Convolution powers of complex functions on $\mathbb{Z}$

Persi Diaconis[*]
Department of Mathematics and Department of Statistics
Stanford University

Laurent Saloff-Coste[†][‡]
Department of Mathematics
Cornell University

Abstract

Repeated convolution of a probability measure on $\mathbb{Z}$ leads to the central limit theorem and other limit theorems. This paper investigates what kinds of results remain *without* positivity. It reviews theorems due to Schoenberg, Greville, and Thomée which are motivated by applications to data smoothing (Schoenberg and Greville) and finite difference schemes (Thomée). Using Fourier transform arguments, we prove detailed decay bounds for convolution powers of finitely supported complex functions on $\mathbb{Z}$. If M is an hermitian contraction, an estimate for the off-diagonal entries of the powers M_k^n of $M_k = I - (I - M)^k$ is obtained. This generalizes the Carne–Varopoulos Markov chain estimate.

1 Introduction

This is the first of a series of papers regarding the following question. What can be said about the convolution powers of a finitely supported function on a countable group?

A lot is known if we are willing to assume that the function, say ϕ, is non-negative and normalized by $\sum \phi = 1$. In this case, the convolution power $\phi^{(n)}$ of ϕ is the probability distribution of the associated random walk after n-steps and estimating $\phi^{(n)}$ is a well studied problem. But what exactly is the role of "positivity" in classical results?

This article focuses mostly on the simplest case, convolutions on the integer group $\mathbb{Z}$. It reviews what is known (and why people worked on this question before) and provides some new results including an extension of a Markov chain estimate due to Carne and Varopoulos. In forthcoming papers, we will consider the case of $\mathbb{Z}^d$, which is significantly different from the 1-dimensional case, and the case of non-commutative groups where completely different techniques must be used and many open question remain.

Let ϕ be a finitely supported probability measure on the integers $\mathbb{Z}$. Define convolution powers of ϕ by

$$\phi^{(n)}(x) = \sum_y \phi^{(n-1)}(y)\phi(x-y).$$

---

[*] Supported in part by NSF grant 0804324.

[†] Supported in part by NSF grant DMS 104771.

[‡] Corresponding author: `lsc@math.cornell.edu`

The asymptotic "shape" of $\phi^{(n)}$ is described by the local central limit theorem (assuming irreducibility and aperiodicity):

$$\phi^{(n)}(x) = \frac{1}{\sqrt{2\pi\sigma^2 n}} e^{-(x-\alpha n)^2/(2\sigma^2 n)} + o\left(\frac{1}{\sqrt{n}}\right). \tag{1.1}$$

In (1.1), $\sigma^2 = \sum x^2\phi(x)$, $\alpha = \sum x\phi(x)$, and the error term is uniform over $x \in \mathbb{Z}$. Proofs and many refinements are in [5, 21].

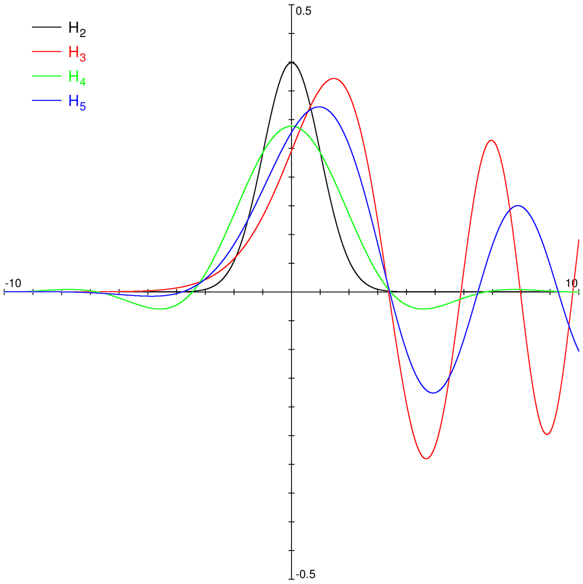

Figure 1:

For this and other basic theorems of probability, could there be similar results if $\phi(x)$ is a more general function? For example, consider

$$\phi(-2) = \phi(2) = -\frac{1}{9}, \quad \phi(1) = \phi(-1) = \frac{4}{9}, \quad \phi(0) = \frac{1}{3} \quad (\phi(x) = 0 \text{ otherwise}). \tag{1.2}$$

The convolution powers $\phi^{(n)}$ are well-defined and results of Schoenberg [24] imply

$$\phi^{(n)}(x) = \frac{1}{(n/9)^{1/4}} H_4\left(\frac{x}{(n/9)^{1/4}}\right) + o\left(\frac{1}{n^{1/4}}\right) \tag{1.3}$$

with H_4 the real even function having Fourier transform e^{-x^4}.

More generally, H_k is defined as having Fourier transform $e^{-\xi^k}$ for even k and $e^{i\xi^k}$ for odd k. A graph of H_4 appears in Figure 1. Also shown are the Gaussian density $H_2(x) = \frac{1}{\sqrt{4\pi}} e^{-x^2/4}$ and H_3, H_5. The function H_3 is the famous Airy function (up to dilation) and it appears as the convolution limit of measures such as

$$\phi(0) = 1 - 3a, \quad \phi(1) = 3a, \quad \phi(-1) = a, \quad \phi(2) = -a \quad (\phi(x) = 0 \text{ otherwise}) \tag{1.4}$$

for $0 < a < 1/4$ fixed. Careful statements of general theorems are in Section 2. We note that there are fundamental differences between the functions H_{2k} and the function H_{2k+1}. An exact formula exists only for H_2 but H_{2k} is even with $H_{2k}(0) > 0$ and satisfies

$$\forall x, \quad |H_{2k}(x)| \leq C_k \exp(-c_k |x|^{2k/(2k-1)}). \tag{1.5}$$

In particular, H_{2k} is in $L^1(\mathbb{R})$. The functions H_{2k+1} satisfy $H_{2k+1}(0) > 0$ but are not even, decay differently along the negative and positive semi-axes and are not absolutely integrable. For any $k > 2$, H_k changes sign infinitely many times.

The following theorem illustrates the main result of this paper. It gives an exponential upper bound of the type (1.5) for the convolution powers of certain complex valued functions on $\mathbb{Z}$. Such bounds are implicit in the literature on finite difference methods. See [28, 29]. The second part of the theorem gives lower bounds on the real and imaginary parts in an appropriate neighborood of 0.

Theorem 1.1. *Let $\phi(x)$ be a finitely supported complex-valued function on $\mathbb{Z}$ with $\sum_x \phi(x) = 1$. Assume that $\hat{\phi}(\theta) = \sum_x \phi(x)e^{ix\theta}$ satisfies*

$$\left|\hat{\phi}(\theta)\right| < 1 \qquad \text{for } \theta \in (-\pi, \pi], \quad \theta \neq 0. \tag{1.6}$$

Assume further that there exist an even integer ν and a complex number γ with $\mathrm{Re}(\gamma) > 0$ *such that*

$$\hat{\phi}(\theta) = e^{-\gamma\theta^\nu(1+o(1))} \qquad \text{as } \theta \to 0. \tag{1.7}$$

Then, there are constants $c, C \in (0, \infty)$ such that, for all $x \in \mathbb{Z}$, $n \in \mathbb{N}^*$,

$$\left|\phi^{(n)}(x)\right| \leq \frac{C}{n^{1/\nu}} \exp\left(-c\left(\frac{|x|}{n^{1/\nu}}\right)^{\nu/(\nu-1)}\right). \tag{1.8}$$

Further, there are constants $c_1, c_2 > 0$ such that, for all $x \in \mathbb{Z}, n \in \mathbb{N}$ with $|x| \leq c_1 n^{1/\nu}$, we have

$$\mathrm{Re}(\phi^{(n)}(x)) \geq c_2 n^{-1/\nu}$$

and

- *if* $\mathrm{Im}(\gamma) \neq 0$,
$$|\mathrm{Im}(\phi(x))| \geq c_2 n^{-1/\nu}$$
with $\mathrm{Im}(\phi^{(n)}(x))$ *having the same sign as* $\mathrm{Im}(\gamma)$*;*
- *if* $\mathrm{Im}(\gamma) \neq 0$,
$$|\mathrm{Im}(\phi(x)|) = o(n^{-1/\nu}).$$

The first part of Theorem 1.1 is proved in Section 3. The second part follows from a local limit theorem derived in Section 2.3. This local theorem gives further information about $\phi^{(n)}$ in the region $|x| \leq c_1 n^{1/\nu}$. Each of these results is obtained using Fourier transform techniques The conclusions of Theorem 1.1 can be described roughly as follows. Under the given hypothesis, $|\phi^{(n)}(x)|$ attains a maximum of order $1/n^{1/\nu}$ and is bounded below by $1/n^{1/\nu}$ in a $n^{1/\nu}$-neighborhood of 0. Further, as x moves away from 0, $|\phi^{(n)}(x)|$ presents a relatively fast decay controlled by

$$\exp\left(-c\left(\frac{|x|}{n^{1/\nu}}\right)^{\nu/(\nu-1)}\right).$$

In modern random walk theory, when ϕ is a centered probability and ν can only take the value $\nu = 2$, the behaviors described in Theorem 1.1 are often discussed under the names of "diagonal behavior" (more precisely, "near diagonal behavior") and "off-diagonal behavior". In particular, in the theory of random walks on graphs, long range off-diagonal upper bounds of the type $e^{-cd(x,y)^2/n}$ are known as Carne–Varopoulos bounds (here $d(x,y)$ denotes the natural graph distance). In Section Section 4, we explain how Theorem 1.1 leads to generalizations of Carne-Varopoulos bounds.

In general, it is not true that the ℓ_1-norm of the convolution powers $\phi^{(n)}$ of a complex valued function ϕ with $\sum \phi = 1$ stay bounded uniformly in n. However, as an immediate corollary of Theorem 1.1 we recover the following known result.

Corollary 1.2. *Under the hypotheses of* Theorem 1.1 *there exists a constant* C *such that*

$$\forall\, n,\ \sum_{x\in\mathbb{Z}} |\phi^{(n)}(x)| \le C < \infty.$$

This important property is studied in detail in [1, 28, 29, 31] because of its close connections to the "stability" of certain approximation schemes for partial differential equations. See Theorem 2.1 below and the short discussion in Section 5.2.

The hypotheses made in Theorem 1.1 are somewhat mysterious. Indeed, it is very unclear how to replace these hypotheses in the context of non-commutative countable groups (for which no viable Fourier transform exists, in general). Even on $\mathbb{Z}$, these hypotheses are not entirely natural since the Fourier transform $\hat{\phi}$ of a function ϕ with $\sum_x \phi(x) = 1$ may well attain its maximum at multiple points and not at 0. Sections 2 and 3 discuss this in detail.

We end this introduction with a brief description of the content of the paper. Section 2 is devoted to local limit theorems for convolution of complex valued functions. In particular, Section 2.3 provides an extension of earlier results of Schoenberg and Greville and Section 2.4 describes illustrative examples. Section 3 is the main section of this paper and treats upper bounds of the type (1.8). See Theorems 3.1 and 3.4. In Section 4, the results of Section 3 and the discrete transmutation formula of Carne are used to obtain long range estimates for the powers of the operator $M_k = I - (I - M)^k$ when M is the infinite matrix of a reversible Markov chain (more generally, an hermitian contraction). Section 5 gives pointers to various earlier works where convolutions of signed measures play an explicit or implicit part.

2 The results of Shoenberg, Greville and Thomé

In this short section, we briefly review results by Schoenberg, Greville and Thomée. Schoenberg and Greville where motivated by the earlier work of de Forest and statistical data smoothing procedures. Thomée's work is motivated by numerical approximation schemes for differential equations. Brief explanations regarding the motivations of these authors and other applications are collected in Section 5.

