction-diffusion equation. Recent studies have considered A_j given by fractional derivatives. For example, [CCL3] considers the following scalar equation (a modification of the Fisher-Kolmogorov equation):

$$\frac{\partial u}{\partial t} = -\psi_\alpha^b(D)u + u(1-u), \quad u(0) = f, \tag{6.4}$$

where $\alpha \in (1,2)$ and ψ_α^b is given by (5.9).

Our next goal is to present an extension of Proposition 4.4 in Chapter 15 of [T], giving a global existence result and some qualitative information on an important class of systems of the form (6.1). Here is the set-up. We assume there is a family $\{K_s : 0 \le s < \infty\}$ of compact subsets of $\mathbb{R}^\ell$ such that each K_s has the invariance property

$$f(x) \in K_s\ \forall\, x \Longrightarrow e^{-tL} f(x) \in K_s\ \forall\, x. \tag{6.5}$$

For example, K_s could be a Cartesian product of intervals, and then (6.3) implies (6.5). Furthermore, we assume that

$$\mathcal{F}_X^t(K_s) \subset K_{s+t}, \quad s,t \in \mathbb{R}^+, \tag{6.6}$$

where $\mathcal{F}_X^t$ is the flow on $\mathbb{R}^\ell$ generated by X. Then we have the following result.

Proposition 6.1. *Under the hypotheses (6.5)–(6.6), if $f(x) \in K_0$ for all x, then (6.1) has a solution for all $t \in [0, \infty)$, and, for each $t > 0$,*

$$u(t,x) \in K_t, \quad \forall\, x. \tag{6.7}$$

The proof is basically the same as the proof of Proposition 4.4 mentioned above. The key behind (6.7) is the nonlinear Trotter product formula:

$$u(t) = \lim_{n\to\infty} \Bigl(e^{-(t/n)L} \mathcal{F}^{t/n}\Bigr)^n f, \tag{6.8}$$

where

$$\mathcal{F}^t f(x) = \mathcal{F}^t_X(f(x)). \tag{6.9}$$

As one application, in case $\ell = 1$, we see that if $0 < a < b < \infty$, and if

$$a \le f(x) \le b, \quad \forall\, x \in \mathbb{R}, \tag{6.10}$$

then (6.4) has a solution for all $t \in [0, \infty)$, and

$$\lim_{t\to\infty} u(t,x) \equiv 1. \tag{6.11}$$

With a little more work, we could allow $a = 0$ and obtain (6.11) as long as f is not identially zero. In [CCL3] there is an intriguing discussion of finer qualitative behavior of moving front solutions to (a variant of) (6.4), based on numerical evidence. See §7 for some more on this.

One can consider various other reaction-diffusion equations, such as the Fitzhugh-Nagumo equations, and variants, with ∂_x^2 replaced by fractional derivatives, to which Proposition 6.1 would be applicable. See Chapter 15, §4 of [T] for other examples, which could be similarly generalized.

7. Numerical attack

Here we discuss numerical results on five linear (fractional) diffusion equations:

$$\frac{\partial u}{\partial t} = -Lu, \quad u(0) = f, \tag{7.1}$$

and five (fractional) diffusion-reaction equations of Fisher-Kolmogorov type:

$$\frac{\partial u}{\partial t} = -Lu + X(u), \quad u(0) = f, \tag{7.2}$$

for $u = u(t,x)$ defined on $[0,\infty) \times S^1$, where $S^1 \approx \mathbb{R}/(2\pi\mathbb{Z})$ is the circle. In (7.2) we take

$$X(u) = 6u(1-u), \tag{7.3}$$

and the five operators L we consider are, respectively,

$$-\frac{d^2}{dx^2}, \quad \Big(-\frac{d^2}{dx^2}\Big)^{1/2}, \quad \Big(-\frac{d^2}{dx^2}\Big)^{1/4}, \quad \psi^b_{3/2}(D), \quad \varphi^\#_{1/2}(D), \tag{7.4}$$

where $\psi^b_\alpha(\xi)$ and $\varphi^\#_\alpha(\xi)$ are given by (5.8)–(5.9). In all cases we take

$$\begin{aligned} f(x) &= 1 \quad \text{if } |x| < \frac{2\pi}{10}, \\ &0 \quad \text{otherwise,} \end{aligned} \tag{7.5}$$

and we picture $S^1 = [-\pi,\pi]$, with the endpoints identified.

To solve (7.1), we represent the solution as a Fourier multiplier, namely Fourier multiplication by $e^{-tL(\xi)}$, where $L(\xi)$ is given, respectively, by

$$\xi^2, \quad |\xi|, \quad |\xi|^{1/2}, \quad \psi^b_{3/2}(\xi), \quad \varphi^\#_{1/2}(\xi). \tag{7.6}$$

In particular, by (5.11)–(5.15), we have

$$e^{-t\psi^b_{3/2}(\xi)} = e^{-(\sqrt{2}/2)t|\xi|^{3/2}} \left[\cos\Big(\frac{\sqrt{2}}{2}t|\xi|^{3/2}\Big) + i\sigma(\xi)\sin\Big(\frac{\sqrt{2}}{2}t|\xi|^{3/2}\Big)\right], \tag{7.7}$$

and

$$e^{-t\varphi^\#_{1/2}(\xi)} = e^{-(\sqrt{2}/2)t|\xi|^{1/2}} \left[\cos\Big(\frac{\sqrt{2}}{2}t|\xi|^{1/2}\Big) + i\sigma(\xi)\sin\Big(\frac{\sqrt{2}}{2}t|\xi|^{1/2}\Big)\right]. \tag{7.8}$$

Our numerical approximation uses a 1024 point discrete Fourier transform, implemented by an FFT.

To solve (7.2) numerically, we use Strang's splitting method, a variant of (6.8) given by

$$u(t) = \lim_{n\to\infty} \big(\mathcal{F}^{t/2n} e^{-(t/n)L} \mathcal{F}^{t/2n}\big)^n f, \tag{7.9}$$

which is formally second order accurate. More precisely, we fix a time step $h = 0.001$ and take

$$u(nh) \approx \big(\mathcal{F}^{h/2} e^{-hL} \mathcal{F}^{h/2}\big)^n f, \tag{7.10}$$

for $0 \le n \le 500$, so $t = nh \in [0, 0.5]$. We evaluate e^{-hL} as above, via Fourier multiplication, and we use a difference scheme to approximate the action of $\mathcal{F}^{h/2}$.

Figures produced by these computations can be found on the author's web site. https://mtaylor.web.unc.edu/notes, item #6.

8. Inhomogeneous fractional diffusion equations

Here we consider equations of the form

$$ {}^c\partial_t^\beta u = -Au + q(t), \quad u(0) = f, \tag{8.1} $$

where A is a positive, self-adjoint operator on a Hilbert space H, $f \in H$, and $q \in C(\mathbb{R}^+, H)$. We assume $0 < \beta \le 1$. The operator ${}^c\partial_t^\beta$ is as in (1.2) if $\beta \in (0,1)$. We have the Laplace transform identity

$$ \mathcal{L}({}^c\partial_t^\beta u)(s) = s^\beta \mathcal{L}u(s) - s^{\beta-1}u(0). \tag{8.2} $$

Hence (8.1) implies

$$ \mathcal{L}u(s) = (s^\beta + A)^{-1}\mathcal{L}q(s) + s^{\beta-1}(s^\beta + A)^{-1}f. \tag{8.3} $$

Recall that the Laplace transform of $E_\beta(-t^\beta A)$ is $s^{\beta-1}(s^\beta + A)^{-1}$, with E_β as in (3.5)–(3.8). In fact, if

$$ e_\beta(t) = E_\beta(-t^\beta), \tag{8.4} $$

we have

$$ \int_0^\infty e_\beta(t)e^{-st}\,dt = \frac{s^{\beta-1}}{s^\beta + 1}. \tag{8.5} $$

It follows that

$$ \int_0^\infty e_\beta(t\gamma)e^{-st}\,dt = \frac{s^{\beta-1}}{s^\beta + \gamma^\beta}, \tag{8.6} $$

which gives

$$ \int_0^\infty E_\beta(-t^\beta A)e^{-st}\,dt = s^{\beta-1}(s^\beta + A)^{-1}. \tag{8.7} $$

