BULLETIN *of the*
Malaysian Mathematical
Sciences Society
http://math.usm.my/bulletin

Bull. Malays. Math. Sci. Soc. (2) **32**(3) (2009), 261–282

Approximation of Gaussian by Scaling Functions and Biorthogonal Scaling Polynomials

S. L. Lee

Department of Mathematics, National University of Singapore,
10 Kent Ridge Road, Singapore 119260
matleesl@nus.edu.sg

Abstract. The derivatives of the Gaussian function, $G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, produce the Hermite polynomials by the relation, $(-1)^m G^{(m)}(x) = H_m(x)G(x)$, $m = 0, 1, \ldots$, where $H_m(x)$ are Hermite polynomials of degree m. The orthonormal property of the Hermite polynomials, $\frac{1}{m!}\int_{-\infty}^{\infty} H_m(x)H_n(x)G(x)dx = \delta_{mn}$, can be considered as a biorthogonal relation between the derivatives of the Gaussian, $\{(-1)^n G^{(n)} : n = 0, 1, \ldots\}$, and the Hermite polynomials, $\{\frac{H_m}{m!} : m = 0, 1, \ldots\}$. These relationships between the Gaussian and the Hermite polynomials are useful in linear scale-space analysis and applications to human and machine vision and image processing. The main objective of this paper is to extend these properties to a family of scaling functions that approximate the Gaussian function and to construct a family of Appell sequences of "scaling biorthogonal polynomials" that approximate the Hermite polynomials.

2000 Mathematics Subject Classification: 41A15, 41A30, 42C05, 42C15

Key words and phrases: Gaussian function, Hermite polynomials, scale-space, normal approximation, B-splines, scaling functions, variation diminishing, causality.

1. Introduction

The Gaussian function, $G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, has many interesting properties and was the genesis of many branches of mathematics, science and engineering. As an icon of statistics, it is the frequency function of the normal distribution. The approximation of the normal distribution by the binomial distributions, an elementary statistical technique that has been used for centuries, is still a motivation for the development of new mathematics. The binomial distributions are defined by the binomial coefficients, $\left\{\frac{1}{2^n}\binom{n}{j} : j = 0, \ldots, n\right\}$. For each n the binomial coefficients also determine, uniquely up to a constant multiple, the uniform B-spline, B_n, of order n, by the

Received: November 25, 2008; *Revised:* March 21, 2009.

equation

$$B_n(x) = 2\sum_{j=0}^{n} \frac{1}{2^n}\binom{n}{j} B_n(2x-j), \quad x \in \mathbb{R}. \tag{1.1}$$

Equation (1.1) is known as a *refinement equation,* B_n is called a *scaling function* and $\{\frac{1}{2^n}\binom{n}{j} : j = 0, \ldots, n\}$ is called its *mask.* Like the binomial distributions, the uniform B-splines, when suitably standardized, approximate the Gaussian function for large n, a process known as *normal approximation* by B-splines. The binomial coefficients and the B-splines, which are related by the refinement equation (1.1), share other properties besides normal approximation. In Physics and Signal Processing, the Gaussian function is known to be optimal in time-frequency localization. More precisely, it is the unique function that attains the uncertainty bound in the Heisenberg's Uncertainty Principle. It is well-known that the scaled Gaussian,

$$G_t(x,y) := \frac{1}{2t} G(x/\sqrt{2t}) G(y/\sqrt{2t}) = \frac{1}{4\pi t} e^{-(x^2+y^2)/4t}, \tag{1.2}$$

is the convolution kernel for the solution of the heat equation, i.e. the convolution, of G_t with a function f,

$$u(x,y,t) := G_t * f(x,y), \tag{1.3}$$

is the solution of the heat equation,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \text{ with initial condition, } u(x,y,0) = f(x,y).$$

This establishes a connection for the heat equation with linear scale-space analysis, an image analysis technique in computer vision that employs the *scale-space operator,* also known as *multiscale operator,* defined by

$$\begin{aligned} M_t f(x,y) &:= G_{t^2/2} * f(x,y) \\ &= \frac{1}{t^2} \int G\left(\frac{y-v}{t}\right) dv \int G\left(\frac{x-u}{t}\right) f(u,v) du, \quad t > 0. \end{aligned} \tag{1.4}$$