2.1 Stability

The largest body of work concerning the problem discussed in this paper can be found in the literature concerning numerical approximation schemes for differential equations with work by John [18], Aronson [1], and Widlund [31]. Thomé's articles [28, 29] give excellent pointers to the literature. However, these papers do not isolate or emphasize the basic convolution aspect of the results. This makes extracting the relevant results somewhat difficult. Nevertheless, the following essential result is explicitly stated in [29]. Given a complex valued absolutely summable function ϕ on $\mathbb{Z}$, consider the property

$$\forall\, n, \qquad \sum_x |\phi^{(n)}(x)| \le C < \infty. \tag{2.1}$$

Theorem 2.1 (See [29, Theorem 7.54])**.** *Condition* (2.1) *is satisfied if and only one or the other of the following two conditions is satisfied:*

1. $\hat{\phi}(\xi) = \zeta e^{iy\xi}$ *for some* $y \in \mathbb{Z}$ *and* $\zeta \in \mathbb{C}$ *with* $|\zeta| = 1$
2. $|\hat{\phi}(\xi)| < 1$ *except for at most a finite number of points* ξ_q*,* $q \in \{1, \dots, Q\}$ *in* $|\xi| \le \pi$ *where* $|\hat{a}(\xi_q)| = 1$*, and there are constants* α_q, β_q, ν_q*,* $q = 1, \dots, Q$*, with* α_q *real,* *Re*$(\beta_q) > 0$ *and* ν_q *an even natural integer, such that*

$$\hat{\phi}(\xi_q + \xi) = \hat{\phi}(\xi_q) \exp\left(i\alpha_q \xi - \beta_q \xi^{\nu_q} (1 + o(1))\right) \qquad \textit{as } \xi \to 0.$$

Thomée also describes what happens when the principal nonlinear term of the expansion at a point is imaginary (of odd or even degree). In particular, if $|\hat{\phi}| \le 1$, $\hat{\phi}(0) = 1$ and 0 is the unique point where $|\hat{\phi}|$ is maximum with

$$\hat{\phi}(\xi) = \exp\left(i\alpha\xi + i\gamma\xi^{\mu} - \beta\xi^{\nu} (1 + o(1))\right) \qquad \text{as } \xi \to 0$$

with α, γ real, $\gamma \ne 0$, $\mathrm{Re}(\beta) > 0$, $0 < \mu \le \nu$ and ν an even integer, [29, (7.9)] gives

$$\sum_y |\phi^{(n)}(y)| \simeq n^{(1-\mu/\nu)/2}. \tag{2.2}$$

Here, $\simeq$ means that the ratio of the two sides stays between two positive constants as n tends to infinity.

Condition (2.1) is called a stability condition. In the case of smoothing, it shows that the iterated smoother is continuous from ℓ^∞ to ℓ^∞; small changes in the input sequence lead to close smoothed sequences, uniformly in the iteration number n. In the case of divided difference schemes, it shows that the scheme applied to bounded data gives bounded output, uniformly in time.

2.2 De Forest local limit theorems

Following de Forest, Schoenberg [24] and Greville [16] explore local limit theorems for real valued ϕ that are allowed to change sign and are normalized by $\sum \phi = 1$. Schoenberg treats the case of symmetric ϕ (i.e., $\phi(x) = \phi(-x)$) under the assumption that $|\hat{\phi}(\theta)| < 1$ for $\theta \in (0, 2\pi)$ and $\phi(\theta) = 1 - \lambda\theta^k + 0(|\theta|^{2k+1})$ (with, by necessity, $\lambda > 0$). Greville observes that symmetry is not essential and also treats the case when $\hat{\phi}(\theta) = 1 + ai\theta^{2k+1} + O(|\theta|^{2k+2})$.

For each positive integer k, let H_k denote the function of one real variable defined by $\hat{H}_k(\xi) = e^{-\xi^k}$ for even k and $\hat{H}_k(\xi) = e^{i\xi^k}$ for odd k.

Theorem 2.2 (Greville [16] and Schoenberg [24], in the spirit of de Forest)**.** *Let* ϕ *be a real summable function on* $\mathbb{Z}$ *with Fourier transform* $\hat{\phi}$ *satisfying* $|\hat{\phi}(\theta)| < 1$ *for all* $\theta \in (0, 2\pi)$ *and*

$$\hat{\phi}(\theta) = 1 + a(i\theta)^m + O(|\theta|^{m+1}) \qquad \textit{with } 0 \ne a \in \mathbb{R} \textit{ and } m \textit{ an integer greater than } 1. \tag{2.3}$$

- *If m is even then $-ai^m = \lambda$ must be positive and we have*

$$\phi^{(n)}(x) = (\lambda n)^{-1/m} H_m \left(x(\lambda n)^{-1/m} \right) + o(n^{-1/m}),$$

uniformly over $x \in \mathbb{Z}$.

- *If m is odd and ϵ denotes the sign of the real ai^{m-1}, $\epsilon = \operatorname{sign}(ai^{m-1})$, then we have*

$$\phi^{(n)}(x) = (|a|n)^{-1/m} H_m \left(\epsilon x(|a|n)^{-1/m} \right) + o(n^{-1/m}).$$

Remark. If we consider the expansion $\hat{\phi}(\theta) = \sum_j A_j (i\theta)^j$ then $A_j = \frac{1}{j!} \sum_x \phi(x) x^j$. Hypothesis (2.3) is simply the assumption that the first $k-1$ moments of ϕ vanish and $a = A_k \neq 0$. Note that a is real here because ϕ is real.

Remark. The hypotheses made in Theorem 2.2 are somewhat ad hoc. Schoenberg and Greville are interested in data smoothing procedures. This explains the basic hypotheses that ϕ is real and satisfies $\sum \phi = 1$ but these properties are not essential, at least in the case when m is even. The assumption that $\max |\hat{\phi}| = 1$ is actually the essential condition. Coupled with the condition $\sum \phi = 1$, it implies that $|\hat{\phi}|$ attains its maximum at $\theta = 0$ and that, in fact, $\hat{\phi}(0) = 1$. Schoenberg and Greville then make the additional assumptions that $|\hat{\phi}(\theta)| < 1$ if $\theta \neq 0$. The condition (2.3) simply captures the vanishing order of $\hat{\phi} - 1$ at 0. It excludes the possibility of a drift. In the next section, we extend the even m result to allow complex ϕ, a drift, and multiple points where $|\hat{\phi}|$ attains it maximum.

Remark. The proof of Theorem 2.2 is significantly more difficult in the case where m is odd than in the case when m is even. Indeed, for odd $m = 2k+1 > 1$,

$$H_{2k+1}(x) = \frac{1}{\pi} \int_0^\infty \cos(u^{2k+1} - xu)\, du.$$

This integral is not absolutely convergent and H_{2k+1} is neither integrable nor square integrable. Nevertheless, it is an entire function.

2.3 An extended local limit theorem in the even case

Theorem 2.2 is taken from [16]. Its form is somewhat constrained, perhaps due to the applications that Greville had in mind. Comparison with Theorem 2.1 leads to the following extension where $H_{2k,b}$ denotes the complex valued Schwartz function whose Fourier transform is $\hat{H}_{2k,b}(\xi) = e^{-(1+ib)\xi^{2k}}$, $b \in \mathbb{R}$.

Theorem 2.3. *Let ϕ be a complex absolutely summable function on $\mathbb{Z}$ with Fourier transform $\hat{\phi}$ satisfying $|\hat{\phi}(\theta)| \leq 1$ for all $\theta \in [-\pi, \pi]$. Assume that $|\hat{\phi}(\xi)| < 1$ except for at most a finite number of points ξ_q, $q \in \{1, \ldots, Q\}$, in $|\xi| \leq \pi$ where $|\hat{\phi}(\xi_q)| = 1$. Assume further that there is an integer $Q_1 \leq Q$, an even integer m and constants $\alpha_q, \beta_q = b_q + ib'_q$, $q = 1, \ldots, Q_1$, with $\alpha_q, b_q, b'_q \in \mathbb{R}$, $b_q = \operatorname{Re}(\beta_q) > 0$ such that*

$$\forall\, q \in \{1, \ldots, Q_1\}, \;\; \hat{\phi}(\xi_q + \xi) = \hat{\phi}(\xi_q) \exp\left(i\alpha_q \xi - \beta_q \xi^m (1 + o(1)) \right) \qquad \text{as } \xi \to 0; \tag{2.4}$$

and

$$\forall\, q \in \{Q_1 + 1, \ldots, Q\}, \;\; |\hat{\phi}(\xi_q + \xi)| \leq \exp\left(-c_q |\xi|^\mu \right) \qquad \text{as } \xi \to 0,$$

for some $\mu \in (0, m)$. Then we have

$$\phi^{(n)}(x) = \sum_{1}^{Q_1} (b_q n)^{-1/m} e^{-ix\xi_q} \hat{\phi}(\xi_q)^n H_{m,b'_q/b_q}((x - \alpha_q n)(b_q n)^{-1/m}) + o(n^{-1/m}),$$

where the error term is uniform in $x \in \mathbb{Z}$.

Remark. To make sense of this result, it is necessary to know something about the functions $H_{2k,\beta}$. By definition,

$$\mathrm{Re}(H_{2k,b}(x)) = \frac{1}{\pi} \int_0^\infty (\cos x\xi)(\cos(b\xi^{2k})) e^{-\xi^{2k}} d\xi$$

and

$$\mathrm{Im}(H_{2k,b}(x)) = \frac{1}{\pi} \int_0^\infty (\cos x\xi)(\sin(b\xi^{2k})) e^{-\xi^{2k}} d\xi.$$

From these formula, one easily sees that $\mathrm{Re}(H_{2k,b}(0))$ is always positive and that $\mathrm{Im}(H_{2k,b}(0))$ is zero, positive or negative if and only if b is zero, positive or negative. Further, it holds that

$$|H_{2k,b}(x)| \le C_{k,b} \exp(-c_{k,b} |x|^{\frac{2k}{2k-1}}).$$

In the case $k = 1$, we have

$$H_{2,b}(x) = \frac{1}{\sqrt{4\pi(1 + ib)}} e^{-\frac{|x|^2}{4(1+ib)}}.$$

This complex valued function of the real variable x is the heat kernel, i.e., the kernel of $e^{z\Delta}$, computed at the complex time $z = 1 + ib$. Here Δ denotes the unique self-adjoint extension of $(d/dx)^2$ originally defined on smooth compactly supported functions.

Proof. Set $I = (-\pi, \pi]$. For each q, let $I_q = [\xi_q - \epsilon, \xi_q + \epsilon]$, $\epsilon > 0$, be a small interval centered around ξ_q and let $J = I \setminus \cup_1^Q I_q$. We have

$$\begin{aligned} 2\pi\phi^{(n)}(x) &= \int_{-\pi}^{\pi} e^{-ix\theta} [\hat{\phi}(\theta)]^n d\theta \\ &= \int_J e^{-ix\theta} [\hat{\phi}(\theta)]^n d\theta + \sum_{q=1}^{Q} \int_{I_q} e^{-ix\theta} [\hat{\phi}(\theta)]^n d\theta = \sum_0^Q \Phi_q \end{aligned}$$

where

$$\Phi_0 = \int_J e^{-ix\theta} [\hat{\phi}(\theta)]^n d\theta \ \text{ and } \ \Phi_q = \int_{I_q} e^{-ix\theta} [\hat{\phi}(\theta)]^n d\theta, \ q = 1, \dots, Q.$$

On J, there exists $\rho = \rho_\epsilon \in (0, 1)$ such that $|\hat{\phi}| \le \rho$ and thus, $|\Phi_0| = O(\rho^n)$. For $q \in \{Q_1 + 1, \dots, Q\}$, we have

$$\begin{aligned} |\Phi_q| &= \left|\int_{-\epsilon}^{\epsilon} e^{-ix(\xi_q + \theta)} [\hat{\phi}(\xi_q + \theta)]^n d\theta\right| \\ &\le \int_{-\epsilon}^{\epsilon} e^{-c_q n|\theta|^\mu} d\theta \\ &\le n^{-1/\mu} \int_{-\infty}^{\infty} e^{-c_q |u|^\mu} du = o(n^{-1/m}). \end{aligned}$$

The main contribution comes from the integrals Φ_q, $q \in \{1, \cdots, Q_1\}$. We set

$$\psi_{q,n}(u) = \hat{\phi}(\xi_q)^{-1} e^{-\alpha_q u (b_q n)^{-1/m}} \hat{\phi}(\xi_q + u(b_q n)^{-1/m})]^n,$$

$$y_{q,n} = \frac{x - \alpha_q n}{(b_q n)^{1/m}},$$

and write

$$\begin{aligned} \Phi_q &= e^{-ix\xi_q} \hat{\phi}(\xi_q)^n \int_{-\epsilon}^{\epsilon} e^{-i\frac{(x-\alpha_q n)}{(b_q n)^{1/m}}[(b_q n)^{1/m}\theta]} [\hat{\phi}(\xi_q)^{-1} e^{-i\alpha_q \theta} \hat{\phi}(\xi_q + \theta)]^n d\theta \\ &= (b_q n)^{-1/m} e^{-ix\xi_q} \hat{\phi}(\xi_q)^n \int_{-(b_q n)^{1/m}\epsilon}^{(b_q n)^{1/m}\epsilon} e^{-iy_{q,n}u} \psi_{q,n}(u) du. \end{aligned}$$