We also have

$$ \begin{aligned} A^{1/\beta} &\int_0^\infty e_\beta'(tA^{1/\beta})e^{-st}\,dt \\ &= s\int_0^t e_\beta(tA^{1/\beta})e^{-st}\,dt + e_\beta(tA^{1/\beta})e^{-st}\Big|_0^\infty \\ &= s^\beta(s^\beta + A)^{-1} - 1 \\ &= -A(s^\beta + A)^{-1}. \end{aligned} \tag{8.8} $$

With this we can apply Laplace inversion to (8.3) and obtain

$$u(t) = e_\beta(tA^{1/\beta})f - A^{-1+1/\beta} \int_0^t e'_\beta(\tau A^{1/\beta}) q(t-\tau)\, d\tau, \tag{8.9}$$

using the fact that

$$g(t) = \int_0^t h(\tau) q(t-\tau)\, d\tau \Longrightarrow \mathcal{L}g(s) = \mathcal{L}h(s)\, \mathcal{L}q(s). \tag{8.10}$$

A formula equivalent to (8.9) is

$$u(t) = E_\beta(-t^\beta A)f + \beta \int_0^t \tau^{\beta-1} E'_\beta(-\tau^\beta A) q(t-\tau)\, d\tau. \tag{8.11}$$

Compare (A.30) of [MPG] for the case $A = 1$, and (7.4) of [D] for the general formula.

Recalling (3.6), we have

$$E'_\beta(-s) = \int_0^\infty M_\beta(r) r e^{-rs}\, ds, \quad s > 0, \tag{8.12}$$

with $M_\beta(r)$ given by (3.7), and also by (3.11)–(3.12). Hence (3.8), i.e.,

$$E_\beta(-t^\beta A) = \int_0^\infty M_\beta(r) e^{-rt^\beta A}\, dr, \tag{8.13}$$

is complemented by

$$E'_\beta(-t^\beta A) = \int_0^\infty M_\beta(r) r e^{-rt^\beta A}\, dr, \tag{8.14}$$

for $t > 0$, $\beta \in (0,1)$. Note that $M_\beta(r)$ and $M_\beta(r)r$ are positive and integrable on $\mathbb{R}^+$. Hence, if $\{e^{-sA} : s > 0\}$ is positivity preserving, on $H = L^2(M)$, so are the operators (8.13) and (8.14).

We desire to obtain some estimates on $E_\beta(-s)$ for $s \in \mathbb{R}^+$, hence on the operators that appear in (8.11). Of course, the formula (3.5) implies this function is smooth on $[0,\infty)$. We want to examine its asymptotic behavior as $s \nearrow +\infty$. We first tackle the behavior of $e_\beta(t)$ as $t \nearrow \infty$. The key tool for is the identity (8.5), which is valid for $\operatorname{Re} s \ge 0$. The evaluation for $s = i\xi$, $\xi \in \mathbb{R}$ gives the Fourier transform of $e_\beta(t)$ (extended to vanish on $\mathbb{R}^-$):

$$\hat{e}_\beta(\xi) = \frac{(i\xi)^{\beta-1}}{(i\xi)^\beta + 1}, \quad 0 < \beta < 1. \tag{8.15}$$

This Fourier transform identity enables us to determine the behavior of $e_\beta(t)$ as $t \nearrow \infty$, due to the (almost) classical conormal singularity of $\hat{e}_\beta$ at $\xi = 0$. We get, as $t \nearrow +\infty$,

$$e_\beta(t) \sim \sum_{k\ge 1} a_{\beta k} t^{-k\beta}, \tag{8.16}$$

$$e'_\beta(t) \sim -\sum_{k\ge 1} k\beta a_{\beta k} t^{-k\beta-1}. \tag{8.17}$$

Equivalently, as $s \nearrow +\infty$,

$$E_\beta(-s) = e_\beta(s^{1/\beta}) \sim \sum_{k\ge 1} a_{\beta k} s^{-k}, \tag{8.18}$$

$$E'_\beta(-s) = \frac{1}{\beta} s^{1/\beta-1} e'_\beta(s^{1/\beta}) \sim -\sum_{k\ge 1} k a_{\beta k} s^{-k-1}, \tag{8.19}$$

assuming $\beta \in (0,1)$. We emphasize the leading terms:

$$E_\beta(-s) \sim a_{\beta 0} s^{-1} + \cdots, \quad E'_\beta(-s) \sim -a_{\beta 0} s^{-2} + \cdots. \tag{8.20}$$

By contrast,

$$E_1(-s) = e^{-s}. \tag{8.21}$$

We now collect some operator estimates on $E_\beta(-t^\beta A)$ and $E'_\beta(-t^\beta A)$. First, suppose B is a Banach space on which e^{-tA} acts as a contraction semigroup:

$$\|e^{-tA} f\|_B \le \|f\|_B, \quad \forall\, t > 0. \tag{8.22}$$

Then (8.13)–(8.14), plus the positivity of $M_\beta(r)$ and the fact that $M_\beta(r)$ and $M_\beta(r)r$ integrate to $E_\beta(0)$ and $E'_\beta(0)$, respectively, give

$$\begin{aligned} &\|E_\beta(-t^\beta A) f\|_B \le \|f\|_B, \\ &\|E'_\beta(-t^\beta A) f\|_B \le \frac{1}{\beta\Gamma(\beta)} \|f\|_B. \end{aligned} \tag{8.23}$$

Next, assume H is a Hilbert space and A is a positive, self-adjoint operator on H. Then (8.20) plus the smoothness of $E_\beta(-s)$ on $[0,\infty)$ imply that $sE_\beta(-s)$, $sE_\beta(-s)$, and $s^2 E'_\beta(-s)$ are bounded on $[0,\infty)$, hence

$$\|t^\beta A E_\beta(-t^\beta A) f\|_H, \quad \|t^\beta A E'_\beta(-t^\beta A) f\|_H, \quad \|t^{2\beta} A^2 E'_\beta(-t^\beta A) f\|_H \le C\|f\|_H, \tag{8.24}$$

for $t \in [0,\infty)$. Interpolation with (8.23) (with $B = H$) yields further estimates, such as

$$\|\tau^{\sigma\beta} E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(H,\mathcal{D}(A^\sigma))} \le C, \quad \tau > 0, \tag{8.25}$$

given $\sigma \in (0,1)$, hence

$$\|\tau^{\beta-1}E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(H,\mathcal{D}(A^\sigma))} \le C\tau^{-1+(1-\sigma)\beta}. \tag{8.26}$$

We begin to specialize. For the rest of this section, we assume M is a compact, smooth Riemannian manifold, without boundary, and

$$A = (-\Delta)^{m/2}, \quad 0 < m \le 2, \tag{8.27}$$

where Δ is the Laplace-Beltrami operator on M. Then (8.22)–(8.23) hold for

$$B = L^p(M), \quad 1 \le p \le \infty, \quad B = C(M), \tag{8.28}$$

and (8.24)–(8.26) hold for

$$H = L^2(M), \quad \mathcal{D}(A^\sigma) = H^{\sigma m,2}(M). \tag{8.29}$$

We can go further, noting that

$$E_\beta(-s) \in S^{-1}_{1,0}([0,\infty)), \quad E'_\beta(-s) \in S^{-2}_{1,0}([0,\infty)), \tag{8.30}$$

where to say $F \in S^\mu_{1,0}([0,\infty))$ is to say $F \in C^\infty([0,\infty))$ and

$$|\partial_s^j F(s)| \le C_j \langle s\rangle^{\mu-j}, \quad \forall\, j \in \mathbb{Z}^+,\ s \in [0,\infty). \tag{8.31}$$

Now (8.27) implies

$$A \in OPS^m(M) \tag{8.32}$$

is elliptic, as well as positive and self-adjoint. Results in Chapter 12 of [T2] then imply that, given $T_0 \in (0,\infty)$,

$$\begin{aligned} &E_\beta(-t^\beta A), \quad t^\beta A E_\beta(-t^\beta A), \\ &E'_\beta(-t^\beta A), \quad t^\beta A E'_\beta(-t^\beta A), \quad t^{2\beta} A^2 E'_\beta(-t^\beta A) \\ &\text{are bounded in } \ OPS^0_{1,0}(M), \quad \text{for } \ t \in (0,T_0]. \end{aligned} \tag{8.33}$$