In scale-space analysis, $M_t f$ represents the object f, which can be an image, a surface or a geometric object at scale t. In (1.4) the Gaussian function is the *filter* of the scale-space process, which is known as *Gaussian scale-space.* This connection, which was pointed out by Koenderink [23], has led to the development of modern nonlinear scale-space analysis (see [27], [22]). Linear scale-space analysis was introduced earlier, independently, by Iijima [19] (see also [38]) and Witkin [39]. Since the Gaussian function is separable, the Gaussian scale-space operator (1.4) is an iterated application of the one-dimensional filtering,

$$T_t f(x) = \frac{1}{t} \int G\left(\frac{x-u}{t}\right) f(u) du. \tag{1.5}$$

The pervasiveness of the Gaussian function in mathematics, science and technology has motivated the study of Gaussian approximation. The central limit theorem is about Gaussian approximation. It is one of the fundamental theorems in probability and statistics. The binomial approximation of normal distribution is an example of the central limit theorem. The normal approximation of B-splines is another example. Since the B-splines approximate the Gaussian function G, they can also be used as a filter in place of the Gaussian filter in (1.5) for linear scale-space. Wang and Lee [37] introduced this idea and developed a linear scale-space using B-splines. The B-spline scale-space approximates the Gaussian scale-space with filtering at both dyadic and rational scales that provide fast parallel algorithms for computations.

It is shown in [5] that a large class of scaling functions, ϕ_n, that satisfy refinement equations of the form,

$$\phi_n(x) = 2 \sum_{j=0}^{n} a_n(j)\phi_n(2x-j), \quad x \in \mathbb{R}, \tag{1.6}$$

where $a_n(j) \geq 0$, $\sum_{j=0}^{n} a_n(j) = 1$, as well as the distributions of their masks $\{a_n(j)\}_{j=0}^{n}$ converge to G, and various forms of convergence are investigated. This class of scaling functions includes the B-splines, which corresponds to the case when the masks $\{a_n\}$ are the binomial coefficients. The scaling functions not only approximate the Gaussian but also inherit many of its properties. Very general conditions on the locations of the roots of the polynomials $A_n(z) := \sum_{j=0}^{n} a_n(j)z^j$ are also found for various forms of convergence. More precise conditions are also known for different orders of convergence and a family of sequences of scaling functions are identified that converge to the Gaussian faster than the B-splines [5]. Motivated by the fact that the Gaussian is optimal in time-frequency localization, it is shown in [8] and [15] that its approximate scaling functions are asymptotically optimal in time-frequency localization. The Gaussian scale-space operator,

$$T_t f(x) = \frac{1}{t} \int G\left(\frac{x-u}{t}\right) f(u)du,$$

which represents the evolution of the function f with scale t, enjoys the causality property, i.e. no new "features" are introduced as the scale t increases. The Gaussian function is the "unique" linear scale-space kernel that enjoys the causality property. Causality is an important scientific and engineering concept, but there were no rigorous mathematical formulation and understanding. In [14] we give a definition of causality using the concept of variation diminishing, extend the definition, and show that the scale-space operators defined by the approximate scaling functions, in particular by the B-splines, also enjoy the causality property in the extended sense. The Gaussian function and the Hermite polynomials are intimately related. Indeed,

$$(-1)^m G^{(m)}(x) = H_m(x)G(x), \quad m = 0, 1, \ldots, \tag{1.7}$$

where $H_m(x)$ are Hermite polynomials of degree m, and

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)G(x)dx = m!\delta_{mn}. \tag{1.8}$$

These relationships between the Gaussian function and Hermite polynomials are useful for scale-space computation and the main objective of this paper is to extend these relationships to the class of scaling functions that approximate the Gaussian.

Equations (1.7) and (1.8) show that the sequences $\{(-1)^m G^{(m)} : m = 0, 1, \ldots\}$ and $\{H_m : m = 0, 1, \ldots\}$ are biorthogonal. We shall construct sequences of polynomials, $\{P_{n,m} : m = 0, 1, \ldots\}$, associated with the approximating scaling functions $, \phi_n$, for each arbitrary fixed n, in the same way as the Hermite polynomials H_m are associated with the Gaussian function, and show that the sequences $\{P_{n,m} : m = 0, 1, \ldots\}$ and $\{(-1)^m \phi_n^{(m)} : m = 0, 1, \ldots\}$ are biorthogonal and that $P_{n,m}(x) \to H_m(x)$ locally uniformly as $n \to \infty$. The derivatives, $\phi_n^{(m)}$, are, in general, taken in the sense of distribution. In the case when ϕ_n are uniform B-splines, the polynomials, $P_{n,m}, m = 0, 1, \ldots,$ are Bernoulli polynomials of order n.

This investigation is an extension of that for the approximation of the Gaussian function by a class of scaling functions studied in [5]. In Section 2 we give an expository account of the approximation of the Gaussian function by a sequence of discrete distributions and their corresponding scaling functions. Section 3 extends the biorthogonal relation between the derivatives of the Gaussian function and Hermite polynomials to scaling functions and a family of Appell sequences of polynomials, which we call *scaling biorthogonal polynomials* because of their biorthogonal relation, with the derivatives of their generating scaling functions. This family of Appell sequences also includes the Bernoulli polynomials and the Kabaya-Iri polynomials ([20], [21]) . A general algorithm for generating these polynomials is also given in Section 3. In Section 4 we study the class of scaling biorthogonal polynomials that are generated by compactly supported scaling functions that converge to the Gaussian. We show that these scaling biorthogonal polynomials converge to the Hermite polynomials locally uniformly.