Next, since

$$2\pi H_{m,b'_q/b_q}(y_{q,n}) = \int_{-\infty}^{\infty} e^{-y_{q,n}u} e^{-(1+ib'_q/b_q)u^m} du.$$

we have

$$\begin{aligned} &\Phi_q - 2\pi(b_q n)^{-1/m} e^{-ix\xi_q} \hat{\phi}(\xi_q)^n H_{m,b'_q/b_q}(y_{q,n})) = \\ &\quad \int_{(b_q n)^{1/m} I_q} e^{-iy_{q,n}u} \psi_{q,n}(u) du - \int_{-\infty}^{\infty} e^{-iy_{q,n}u} e^{-(1+ib'_q/b_q)u^m} du \\ &= \int_{|u| \le \epsilon (b_q n)^{1/m}} e^{-iy_{q,n}u} [\psi_{q,n}(u) - e^{-(1+ib'_q/b_q)u^m}] du \\ &\quad - \int_{|u| > \epsilon (b_q n)^{1/m}} e^{-iy_{q,n}u} e^{-(1+ib'_q/b_q)u^m} du = \mathcal{I}_1 - \mathcal{I}_2. \end{aligned}$$

The integral $\mathcal{I}_2$ can be estimated brutally by

$$|\mathcal{I}_2| \le 2 \int_{u > \epsilon (b_q n)^{1/m}} e^{-u^m} du = O(e^{-\epsilon^m b_q n}).$$

To estimate $\mathcal{I}_1$, note that (2.4) shows that for any $\eta > 0$ there exists $\epsilon > 0$ such that

$$|\psi_{q,n}(u) - e^{-(1+ib'_q/b_q)u^m}| \le e^{-u^m} |e^{\eta u^m} - 1| \le C\eta e^{-u^m/4}$$

for all $|u| \le \epsilon (b_q n)^{1/m}$. It follows that

$$|\mathcal{I}_1| \le 2C\eta \int_0^{\infty} e^{-u^m/4} du.$$

Putting all these estimates together, we find that

$$\phi^{(n)}(x) - (b_q n)^{-1/m} e^{-ix\xi_q} \hat{\phi}(\xi_q)^n H_{m,b'_q/b_q}(y_{q,n})) = o(1/n^{1/m}).$$

□

2.4 Examples

One of the simplest example that can be used to illustrate the results discussed in this paper appears in the following proposition.

Proposition 2.4. *Assume that* ϕ *is real symmetric,* $|\hat{\phi}| \leq 1$ *and there exists* $a > 0$ *such that*

$$\hat{\phi}(\theta) = 1 - a\theta^{2k}(1 + o(1)) \text{ at } \theta = 0. \tag{2.5}$$

Then $\phi(x) \neq 0$ *for some* $x \geq k$. *If we assume further that* ϕ *is supported on* $\{-k, \ldots, k\}$ *then* $a \in (0, 2^{-2k+1}]$ *and*

$$\begin{aligned} \phi &= \delta_0 - \lambda(\delta_0 - \beta)^k = \delta_0 - (\delta_0 - \beta_{\lambda^{1/k}})^k \qquad (2.6) \\ &= \sum_{j=1}^{k} (-1)^j \binom{k}{j} \beta^{(j)}_{\lambda^{1/k}} \end{aligned}$$

where $\lambda = a2^k \in (0, 2^{-k+1}]$, $\beta = \frac{1}{2}(\delta_{-1} + \delta_1)$ *and* $\beta_s = (1 - s)\delta_0 + s\beta$, $s \in (0, 1)$.

Remark. The function ϕ defined at (2.6) satisfies

$$\hat{\phi}(\theta) = 1 - \lambda(1 - \cos\theta)^k.$$

It follows that $\max\{|\hat{\phi}|\} = 1$ as well as (2.5) for any $\lambda = a2^k \in (0, 2^{-k+1}]$.The maximum $1 = \max\{|\hat{\phi}|\}$ is attained solely at 0 if and only if $\lambda \in (0, 2^{-k+1})$. Note that any parameter λ in the range $(0, 1/2]$ is admissible for all values of k.

Remark. If convolution by $\delta_0 - \beta_s$ (for some fixed $s \in (0, 1)$, say $s = 1/2$) is interpreted as the discrete analog of the (positive) Laplacian $-\partial_x^2$ then convolution by $\delta_0 - \phi = (\delta_0 - \beta_s)^k$ is analogous to the higher even powers of the Laplacian, that is, $(-1)^k \partial_x^{2k}$.

Proof. By assumption, if p is the largest integer such that $\phi(p) \neq 0$, $\hat{\phi}$ is a polynomial Q in $\cos\theta$ of degree p. If (2.5) holds then the polynomial $1 - Q$ vanishes of order k at 1. Hence, we must have $p \geq k$. If we assume that ϕ is supported on $\{-k, \ldots, k\}$, then we must must have $1 - Q(u) = \lambda(1 - u)^k$ and $a = 2^k\lambda$ with a as in (2.5). That is,

$$\hat{\phi}(\theta) = 1 - \lambda(1 - \cos\theta)^k = 1 - \left(1 - \left(1 - \lambda^{1/k}\right) + \lambda^{1/k}\cos\theta\right)^k.$$

The condition that $|\hat{\phi}| < 1$ on $(0, 2\pi)$ translates into $1 - \lambda 2^k > -1$, that is $\lambda < 2^{-k+1}$. Further, since $\cos\theta = \hat{\beta}(\theta)$ where $\beta(1) = \beta(-1) = 1/2$ and $\beta = 0$ otherwise (i.e., β =Bernoulli(1/2)), we see that we must have

$$\phi = -\sum_{i=1}^{k} (-1)^i \binom{k}{i} \beta^{(i)}_{\lambda^{1/k}}$$

where $\beta_s = (1 - s)\delta_0 + s\beta$. ☐

Example 2.1. In Proposition 2.4, consider the case when $k = 1$ and $\lambda = 1$ so that $\phi = \beta$. We have $\hat{\phi}(\theta) = \cos\theta$. In this classical case, Theorem 2.3 yields

$$\phi^{(n)}(x) = (1 + (-1)^n e^{-ix\pi})(n/2)^{-1/2} H_2(x/(n/2)^{1/2}) + o(n^{-1/2}).$$

This captures the periodicity of the Bernoulli walk.

Example 2.2. In Proposition 2.4, consider the case when $k = 2$, that is $\phi = 2\beta_s - \beta_s^{(2)}$ with $s \in (0, 1/\sqrt{2})$. We have $\hat{\phi}(\theta) = 1 - s^2(1 - \cos\theta)^2$. Theorem 2.2 yields

$$\phi^{(n)}(x) \sim (4/s^2 n)^{1/4} H_4(x(4/(s^2 n)^{1/4}).$$

The same result holds true in the limit case where $s = 1/\sqrt{2}$ and $\phi = 2\beta_{1/\sqrt{2}} - \beta^{(2)}_{1/\sqrt{2}}$. However, in this case, $|\hat{\phi}| \leq 1$ and $\hat{\phi}(\theta) = 1$ if and only if $\theta = 0$ or $\theta = \pi$. At 0, $\phi(\theta) = 1 - \frac{1}{8}|\theta|^4 + O(|\theta|^6)$. At π, $\hat{\phi}(\theta) = -1 + (\pi - \theta)^2 + O(|\pi - \theta|^4)$. To obtain the desired asymptotic, apply Theorem 2.3. Compare to the previous example.

Example 2.3. Let ϕ be defined by $\phi(0) = 5/8$, $\phi(\pm 2) = -1/4$, $\phi(\pm 4) = -1/16$ and $\phi(x) = 0$ otherwise. We have

$$\begin{aligned} \hat{\phi}(\theta) &= \frac{5}{8} - \frac{1}{2}\cos 2\theta - \frac{1}{8}\cos 4\theta \\ &= \frac{5}{4} - \cos^2\theta - \frac{1}{4}\cos^2 2\theta \\ &= 1 - \cos^4\theta. \end{aligned}$$

Hence $|\hat{\phi}| \leq 1$ and $\hat{\phi}(\theta) = 1$ if and only if $\theta = \pm\pi/2$. Further, at $\theta_\pm = \pm\pi/2$,

$$\hat{\phi}(\theta) = 1 - |\theta - \theta_\pm|^4 + O(|\theta - \theta_\pm|^5).$$

Theorem 2.3 applies and gives

$$\begin{aligned} \phi^{(n)}(x) &= (e^{-ix\frac{\pi}{2}} + e^{ix\frac{\pi}{2}}) n^{-1/4} H_4(xn^{-1/4}) + o(n^{-1/4}) \\ &= 2\cos(\pi x/2) n^{-1/4} H_4(xn^{-1/4}) + o(n^{-1/4}). \end{aligned}$$

Example 2.4. In this example, we consider the convolution powers of an arbitrary real valued function supported on $\{-1, 0, +1\}$ (except for some trivial cases). For $a_0, a_+, a_- \in \mathbb{R}$, let ϕ be given by

$$\phi(0) = a_0, \quad \phi(\pm 1) = a_\pm \text{ and } \phi = 0 \text{ otherwise.}$$

We assume $a_0 > 0$ and that either $a_+ \neq 0$ or $a_- \neq 0$ to avoid trivialities. We do not assume any normalization. In particular,

$$\sum \phi = a_0 + a_+ + a_-$$

is an arbitrary real number. The following proposition shows that there are essentially 3 different "generic" cases (each occurring on an open subset of the parameter space). In only one of these 3 cases is the normalization $\sum \phi = 1$ the correct normalization giving $\max\{|\hat{\phi}|\} = 1$. In the other two cases, different normalizations are needed to insure $\max\{|\hat{\phi}|\} = 1$. Because we do not incorporate any normalization, the asymptotic described below for $\phi^{(n)}$ contain an exponential term A^n. In each case, the constant A satisfies $A = |\max\{|\hat{\phi}|\}$ and is given explicitly.

Proposition 2.5. *Referring to the function ϕ defined above, the following asymptotics hold true:*

- *Assume that* $a_+a_- \geq 0$ *or that* $a_+a_- < 0$ *and* $4|a_+a_-| < a_0|a_+ + a_-|$*. Set*

$$A = a_0 + |a_+| + |a_-|, \quad \alpha = \frac{a_+ - a_-}{A} \quad \textit{and} \quad \gamma = \frac{|a_+| + |a_-|}{2A} - \frac{\alpha^2}{2}.$$

  Then

$$\phi^{(n)}(x) = \left(\frac{a_+ + a_-}{|a_+ + a_-|}\right)^x A^n (\gamma n)^{-1/2} e^{-|x-\alpha n|^2/\gamma n} + o(A^n n^{-1/2})$$

  where the error term is uniform in x*.*

- *Assume that* $a_+a_- < 0$ *and* $4|a_+a_-| > a_0|a_+ + a_-|$*. Set*

$$A = |a_+ - a_-|(1 + a_0^2/4|a_+a_-|)^{1/2}, \quad \alpha = \frac{a_+ + a_-}{a_+ - a_-}.$$

  Let $\theta_0 \in (0, \pi)$ *be defined by* $\cos\theta_0 = -a_0(a_+ + a_-)/4a_+a_-$ *and set*

$$b = \frac{4a_0|a_+a_-|\sin\theta_0}{(a_+ - a_-)(4|a_+a_-| + a_0^2)}, \quad \gamma = \frac{16|a_+a_-|^2 - a_0^2(a_+ + a_-)^2}{2(a_0^2 + 4|a_+a_-|)(a_+ - a_-)^2}.$$

  Let ω_0 *be the argument of* $\hat{\phi}(\theta_0)$*. Then*

$$\begin{aligned}\phi^{(n)}(x) &= (n/\gamma)^{-1/2} A^n e^{-ix\theta_0 + in\omega_0} H_{2,b/\gamma}((x - \alpha n)/(\gamma n)^{1/2}) \\ &\quad + (n/\gamma)^{-1/2} A^n e^{ix\theta_0 - in\omega_0} H_{2,-b/\gamma}((x - \alpha n)/(\gamma n)^{1/2}) + o(A^n n^{-1/2})\end{aligned}$$

  where the error term is uniform in x*.*

Remark. The two principal terms on the right-hand side of the last equation are complex conjugate so that their sum is real (as it should be since ϕ is real valued).

Remark. This example can also be used to illustrate the stability theorem, Theorem 2.1. Indeed, all cases with $a_0 > 0$ and either a_+ or a_- non-zero are considered in Proposition 2.5 except for the very special case when $4|a_+a_-| = a_0|a_+ + a_-|$. Let $A = A(a_0, a_+, a_-)$ be as defined in Proposition 2.5. As a corollary of the proof given below, it follows that the normalized function $\phi_0 = \phi/A$ satisfies the stability condition (2.1) in all cases but the special case $4|a_+a_-| = a_0|a_+ + a_-|$ for which it actually fails.