Boundedness of elements of $OPS^0_{1,0}(M)$ on $L^p(M)$ for $1 < p < \infty$ yield the following estimates, for such p:

$$\|E_\beta(-t^\beta A)f\|_{L^p}, \quad t^\beta \|AE_\beta(-t^\beta A)f\|_{L^p} \le C\|f\|_{L^p}, \tag{8.34}$$

and

$$\|E'_\beta(-t^\beta A)f\|_{L^p}, \ \ t^\beta\|AE'_\beta(-t^\beta A)f\|_{L^p}, \ \ t^{2\beta}\|A^2E'_\beta(-t^\beta A)f\|_{L^p} \le C\|f\|_{L^p}. \tag{8.35}$$

Then elliptic regularity yields

$$\|E_\beta(-t^\beta A)\|_{\mathcal{L}(L^p,H^{m,p})},\quad \|E'_\beta(-t^\beta A)\|_{\mathcal{L}(L^p,H^{m,p})} \le Ct^{-\beta}, \tag{8.36}$$

and

$$\|E'_\beta(t^{-\beta} A)\|_{\mathcal{L}(L^p,H^{2m,p})} \le Ct^{-2\beta}, \tag{8.37}$$

uniformly for $t \in (0, T_0]$. As in (8.25), interpolation of (8.36) with some of the estimates in (8.35) gives, for $\sigma \in (0,1)$, $p \in (1,\infty)$,

$$\|\tau^{\sigma\beta} E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(L^p,H^{\sigma m,p})} \le C, \tag{8.38}$$

hence

$$\|\tau^{\beta-1} E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(L^p,H^{\sigma m,p})} \le C\tau^{-1+(1-\sigma)\beta}, \tag{8.39}$$

uniformly for $\tau \in (0, T_0]$. We also get estimates on Zygmund spaces, such as

$$\|\tau^{\beta-1} E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(C^0_*,C^{\sigma m}_*)} \le C\tau^{-1+(1-\sigma)\beta}, \tag{8.40}$$

and similar estimates on other familite of Besov spaces.

We can produce another demonstration of (8.34)–(8.35), and extend the scope of these estimates, using (8.13)–(8.14) in concert with the following result, which for $\beta = 1/2$ and $\beta = 1/3$ is illustrated by (3.9).

Proposition 8.1. *For $\beta \in (0,1)$, the function $M_\beta(r)$ in (8.13)–(8.14) satisfies*

$$M_\beta \in \mathcal{S}([0,\infty)), \tag{8.41}$$

i.e., M_β is smooth on $[0,\infty)$ and rapidly decreasing, with all its derivatives, at infinity.

Proof. We make use of the identity (3.12),

$$\beta r^{-\beta-1} M_\beta(r^{-\beta}) = \Phi_{1,\beta}(r),$$

plus the results about $\Phi_{1,\beta}$ established at the end of §2. The fact that $\Phi_{1,\beta}(s)$ is smooth on $[0,\infty)$ and vanishes to all orders as $s \to 0$ implies M_β is smooth on $(0,\infty)$ and vanishes rapidly, with all derivatives, at ∞.

It remains to show that $M_\beta(r)$ is smooth up to $r = 0$. For this, we use the asymptotic expansion (2.10), which implies

$$M_\beta(r) = \frac{1}{\beta} r^{-1-1/\beta} \Phi_{1,\beta}(r^{-1/\beta}) \sim \frac{1}{\beta} \sum_{k\ge 1} \gamma_{\beta k} r^{k-1},$$

as $r \searrow 0$.

We can exploit Proposition 8.1 as follows. Given (8.41), we can write

$$\begin{aligned} sE_\beta(-s) &= -\int_0^\infty M_\beta(r)\frac{\partial}{\partial r}e^{-rs}\,dr \\ &= \int_0^\infty M_\beta'(r)e^{-rs}\,dr + M_\beta(0), \end{aligned} \tag{8.42}$$

and deduce that, whenever B is a Banach space such that (8.22) holds, or more generally

$$\|e^{-tA}f\|_B \le C\|f\|_B, \quad \forall\, t > 0, \tag{8.43}$$

then

$$\|t^\beta A E_\beta(-t^\beta A)f\|_B \le C\|f\|_B. \tag{8.44}$$

Similarly,

$$\|t^\beta A E_\beta'(-t^\beta A)f\|_B, \ \ \|t^{2\beta}A^2 E_\beta'(-t^\beta A)f\|_B \le C\|f\|_B. \tag{8.45}$$

As advertised, this provides another demonstration of (8.34)–(8.35), and extends the scope of these estimates.

9. Fractional diffusion-reaction equations – local existence

Here we study the initial-value problem

$${}^c\partial_t^\beta u = -Au + F(u), \quad u(0) = f, \tag{9.1}$$

on $[0, T_0] \times M$, given a suitable f on M (perhaps with values in $\mathbb{R}^k$). We assume $\beta \in (0, 1)$. As in the end of §8, we assume M is a compact Riemannian manifold, and

$$A = (-\Delta)^{m/2}, \quad 0 < m \le 2, \tag{9.2}$$

where Δ is the Laplace-Beltrami operator on M.

Using (8.11), we rewrite (9.1) as

$$u(t) = E_\beta(-t^\beta A)f + \beta\int_0^t \tau^{\beta-1}E_\beta'(-\tau^\beta A)F(u(t-\tau))\,d\tau. \tag{9.3}$$

Hence we desire to solve

$$\Phi u = u, \tag{9.4}$$

where

$$\Phi u(t) = E_\beta(-t^\beta A)f + \beta \int_0^t \tau^{\beta-1} E'_\beta(-\tau^\beta A) F(u(t-\tau))\, d\tau. \tag{9.5}$$

Thus we seek a fixed point of

$$\Phi : \mathfrak{X} \longrightarrow \mathfrak{X}, \tag{9.6}$$

where $\mathfrak{X}$ is a suitably chosen complete metric space.

To begin, we assume $f \in C(M)$. We pick $a \in (0,\infty)$ and set

$$\mathfrak{X} = \{u \in C(I, C(M)) : u(0) = f,\ \sup_{t\in I} \|u(t) - f\|_{L^\infty} \le a\}, \quad I = [0,\delta], \tag{9.7}$$

where $\delta > 0$ will be specified below. We assume u takes values in $\mathbb{R}^k,\ F : \mathbb{R}^k \to \mathbb{R}^k$, and

$$u \in \mathbb{R}^k,\ |u| \le A \Longrightarrow |F(u)| \le K,\ |DF(u)| \le L. \tag{9.8}$$

Here $|\cdot|$ denotes some convenient norm on $\mathbb{R}^k$ and also the associated operator norm on $\mathrm{End}(\mathbb{R}^k)$. Now $t \mapsto E_\beta(-t^\beta A)$ is strongly continuous on $C(M)$ (by (8.13)), and $E_\beta(0) = I$, so we can pick $\delta > 0$ so small that

$$t \in (0,\delta] \Longrightarrow \|E_\beta(-t^\beta A)f - f\|_{L^\infty} \le \frac{1}{2}a. \tag{9.9}$$

To get $\Phi : \mathfrak{X} \to \mathfrak{X}$, it suffices to ensure that

$$t \in I,\ u \in \mathfrak{X} \Rightarrow \beta \int_0^t \tau^{\beta-1} \|E'_\beta(-\tau^\beta A) F(u(t-\tau))\|_{L^\infty}\, d\tau \le \frac{1}{2}a. \tag{9.10}$$

By (9.8), $u \in \mathfrak{X} \Rightarrow \|F(u(t-\tau))\|_{L^\infty} \le K$. Then (8.23), with $B = C(M)$, implies $\|E'_\beta(-\tau^\beta A)F(u(t-\tau))\|_{L^\infty} \le K/\beta\Gamma(\beta)$, so (9.10) holds provided

$$t \in I \Longrightarrow \frac{K}{\Gamma(\beta)} \int_0^t \tau^{\beta-1}\, d\tau \le \frac{a}{2}, \tag{9.11}$$

i.e., provided

$$\delta^\beta \le \frac{\beta\Gamma(\beta)}{2}\, \frac{a}{K}. \tag{9.12}$$

Hence $\Phi : \mathfrak{X} \to \mathfrak{X}$ whenever (9.9) and (9.12) hold.