2. Normal approximation by probability measures and their scaling functions

2.1. Binomial distributions and uniform B-splines

The uniform B-spline, B_1, of order 1 or degree 0, is the characteristic function of the interval $[0, 1)$, i.e.

$$B_1(x) = \begin{cases} 1, & 0 \le x < 1, \\ 0, & \text{otherwise} . \end{cases}$$

The uniform B-splines, B_n, of order n or degree $n - 1$, is defined recursively for $n = 2, 3, \ldots,$ by convolution:

$$B_n(x) = B_{n-1} * B_1 := \int_{-\infty}^{\infty} B_{n-1}(x - t) B_1(t) dt. \tag{2.1}$$

From the probabilistic point of view, B_1 is the frequency function of a random variable X that is uniformly distributed on the interval $[0, 1)$ and B_n is the frequency function of the sum of n independent copies of X. The function B_n is equal to a polynomial on each interval between two consecutive integers and vanishes outside the interval $[0, n]$. It has continuous derivatives up to order $n - 2$ and its $(n - 1)$-th order derivative is not continuous at the *knots,* $0, 1, \ldots, n$. The uniform B-splines, B_n, were introduced in 1946 by Schoenberg for smoothing and interpolation [29]. They were rediscovered in the early 1970s for computer aided design of free-form curves and surfaces [3] and resurrected in the 1990s in conjunction with wavelets and digital signal processing ([6], [33]). Their discovery and rediscoveries at various stages of technological development are testimonies to their usefulness.

Fourier transform plays an important role in working with B-splines. The Fourier transform of B_1 is

$$\widehat{B}_1(u) = \left(\frac{1-e^{-iu}}{iu}\right), \quad u \in \mathbb{R},$$

and for $n = 2, 3, \ldots,$

$$\widehat{B}_n(u) = \left(\frac{1-e^{-iu}}{iu}\right)^n, \quad u \in \mathbb{R}. \tag{2.2}$$

Expressing

$$\begin{aligned}\widehat{B}_n(u) &= \left(\frac{1+e^{-iu/2}}{2}\right)^n \left(\frac{1-e^{-iu/2}}{i\ u/2}\right)^n \\ &= \left(\frac{1+e^{-iu/2}}{2}\right)^n \widehat{B}_n(u/2),\end{aligned}$$

and taking inverse Fourier transform leads to the refinement equation (1.1), which is the basis for the construction of semi-orthogonal spline wavelets by Chui and Wang [7]. The uniform B-splines B_n are solutions of the refinement equations of the form (1.1) and their distribution functions

$$F_n(x) := \int_{-\infty}^{x} B_n(t)dt, \quad x \in \mathbb{R}$$

satisfy the refinement equations

$$F_n(x) = \sum_{j=0}^{n} \frac{1}{2^n}\binom{n}{j} F_n(2x-j), \quad x \in \mathbb{R}, \tag{2.3}$$

which is slightly different from (1.1). One can work with refinement equations of the form (2.3), but in this paper we work only with (1.1).

It is well-known that the binomial distributions converge to the normal distribution, in the sense that

$$\lim_{n\to\infty} \sum_{k=0}^{[x_n]} \frac{1}{2^n}\binom{n}{k} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\ dt, \tag{2.4}$$

where $x_n = \sqrt{n}x/2 + n/2$. The convergence in (2.4) is known as convergence in distribution. By the central limit theorem, the uniform B-splines also converges in distribution to the normal distribution, i.e.

$$\lim_{n\to\infty} \int_{-\infty}^{x'_n} B_n(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\ dt, \tag{2.5}$$

where $x'_n := \sqrt{n}x/2\sqrt{3} + n/2$. Further, the standardized B-splines,

$$\widetilde{B}_n(x) := \frac{1}{2}\sqrt{\frac{n}{3}} B_n\left(\frac{1}{2}\sqrt{\frac{n}{3}}x + \frac{n}{2}\right).$$

converge uniformly on $\mathbb{R}$ to the Gaussian function (see [4] and [35]). In fact Curry and Schoenberg [4] considered the more general class of Polya frequency functions as limits of non-uniform B-splines with arbitrary knots. On the other hand, Unser, Aldroubi and Eden [35] proved L^p-convergence of $\widetilde{B}_n$ to G for $1 \le p \le \infty$.