Proof. Obviously

$$\hat{\phi}(\theta) = a_0 + (a_+ + a_-)\cos\theta + i(a_+ - a_-)\sin\theta$$

and

$$|\hat{\phi}(\theta)|^2 = a_0^2 + a_+^2 + a_-^2 + 2a_+a_-(2\cos^2\theta - 1) + 2a_0(a_+ + a_-)\cos\theta.$$

Further

$$\hat{\phi}(0) = a_0 + a_+ + a_-, \quad \hat{\phi}(\pi) = a_0 - (a_+ + a_-).$$

If $a_+a_- \geq 0$ then $|\hat{\phi}|$ has a maximum which is attained only at 0 if $a_+ + a_- > 0$ and only at π if $a_+ + a_- < 0$.

Assume first that $a_+a_- \geq 0$ and set

$$A = a_0 + |a_+| + |a_-|, \quad \alpha = \frac{a_+ - a_-}{A} \text{ and } \gamma = \frac{|a_+| + |a_-|}{2A} - \frac{(a_+ - a_-)^2}{2A^2} > 0.$$

Considering separately the two cases $a_+ + a_- > 0$ and $a_+ + a_- < 0$, we obtain

$$\begin{aligned} e^{-i\alpha\theta}\hat{\phi}(0)^{-1}\hat{\phi}(\theta) &= (1 - i\alpha\theta - \frac{\alpha^2}{2}\theta^2 + o(|\theta|^2))(1 + i\alpha\theta - (\gamma + \alpha^2/2)\theta^2 + o(|\theta|^2)) \\ &= 1 - \gamma\theta^2(1 + o(1)) = e^{-\gamma\theta^2(1+o(1))}. \end{aligned}$$

Hence, if $a_+ + a_- > 0$ we have

$$\phi^{(n)}(x) = A^n(\gamma n)^{-1/2}H_2((x - \alpha n)/(\gamma n)^{1/2}) + o(A^n n^{-1/2}).$$

If instead $a_+ + a_- < 0$ then we have

$$\phi^{(n)}(x) = (-1)^x A^n(\gamma n)^{-1/2}H_2((x - \alpha n)/(\gamma n)^{1/2}) + o(A^n n^{-1/2}).$$

Next we consider what happens when $a_+a_- < 0$. Computing the derivative of $f(\theta) = |\hat{\phi}(\theta)|^2$ gives

$$f'(\theta) = -2(4a_+a_- \cos\theta + a_0(a_+ + a_-))\sin\theta.$$

If $4|a_+a_-| < a_0|a_+ + a_-|$, then $|\hat{\phi}|^2$ attains its maxima at 0 if $a_+ + a_- > 0$ and at π if $a_+ + a_- < 0$. In each case, the asymptotic is the same as described above.

If $4|a_+a_-| \geq a_0|a_+ + a_-|$, then we set θ_0 to be the point in $(0, \pi)$ such that

$$\cos\theta_0 = -\frac{a_0(a_+ + a_-)}{4a_+a_-}$$

and note that $|\hat{\phi}|^2$ has twin maxima at $\theta = \pm\theta_0$ where

$$\begin{aligned} A^2 &= |\hat{\phi}(\pm\theta_0)|^2 = a_0^2 + a_+^2 + a_-^2 + 2|a_+a_-| + a_0^2\frac{(a_+ + a_-)^2}{4|a_+a_-|} \\ &= (a_+ - a_-)^2\left(1 + \frac{a_0^2}{4|a_+a_-|}\right). \end{aligned}$$

Further

$$\begin{aligned} \mathrm{Re}(\overline{\hat{\phi}(\theta_0)}\hat{\phi}(\theta)) &= (a_0 + (a_+ + a_-)\cos\theta_0)(a_0 + (a_+ + a_-)\cos\theta) + (a_+ - a_-)^2\sin\theta_0\sin\theta \\ &= (a_+ - a_-)^2\left(\frac{a_0^2}{4|a_+a_-|} + \cos\theta_0\cos\theta + \sin\theta_0\sin\theta\right) \\ &= (a_+ - a_-)^2\left(\frac{a_0^2}{4|a_+a_-|} + \cos(\theta_0 - \theta)\right) \\ &= |\hat{\phi}(\theta_0)|^2\left(1 - \frac{2|a_+a_-|}{a_0^2 + 4|a_+a_-|}(\theta - \theta_0)^2(1 + o(1))\right) \end{aligned}$$

and

$$\begin{aligned}
\mathrm{Im}(\overline{\hat{\phi}(\theta_0)}\hat{\phi}(\theta)) &= (a_+ - a_-)((a_0 + (a_+ + a_-)\cos\theta_0)\sin\theta - (a_0 + (a_+ + a_-)\cos\theta)\sin\theta_0) \\
&= a_0(a_+ - a_-)(\sin\theta - \sin\theta_0) + (a_+^2 - a_-^2)\sin(\theta - \theta_0) \\
&= (a_+ - a_-)(a_0\cos\theta_0 + a_+ + a_-)(\theta - \theta_0) \\
&\quad -\frac{a_0(a_+ - a_-)}{2}\sin\theta_0(\theta - \theta_0)^2(1 + o(1)) \\
&= (a_+^2 - a_-^2)\left(1 + \frac{a_0^2}{4|a_+a_-|}\right)(\theta - \theta_0) \\
&\quad -\frac{a_0(a_+ - a_-)}{2}\sin\theta_0(\theta - \theta_0)^2(1 + o(1)) \\
&= |\hat{\phi}(\theta_0)|^2\frac{a_+ + a_-}{a_+ - a_-}(\theta - \theta_0) - \\
&\quad |\hat{\phi}(\theta_0)|^2\frac{4a_0|a_+a_-|\sin\theta_0}{(a_+ - a_-)(4|a_+a_-| + a_0^2)}(\theta - \theta_0)^2(1 + o(1))
\end{aligned}$$

Set

$$\alpha = \frac{a_+ + a_-}{a_+ - a_-}, \quad b = \frac{4a_0|a_+a_-|\sin\theta_0}{(a_+ - a_-)(4|a_+a_-| + a_0^2)}$$

and

$$\begin{aligned}
\gamma &= \frac{2|a_+a_-|}{a_0^2 + 4|a_+a_-|} - \frac{1}{2}\left(\frac{a_+ + a_-}{a_+ - a_-}\right)^2 \\
&= \frac{16|a_+a_-|^2 - a_0^2(a_+ + a_-)^2}{2(a_0^2 + 4|a_+a_-|)(a_+ - a_-)^2}.
\end{aligned}$$

Note that γ is (strictly) positive if $4|a_+a_-| > a_0|a_+ + a_-|$. With this notation, assuming that $4|a_+a_-| > a_0|a_+ + a_-|$, we have

$$\hat{\phi}(\theta_0 + \theta) = \phi(\theta_0)e^{i\alpha\theta - (\gamma + i\beta)\theta^2(1 + o(1))}.$$

Similarly,

$$\hat{\phi}(-\theta_0 + \theta) = \phi(-\theta_0)e^{i\alpha\theta - (\gamma - i\beta)\theta^2(1 + o(1))}.$$

Let ω_0 be the argument of $\hat{\phi}(\theta_0)$. Then Theorem 2.3 gives

$$\begin{aligned}
\phi^{(n)}(x) &= (n/\gamma)^{-1/2}A^n e^{-ix\theta_0 + in\omega_0}H_{2,b/\gamma}((x - \alpha n)/(\gamma n)^{1/2}) \\
&\quad +(n/\gamma)^{-1/2}A^n e^{ix\theta_0 - in\omega_0}H_{2,-b/\gamma}((x - \alpha n)/(\gamma n)^{1/2}) + o({}^{n}n^{-1/2}).
\end{aligned}$$

Finally, if $4a_+a_- = -a_0(a_+ + a_-)$ (resp. $4a_+a_- = a_0(a_+ + a_-)$), $|\hat{\phi}|$ is maximum at 0 (resp. at π). The two case are similar and we treat only the case when $4a_+a_- = -a_0(a_+ + a_-)$. In this case we have

$$\hat{\phi}(0)^{-1}\hat{\phi}(\theta) = e^{i\alpha\theta - i\frac{1}{6}(\alpha - \alpha^3)\theta^3 - \frac{1}{8}(\alpha^2 - \alpha^4)\theta^4(1 + o(1))}.$$

□

By (2.2), in this case we have $\sum_x |\phi^{(n)}(x)| \simeq A^n n^{1-3/4}$.

3 Bounds on convolution powers of normalized complex functions on $\mathbb{Z}$

The goal of this section is to give good upper bounds for the convolution powers $\phi^{(n)}$ of a given complex valued function that is finitely supported on $\mathbb{Z}$. Let $\hat{\phi}$ be the Fourier transform of ϕ so that

$$\phi^{(n)}(x) = \frac{1}{2\pi}\int_I e^{-ix\theta}[\hat{\phi}(\theta)]^n d\theta.$$

Obviously the function ϕ can be normalized in some appropriate way and it is very reasonable to chose the normalization

$$\max_{\theta\in I}\{|\hat{\phi}|\} = 1.$$

Note that this is the same as saying that the operator norm of the convolution operator by ϕ acting on $\ell^2(\mathbb{Z})$ is 1.

As we assume that ϕ is finitely supported, either $|\hat{\phi}| = 1$ on I, that is, $\hat{\phi}(\theta) = ce^{ij\theta}$ with $|c| = 1$, j an integer, or $\max\{|\hat{\phi}|\}$ is attained at at most finitely many points in I (the zeros of the trigonometric polynomial $|\hat{\phi}|^2 - 1$).

In the first subsection of this section, we concentrate on the case when $\max|\hat{\phi}|$ is attained at only one point $\theta_0 \in I$ and $\phi(\theta) - \phi(\theta_0)$ vanishes up to some even order.

The case when the maximum of $\hat{\phi}$ is attained at more than one point will be considered in the second subsection below but our results are much less precise in that case.

3.1 Generalized Gaussian bounds, I

In this section, we assume that $\max|\hat{\phi}|$ is attained at only one point $\theta_0 \in I$. By replacing ϕ by $x \mapsto \phi_{\theta_0}(x) = \hat{\phi}(\theta_0)^{-1}e^{ix\theta_0}\phi(x)$, it is enough to consider the case when $\theta_0 = 0$ and $\hat{\phi}(0) = 1$. So assume from now on that ϕ is finitely supported with

$$\{x : \phi(x) \neq 0\} \subset [-K, K], \quad \|\phi\|_1 = \sum|\phi| < \infty$$

and

$$\hat{\phi}(0) = 1, \quad |\hat{\phi}(\theta)| < 1 \text{ for } \theta \in I^* = I \setminus \{0\}. \tag{3.1}$$

Since ϕ is finitely supported, $\hat{\phi}$ is actually an entire function of the complex variable $z = u + iv \in \mathbb{C}$, namely,

$$\hat{\phi}(z) = \sum_{x\in\mathbb{Z}}\phi(x)e^{izx}.$$

As noticed in [28, 29] (and many other places), it is most efficient to expand $\log\phi$ and write

$$\hat{\phi}(\theta) = \exp(\sum_{j=1}^{\infty}c_j\theta^j) \text{ near } 0$$

to study the convolution powers $\phi^{(n)}$. Indeed, the conditions $\hat{\phi}(0) = 1$ and $|\hat{\phi}| < 1$ in I^* implies that only two cases may arise. Namely, either there exist reals α, β, two integers $0 < \mu < \nu$, ν even, a real polynomial q with $q(0) \neq 0$, and a complex number γ with $\mathrm{Re}(\gamma) > 0$ such that

$$\hat{\phi}(\theta) = e^{i\alpha\theta + i\theta^\mu q(\theta) - \gamma\theta^\nu(1+o(1))} \tag{3.2}$$

or there exists a real α, an even integer ν and a complex number γ with $\mathrm{Re}(\gamma) > 0$ such that

$$\hat{\phi}(\theta) = e^{i\alpha\theta - \gamma\theta^{\nu}(1+o(1))}. \tag{3.3}$$

When (3.2) holds, [28] shows that $\|\phi^{(n)}\|_1$ tends to infinity with n and in fact

$$\|\phi^{(n)}\|_1 \simeq n^{(1-\mu/\nu)/2} \text{ as } n \to \infty.$$

See [29, Sect. 7] and the references therein. Hence we focus on the case when (3.3) holds. The simplest form of our main result is stated in the following theorem.