We next produce a condition that guarantees Φ is a contraction on $\mathfrak{X}$. Given $u, v \in \mathfrak{X}$, $t \in I$, we have

$$(9.13)\quad \|\Phi u(t) - \Phi v(t)\|_{L^\infty} \le \beta \int_0^t \tau^{\beta-1} \big\| E_\beta'(-\tau^\beta A)[F(u(t-\tau) - F(v(t-\tau))] \big\|_{L^\infty} \, d\tau.$$

Now

$$(9.14)\qquad F(u) - F(v) = \int_0^1 \frac{d}{ds} F(su + (1-s)v)\, ds = G(u,v)(u-v),$$

with

$$(9.15)\qquad G(u,v) = \int_0^1 DF(su + (1-s)v)\, ds.$$

Hence $t \in I$, $\tau \in [0,t]$ imply, via (9.8),

$$(9.16)\qquad \|F(u(t-\tau)) - F(v(t-\tau))\|_{L^\infty} \le L \|u(t-\tau) - v(t-\tau)\|_{L^\infty},$$

so, again by (8.23), the right side of (9.13) is bounded by

$$(9.17)\qquad \begin{aligned} &\frac{L}{\Gamma(\beta)} \int_0^t \tau^{\beta-1} \|u(t-\tau) - v(t-\tau)\|_{L^\infty}\, d\tau \\ &\qquad \le \frac{L}{\beta\Gamma(\beta)} \sup_{0\le\tau\le t} \|u(\tau) - v(\tau)\|_{L^\infty}\, t^\beta. \end{aligned}$$

Thus we get

$$(9.18)\qquad \sup_{t\in I} \|\Phi u(t) - \Phi v(t)\|_{L^\infty} \le \theta \sup_{t\in I} \|u(t) - v(t)\|_{L^\infty},$$

provided

$$(9.19)\qquad \delta^\beta \le \beta\Gamma(\beta)\, \frac{\theta}{L}.$$

Hence, as long as δ satisfies (9.9), (9.12), and (9.18), with $\theta \in (0,1)$, Φ is a contraction on $\mathfrak{X}$, given by (9.7). We record the local existence result.

Proposition 9.1. *Assume M is a compact Riemannian manifold and A is given by (9.2). Assume $F : \mathbb{R}^k \to \mathbb{R}^k$ satisfies (9.8). Take $f \in C(M)$. Then (9.1) has a solution in $C([0,\delta], C(M))$ provided $\delta > 0$ satisfies (9.9), (9.12), and (9.18), with $\theta < 1$.*

We look at situations with more singular initial data. Not to get too general, we assume

$$f \in L^6(M). \tag{9.20}$$

The analysis will be dimension dependent; say

$$\dim M = n. \tag{9.21}$$

We again take $a \in (0, \infty)$ and set

$$\mathfrak{X} = \{u \in C(I, L^6(M)) : u(0) = f,\ \sup_{t \in I} \|u(t) - f\|_{L^6} \le a\}, \quad I = [0, \delta], \tag{9.22}$$

with $\delta > 0$ to be specified below. This time, we assume

$$|F(u)| \le K(1 + |u|^3), \quad |DF(u)| \le L(1 + |u|^2), \tag{9.23}$$

which holds if $F(u)$ is a cubic polynomial in u. Again Φ is given by (9.5). We desire to show that if $\delta > 0$ is small enough, $\Phi : \mathfrak{X} \to \mathfrak{X}$ and is a contraction. We will succeed in case

$$\frac{n}{3} < m \le 2, \tag{9.24}$$

with m as in (9.2). Note that this requires $n \le 5$.

To start, $t \mapsto E_\beta(-t^\beta A)$ is strongly continuous on $L^6(M)$, again by (8.13), so we can pick $\delta > 0$ so small that

$$t \in (0, \delta] \Longrightarrow \|E_\beta(-t^\beta A) f - f\|_{L^6} \le \frac{a}{2}. \tag{9.25}$$

To get $\Phi : \mathfrak{X} \to \mathfrak{X}$, it suffices to show that

$$t \in I,\ u \in \mathfrak{X} \Rightarrow \beta \int_0^t \tau^{\beta-1} \|E_\beta'(-\tau^\beta A) F(u(t-\tau))\|_{L^6}\, d\tau \le \frac{a}{2}. \tag{9.26}$$

By (9.23),

$$u \in \mathfrak{X} \Longrightarrow \|F(u(t-\tau))\|_{L^2} \le C(a, K). \tag{9.27}$$

The estimate (8.26), with $H = L^2(M)$ (or (8.39), with $p = 2$) gives

$$\tau^{\beta-1} \|E'(-\tau^\beta A) F(u(t-\tau))\|_{H^{\sigma m, 2}} \le C \tau^{-1+(1-\sigma)\beta}, \tag{9.28}$$

for $\sigma \in (0, 1)$. Sobolev embedding theorems give

$$H^{\sigma m, 2}(M) \subset L^6(M), \quad \text{for some } \ \sigma < 1, \tag{9.29}$$

provided (9.24) holds. We mention parenthetically that $H^{\sigma m,2}(M) \subset L^\infty(M)$ for some $\sigma < 1$ provided $n/2 < m \leq 2$. Consequently, if (9.24) holds, we have the integral in (9.26) bounded by

$$Ct^{(1-\sigma)\beta}, \tag{9.30}$$

which is $\leq a/2$ for all $t \in (0, \delta]$ if δ is small enough. This gives $\Phi : \mathfrak{X} \to \mathfrak{X}$.

We next want to show that Φ is a contraction on $\mathfrak{X}$ if $\delta > 0$ is small enough. This would follow if we could show that, for $u, v \in \mathfrak{X}$,

$$\begin{aligned} \beta \int_0^t \tau^{\beta-1} \|E'_\beta(-\tau^\beta A)[F(u(t-\tau)) - F(v(t-\tau))]\|_{L^6}\, d\tau \\ \leq Ct^{(1-\sigma)\beta} \sup_{0\leq\tau\leq t} \|u(\tau) - v(\tau)\|_{L^6}, \end{aligned} \tag{9.31}$$

since this would yield

$$\sup_{t\in I} \|\Phi u(t) - \Phi v(t)\|_{L^6} \leq \theta \sup_{t\in I} \|u(t) - v(t)\|_{L^6}, \tag{9.32}$$

for $u, v \in \mathfrak{X}$, for some $\theta < 1$, if $I = [0, \delta]$ and $\delta > 0$ is small enough.

To proceed, with notation as in (9.14)–(9.15), we have, for $t, \tau \in I$, $u = u(t-\tau)$, $v = v(t-\tau)$, elements of $\mathfrak{X}$,

$$\begin{aligned} \|F(u) - F(v)\|_{L^2} &= \|G(u, v)(u - v)\|_{L^2} \\ &\leq \|G(u, v)\|_{L^3} \|u - v\|_{L^6} \\ &\leq C(a)\|u - v\|_{L^6}, \end{aligned} \tag{9.33}$$

the last inequality by the hypothesis (9.23) on $DF(u)$. Hence the left side of (9.31) is

$$\leq C(A)\beta \int_0^t \tau^{\beta-1} \|E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(L^2,L^6)} \|u(t-\tau) - v(t-\tau)\|_{L^6}\, d\tau. \tag{9.34}$$

Via (8.26) or (8.39), plus (9.29), we have

$$\tau^{\beta-1} \|E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(L^2,L^6)} \leq C\tau^{-1+(1-\sigma)\beta}, \tag{9.35}$$

provided (9.24) holds. Hence (9.34)–(9.35) yield the desired estimate (9.31), and we have the contraction property. We record the result.