The scaling relations (1.1) that B_n enjoy pass over in the limit to the Gaussian function in the integral form:

$$G(x) = \int_{\mathbb{R}} \alpha G(\alpha x - y) dg(y), \quad x \in \mathbb{R}, \tag{2.6}$$

where $\alpha > 1$ is a scale or scaling constant and g is the absolutely continuous measure given by

$$dg(y) = \frac{1}{\sqrt{2\pi}(\alpha^2 - 1)} e^{-y^2/2(\alpha^2-1)} dy.$$

We shall call (2.6) a *continuous refinement equation.* The Gaussian function and its derivatives and the modulated Gaussian have been used extensively in many applications such as scale-space analysis, computer vision and image processing ([1], [24]). The normal approximations of the binomial distributions and the uniform B-splines enable the binomial coefficients and B-splines to replace the Gaussian function in the scale-space representation and vice versa ([24], [36], [37]). The Gaussian function is optimal in time-frequency localization, amenable to statistical analysis, and provides an accurate model of human vision ([13], [40]). This motivates the study of approximation of the Gaussian function by scaling functions that satisfy both the discrete and continuous refinement equations.

2.2. Sequences of refinement equations

In order to integrate the continuous refinement equations with the discrete ones (1.6), we consider a sequence of refinement equations,

$$\phi_n(x) = \int_{\mathbb{R}} \alpha \phi_n(\alpha x - y) dm_n(y), \quad x \in \mathbb{R},\ n = 1, 2, \ldots, \tag{2.7}$$

where $\alpha > 1$ and $\{m_n\}$ is a sequence of probability measures with finite first and second moments. Equivalently, (2.7) can be expressed in terms of Fourier transforms in the frequency domain in the form:

$$\widehat{\phi}_n(u) = \widehat{m}_n \left(\frac{u}{\alpha}\right) \widehat{\phi}_n \left(\frac{u}{\alpha}\right), \quad u \in \mathbb{R}. \tag{2.8}$$

It is known [5] that if m_n has finite mean $\mu(m_n)$ and variance $\sigma(m_n)^2$ for each n, then (2.8) has a unique solution,

$$\widehat{\phi}_n(u) := \prod_{j=1}^{\infty} \widehat{m}_n(u/\alpha^j), \quad u \in \mathbb{R}, \tag{2.9}$$

and ϕ_n is also a probability measure with finite mean $\mu(\phi_n) = (\alpha - 1)^{-1}\mu(m_n)$ and variance $\sigma(\phi_n)^2 = (\alpha^2 - 1)^{-1}\sigma(m_n)^2$. The infinite product in (2.9) converges uniformly on compact sets. We shall call ϕ_n the *m_n-scaling function*, and m_n its *mask*. Note that although ϕ_n is called a function, it is, in general, a probability measure. The problem of determining the nature of ϕ_n for arbitrary scale is a difficult one (see [25]).

If m_n is a discrete measure concentrated on the integers $\mathbb{Z}$ with mass $a_n(j)$ at $j \in \mathbb{Z}$, then (2.7) becomes the discrete scaling equation,

$$\phi_n(x) = \sum_{j\in\mathbb{Z}} \alpha a_n(j)\phi_n(\alpha x - j), \quad x \in \mathbb{R}. \tag{2.10}$$

In particular, if m_n is the discrete measure concentrated on the set $\{0, 1, \ldots, n\}$ with mass $\frac{1}{2^n}\binom{n}{j}$ at $j = 0, 1, \ldots, n$, and scale $\alpha = 2$, then (2.10) reduces to the refinement equation (1.1) for the B-spline of order n.

Suppose that $\{m_n\}$ is a sequence of probability measures on $\mathbb{R}$ with finite mean $\mu(m_n) = \mu_n$ and standard deviation $\sigma(m_n) = \sigma_n$. The standardized measure, $\widetilde{m}_n$, of m_n is defined by

$$\widetilde{m}_n(S) := m_n(\sigma_n S + \mu_n), \quad \text{for measurable } S \subset \mathbb{R},$$

or equivalently,

$$\widehat{\widetilde{m}}_n(u) = e^{iu\mu_n/\sigma_n}\widehat{m}_n(u/\sigma_n), \quad u \in \mathbb{R}. \tag{2.11}$$

The sequence $\{m_n\}$ is said to be *asymptotically normal* if for all $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \int_{-\infty}^{x} d\widetilde{m}_n(t) = \int_{-\infty}^{x} G(t)dt. \tag{2.12}$$

If m_n is absolutely continuous and $dm_n(t) = f_n(t)dt$ with probability density function f_n, we define $\widetilde{f}_n$ by $d\widetilde{m}_n(t) = \widetilde{f}_n(t)dt$, so that

$$\widetilde{f}_n(t) = \sigma_n f_n(\sigma_n t + \mu_n).$$

It is known (see [11], p 249) that asymptotic normality of $\{m_n\}$ is equivalent to

$$\widehat{\widetilde{m}}_n(u) \to e^{-u^2/2} \text{ locally uniformly on } \mathbb{R}. \tag{2.13}$$

The probability measures, m_n, and the corresponding m_n-scaling functions have many properties in common. Here is an example to illustrate their common behaviour.