Theorem 3.1. *Let $\phi : \mathbb{Z} \to \mathbb{C}$ be a finitely supported complex valued function such that (3.3) holds true and $|\hat{\phi}(\theta)| < 1$ on I^*. Then there are constants $C, c \in (0, \infty)$ such that for any $x \in \mathbb{Z}$ and $n \in \mathbb{N}^*$, we have*

$$|\phi^{(n)}(x)| \leq \frac{C}{n^{1/\nu}} \exp\left(-c\left(\frac{|x-\alpha n|}{n^{1/\nu}}\right)^{\frac{\nu}{\nu-1}}\right).$$

Proof. Write

$$\phi^{(n)}(x) = \frac{1}{2\pi}\int_I e^{-ix\theta}\hat{\phi}(\theta)^n d\theta = \frac{1}{2\pi}\int_I e^{-i(x-\alpha n)\theta}P(\theta)^n d\theta$$

where

$$P(\theta) = e^{-i\alpha\theta}\hat{\phi}(\theta).$$

Condition (3.3) together with $|\hat{\phi}| < 1$ on I^* implies that

$$P(\theta) = \exp(-\gamma\theta^{\nu}(1+o(1))), \quad |P| < 1 \text{ on } I^* \tag{3.4}$$

with γ, ν as in (3.3), that is $\mathrm{Re}(\gamma) > 0$ and ν is an even integer. In addition P is obviously an entire function of the complex variable $z = u + iv$ and (3.4) implies that there are constants $\gamma_1, \gamma_2 \in (0, \infty)$ such that

$$|P(z)| \leq \exp(-\gamma_1 u^{\nu} + \gamma_2 v^{\nu}) \text{ on } \{z = u + iv : |u| \leq 3\pi/2\} \tag{3.5}$$

and

$$|P(z)| \leq \exp(\gamma_2 v^{\nu}) \text{ on } \mathbb{C}. \tag{3.6}$$

To see why (3.6) holds, at the origin, use the assumed expansion $P(u) = e^{i\alpha u - \gamma u^{\nu}(1+o(1))}$. Away from the origin, note that $\hat{\phi}(z)$ is a trigonometric polynomial and thus is periodic in u with growth at most exponential in $|v|$ for large v. The estimate (3.5) improves significantly upon (3.6) only in the sector

$$\{z = u + iv : |v| \leq \frac{\gamma_2}{\gamma_1}|u|; |u| \leq 3\pi/2\}$$

where the the two parts of the hypothesis (3.4) together give the desired result.

Next, write

$$\phi^{(n)}(x) = \frac{1}{2\pi n^{1/\nu}}\int_{n^{1/\nu}I} e^{-i\frac{(x-\alpha n)}{n^{1/\nu}}\theta}P_n(\theta)\, d\theta$$

with

$$n^{1/\nu}I = (-n^{1/\nu}\pi, n^{1/\nu}\pi] \text{ and } P_n(\theta) = P(\theta/n^{1/\nu})^n.$$

It follows that

$$\left(\frac{|x-\alpha n|}{n^{1/\nu}}\right)^q |\phi^{(n)}(x)| = \frac{1}{2\pi n^{1/\nu}}\left|\int_{n^{1/\nu}I} e^{-i\frac{(x-\alpha n)}{n^{1/\nu}}\theta}\partial_\theta^q P_n(\theta)\ d\theta\right|. \tag{3.7}$$

This step requires some explanation because when αn is not an integer, neither $e^{-i\frac{(x-\alpha n)}{n^{1/\nu}}\theta}$ nor $P_n(\theta)$ are periodic function of period $2\pi n^{1/\nu}$ so that, a priori, the repeated integration by parts used to obtain (3.7) should produce boundary terms. However, introducing the operator δ defined on smooth functions by

$$\delta f(\theta) = e^{-i\frac{\alpha n}{n^{1/\nu}}\theta}\partial_\theta[e^{i\frac{\alpha n}{n^{1/\nu}}\theta}f(\theta)],$$

write

$$\begin{aligned}\int_{n^{1/\nu}I}[\partial_\theta^q e^{-i\frac{(x-\alpha n)}{n^{1/\nu}}\theta}]P_n(\theta)\ d\theta &= \int_{n^{1/\nu}I} e^{-i\frac{\alpha n}{n^{1/\nu}}\theta}\partial_\theta^q[e^{i\frac{\alpha n}{n^{1/\nu}}\theta}e^{-i\frac{x}{n^{1/\nu}}\theta}](\hat{\phi}(\theta/n^{1/\nu}))^n\ d\theta\\ &= \int_{n^{1/\nu}I}\delta^r[e^{-i\frac{x}{n^{1/\nu}}\theta}](\hat{\phi}(\theta/n^{1/\nu}))^n\ d\theta.\end{aligned}$$

Next, observe that δ preserves the periodicity of the function f. In particular, if f is periodic of period $2\pi n^{1/\nu}$ then δf has the same property. Hence the formal adjoint of δ on $2\pi n^{1/\nu}$-periodic functions is given by

$$\delta^* f(\theta) = -e^{i\frac{\alpha n}{n^{1/\nu}}\theta}\partial_\theta[e^{-i\frac{\alpha n}{n^{1/\nu}}\theta}f(\theta)].$$

It follows that

$$\begin{aligned}\int_{n^{1/\nu}I}\delta^r[e^{-i\frac{x}{n^{1/\nu}}\theta}](\hat{\phi}(\theta/n^{1/\nu}))^n\ d\theta &= (-1)^r\int_{n^{1/\nu}I} e^{-i\frac{x}{n^{1/\nu}}\theta}\delta^{*r}(\hat{\phi}(\theta/n^{1/\nu}))^n\ d\theta\\ &= (-1)^r\int_{n^{1/\nu}I} e^{-i\frac{(x-\alpha n)}{n^{1/\nu}}\theta}\partial_\theta^q P_n(\theta)\ d\theta.\end{aligned}$$

This justifies (3.7). We note that a formula similar to (3.7) appears on page 124, equation (2.2), of the classic paper by Ney and Spitzer [20] without warning or detailed explanations.

Observe that (3.5)-(3.6) translate immediately into

$$|P_n(z)| \le \exp(-\gamma_1 u^\nu + \gamma_2 v^\nu) \text{ on } \{z = u + iv : |u| \le 3n^{1/\nu}\pi/2\} \tag{3.8}$$

and

$$|P_n(z)| \le \exp(\gamma_2 v^\nu) \text{ on } \mathbb{C}. \tag{3.9}$$

The crucial estimate is given by the following proposition.

Proposition 3.2. *There are constants $C_1, A_1, a_1 \in (0,\infty)$ such that for any $q, n = 0, 1, 2\ldots$, and $\theta \in n^{1/\nu}I$,*

$$|\partial_\theta^q P_n(\theta)| \le C_1^{1+q}q!q^{-q/\nu}\exp\left(-a_1\theta^\nu\right).$$

Assuming that this proposition has been proved, we obtain

$$\left(\frac{|x-\alpha n|}{n^{1/\nu}}\right)^q |\phi^{(n)}(x)| \le n^{-1/\nu} C_2^{1+q} q! q^{-q/\nu}.$$

This is of the form

$$|\phi^{(n)}(x)| \le C_2 n^{-1/\nu} M^{-q} q^{(1-1/\nu)q}, \;\; M = \frac{|x-\alpha n|}{C_2 n^{1/\nu}}.$$

Elementary calculus shows that

$$\inf_{q=0,1,2,\ldots} \{M^{-q} q^{(1-1/\nu)q}\} \le C \exp(-cM^{\frac{\nu}{\nu-1}}), \;\; c = \frac{\nu}{e(\nu-1)}.$$

This gives the upper bound stated in Theorem 3.1.

It remains to prove Proposition 3.2. By Cauchy's formula,

$$\partial_\theta^q P_n(\theta) = \frac{q!}{2\pi i} \int_{|\xi|=r} \frac{P_n(z)}{(\zeta-\theta)^{q+1}} d\zeta, \theta \in n^{1/\nu} I.$$

Consider two cases.

If $q \le \theta^\nu$, pick $r = \epsilon q^{1/\nu}$ with $\epsilon > 0$ small enough (depending on γ_1, γ_2 in (3.8)) so that $|P_n(z)| \le \exp(-\gamma_3 \theta^\nu)$ on $|z-\theta| = r$. This easily gives the inequality of Proposition 3.2 when $q \le \theta^\nu$.

If, instead, $q > \theta^\nu$ then pick $r = q^{1/\nu}$ and observe that, on $|z-\theta| = r$,

$$|P_n(z)| \le \exp(2\gamma_2 r^\nu) \le \exp(-a_1 \theta^\nu + (2\gamma_2 + a_1) q).$$

This yields the desired estimate when $q > \theta^\nu$. □

A useful complement to Theorem 3.1 involves "regularity" estimates for $\phi^{(n)}(x)$. Namely, for any integer y and function $f : \mathbb{Z} \to \mathbb{C}$, set $\partial_y f(x) = f(x+y) - f(x)$.

Theorem 3.3. *Let $\phi : \mathbb{Z} \to \mathbb{C}$ be a finitely supported function such that* (3.3) *holds true and $|\hat{\phi}(\theta)| < 1$ on I^*. Then there are constants $A, C, c \in (0, \infty)$ such that for any $x \in \mathbb{Z}$, $n \in \mathbb{N}^*$ and any $y_1, \ldots, y_m \in \mathbb{Z}$ with $|y_j| \le An^{1/\nu}$, $j = 1, \ldots, m$, we have*

$$|\partial_{y_1} \cdots \partial_{y_m} \phi^{(n)}(x)| \le \frac{C^m \prod_1^m |y_j|}{n^{(1+m)/\nu}} \exp\left(-c\left(\frac{|x-\alpha n|}{n^{1/\nu}}\right)^{\frac{\nu}{\nu-1}}\right).$$

In particular, there exists C_1 such that for all $n \ge 1$ and $y_1, \ldots, y_m \in \mathbb{Z}$ with $|y_j| \le An^{1/\nu}$, $j = 1, \ldots, m$, we have

$$\sum_{x \in \mathbb{Z}} |\partial_{y_1} \cdots \partial_{y_m} \phi^{(n)}(x)| \le C_1 \frac{C^m \prod_1^m |y_j|}{n^{m/\nu}}.$$

Proof. Observe that

$$\partial_{y_1}\cdots\partial_{y_m}\phi^{(n)}(x) = \frac{1}{2\pi n^{1/\nu}}\int_{n^{1/\nu}I} e^{-i\frac{(x-\alpha n)}{n^{1/\nu}}\theta} Q_n(\theta)d\theta$$

where

$$Q_n(z) = \prod_{j=1}^{m}(e^{iy_jz/n^{1\nu}} - 1)P_n(z).$$

Clearly, for $\epsilon > 0$ small enough, if $z = u + iv$ with $|u| \le 3n^{1/\nu}\pi/2$ and $|v| \le \epsilon|u|$, (3.5) gives

$$|Q_n(z)| \le C_1^m \left(\prod_{j=1}^{m}\frac{|y_i|}{n^{1/\nu}}\right) e^{-\gamma_4 u^\nu}.$$

Further for all $z \in \mathbb{C}$, (3.6) gives

$$|Q_n(z)| \le C_1^m \left(\prod_{j=1}^{m}\frac{|y_i|}{n^{1/\nu}}\right) e^{\gamma_5 v^\nu}.$$

Using these estimates in place of (3.5)–(3.6) and the line of reasoning of the proof of Theorem 3.1 gives the desired estimate. □

We end by stating explicitly the most general result obtained in this section.

Theorem 3.4. *Let $\phi : \mathbb{Z} \to \mathbb{C}$ be a finitely supported function such that there exists $\theta_0 \in I = (-\pi, \pi]$ such that*

$$|\hat{\phi}(\theta_0)| = 1 \text{ and } |\hat{\phi}(\theta)| < 1 \text{ on } I \setminus \{\theta_0\}.$$

Assume further that there exist $\alpha \in \mathbb{R}$, $\gamma \in \mathbb{C}$ with $Re(\gamma) > 0$, and an even integer ν such that

$$\hat{\phi}(\theta_0 + \theta) = \hat{\phi}(\theta_0)e^{i\alpha\theta - \gamma\theta^\nu(1+o(1))} \text{ at } \theta = 0. \tag{3.10}$$

Then there are constants $A, C, c \in (0, \infty)$ such that for any $x \in \mathbb{Z}$, $n \in \mathbb{N}^*$, $m = 0, 1, \ldots,$ and $y_1, \ldots, y_m \in \mathbb{Z}$ with $|y_j| \le An^{1/\nu}$, $j = 1, \ldots, m$, we have

$$|\partial_{y_1}\cdots\partial_{y_m}[\hat{\phi}(\theta_0)^{-n}e^{ix\theta_0}\phi^{(n)}(x)]| \le \frac{C^m\prod_1^m |y_j|}{n^{(1+m)/\nu}} \exp\left(-c\left(\frac{|x-\alpha n|}{n^{1/\nu}}\right)^{\frac{\nu}{\nu-1}}\right).$$

Further, there exists $\epsilon, \eta > 0$ such that, on $|x - \alpha n| \le \epsilon n^{1/\nu}$,

$$\mathrm{Re}\left(\hat{\phi}(\theta_0)^{-n}e^{ix\theta_0}\phi^{(n)}(x)\right) \ge \eta n^{-1/\nu}.$$

Proof. This follows from Theorems 3.1-3.3 applied to $\phi_{\theta_0}(x) = \hat{\phi}(\theta_0)^{-1}e^{ix\theta_0}\phi(x)$. The last assertion follows from Theorem 2.3. □

3.2 Sub-Gaussian bounds

Examples such as 2.4 show it is very natural to allow the Fourier transform $\hat{\phi}$ to attain its maximum of 1 at more than one point on $(-\pi, \pi]$. This section explores briefly what global estimates can be obtained in such cases. In the following theorem, $\hat{\phi}$ is allowed to attain its maximum, 1, at finitely many points. At each of this points, we assume that the dominant non-linear term in the expansion of $\log \hat{\phi}$ is of an even degree.