Proposition 9.2. *Assume M is a compact Riemannian manifold of dimension n, A is given by (9.2), and m satisfies (9.24). Let $F : \mathbb{R}^k \to \mathbb{R}^k$ satisfy (9.23). Take $f \in L^6(M)$. Then (9.1) has a solution $u \in C([0, \delta], L^6(M))$ provided $\delta > 0$ is sufficiently small.*

10. More local existence results

Here we seek other complete metric spaces $\mathfrak{X}$ for which $\Phi : \mathfrak{X} \to \mathfrak{X}$ is a contraction, given Φ as in (9.5), i.e.,

$$\Phi u(t) = E\beta(-t^\beta A)f + \beta \int_0^t \tau^{\beta-1} E'_\beta(-\tau^\beta A)F(u(t-\tau))\, d\tau. \tag{10.1}$$

We continue to assume $\beta \in (0,1),\ A = (-\Delta)^{m/2},\ m \in (0,2]$, and $F : \mathbb{R}^k \to \mathbb{R}^k$ satisfies (9.23), i.e.,

$$|F(u)| \le K(1+|u|^3), \quad |DF(u)| \le L(1+|u|^2), \tag{10.2}$$

which holds if F is a cubic polnomial in u. We also continue to assume M is a compact Riemannian manifold of dimension n. Generalizing (9.20)–(9.22), we pick $q \in (1,\infty),\ a \in (0,\infty)$,

$$f \in L^{3q}(M), \tag{10.3}$$

and set

$$\mathfrak{X} = \{u \in C(I, L^{3q}(M)) : u(0) = f,\ \sup_{t\in I} \|u(t) - f\|_{L^{3q}} \le a\}, \quad I = [0,\delta], \tag{10.4}$$

with $\delta > 0$ to be specified.

Parallel to (9.25), since $t \mapsto E_\beta(-t^\beta A)$ is strongly continuous on $L^{3q}(M)$, we can pick $\delta > 0$ so small that

$$t \in (0,\delta] \Longrightarrow \|E_\beta(-t^\beta A)f - f\|_{L^{3q}} \le \frac{a}{2}. \tag{10.5}$$

To get $\Phi : \mathfrak{X} \to \mathfrak{X}$, it suffices to show that

$$t \in I,\ u \in \mathfrak{X} \Rightarrow \beta \int_0^t \tau^{\beta-1} \|E'_\beta(-\tau^\beta A)F(u(t-\tau))\|_{L^{3q}}\, d\tau \le \frac{a}{2}. \tag{10.6}$$

By (10.2),

$$u \in \mathfrak{X} \Longrightarrow \|F(u(t-\tau))\|_{L^q} \le C(a,K). \tag{10.7}$$

The estimate (8.39) gives

$$\tau^{\beta-1} \|E'_\beta(-\tau^\beta A)F(u(t-\tau))\|_{H^{\sigma m,q}} \le C\tau^{-1+(1-\sigma)\beta}, \tag{10.8}$$

for $\sigma \in (0,1)$. We seek a condition implying

$$H^{\sigma m,q}(M) \subset L^{3q}(M). \tag{10.9}$$

for some $\sigma \in (0,1)$. If $n = \dim M$, Sobolev embedding results imply

$$\begin{aligned} H^{\sigma m,q}(M) \subset L^\infty(M), &\quad \text{for some } \sigma < 1, \text{ if } \ mq > n, \\ L^{nq/(n-\sigma mq)}(M), &\quad \text{if } \ mq \le n. \end{aligned} \tag{10.10}$$

Thus (10.9) holds provided either $mq > n$ or $mq \le n$ and $nq/(n-\sigma mq) \ge 3q$ for some $\sigma \in (0,1)$. Hence (10.9) holds provided

$$3q > \frac{2n}{m}. \tag{10.11}$$

As for how this constrains m, recalling that $m \le 2$, we require

$$\frac{2n}{3q} < m \le 2. \tag{10.12}$$

This requires $n < 3q$. For $q = 2,\ 3q = 6$, this is (9.24). If (10.9) holds, (10.8) yields

$$\tau^{\beta-1} \|E'_\beta(-\tau^\beta A) F(u(t-\tau))\|_{L^{3q}} \le C\tau^{-1+(1-\sigma)\beta}, \tag{10.13}$$

and we get (10.6), as long as $\delta > 0$ is small enough. Hence $\Phi : \mathfrak{X} \to \mathfrak{X}$.

We next want to show that Φ is a contraction on $\mathfrak{X}$ if $\delta > 0$ is small enough. Parallel to (9.31), this would follow if we could show that, for $u, v \in \mathfrak{X}$,

$$\begin{aligned} \beta \int_0^t \tau^{\beta-1} \|E'_\beta(-\tau^\beta A)[F(u(t-\tau)) - F(v(t-\tau))]\|_{L^{3q}}\, d\tau \\ \le C t^{(1-\sigma)\beta} \sup_{0\le\tau\le t} \|u(\tau) - v(\tau)\|_{L^{3q}}. \end{aligned} \tag{10.14}$$

To proceed, with notation as in (9.14)–(9.15), and parallel to (9.33), we have, for $t, \tau \in I,\ u = u(t-\tau),\ v = v(t-\tau)$, elements of $\mathfrak{X}$,

$$\begin{aligned} \|F(u) - F(v)\|_{L^q} &= \|G(u,v)(u-v)\|_{L^q} \\ &\le \|G(u,v)\|_{L^{3q/2}} \|u - v\|_{L^{3q}} \\ &\le C(a) \|u - v\|_{L^{3q}}, \end{aligned} \tag{10.15}$$

the last inequality by the hypothesis (10.2) on $DF(u)$. Hence the left side of (10.14) is

$$\le C(a)\beta \int_0^t \tau^{\beta-1} \|E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(L^q, L^{3q})} \|u(t-\tau) - v(t-\tau)\|_{L^{3q}}\, d\tau. \tag{10.16}$$

Via (10.8)–(10.10), we have

$$\tau^{\beta-1} \|E'_\beta(-\tau^\beta A)\|_{\mathcal{L}(L^q, L^{3q})} \le C\tau^{-1+(1-\sigma)\beta}, \tag{10.17}$$

provided (10.12) holds. Hence (10.16)–(10.17) yield the desired estimate (10.14), and we have the contraction property. We record the result.

Proposition 10.1. *Assume M is a compact, n-dimensional, Riemannian manifold, $A = (-\Delta)^{m/2}$, $f \in L^{3q}(M)$, and m and q satisfy $q > 1$ and (10.12). Let $F : \mathbb{R}^k \to \mathbb{R}^k$ satisfy (10.2). Then (9.1) has a solution $u \in C([0, \delta], L^{3q}(M))$ provided $\delta > 0$ is small enough.*

11. Further variants

Let us write the putative solution of (9.1) as

$$u(t) = u_0(t) + v(t), \quad u_0(t) = E_\beta(-t^\beta A)f. \tag{11.1}$$

Then the integral equation (9.3) is equivalent to

$$v(t) = \beta \int_0^t \tau^{\beta-1} E'_\beta(-\tau^\beta A) F(u_0(t-\tau) + v(t-\tau))\, d\tau, \tag{11.2}$$

or

$$\Psi v = v, \tag{11.3}$$

where

$$\Psi v(t) = \beta \int_0^t \tau^{\beta-1} E'_\beta(-\tau^\beta A) F(u_0(t-\tau) + v(t-\tau))\, d\tau. \tag{11.4}$$

Thus we seek a complete metric space $\mathfrak{Z}$ for which

$$\Psi : \mathfrak{Z} \longrightarrow \mathfrak{Z} \tag{11.5}$$

is a contraction.

For example, picking $q \in (1, \infty)$, $a \in (0, \infty)$, we can take

$$\mathfrak{Z} = \{v \in C(I, L^{3q}(M)) : v(0) = 0, \sup_{t \in I} \|v(t)\|_{L^{3q}} \le a\}, \quad I = [0, \delta]. \tag{11.6}$$

We assume F satisfies (10.2). We assume $f \in L^{3q}(M)$, so $u_0 \in C(I, L^{3q}(M))$. Estimates parallel to those given in §10 show that if (10.12) holds, then, for $\delta > 0$ small enough, (11.5) holds and Ψ is a contraction.