Theorem 2.1. [5] *Let $\{m_n\}$ be a sequence of probability measures on $\mathbb{R}$ with finite first and second moments and $\{\frac{d}{du}\widehat{\widetilde{m}}_n(u)\}$ be uniformly bounded in a neighbourhood of the origin. Then $\{m_n\}$ is asymptotically normal if and only if the corresponding sequence of m_n-scaling functions, $\{\phi_n\}$, is asymptotically normal.*

Theorem 2.1 can be proved using (2.8) and (2.9) and the fact that asymptotic normality of a sequence of probability measures is equivalent to the uniform convergence on compact sets of their Fourier transforms.

Because of Theorem 2.1, to study the asymptotic normality of scaling functions, we only need to study the asymptotic normality of their masks. We first consider masks that are discrete probability measures. We remark that if $\{m_n\}$ is a sequence of discrete probability measures on $\mathbb{Z}$ with finite first and second moments, then the condition that $\{\frac{d}{dx}\widehat{\widetilde{m}}_n(u)\}$ is uniformly bounded in a neighbourhood of 0 is automatically satisfied. Therefore, $\{m_n\}$ is asymptotically normal if and only if the corresponding sequence of m_n-scaling functions, $\{\phi_n\}$ for any scale $\alpha > 1$, is asymptotically normal.

We want to focus on probability measures with finite support. For each $n = 1, 2, \dots$, let m_n be a discrete probability measure supported on $\{0, 1, \dots, n\}$, defined by $m_n(\{k\}) = a_{n,k}$, $k = 0, 1, \dots, n$. Let

$$A_n(z) := \sum_{k=0}^{n} a_{n,k} z^k = \prod_{j=1}^{n} (z + r_{n,j})/(1 + r_{n,j}), \tag{2.14}$$

be the Z-transform of $\{a_{n,k} : k = 0, 1, \dots, n\}$ or m_n. Note that the roots of $A_n(z)$ are $-r_{n,j}$ and $A_n(1) = 1$ for all n. Taking logarithms of the expressions on the left and right of (2.14) and then differentiating, gives

$$\frac{A_n'(z)}{A_n(z)} = \sum_{j=1}^{n} \frac{1}{z + r_{n,j}}$$

and

$$\frac{A_n''(z)A_n(z) - A_n'(z)^2}{A_n(z)^2} = -\sum_{j=1}^{n} \frac{1}{(z + r_{n,j})^2},$$

from which we have

$$\sum_{k=0}^{n} k a_{n,k} = A_n'(1) = \sum_{j=1}^{n} \frac{1}{1 + r_{n,j}}$$

and

$$\sum_{k=0}^{n} k(k-1) a_{n,k} = A_n''(1) = A_n'(1)^2 - \sum_{j=1}^{n} \frac{1}{(1 + r_{n,j})^2}.$$

Therefore, the mean of m_n is

$$\mu_n = A_n'(1) = \sum_{j=1}^{n} \frac{1}{1 + r_{n,j}}.$$

Its variance,

$$\begin{aligned}
\sigma_n^2 &= \sum_{k=0}^{n} (k - \mu_n)^2 a_{n,k} \\
&= \sum_{k=0}^{n} k^2 a_{n,k} - \mu_n^2 \\
&= \sum_{k=0}^{n} k(k-1) a_{n,k} + \mu_n - \mu_n^2 \\
&= \sum_{j=1}^{n} \frac{1}{1 + r_{n,j}} - \sum_{j=1}^{n} \frac{1}{(1 + r_{n,j})^2} \\
&= \sum_{j=1}^{n} \frac{r_{n,j}}{(1 + r_{n,j})^2}.
\end{aligned}$$

The objective here is to find the distribution of the roots, $-r_{n,j}$, $j = 1, 2, \dots, n$, of $A_n(z)$ for $\{m_n\}$ to be asymptotically normal. This problem is not new. A brief history is given in the next subsection.

2.3. A brief history of asymptotic normality of sequences of combinatorial numbers

The asymptotic normality of sequences of well-known combinatorial numbers, such as Stirling numbers and Eulerian nunbers, have been studied by many authors. A short historical development and descriptions are given below.

(1) Binomial distributions: $m_n(k) = \frac{1}{2^n}\binom{n}{k}$, $k = 0, 1, \ldots, n$, is well-known to be asymptotically normal. Its Z-transform, $A_n(z) = \left(\frac{z+1}{2}\right)^n$ and its roots, $r_{n,j} = -1$ for all $j = 1, 2, \ldots, n$.