Theorem 3.5. *Let $\phi : \mathbb{Z} \to \mathbb{C}$ be a finitely supported function such that there exists $\theta_q \in I = (-\pi, \pi]$, $q \in \{1, \dots, Q\}$ such that*

$$|\hat{\phi}(\theta_q)| = 1 \text{ and } |\hat{\phi}(\theta)| < 1 \text{ on } I \setminus \{\theta_1, \dots, \theta_Q\}.$$

Assume further that for each $q \in \{1, \dots, Q\}$, there exist $\alpha_q \in \mathbb{R}$, $\gamma_q \in \mathbb{C}$ with $Re(\gamma_q) > 0$, and an even integer ν_q such that

$$\hat{\phi}(\theta_q + \theta) = \hat{\phi}(\theta_q) e^{i\alpha_q \theta - \gamma_q \theta^{\nu_q}(1+o(1))} \text{ at } \theta = 0. \tag{3.11}$$

Then, for each N, there is a constant C_N such that for any $x \in \mathbb{Z}$, $n = 1, 2, \dots$, we have

$$|\phi^{(n)}(x)| \le C_N \sum_1^Q \frac{1}{n^{1/\nu_q}} \left(1 + \frac{|x - \alpha_q n|}{n^{1/\nu_q}}\right)^{-N}. \tag{3.12}$$

Remark. An immediate application of Theorem 3.5 is that there exists a constant C such that, for all n,

$$\sum_x |\phi^{(n)}(x)| \le C.$$

That is, we recover the positive part of the stability theorem, Theorem 2.1. Indeed, estimates such as (3.12) are more or less implicit in the proof of Theorem 2.1 given in [28, 29].

Proof. We need to introduce smooth non-negative cut-off functions ψ_q such that each ψ_q vanishes outside an interval $\theta_q + I_q = [\theta_q - \epsilon_q, \theta_q + \eta_q]$, $\psi_q \equiv 1$ on $[\theta_q - \epsilon_q/2, \theta_q + \eta_q/2]$, $\theta_{q'} \notin [\theta_q - 3\epsilon_q/2, \theta_q + 3\eta_q/2]$, $q' \neq q$, and $\sum_1^Q \psi_q \equiv 1$ (i.e., the ψ_q's form a partition of unity). Using these cut-off functions, we write

$$\begin{aligned}
\phi^{(n)}(x) &= \frac{1}{2\pi} \int_I e^{-ix\theta} \hat{\phi}(\theta)^n d\theta = \sum_1^Q \frac{1}{2\pi} \int_I e^{-ix\theta} \hat{\phi}(\theta)^n \psi_q(\theta) d\theta \\
&= \sum_1^Q \frac{e^{-ix\theta_q} \hat{\phi}(\theta_q)^n}{2\pi} \int_I e^{-i(x-\alpha_q n)\theta} [\hat{\phi}(\theta_q)^{-1} e^{-\alpha_q \theta} \hat{\phi}(\theta_q + \theta)]^n \psi_q(\theta_q + \theta) d\theta \\
&= \sum_1^Q \frac{e^{-ix\theta_q} \hat{\phi}(\theta_q)^n}{2\pi} \int_{I_q} e^{-i(x-\alpha_q n)\theta} P_q(\theta)^n \psi_q(\theta_q + \theta) d\theta \\
&= \sum_1^Q \frac{e^{-ix\theta_q} \hat{\phi}(\theta_q)^n}{2\pi n^{1/\nu_q}} \int_{n^{1/\nu_q} I_q} e^{-i\frac{(x-\alpha_q n)}{n^{1/\nu_q}}\theta} P_{q,n}(\theta) \psi_{q,n}(\theta) d\theta
\end{aligned}$$

where

$$\begin{aligned} P_q(\theta) &= \hat{\phi}(\theta_q)^{-1}e^{-\alpha_q\theta}\hat{\phi}(\theta_q+\theta), \\ P_{q,n}(\theta) &= P_q(\theta/n^{1/\nu_q})^n, \\ \psi_{q,n}(\theta) &= \psi_q(\theta_q+\theta/n^{1/\nu_q}). \end{aligned}$$

By hypothesis,

$$P_q(\theta) = e^{-\gamma\theta^{\nu_q}(1+o(1))} \text{ at } 0.$$

The function P_q can be viewed as an entire function of $z = u + iv \in \mathbb{C}$ which is periodic in u and has at most exponential growth in $|v|$ for large $|v|$. Hence there are constants $\gamma_1, \gamma_2 \in (0,\infty)$ such that

$$|P_q(z)| \le \exp(-\gamma_1 u^{\nu_q} + \gamma_2 v^{\nu_q}) \text{ on } \{z = u + iv : u \in (-3\epsilon_q/2, 3\eta_q/2)\} \tag{3.13}$$

and

$$|P_q(z)| \le \exp(\gamma_2 v^{\nu_q}) \text{ on } \mathbb{C}. \tag{3.14}$$

A priori, the constants γ_1, γ_2 depend on q but since there are only finitely many q, one can assume that they are the same for all q.

Next, for each $q \in \{1,\dots,Q\}$, consider

$$\mathcal{I}_q = \frac{1}{2\pi n^{1/\nu_q}} \int_{n^{1/\nu_q}I} e^{-i\frac{(x-\alpha_q n)}{n^{1/\nu_q}}\theta} P_{q,n}(\theta)\psi_{q,n}(\theta)d\theta.$$

Observe that

$$\begin{aligned} \left(-i\frac{(x-\alpha_q n)}{n^{1/\nu_q}}\right)^k \mathcal{I}_q &= \frac{1}{2\pi n^{1/\nu_q}} \int_{n^{1/\nu_q}I} \partial_\theta^k \left[e^{-i\frac{(x-\alpha_q n)}{n^{1/\nu_q}}\theta}\right] P_{q,n}(\theta)\psi_{q,n}(\theta)d\theta \\ &= \frac{1}{2\pi n^{1/\nu_q}} \int_{n^{1/\nu_q}I_q} e^{-i\frac{(x-\alpha_q n)}{n^{1/\nu_q}}\theta} \partial_\theta^k [P_{q,n}(\theta)\psi_{q,n}(\theta)]d\theta. \end{aligned} \tag{3.15}$$

The iterated integration by parts performed to obtain the second equality does not produce boundary terms thanks to the cutoff function $\psi_{q,n}$. The following proposition is analogous to Proposition 3.2 and the proof is the same.

Proposition 3.6. *There are constants $C_1, A_1, a_1 \in (0,\infty)$ such that for any $q, n = 0, 1, 2\dots$, and $\theta \in (-n^{1/\nu_q}\epsilon_q, n^{1/\nu_q}\eta_q)$,*

$$|\partial_\theta^k P_{q,n}(\theta)| \le C_1^{1+k} k! k^{-k/\nu} \exp\left(-a_1\theta^{\nu_q}\right).$$

Using (3.15) and Proposition 3.6, we obtain

$$\left|\frac{(x-\alpha_q n)}{n^{1/\nu_q}}\right|^k \mathcal{I}_q \le \frac{C_k}{n^{1/\nu_q}}. \tag{3.16}$$

The desired result follows. The necessity to separate the contributions of the different θ_q via the use of cutoff functions prevents us to obtain more precise upper bound in this case. □

4 Carne's transmutation formula

This section develops a version of Carne's transmutation formula to obtain universal long range upper bound on $M_k^n(x,y)$ where M is a (finite range) normal contraction acting on $\ell^2(X,\pi)$. Here, X is a countable space, π a positive measure, M acts on $\ell^2(X,\pi)$ by

$$Mf(x)=\sum_y M(x,y)f(y),$$

k is a fixed integer and

$$M_k=I-(I-M)^k=\sum(-1)^j\binom{k}{j}M^j.$$

If we think of $I-M$ as a "Laplacian", then M_k^n is the discrete semigroup of operators associated with the k-th power of this Laplacian, namely, $(I-M)^k$.

Assume that X is equipped with a metric d such that

$$d(x,y)>1\Longrightarrow M(x,y)=0. \tag{4.1}$$

This is a finite range condition in terms of the metric d. If the matrix M is finite range in the sense that for each x there are finitely many y such that $M(x,y)\neq 0$ then we can define the metric d_M by

$$d_M(x,y)=\inf\{k: M^k(x,y)\neq 0\}.$$

This metric obviously has property (4.1) (we do not have to assume that d_M is finite for all x,y).

Our goal is to obtain bounds showing that $M_k^n(x,y)$ is small when $d(x,y)$ is large when compared to $n^{1/(2k)}$. In the classical case where $k=1$ and M is the transition matrix of a reversible Markov chain, the Carne-Varopoulos estimate states that

$$|M^n(x,y)|\le 2\left(\frac{\pi(x)}{\pi(y)}\right)^{1/2}\exp\left(-\frac{d(i,j)^2}{2n}\right).$$

This bound is remarkable for its generality and explicitness. To see it in action for Markov chains on $\mathbb{Z}^d$, see [3]. For an extension to non-reversible chains, see [19]. For a probabilistic interpretation, see [22].

Our result reads as follows.

Theorem 4.1. *Let (X,d) be a countable metric space equipped with a positive measure π. Fix k and $s_0\in(0,2^{-1+1/k})$. There exist constants $C,c\in(0,\infty)$ (depending on k and s_0) such that, for any normal contraction M on $\ell^2(X,\pi)$ satisfying* (4.1) *and whose spectrum is contained in $[a,1]$ with $a\in[-1,1)$ and $1-a\le 2s_0$,*

$$\forall x,y\in X,\ n=1,\dots,\ \ |M_k^n(x,y)|\le C\left(\frac{\pi(x)}{\pi(y)}\right)^{1/2}\exp\left(-c\left(\frac{d(x,y)}{n^{1/(2k)}}\right)^{\frac{2k}{2k-1}}\right).$$

Remark. Proving that an operator is normal and computing its spectrum is not an easy task, in general. Most application of Theorem 4.1 are likely to involve cases when M is a hermitian contraction. In this case the hypothesis that the spectrum is contained in $[a, 1]$ is a very natural hypothesis since the spectrum is real and contained in $[-1, 1]$.