In a search for other candidates for the space $\mathfrak{Z}$, we investigate the behavior of $v_1 = \Psi 0$, i.e., of

$$v_1(t) = \beta \int_0^t \tau^{\beta-1} E'_\beta(-\tau^\beta A) F(u_0(t-\tau))\, d\tau. \tag{11.7}$$

To start, let us take

$$n = \dim M = 2, \quad f \in L^2(M). \tag{11.8}$$

Then, for $\sigma \in (0,1]$,

$$\|u_0(t-\tau)\|_{H^{\sigma m,2}} \le C(t-\tau)^{-\sigma\beta}. \tag{11.9}$$

We have

$$\begin{aligned} H^{\sigma m,2}(M) \subset L^\infty(M), &\quad \text{if } \ \sigma m > 1, \\ L^{4/(2-2\sigma m)}, &\quad \text{if } \ \sigma m < 1. \end{aligned} \tag{11.10}$$

In particular, $4/(2-2\sigma m) = 6$ if $\sigma m = 2/3$, so

$$\|u_0(t-\tau)\|_{L^6} \le C(t-\tau)^{-\sigma\beta} \ \text{ if } \ \sigma m \ge \frac{2}{3}, \tag{11.11}$$

hence

$$\|F(u_0(t-\tau))\|_{L^2} \le C(t-\tau)^{-3\sigma\beta}, \tag{11.12}$$

for $0 < \tau < t \le T_0$, if $\sigma m \ge 2/3$, while also $\sigma \le 1$, i.e., if

$$\frac{2}{3m} \le \sigma \le 1, \tag{11.13}$$

which is possible provided

$$\frac{2}{3} \le m \le 2. \tag{11.14}$$

In such a case,

$$\|v_1(t)\|_{L^2} \le C \int_0^t \tau^{\beta-1} (t-\tau)^{-3\sigma\beta}\, d\tau, \tag{11.15}$$

which is finite provided

$$3\sigma\beta < 1. \tag{11.16}$$

This is consistent with (11.13) if

$$\frac{2}{m}\beta < 1, \quad \text{i.e.,} \ \ 2\beta < m, \quad \text{or} \ \ \beta < \frac{m}{2}. \tag{11.17}$$

In such a case we can take

$$\sigma = \frac{2}{3m}, \quad \text{so} \;\; 3\sigma\beta = \frac{2\beta}{m}, \tag{11.18}$$

and (11.15) yields

$$\begin{aligned} \|v_1(t)\|_{L^2} &\le C \int_0^t \tau^{\beta-1}(t-\tau)^{-2\beta/m}\, d\tau \\ &= \widetilde{C} t^{-(2-m)\beta/m}. \end{aligned} \tag{11.19}$$

In particular,

$$\|v_1(t)\|_{L^2} \le \widetilde{C} \;\; \text{if} \;\; m = 2. \tag{11.20}$$

So let's assume

$$n = 2, \quad f \in L^2(M), \quad m = 2, \quad \sigma = \frac{1}{3}, \quad \beta \in (0,1). \tag{11.21}$$

In such a case, we have the conclusion (11.20). Under the hypotheses of (11.21), let us pick $a, b \in (0,\infty)$ and set

$$\begin{aligned} \mathfrak{Z} = \{v \in C(I, L^2(M)) : v(0) = 0,\ \|v(t)\|_{L^2} \le a, \\ \|v(t)\|_{L^6} \le b t^{-\sigma\beta},\ \forall\, t \in I\}, \end{aligned} \tag{11.22}$$

with $I = [0, \delta]$. Then

$$\begin{aligned} v \in \mathfrak{Z} &\Rightarrow \|u_0(t-\tau) + v(t-\tau)\|_{L^6} \le C(t-\tau)^{-\sigma\beta} \\ &\Rightarrow \|F(u_0(t-\tau) + v(t-\tau))\|_{L^2} \le C(t-\tau)^{-3\sigma\beta} \\ &\Rightarrow \|\Psi v(t)||_{L^2} \le C \int_0^t \tau^{\beta-1}(t-\tau)^{-3\sigma\beta}\, d\tau = \widetilde{C}. \end{aligned} \tag{11.23}$$

However, we cannot guarantee that $\widetilde{C} \le a$, even if we shrink I.

Nevertheless, we proceed to estimate $\|\Psi v(t)\|_{L^6}$. We have

$$\|\tau^{\beta-1} E_\beta'(-\tau^\beta A) F\|_{H^{\sigma m,2}} \le C\tau^{-1+(1-\sigma)\beta} \|F\|_{L^2}. \tag{11.24}$$

Hence, from the L^2 estimate of F in (11.23), if $v \in \mathfrak{Z}$,

$$\begin{aligned} \tau^{\beta-1} \|E_\beta'(-\tau^\beta A) F(u_0(t-\tau) + v(t-\tau))\|_{H^{\sigma m,2}} \\ \le C\tau^{-1+(1-\sigma)\beta}(t-\tau)^{-3\sigma\beta}, \end{aligned} \tag{11.25}$$

under hypothesis (11.21), hence

$$\begin{aligned}
\|\Psi v(t)\|_{L^6} &\le C\|\Psi v(t)\|_{H^{\sigma m,2}} \\
&\le C\int_0^t \tau^{-1+(1-\sigma)\beta}(t-\tau)^{-3\sigma\beta}\,d\tau \qquad (11.26)\\
&= \widetilde{C}t^{-\beta/3}.
\end{aligned}$$

Again we get an estimate of $\widetilde{C}t^{-\sigma\beta}$, since $\sigma = 1/3$, but we cannot establish that $\widetilde{C} \le b$. In other words, the hypothesis (11.21) seems to be of "critical" type.

We will try again, with the hypothesis $f \in L^2(M)$ replaced by

$$f \in L^p(M), \quad \text{for some} \ \ p > 2. \tag{11.27}$$

We already know that things work out if

$$p = 3q > \frac{2n}{m} = 2, \ \text{when} \ \ n = m = 2, \ \text{provided also} \ \ q > 1, \ \text{i.e.,} \ \ p > 3. \tag{11.27A}$$

Now we want to take p closer to 2, when $n = m = 2$. We need further estimates on $v_1(t)$, in order to set up a replacement for the space (11.22).

To start, we need an estimate on

$$\|u_0(t-\tau)\|_{L^{3p}}, \tag{11.28}$$

parallel to that in (11.11). Parallel to (11.9), we have

$$\|u_0(t-\tau)\|_{H^{\sigma m,p}} \le C(t-\tau)^{-\sigma\beta}, \tag{11.29}$$

and, parallel to (11.10), we have (when $n = 2$)

$$\begin{aligned}
H^{\sigma m,p}(M) \subset L^\infty(M), &\quad \text{if} \ \ \sigma m > \frac{2}{p}, \\
L^{2p/(2-\sigma m p)}, &\quad \text{if} \ \ \sigma m < \frac{2}{p}.
\end{aligned} \tag{11.30}$$

In particular, $2p/(2-\sigma m p) = 3p$ if $\sigma m = 4/3p$, so

$$\|u_0(t-\tau)\|_{L^{3p}} \le C(t-\tau)^{-\sigma\beta} \ \ \text{if} \ \ \sigma m \ge \frac{4}{3p}, \tag{11.31}$$

hence

$$\|F(u_0(t-\tau))\|_{L^p} \le C(t-\tau)^{-3\sigma\beta}, \tag{11.32}$$

for $0 < \tau < t \le T_0$, if $\sigma m \ge 4/3p$, while also $\sigma \le 1$, i.e., if

$$\frac{4}{3pm} \le \sigma \le 1, \tag{11.33}$$

or, assuming $m = 2$, if

$$\frac{2}{3p} \le \sigma \le 1, \tag{11.34}$$

which of course is true if $p > 2$, so we can take

$$\sigma = \frac{2}{3p}, \quad \text{so} \ \ 3\sigma\beta = \frac{2\beta}{p}, \tag{11.35}$$

and we have

$$\begin{aligned} \|v_1(t)\|_{L^p} &\le C \int_0^t \tau^{\beta-1} (t-\tau)^{-2\beta/p} \, d\tau \\ &= \widetilde{C} t^{\beta(1-2/p)}. \end{aligned} \tag{11.36}$$

Also (11.32) and the analogue of (11.25) with $H^{\sigma m,2}$ replaced by $H^{\sigma m,p}$, give

$$\begin{aligned} \|v_1(t)\|_{H^{\sigma m,p}} &\le C \int_0^t \tau^{-1+(1-\sigma)\beta} (t-\tau)^{-2\beta/p} \, d\tau \\ &= \widetilde{C} t^{-\beta(8/3p-1)} \\ &= \widetilde{C} t^{-\beta(4\sigma-1)}. \end{aligned} \tag{11.37}$$

Compare (11.29). Note that $4\sigma - 1 < \sigma \Leftrightarrow \sigma < 1/3$, which by (11.35) holds if $p > 2$. Hence $\|v_1(t)\|_{H^{\sigma m,p}}$ has a gentler blow-up as $t \searrow 0$ than $\|u_0(t)\|_{H^{\sigma m,p}}$ does (given $m = 2$).