(2) Stirling numbers of the first kind: $s(n, k)$ is defined by $s(0, 0) = 1$ and for $n = 1, 2, \ldots,$

$$s(n,k) = s(n-1, k-1) + ns(n-1, k) \quad , \ k = 0, 1, \ldots, n. \tag{2.15}$$

Let $s_n = \sum_{k=0}^{n} s(n, k)$ and define $m_n(k) = s(n, k)/s_n$, $k = 0, 1, \ldots, n$. Harper [17] proved that $\{m_n\}$ is asymptotically normal. It follows from (2.15) that

$$\sum_{k=0}^{n} s(n,k)z^k = (z+n)\sum_{k=0}^{n-1} s(n-1,k)z^k,$$

so that the Z-transform of m_n,

$$A_n(z) = (z+n)(z+n-1)\cdots(z+1)/(n+1)!$$

and its roots are $-1, -2, \ldots, -n$.

(3) Stirling numbers of the second kind: $S(n, k)$ is defined by $S(0, 0) = 1$ and for $n = 1, 2, \ldots,$

$$S(n,k) = S(n-1, k-1) + (k+1)S(n-1, k) \quad , \ k = 0, 1, \ldots, n. \tag{2.16}$$

Again, $m_n(k) := S(n, k)/\sum_{j=0}^{n} S(n, j)$, $k = 0, 1, \ldots, n$, is asymptotically normal. If we let $P_n(z) := \sum_{k=0}^{n} S(n, k)z^k$, then the Z-transform of $\{m_n\}$ is $A_n(z) = P_n(z)/P_n(1)$. It follows from (2.16) that

$$P_n(z) = (1+z)P_{n-1}(z) + zP'_{n-1}(z). \tag{2.17}$$

Using (2.17) and induction, it is easy to prove that all the roots of $P_n(z)$ are real and negative and they interlace with the roots of $P_{n-1}(z)$ for $n = 1, 2, \ldots$.

(4) Eulerian numbers: $A(n, k)$ is defined by $A(0, 0) = 1$ and for $n = 1, 2, \ldots,$

$$A(n,k) = (n-k+1)A(n-1,k-1) + (k+1)A(n-1,k), \qquad k = 0, 1, \ldots, n. \tag{2.18}$$

These numbers have interesting properties. They are probably as old as the binomial coefficients and they feature prominently in the study of cardinal spline functions [30] and the construction of semi-orthogonal spline wavelets [7] because of their connection with the uniform B-splines, B_n. Indeed, the Eulerian numbers are the values of the uniform B-splines at the integers:

$$\frac{A(n,k)}{n!} = B_{n+1}(k), \quad k \in \mathbb{Z}.$$

It was shown by Carlitz, Kurtz, Scoville and Stackelberg [2] that the sequence $\{m_n\}$, where $m_n(k) := A(n,k)/\sum_{j=0}^{n} A(n,j)$, $k = 0,1,\ldots,n$, is asymptotically normal. A similar argument as in the case of Stirling numbers of the second kind shows that the Z-transform, $A_n(z)$, of $\{m_n\}$ satisfies the relation:

$$(n+1)A_n(z) = (1+nz)A_{n-1}(z) + z(1-z)A'_{n-1}(z). \tag{2.19}$$

The polynomials, $A_n(z)$, are known as the *Euler-Frobenius polynomials.* Using (2.18), it is easy to show by induction that $A(n,n-k) = A(n,k)$ for all $n = 1,2,\ldots,$ and $k = 0,1,\ldots,n$, from which we conclude that $A_n(z)$ are reciprocal polynomials, i.e. they satisfy the relation $A_n(z) = z^n A_n(1/z)$, $z \neq 0$. In particular its roots occur in reciprocal pairs, r and $\frac{1}{r}$. Again using (2.19) and induction, it is easy to see that all the roots of $A_n(z)$ are real and negative and they interlace with the roots of $A_{n-1}(z)$ for $n = 1,2,\ldots$. Therefore -1 is a root if and only if n is odd.

(5) Note that in all the examples above the roots of the Z-transforms of $\{m_n\}$ are all real and negative. This observation has led Szekely [32] to prove a more general result that $\{m_n\}$ is indeed asymptotically normal if all the roots of the Z-transforms of $\{m_n\}$ are real and negative for all n.