Proof. Following [6], consider the Chebyshev polynomials

$$Q_m(z) = \frac{1}{2}\left((z + (z^2 - 1)^{1/2})^m + (z - (z^2 - 1)^{1/2})^m\right), \quad m \in \mathbb{Z}.$$

Each Q_m is in fact a polynomial of degree $|m|$. For $a \in [-1, 1]$, let

$$Q_{a,m}(z) = Q_k\left(\frac{2z - 1 - a}{1 - a}\right).$$

These are the Chebyshev polynomials for the interval $[a, 1] \subset [-1, 1]$. See [6, Theorem 2']. As explained in [6, Sect. 2], assuming that the spectrum of M on $\ell^2(X, \pi)$ is contained in $[a, 1]$, it holds that

$$Q_{a,m}(M) : \ell^2(X, \pi) \to \ell^2(X, \pi) \text{ is a contraction}$$

and

$$M^n = \sum_m \mathbf{P}_0(X^a_n = m) Q_{a,m}(M) \tag{4.2}$$

where X^a_n is the simple random walk on $\mathbb{Z}$ with transition probabilities

$$\mathbb{P}(X_n = \pm 1 | X_{n-1}) = (1 - a)/4, \;\; \mathbb{P}(X_n = 0 | X_{n-1}) = (1 + a)/2.$$

Consider the measures β_s on $\mathbb{Z}$ where β is the Bernoulli measure $\beta(\pm 1) = 1/2$ and

$$\beta_s = (1 - s)\delta_0 + s\beta, \;\; s \in [0, 1].$$

Further, set

$$\Delta_s f = f * (\delta_0 - \beta_s), \;\; \psi_{s,k} = \delta_0 - (\delta_0 - \beta_s)^{*k}.$$

By definition,

$$\mathbb{P}_0(X^a_n = m) = \beta^{(n)}_s(m), \;\; s = (1 - a)/2.$$

Also, by (4.2),

$$M^n_k = (I - (I - M)^k)^n = \sum_{m \in \mathbb{Z}} \psi^{(n)}_{s,k}(m) Q_{a,m}(M). \tag{4.3}$$

Similarly,

$$(I - M)^\ell M^n_k = \sum_{m \in \mathbb{Z}} \Delta^\ell_s \psi^{(n)}_{s,k}(m) Q_{a,m}(M). \tag{4.4}$$

Now, let ϕ_i, $i = 1, 2$, be two functions on X with support in $B_i = B(x_i, r_i)$ and assume that $d(B_1, B_2) \geq r$. Then

$$\langle (I - M)^\ell M^n_k \phi_1, \phi_2 \rangle_\pi = \sum_{m \in \mathbb{Z}} \Delta^\ell_s \psi^{(n)}_{s,k}(m) \langle Q_{a,m}(M)\phi_1, \phi_2 \rangle_\pi.$$

Further, since $Q_{a,m}(M)$ is an $\ell^2(X,\pi)$ contraction and a polynomial in M of degree $|m|$,

$$|\langle Q_{a,m}(M)\phi_1,\phi_2\rangle_\pi| \le \|\phi_1\|_2\|\phi_2\|_2$$

and

$$\langle Q_{a,m}(M)\phi_1,\phi_2\rangle_\pi = 0 \text{ if } |m|<r.$$

This last property follows from the hypothesis that $M(x,y)=0$ if $d(x,y)>1$ which implies $M^i(x,y)=0$ is $d(x,y)>i\ge 0$. Putting these properties together, we obtain

$$|\langle (I-M)^\ell M_k^n\phi_1,\phi_2\rangle_\pi| \le \|\phi_1\|_2\|\phi_2\|_2 \sum_{|m|\ge r} |\Delta^\ell \psi_{s,k}^{(n)}(m)|,\ s=(1-a)/2.$$

For $0<s=(1-a)/2\le s_0<2^{-1+1/k}$, the Fourier transform

$$\hat{\psi}_{s,k}(\theta) = 1 - s^k(1-\cos\theta)^k)$$

satisfies $|\hat{\psi}_{s,k}|<1$ on I^* and (3.3) with $\nu=2k$ and $\gamma=(s/2)^k$. One can conclude that there are constants $C,c\in(0,\infty)$ depending only on s such that

$$|\langle (I-M)^\ell M_k^n\phi_1,\phi_2\rangle_\pi| \le \frac{C^\ell\|\phi_1\|_2\|\phi_2\|_2}{(1+n)^{\ell/k}} \exp\left(-c\left(\frac{r}{n^{1/(2k)}}\right)^{\frac{2k}{2k-1}}\right). \tag{4.5}$$

The inequality stated in Theorem 4.1 follows by taking $\phi_1=\delta_x$, $\phi_2=\delta_{x_2}$, $x_1=x, x_2=y$, $r_1,r_2=0$, $r=d(x,y)$. □

Inequality (4.5) contains more information than captured in Theorem 4.1. In particular, we get the following result.

Theorem 4.2. *Let (X,d) be a countable metric space equipped with a positive measure π. Fix k and $s_0\in(0,2^{-1+1/k})$. There exist constants $C,c\in(0,\infty)$ (depending on k and s_0) such that, for any normal contraction M on $\ell^2(X,\pi)$ satisfying* (4.1) *and whose spectrum is contained in $[a,1]$ with $a\in[-1,1)$ and $1-a\le 2s_0$, we have*

$$|(I-M)^\ell M_k^n(x,y)| \le C^\ell\left(\frac{\pi(x)}{\pi(y)}\right)^{1/2}(1+n)^{-\ell/k}\exp\left(-c\left(\frac{d(x,y)}{n^{1/(2k)}}\right)^{\frac{2k}{2k-1}}\right).$$

Further, if $A_x(r,R)=\{y: r<d(x,y)\le R\}$ then

$$\sum_{y\in A_x(r,R)} |(I-M)^\ell M_k^n(x,y)|^2\pi(y) \le \frac{C^\ell \pi(x)^{1/2}}{(1+n)^{\ell/k}}\exp\left(-c\left(\frac{r}{n^{1/(2k)}}\right)^{\frac{2k}{2k-1}}\right).$$

Although these results are far from optimal in the sense that they do not capture the decay of $M_k^n(x,x)$, they do contain some useful information as demonstrated in the following corollary which provide a highly non-trivial lower bound on $M_k^{2n}(x,x)$. This follows closely one of the result obtained in [7] in the reversible Markov case and with $k=1$.

Theorem 4.3. *Let (X,d) be a countable metric space equipped with a positive measure π. Fix k and $s_0 \in (0, 2^{-1+1/k})$. Let M be a hermitian contraction satisfying (4.1), $\sum_y M(x,y) = 1$ and with spectrum contained in $(-2s_0+1, 1]$. Fix $x \in X$, set $V(x,t) = \pi(B(x,t))$ and assume that there exists a positive increasing function v such that*

$$V(x,2t) \le \pi(x)v(t), \;\; v(0) \ge (1+2C)^{1/2}, \tag{4.6}$$

and with

$$t \mapsto t^{\frac{2k}{2k-1}} / \log v(t) \text{ increasing to infinity.} \tag{4.7}$$

Given n, let $r(n)$ be the smallest integer such that

$$5n^{1/(2k-1)} \le c r^{\frac{2k}{2k-1}} / \log v(r). \tag{4.8}$$

Then

$$M_k^{2n}(x,x) \ge \frac{1}{4} \frac{\pi(x)}{V(x,r(n))}.$$

In particular,

1. *If $V(x,t) \le A\pi(x)(1+t)^D$ then*
$$M_k^{2n}(x,x) \ge a\pi(x)[(1+n)\log(1+n)]^{-D/2k}.$$
2. *If $V(x,t) \le \pi(x)\exp(At^\beta)$ with $\beta \in (0, \frac{2k}{2k-1})$,*
$$M_k^{2n}(x,x) \ge \pi(x)\exp\left(-A_0 n^{\frac{\beta(2k-1)}{2k(1-\beta)+\beta}}\right).$$

Remark. The last inequality is informative only when $\beta < 1/(2k-1)$ because $M_k^n(x,x) \ge \pi(x)V(x,n)^{-1}$, always.

Proof. Set $B(x,r) = \{z : d(x,z) \le r\}$ and $m_k^n(x,y) = M_k^n(x,y)/\pi(y)$. Since M is hermitian, that is, $M(x,y)/\pi(y) = \overline{M(y,x)}/\pi(x)$, and $\sum_y M(x,y) = 1$, we have (the first step uses that M is hermitian)

$$\begin{aligned} m_k^{2n}(x,x) &= \sum_z |m_k^n(x,z)|^2\pi(z) \ge \sum_{z\in B(x,r)} |m_k^n(x,z)|^2\pi(z) \\ &\ge \frac{1}{\pi(B(x,r))}\left|\sum_{z\in B(x,r)} m_k^n(x,z)\pi(z)\right|^2 \\ &\ge \frac{1}{\pi(B(x,r))}\left(1 - \sum_{z\in B(x,r)^c} |m_k^n(x,z)|\pi(z)\right)^2. \end{aligned}$$

Next, set $A_q = A(r2^q, r2^{q+1})$ and, using Theorem 4.2, write

$$\begin{aligned}\sum_{z \in B(x,r)^c} |m_k^n(x,z)|\pi(z) &= \sum_{q=0}^{\infty} \sum_{z \in A_q} |M_k^n(x,z)| \\ &\leq C\pi(x)^{1/2} \sum_q \pi(A_q)^{1/2} e^{-c(2^q r)^{\frac{2k-1}{2k}} / n^{1/(2k-1)}} \\ &\leq C\pi(x)^{1/2} \sum_q e^{-c \frac{(2^q r)^{\frac{2k}{2k-1}}}{n^{1/(2k-1)}} + \frac{1}{2} \log[\frac{\pi(B(x, 2^{q+1} r))}{\pi(x)}]}\end{aligned}$$

Using the function v given by (4.6)-(4.7) and any r such that

$$5n^{1/(2k)} \leq cr^{\frac{2k}{2k-1}} / \log v(r),$$

the first term in the series above is bounded by

$$e^{-c \frac{r^{\frac{2k}{2k-1}}}{n^{1/(2k)}} + \frac{1}{2} \log[\frac{\pi(B(x,2r))}{\pi(x)}]} \leq e^{-(9/2) \log v(r)}.$$

Further the ratio of two consecutive terms in the series is bounded by

$$e^{-\frac{c}{2} \frac{(2^{q+1} r)^{\frac{2k}{2k-1}}}{n^{1/(2k)}} + \frac{1}{2} \log v(2^{q+1} r)} \leq e^{-2 \log v(r)}.$$

Hence,

$$\begin{aligned}\sum_{z \in B(x,r)^c} |m_k^n(x,z)|\pi(z) &\leq \frac{Ce^{-(9/2) \log v(r) + 2 \log v(r)}}{e^{2 \log v(r)} - 1} \\ &\leq \frac{C}{v(r)^{1/2}(v(r)^2 - 1)} \leq \frac{C}{v(r)^{1/2}(v(0)^2 - 1)} \leq \frac{1}{2}.\end{aligned}$$

It follows that

$$M_k^{2n}(x,x) \geq \frac{1}{4} \frac{\pi(x)}{\pi(B(x, r(n)))}$$

where $r(n)$ is the smallest integer such that

$$5n^{1/(2k)} \leq cr^{\frac{2k}{2k-1}} / \log v(r).$$

□

5 Applications and history

Iterations of signed kernels are useful in data smoothing, image analysis, density estimation, and summability. There are a variety of applications to partial differential equations via divided difference schemes for numerical computation and asymptotics of higher-order equations. These applications served as motivation behind most of the previously developed theory. In this section, we briefly give pointers to the literature and describe the basic ideas behind these applications.

5.1 Data smoothing, de Forest's problem, and density estimation

If $y_1, \ldots, y_n, \ldots$ is a sequence of numbers, it is standard practice to smooth this sequence by local averaging to eliminate noise. An early instance of this, due to C.S. Pierce, replaced y_i by $y_i' = (y_i + y_{i+1})/2$. This was iterated four times giving $y_i^{(4)} = 2^{-4} \sum_{j=1}^{4} \binom{4}{j} y_{i+j}$. Because the binomial distribution approximates a normal distribution, this is an early version of smoothing with a Gaussian kernel. This story is told by Stigler [26]. The next few paragraphs constitute a review of some more sophisticated examples.

de Forest's problem

It is natural to insist that a smoothing method applied to a constant sequence returns the same constant sequence. For convolution smoothers

$$y_i' = \sum f(i-j) y_j$$

(applied to two-sided sequences) this forces $\sum f(i) = 1$. In early work on smoothers, de Forest asked about smoothers that would preserve low order polynomials. To achieve this, one must allow f to change sign. For example, convolution by f supported on $\{0, \pm 1, \pm 2\}$ with $f(\pm 2) = -1/9$, $f(\pm 1) = 4/9$ and $f(0) = 1/3$ preserves all quadratic polynomials. de Forest's problem refers to the study of the behavior of the iteration of such smoothers. See (5.4) below for this example. A charming history of de Forest with good references is in [25] and hard to find papers are reprinted in [26]. Schoenberg [24] studies de Forest's problem melding it with a discussion of variation diminishing and total positivity. This work is followed up by Greville [15, 16].

Other smoothing techniques

In recent years, a host of techniques using iteration and negative weights have emerged. An early iterative technique called "3RSSH" smooths a series by replacing y_i by the median of y_{i-1}, y_i, y_{i+1} [30]. This is iterated until it is stable. Tukey introduced a smoothing method called "twicing" (see, e.g., [4, Sect. 11] and [12, 27]). This is a linear method in which a series y_i is smoothed with a kernel K and then the residuals $y_i - y_i'$ are smoothed by K and added in. This amounts to smoothing with $2K - K * K$. A further iteration leads to $3K - 3K * K + K * K * K$. More generally, $I - (I - K)^{*k}$ gives the k-fold iterate. This fits exactly into some of questions discussed in Proposition 2.4. Indeed, Tukey's smoothing is based on the Bernoulli measure β (with no holding). Proposition 2.4 shows that the n-th convolution power of $\phi = 2\beta - \beta^{(2)}$ blows up with n. The Bernoulli measure $\beta_{1/2}$ (which has holding equals to 1/2) provides a measure such that the convolution powers $\phi_k^{(n)}$ of $\phi_k = \delta_0 - (\delta_0 - \beta_{1/2})^k$ behaves well for all k.