In light of these observations, under hypothesis (11.27), plus

$$n = m = 2, \tag{11.38}$$

and with σ as in (11.35), it is natural to take $a, b \in (0, \infty)$, and set

$$\begin{aligned} \mathfrak{Z} = \{ v \in C(I, L^p(M)) : v(0) = 0, \ \|v(t)\|_{L^p} \le a, \\ \|v(t)\|_{L^{3p}} \le b t^{-\sigma\beta}, \ \forall \, t \in I \}, \end{aligned} \tag{11.39}$$

with $I = [0, \delta]$. We desire to show that, for $\delta > 0$ small enough, Ψ, given by (11.4), maps $\mathfrak{Z}$ to itself, as a contraction.

To start, under the hypotheses (11.27) and (11.38), and taking σ as in (11.35), we have

$$\begin{aligned} v \in \mathfrak{Z} &\Rightarrow \|u_0(t-\tau) + v(t-\tau)\|_{L^{3p}} \le C(t-\tau)^{-\sigma\beta} \\ &\Rightarrow \|F(u_0(t-\tau) + v(t-\tau))\|_{L^p} \le C(t-\tau)^{-3\sigma\beta} \\ &\Rightarrow \|\Psi v(t)\|_{L^p} \le C \int_0^t \tau^{\beta-1}(t-\tau)^{-2\beta/p}\, d\tau = \widetilde{C} t^{\beta(1-2/p)}. \end{aligned} \tag{11.40}$$

We require of δ that

$$\widetilde{C}\delta^{\beta(1-2/p)} \le a, \tag{11.41}$$

which is possible since $p > 2$.

Next we estimate $\|\Psi v(t)\|_{H^{\sigma m,p}}$, which leads to an estimate of $\|\Psi v(t)\|_{L^{3p}}$. We have

$$\tau^{\beta-1}\|E_\beta'(-\tau^\beta A)F\|_{H^{\sigma m,p}} \le C\tau^{-1+(1-\sigma)\beta}\|F\|_{L^p}, \tag{11.42}$$

hence, from the L^p estimates of F in (11.40), if $v \in \mathfrak{Z}$,

$$\begin{aligned} \tau^{\beta-1}\|E_\beta'(-\tau^\beta A)F(u_0(t-\tau) + v(t-\tau))\|_{H^{\sigma m,p}} \\ \le C\tau^{-1+(1-\sigma)\beta}(t-\tau)^{-3\sigma\beta}, \end{aligned} \tag{11.43}$$

under hypotheses (11.27) and (11.38). Hence, bringing in (11.35),

$$\begin{aligned} \|\Psi v(t)\|_{L^{3p}} &\le C\|\Psi v(t)\|_{H^{\sigma m,p}} \\ &\le C\int_0^t \tau^{-1+(1-\sigma)\beta}(t-\tau)^{-2\beta/p}\, d\tau \\ &= \widetilde{C} t^{-\beta(4\sigma-1)}, \end{aligned} \tag{11.44}$$

parallel to (11.37). We require of δ that

$$\widetilde{C}\delta^{-\beta(4\sigma-1)} \le b\delta^{-\beta\sigma}, \tag{11.45}$$

which is possible since $4\sigma - 1 < \sigma$. Then $\Psi : \mathfrak{Z} \to \mathfrak{Z}$.

Similar estimates show that, with δ perhaps further shrunk, Ψ is a contraction on $\mathfrak{Z}$. We omit the details. We record the resulting existence theorem.

Proposition 11.1. *Let M be a compact, 2-dimensional Riemannian manifold, $A = -\Delta$, and $\beta \in (0,1)$. Assume F satisfies (10.2). Assume $f \in L^p(M)$ for some $p > 2$. Then, for some $\delta > 0$, the initial value problem (9.1) has a unique solution $u \in C(I, L^p(M))$ of the form $u = u_0 + v$, as in (11.1), such that v belongs to $\mathfrak{Z}$, given by (11.39), with $\sigma = 2/3p$. Furthermore,*

$$\|v(t)\|_{H^{2\sigma,p}} \le Ct^{-\beta(4\sigma-1)}. \tag{11.46}$$

NOTE. For $n = 2,\ m = 2$, Proposition 10.1 requires $p = 3q > 3$, so Proposition 11.1 is an improvement.

A. Riemann-Liouville fractional integrals and Caputo fractional derivatives

For $\beta > 0$, the Riemann-Liouville fractional integral J^β is defined by

$$J^\beta f(t) = \frac{1}{\Gamma(\beta)} \int_0^t (t-\tau)^{\beta-1} f(\tau)\, d\tau, \tag{A.1}$$

for $t \geq 0$, where f is a suitable function on $[0,\infty)$, say continuous on $[0,\infty)$ and polynomially bounded. We mention that

$$J^\beta 1(t) = \frac{1}{\Gamma(\beta+1)} t_+^\beta. \tag{A.2}$$

With the Laplace transform given by

$$\mathcal{L}f(s) = \int_0^\infty f(t) e^{-st}\, dt, \quad \operatorname{Re} s > 0, \tag{A.3}$$

we have

$$u(s) = \int_0^t g(t-\tau) f(\tau)\, d\tau \Longrightarrow \mathcal{L}u(s) = \mathcal{L}g(s)\, \mathcal{L}f(s), \tag{A.4}$$

and

$$g_\beta(t) = t_+^{\beta-1},\ \beta > 0 \Longrightarrow \mathcal{L}g_\beta(s) = \Gamma(\beta) s^{-\beta}. \tag{A.5}$$

Hence

$$\mathcal{L}J^\beta f(s) = s^{-\beta} \mathcal{L}f(s). \tag{A.6}$$

For $\beta \in (0,1)$, the Riemann-Liouville fractional derivative is given by

$${}^r\partial_t^\beta f = \partial_t J^{1-\beta} f, \tag{A.7}$$

and the Caputo fractional derivative is given by

$${}^c\partial_t^\beta f = J^{1-\beta} \partial_t f. \tag{A.8}$$

One has

$$ {}^r\partial_t^\beta J^\beta f = f \quad \text{and} \quad {}^c\partial_t^\beta J^\beta f = f. \tag{A.9} $$

However, ${}^r\partial_t^\beta$ and ${}^c\partial_t^\beta$ are not identical. For example, given $\beta \in (0,1)$,

$$ {}^c\partial_t^\beta 1 \equiv 0, \quad {}^r\partial_t^\beta 1 = \frac{1}{\Gamma(\beta)} t_+^{\beta-1}. \tag{A.10} $$

We next consider how the Laplace transform interacts with these two fractional derivatives. Note that

$$ \begin{aligned} \mathcal{L}\partial_t f(s) &= \int_0^\infty f'(t)e^{-s}\,dt \\ &= s\mathcal{L}f(s) - f(0), \end{aligned} \tag{A.11} $$

the last identity by integration by parts. It follows that, for $\beta \subset (0,1)$,

$$ \mathcal{L}\,{}^r\partial_t^\beta f(s) = s^\beta \mathcal{L}f(s) - J^{1-\beta}f(0), \tag{A.12} $$

and

$$ \mathcal{L}\,{}^c\partial_t^\beta f(s) = s^\beta \mathcal{L}f(s) - s^{\beta-1}f(0). \tag{A.13} $$

Consequently, one can apply Laplace transform techniques conveniently to initial value problems for fractional differential equations involving the Caputo fractional derivative ${}^c\partial_t^\beta$, but not so well for those involving the Riemann-Liouville fractional derivative ${}^r\partial_t^\beta$.

For application in Appendix C, we compute ${}^c\partial_t^\beta t^\gamma$, for $\beta \in (0,1)$, $\gamma \geq \beta$. We have

$$ \begin{aligned} {}^c\partial_t^\beta t^\gamma &= J^{1-\beta}\partial_t t^\gamma \\ &= \gamma J^{1-\beta} t^{\gamma-1} \\ &= \gamma\Gamma(\gamma) J^{1-\beta} J^{\gamma-1} 1(t), \end{aligned} \tag{A.14} $$

the last identity by (A.2). Now (A.6) implies $J^{1-\beta}J^{\gamma-1} = J^{\gamma-\beta}$, so

$$ \begin{aligned} {}^c\partial_t^\beta t^\gamma &= \gamma\Gamma(\gamma) J^{\gamma-\beta} 1(t) \\ &= \frac{\Gamma(\gamma+1)}{\Gamma(\gamma-\beta+1)} t^{\gamma-\beta}, \end{aligned} \tag{A.15} $$

invoking (A.2) again. In particular, for $k \in \mathbb{N}$,

$$ {}^c\partial_t^\beta t^{k\beta} = \frac{\Gamma(k\beta+1)}{\Gamma(k\beta-\beta+1)} t^{(k-1)\beta}. \tag{A.16} $$

(Recall (A.10) for the case $k = 0$.)