(6) We now give an example of a sequence of asymptotically normal finitely supported discrete probability measures whose Z-transforms do not have all negative roots. Let us look at the uniform B-splines again for inspiration. Now, we take an arbitrary fixed integer scale $\alpha > 1$ and express

$$\begin{aligned}\widehat{B}_n(u) &= \left(\frac{1-e^{-iu}}{iu}\right)^n \\ &= \left(\frac{1+e^{-iu/\alpha}+\cdots+e^{-i(\alpha-1)u/\alpha}}{\alpha}\right)^n \left(\frac{1-e^{-iu/\alpha}}{i\;u/\alpha}\right)^n \\ &= \left(\frac{1+e^{-iu/\alpha}+\cdots+e^{-i(\alpha-1)u/\alpha}}{\alpha}\right)^n \widehat{B}_n(u/\alpha).\end{aligned}$$

This shows that the uniform B-splines, B_n, are also scaling functions with masks, $\{b_{n,k} : k = 0,1,\ldots,n(\alpha-1)\}$, defined by the multinomial expansion:

$$\left(\frac{1+z+\cdots+z^{\alpha-1}}{\alpha}\right)^n = \sum_{k=0}^{n(\alpha-1)} b_{n,k}z^k, \tag{2.20}$$

for any integer scale $\alpha > 1$. Since the uniform B-splines are asymptotically normal, it follows from Theorem 2.1 that the sequence of probability measures $\{m_n\}$ defined by $m_n(k) = b_{n,k}$, $k = 0,1,\ldots,n(\alpha-1)$, is also asymtotically normal. In this case the Z-transforms of $\{m_n\}$ are the polynomials on the left of (2.20). Note that the roots of these polynomials are the complex α-th roots of unity that are not equal to 1.

2.4. A brief survey of convergence of scaling functions to the Gaussian

The above examples show asymptotically normal sequences of finitely supported discrete probability measures whose Z-transforms have all negative roots or roots that lie on the unit circle and bounded away from 1. The next theorem gives the general result on the distribution of the roots of the Z-transforms for a sequence of finitely supported probability measures to be asymptotically normal. For any $\gamma \in [0, \pi/2)$, we define a region D_γ in the complex plane $\mathbb{C}$, which comprises all $z \in \mathbb{C}$ that satisfy the inequality

$$\left|\mathrm{Im}\left\{\frac{z}{(1+z)^2}\right\}\right| \leq \tan\gamma \ \mathrm{Re}\left\{\frac{z}{(1+z)^2}\right\}. \tag{2.21}$$

Note that for any γ, D_γ contains the positive x-axis and the unit circle except the point $(-1, 0)$. Figure 1 shows the plots of D_γ for $\tan\gamma = 1, 2$ and 8.

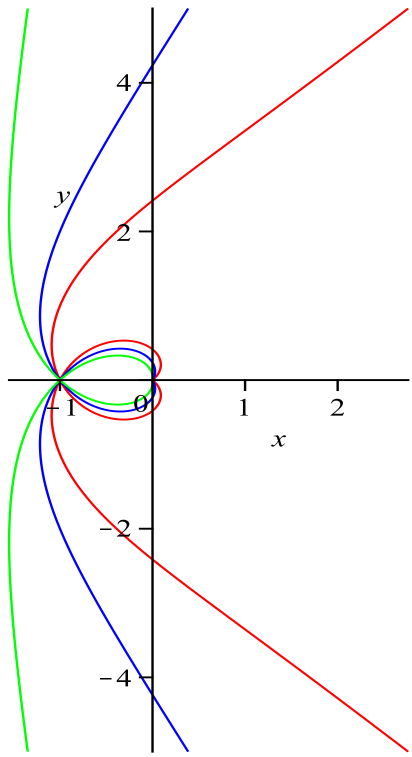

Figure 1. Region D_γ, $\tan\gamma = 1, 2, 8$.

Theorem 2.2. [5] *If for some* $\gamma \in [0, \pi/2)$, $r_{n,j} \in D_\gamma$ *for* $n = 1, 2, \ldots,$ $j = 1, 2, \ldots, n$, *and are bounded away from* -1, *and further*

$$\sigma_n^2 = \sum_{j=1}^{n} r_{n,j}/(1 + r_{n,j})^2 \to \infty \ \textit{as}\ n \to \infty\ , \tag{2.22}$$

then $\{m_n\}$ *is asymptotically normal.*

The proof of Theorem 2.2 in [5] is accomplished by showing the local uniform convergence of the Fourier transforms of m_n. The proof in fact gives estimates of the rate of convergence of $\widehat{m}_n(u)$ to $e^{-u^2/2}$ on compact sets. Because of the relationship (2.9), it also provides estimates on the rates of convergence of the Fourier transforms of the corresponding m_n-scaling functions for any scale $\alpha > 0$. The results are summarized in the following

Theorem 2.3. [5] *Assume the conditions of Theorem 2.2.*

(a) *Then*

$$\|\widehat{\widetilde{\phi}}_n - e^{-(\cdot)^2/2}\|_\infty = O(\sigma_n^{-1}).$$