Density estimation

Given a random sample $y_1, \ldots, y_n$ from an unknown probability density g, the kernel density estimator of g is

$$\hat{g}(x) = \frac{1}{nh} \sum_{1}^{n} K\left(\frac{x - y_i}{h}\right).$$

A huge statistical literature studies the properties of such estimators, their behavior when n is large, choice of h and choice of the kernel K. A standard choice is the Gaussian kernel $K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. A recurring theme known variously as "higher-order kernels" or "super kernels" allows kernels that change sign. A kernel is of order k if

$$\int x^m K(x)dx = \begin{cases} 1 & \text{if } m = 1 \\ 0 & \text{if } m = 1, \dots, k-1 \\ c & \text{if } m = k \quad (c \neq 0). \end{cases}$$

Equivalently (assuming K has high moments), the Fourier transform $\hat{K}$ satisfies $\hat{K}(\xi) = 1 + c\xi^k + o(|\xi|^k)$. The paper [17] develops these ideas in modern language and focuses on

$$K(x) = (2\pi)^{-1} \int \cos(tx) e^{-|t|^k} dt, \qquad \hat{K}(\xi) = e^{-|\xi|^k}, \qquad k \geq 2.$$

For even k, these are exactly the de Forest–Schoenberg–Greville limiting kernels. Many variations of these ideas have been explored and [4, 13] are good sources for this material. The associated density estimators have better mean-square error but, because higher order kernels must have negative values, may give negative estimators for g.

5.2 Partial differential equations

Higher-order evolution equations

Natural scientific problems give rise to higher-order equations such as the initial value problem of solving for $u(x,t)$ satisfying

$$\frac{\partial u}{\partial t} = Hu, \qquad \text{given } u(x,0) = v(x) \text{ for } x \in \mathbb{R}^n,$$

with H, say, a uniformly elliptic operator of order $2m$ densely defined on $L^2(\mathbb{R}^n)$. Under conditions on H, $u(t,x)$ can be expressed as

$$u(t,x) = (e^{Ht}v)(x) = \int_{\mathbb{R}^n} K(t,x,y)v(y)\,dy.$$

If H has constant coefficients, $K(t,x,y) = P_t(x-y) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i(x-y)\xi - p(\xi)t}\,d\xi$ with the polynomial $p(\xi)$ the symbol of H. In the special case that $p(\xi)$ is homogeneous of order $2m$ (so $p(a\xi) = a^{2m}p(\xi)$), $P_t(x) = t^{-n/2m}P(xt^{-1/2m})$ with $P = P_1$. Thus estimates on $P(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix\xi - p(\xi)}\,d\xi$ determine the long-term behavior of K and so $u(t,x)$. This program is explained and carried out in a series of papers by Brian Davies and coauthors [2, 8, 9, 10, 11].

Consider the case where $n = 1$ and $Hu = \frac{\partial^{2k}}{\partial x^{2k}}u$. Then $p(\xi) = \xi^{2k}$ and $P(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi - \xi^{2k}}\,d\xi$. This is exactly $H_{2k}(x)$ of (1.3). It is well known that

$$H_{2k}(x) = \frac{1}{2kn} \sum_{n=0}^{\infty} \frac{(-x)^2}{(2n)!} \frac{\Gamma(2n+1)}{2k}, \quad H_{2k}(0) = \frac{\Gamma\left(\frac{1}{2k}\right)}{2k\pi}, \quad \int_{-\infty}^{\infty} H_{2k}(x) = 1.$$

Further there exists $a, b, c \in (0, \infty)$ such that for all $m = 1, 2, 3, \dots$ and real u,

$$\left| \left(\frac{d}{du} \right)^m H_{2k}(u) \right| \leq c \left(\frac{m}{b} \right)^m e^{-a|u|^{2k/(2k-1)}}.$$

See [14, Chap. 4]. This is also true (with $0^0 = 1$) with $m = 0$. Pictures of H_2 and H_4 are in Figure 1 of Section 1.

Divided difference schemes

Iterated convolutions of complex kernels arise in finite difference approximations to the solutions of partial differential equations. Consider first the simple equation

$$\frac{\partial u}{\partial t} = a \frac{\partial u}{\partial x}; \qquad 0 \leq t \leq T, \quad 0 \leq x \leq 1, \quad u(x, 0) = v(x) \text{ given.} \tag{5.1}$$

In (5.1) a is a given constant and $u(x, t)$ is to be found given v. A finite difference scheme for this equation discretizes time into multiples of Δt and space into multiples of Δx. If $x_i = i\Delta x,\ t_j = j\Delta t$, one simple scheme is

$$\frac{u(x_i, t_{j+1}) - u(x_i, t_j)}{\Delta t} = \frac{a(u(x_{i+1}, t_j) - u(x_i, t_j))}{\Delta x} \tag{5.2}$$

or, with $\lambda = \Delta t / \Delta x$,

$$u(x_i, t_{j+1}) = (1 - a\lambda) u(x_i, t_j) + a\lambda u(x_{i+1}, t_j) = \phi^{(j)} * v(x_i) \tag{5.3}$$

with $\phi(0) = (1 - a\lambda),\ \phi(-1) = a\lambda$. Thus, "running" the differential equation corresponds to repeated convolution of the (perhaps complex) probability measure ϕ. Higher-order constant coefficient equations and more sophisticated difference schemes lead to more complex functions ϕ. A readable introduction to finite difference equations is in [23].

One may approximate even-order derivatives by symmetric differences; writing $h = \Delta x$,

$$\begin{aligned}
f^{(2)}(x) &\doteq \frac{1}{h^2}\{f(x-h) - 2f(x) + f(x+h)\} = \frac{2}{h^2}\left\{\frac{f(x-h)}{2} - f(x) - \frac{f(x+h)}{2}\right\} \\
f^{(4)}(x) &\doteq \frac{1}{h^4}\{f(x-2h) - 4f(x-h) + 6f(x) - 4f(x+h) + f(x+h)\} \\
f^{(2m)}(x) &= \frac{1}{h^{2m}} \sum_{j=0}^{2m} (-1)^j \binom{2m}{j} f(x + (j-m)h).
\end{aligned}$$

Similarly, when $\mu(1) = \mu(-1) = \frac{1}{2},\ \mu(x) = 0$ otherwise, we have

x	-1	0	1
$\delta_0 - \mu$	$-\frac{1}{2}$	1	$-\frac{1}{2}$

and

x	-2	-1	0	1	2
$(\delta_0 - \mu)^{(2)}$	$\frac{1}{4}$	-1	$\frac{3}{2}$	-1	$\frac{1}{4}$

Thus,

$$f^{(2)} = \frac{-2}{h^2}(I - \mu) * f, \quad f^{(4)} = \frac{2^2}{h^4}(I - \mu)^{(2)} * f, \dots, f^{(2m)} = \frac{(-2)^m}{h^m}(\delta_0 - \mu)^{(m)} * f.$$

Now consider the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^{2m}}{\partial x^{2m}} u \qquad \text{with } u(x,0) = v(x). \tag{5.4}$$

Discretizing,

$$\frac{u(x_i, t_{j+1}) - u(x_i, t_j)}{\Delta t} = \frac{(-2)^m}{h^m} (\delta_0 - \mu)^{(m)} * v$$

or

$$u(x_i, t_{j+1}) = \Big(\delta_0 + \lambda(\delta_0 - \mu)^{(m)}\Big) * u(x_i, t_j), \qquad \lambda = (-2)^m \Delta t / h^m.$$

If $h = 2(\Delta t)^{1/m}$, this gives again $\delta_0 - (\delta_0 - \mu)^{(m)} = \phi$, thus again, running the differential equation (5.4) corresponds to repeated convolution by the signed probability ϕ.

A more sophisticated discussion of these ideas together with assorted results can be found in Thomée's survey [29].

Acknowledgement

We thank James Zhao for the useful Figure 1 and Marty Isaacs for useful conversations.

References

[1] D. G. Aronson. On the stability of certain finite difference approximations to parabolic systems of differential equations. *Numer. Math. 5 (1963), 118-137; correction, ibid.*, 5:290, 1963.

[2] G. Barbatis and E. B. Davies. Sharp bounds on heat kernels of higher order uniformly elliptic operators. *J. Operator Theory*, 36(1):179–198, 1996.

[3] Martin T. Barlow and Edwin A. Perkins. Symmetric Markov chains in $\mathbb{Z}^d$: how fast can they move? *Probab. Theory Related Fields*, 82(1):95–108, 1989.

[4] Alain Berlinet and Christine Thomas-Agnan. *Reproducing kernel Hilbert spaces in probability and statistics.* Kluwer Academic Publishers, Boston, MA, 2004. With a preface by Persi Diaconis.

[5] R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions.* John Wiley & Sons, New York-London-Sydney, 1976. Wiley Series in Probability and Mathematical Statistics.

[6] Thomas Keith Carne. A transmutation formula for Markov chains. *Bull. Sci. Math. (2)*, 109(4):399–405, 1985.

[7] Thierry Coulhon and Alexander Grigor'yan. On-diagonal lower bounds for heat kernels and Markov chains. *Duke Math. J.*, 89(1):133–199, 1997.

[8] E. B. Davies. Heat kernel bounds for higher order elliptic operators. In *Journées "Équations aux Dérivées Partielles" (Saint-Jean-de-Monts, 1995)*, pages Exp. No. III, 11. École Polytech., Palaiseau, 1995.

[9] E. B. Davies. Long time asymptotics of fourth order parabolic equations. *J. Anal. Math.*, 67:323–345, 1995.

[10] E. B. Davies. L^p spectral theory of higher-order elliptic differential operators. *Bull. London Math. Soc.*, 29(5):513–546, 1997.

[11] E. B. Davies. Pointwise lower bounds on the heat kernels of higher order elliptic operators. *Math. Proc. Cambridge Philos. Soc.*, 125(1):105–111, 1999.

[12] Luc Devroye. The double kernel method in density estimation. *Ann. Inst. H. Poincaré Probab. Statist.*, 25(4):533–580, 1989.

[13] Luc Devroye and Gábor Lugosi. *Combinatorial methods in density estimation.* Springer Series in Statistics. Springer-Verlag, New York, 2001.

[14] I. M. Gel′fand and G. E. Shilov. *Generalized functions. Vol. 3: Theory of differential equations.* Translated from the Russian by Meinhard E. Mayer. Academic Press, New York, 1967.

[15] T. N. E. Greville. On stability of linear smoothing formulas. *SIAM J. Numer. Anal.*, 3(1):157–170, 1966.

[16] T. N. E. Greville. On a problem of E. L. De Forest in iterated smoothing. *SIAM J. Math. Anal.*, 5:376–398, 1974.

[17] Peter Hall and J. S. Marron. Choice of kernel order in density estimation. *Ann. Statist.*, 16(1):161–173, 1988.

[18] Fritz John. On integration of parabolic equations by difference methods. I. Linear and quasi-linear equations for the infinite interval. *Comm. Pure Appl. Math.*, 5:155–211, 1952.

[19] Pierre Mathieu. Carne-Varopoulos bounds for centered random walks. *Ann. Probab.*, 34(3):987–1011, 2006.

[20] P. Ney and F. Spitzer. The Martin boundary for random walk. *Trans. Amer. Math. Soc.*, 121:116–132, 1966.

[21] V. V. Petrov. *Sums of independent random variables.* Springer-Verlag, New York, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[22] Rémi Peyre. A probabilistic approach to Carne's bound. *Potential Anal.*, 29(1):17–36, 2008.

[23] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical mathematics*, volume 37 of *Texts in Applied Mathematics.* Springer-Verlag, Berlin, second edition, 2007.

[24] I. J. Schoenberg. On smoothing operations and their generating functions. *Bull. Amer. Math. Soc.*, 59:199–230, 1953.

[25] S. Stigler. *American Contributions to Mathemtical Statistics in the Nineteen-Century, Vol I, II.* Arno Press, 1980.

[26] Stephen M. Stigler. Mathematical statistics in the early states. *Ann. Statist.*, 6(2):239–265, 1978.

[27] Werner Stuetzle and Yashaswini Mittal. Some comments on the asymptotic behavior of robust smoothers. In *Smoothing techniques for curve estimation (Proc. Workshop, Heidelberg, 1979)*, volume 757 of *Lecture Notes in Math.*, pages 191–195. Springer, Berlin, 1979.

[28] Vidar Thomée. Stability of difference schemes in the maximum-norm. *J. Differential Equations*, 1:273–292, 1965.

[29] Vidar Thomée. Stability theory for partial difference operators. *SIAM Rev.*, 11:152–195, 1969.

[30] J. W. Tukey. *Exploratory Data Analysis.* Addison-Wesley, 1977.

[31] Olof B. Widlund. On the stability of parabolic difference schemes. *Math. Comp.*, 19:1–13, 1965.