Remark. One can extend the conclusion of (A.15) to $\gamma > 0$ by a direct computation of $J^{1-\beta}t^{\gamma-1}$, using (A.1).

B. Finite-dimensional linear fractional differential systems

Here we briefly discuss linear systems

$$ {}^c\partial_t^\beta u = Lu, \quad u(0) = f, \tag{B.1}$$

when L is not necessarily a negative self adjoint operator on a Hilbert space, but rather

$$ f \in V, \quad L \in \mathrm{End}(V), \tag{B.2}$$

and V is a complex vector space of dimension $k < \infty$. For more details, see [D].

Parallel to (3.4), the solution to (B.1) is given by

$$ u(t) = E_\beta(t^\beta L)f. \tag{B.3}$$

Now we can write

$$ V = \bigoplus_j V_j, \tag{B.4}$$

where, for λ_j in the spectrum of L,

$$ L\Big|_{V_j} = \lambda_j I + N_j, \tag{B.5}$$

with N_j nilpotent on V_j. Then, in the obvious sense,

$$ E_\beta(t^\beta L) = \bigoplus_j E_\beta(t^\beta(\lambda_j I + N_j)). \tag{B.6}$$

Furthermore, standard holomorphic functional calculus gives, for nilpotent N and $\lambda \in \mathbb{C}$,

$$ E_\beta(\lambda I + N) = \sum_{k\ge 0} \frac{1}{k!} E_\beta^{(k)}(\lambda) N^k, \tag{B.7}$$

the sum being finite if N is nilpotent. Hence

$$ E_\beta(t^\beta(\lambda_j I + N_j)) = \sum_{k\ge 0} \frac{1}{k!} E_\beta^{(k)}(t^\beta \lambda_j) t^{k\beta} N_j^k. \tag{B.8}$$

Note that (8.20) extends to

$$E_\beta^{(k)}(-s) \sim a_\beta^k s^{-k-1} + \cdots, \quad s \nearrow +\infty. \tag{B.9}$$

This implies decay of (B.8) as $t \to +\infty$, when $\lambda_j < 0$, though only at a rate $O(t^{-\beta})$, when $\beta \in (0,1)$, not at an exponential rate, as for $\beta = 1$.

To go further, one can extend the scope of (B.9), by extending that of (8.15)–(8.19). With

$$\eta_\beta(\xi) = \frac{(i\xi)^{\beta-1}}{(i\xi)^\beta + 1}, \tag{B.10}$$

as in (8.15), we have, up to a constant factor,

$$\hat{\eta}_\beta(t) = e_\beta(t). \tag{B.11}$$

Analytic continuation arguments give

$$E_\beta^{(k)}(z) \sim a_\beta^k (-z)^{-k-1} + \cdots, \quad \text{as } |z| \to \infty, \quad \text{for } |\text{Arg } z| > \frac{\pi\beta}{2}. \tag{B.12}$$

See [D]. Hence

$$E_\beta(t^\beta(\lambda_j I + N_j)) \longrightarrow 0 \quad \text{as} \quad t \nearrow +\infty, \tag{B.13}$$

provided

$$|\text{Arg}\, \lambda_j| > \frac{\pi\beta}{2}. \tag{B.14}$$

C. Derivation of power series for $E_\beta(t)$

We approach the solution to

$${}^c\partial_t^\beta u = au, \quad u(0) = 1, \tag{C.1}$$

given $\beta \in (0,1)$, $a \in \mathbb{C}$, taking a cue from (A.16), which suggests trying

$$u(t) = \sum_{k \geq 0} c_k t^{k\beta}. \tag{C.2}$$

In fact, granted appropriate convergence, applying (A.16) to (C.2) yields

$$
\text{(C.3)} \qquad \begin{aligned} {}^c\partial_t^\beta u &= \sum_{k\geq 1} \frac{\Gamma(k\beta+1)}{\Gamma(k\beta-\beta+1)} c_k t^{(k-1)\beta} \\ &= \sum_{\ell\geq 0} \frac{\Gamma(\ell\beta+\beta+1)}{\Gamma(\ell\beta+1)} c_{\ell+1} t^{\ell\beta}. \end{aligned}
$$

Comparison with the series for au, given by multiplying (C.2) by a, yields

$$
\text{(C.4)} \qquad c_{\ell+1} = a\frac{\Gamma(\ell\beta+1)}{\Gamma(\ell\beta+\beta+1)} c_\ell.
$$

Given $c_0 = 1$, we have

$$
\text{(C.5)} \qquad c_1 = \frac{a}{\Gamma(\beta+1)}, \quad c_2 = \frac{a^2}{\Gamma(\beta+1)} \frac{\Gamma(\beta+1)}{\Gamma(2\beta+1)}, \dots,
$$

and inductively,

$$
\text{(C.6)} \qquad c_k = \frac{a^k}{\Gamma(k\beta+1)}.
$$

Hence we arrive at

$$
\text{(C.7)} \qquad u(t) = E_\beta(t^\beta a),
$$

where

$$
\text{(C.8)} \qquad E_\beta(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k\beta+1)},
$$

as the solution to (C.1).

To go backwards, note that, for $\beta \in (0,1)$,

$$
\text{(C.9)} \qquad \begin{aligned} J^\beta\, {}^c\partial_t^\beta u(t) &= J\partial_t u(t) \\ &= u(t) - u(0), \end{aligned}
$$

so (C.1) implies

$$
\text{(C.10)} \qquad u(t) = 1 + aJ^\beta u(t),
$$

and in fact, by (A.9)–(A.10), (C.1) and (C.10) are equivalent. This suggests another approach. Write (C.10) as

$$
\text{(C.11)} \qquad (I - aJ^\beta)u(t) = 1,
$$

and then

$$u(t) = \sum_{k\geq 0} a^k J^{k\beta} 1(t), \tag{C.12}$$

which via (A.2) again leads to (C.7)–(C.8).

References

[CCL] D. del-Castillo-Negrete, B. Carreras, and V. Lynch, Non-diffusive transport in plasma turbulence: a fractional diffusion approach, Preprint, 2004.

[CCL2] D. del-Castillo-Negrete, B. Carreras, and V. Lynch, Fractional diffusion in plasma turbulence, Physics of Plasmas, to appear.

[CCL3] D. del-Castillo-Negrete, B. Carreras, and V. Lynch, Front dynamics in reaction-diffusion equations with Levy flights: a fractional diffusion approach, Phys. Rev. Lett. 91 (2003).

[D] K. Diethelm, The Analysis of Fractional Differential Equations, LNM #2004, Springer-Verlag, New York, 2010.

[J] N. Jacob, Pseudodifferential Operators and Markov Processes, Vol. 1, Imperial College Press, London, 2001.

[LCCFH] V. Lynch, B. Carreras, D. del-Castillo-Negrete, K. Ferreira-Mejias, and H. Hicks, Numerical methods for the solution of partial differential equations of fractional order, J. Comp. Phys. 192 (2003), 406–421.

[MPG] F. Mainardi, P. Paradisi, and R. Gorenflo, Probability distributions generated by fractional diffusion equations, http://arxiv.org/abs/0704.0320.

[P] H. Pollard, The representation of $\exp(-x^\lambda)$ as a Laplace integral, Bull. AMS 52 (1946), 908–910.

[ST] G. Samorodnitsky and M. Taqqu, Stable Non-Gaussian Random Processes, Chapman and Hall, New York, 1994.

[T] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer-Verlag, New York, 1996.

[T2] M. Taylor, Pseudodifferential Operators, Princeton Univ. Press, Princeton NJ, 1981.

[Y] K. Yosida, Functional Analysis, Springer-Verlag, Berlin, 1965.