(b) *If* $\sum_{k=0}^n a_{n,k}z^k$ *is a reciprocal polynomial, i.e.* $a_{n,0} \neq 0$ *and* $a_{n,k} = a_{n,n-k}$, $k = 0, 1, \ldots, n$, *then*

$$\|\widehat{\widetilde{\phi}}_n - e^{-(\cdot)^2/2}\|_\infty = O(\sigma_n^{-2}).$$

(c) *If in addition to the condition in (b),*

$$\sigma_n^{-1} \sum_{j=1}^{n} r_{n,j}(r_{n,j}^2 - 4r_{n,j} + 1)/(1 + r_{n,j})^4 \text{ is bounded,} \tag{2.23}$$

then

$$\|\widehat{\widetilde{\phi}}_n - e^{-(\cdot)^2/2}\|_\infty = O(\sigma_n^{-3}).$$

Note that the m_n-scaling functions ϕ_n for any scale $\alpha > 0$ are, in general, probability measures. Now, we are interested in the case where ϕ_n are L^1-functions. The problem of determining whether they are L^1-functions is a difficult problem (see [25] and the references therein). For the case with scale 2, the conditions for ϕ_n to be an integrable function is known (see [16]). A further restriction that $r_{n,j}$ lie in right half plane, not only ensures that ϕ_n is a sequence of integrable functions, but also provides estimates for the rate of convergence to the Gaussian function in conjunction with Theorem 2.3.

Theorem 2.4. [5] *Assume the conditions of Theorem 2.2, that* $r_{n,j}$ *include* 1 *and all* $Re(r_{n,j}) \geq 0$. *For* $n = 1, 2, \ldots$, *let* ϕ_n *denote the scaling function corresponding to the measure* $m_n(k) = a_{n,k}$, $k = 0, 1, \ldots, n$, *with scale* 2, *and define*

$$\widetilde{\phi}_n(x) = \sigma(\phi_n)\phi_n(\sigma(\phi_n)x + \mu(\phi_n)), \quad x \in \mathbb{R}.$$

(a) *Then*

$$\|\widetilde{\phi}_n - G\|_\infty = O(\sigma_n^{-\frac{1}{2}}).$$

(b) *If* $\sum_{k=0}^n a_{n,k}z^k$ *is reciprocal for large enough* n, *then*

$$\|\widetilde{\phi}_n - G\|_\infty = O(\sigma_n^{-1}).$$

(c) *If, in addition, (2.23) is satisfied, then*

$$\|\widetilde{\phi}_n - G\|_\infty = O(\sigma_n^{-\frac{3}{2}}).$$

Remark 2.1. The uniform B-splines, B_n, satisfy the condition of (b) but not condition (2.23) of (c), because in this case, $r_{n,j} = 1$ for all $j = 1, 2, \ldots, n$ and $\sigma_n = \sqrt{n/12}$. Therefore, sequences of scaling functions that satisfy condition (2.23) converge faster to the Gaussian than the uniform B-splines.

3. Scaling functions and biorthogonal polynomials

3.1. Appell sequences of biorthogonal polynomials generated by scaling functions

Recall that the Gaussian function G is a refinable function that satisfies the continuous refinement equation (2.6) and its m-th order derivative is related to the Hermite polynomial H_m of degree m by the relation $(-1)^m G^{(m)}(x) = H_m(x)G(x)$. Further, the orthogonality of the Hermite polynomials with respect to weight G, can be written in the form:

$$\langle (-1)^m G^{(m)}, H_n(x)/n! \rangle = \delta_{mn}, \tag{3.1}$$

where $\langle \cdot, \cdot \rangle$ denotes inner product in $L^2(\mathbb{R})$. We refer to (3.1) by saying that the sequences $\{(-1)^m G^{(m)} : m = 0, 1, \ldots\}$ and $\{H_m/m! : m = 0, 1, \ldots\}$ are biorthogonal. We want to extend this relationship to scaling functions with compact support that approximate the Gaussian and construct a family of Appell sequences of polynomials that approximate the Hermite polynomials. To do this we need to consider distributions with compact support (see [41]). These are continuous linear functionals defined on the space, $C^\infty(\mathbb{R})$, of infinitely differentiable functions. If $\phi : C^\infty(\mathbb{R}) \to \mathbb{R}$ is a linear functional we shall write $\langle \phi, v \rangle = \phi(v)$, $v \in C^\infty(\mathbb{R})$. The linear functional ϕ is continuous if and only if there is a compact subset K of $\mathbb{R}$, a constant $C > 0$ and an integer $k \geq 0$ such that

$$|\langle \phi, v \rangle| \leq C \max_{j \leq k} \sup_{x \in K} \left| v^{(j)}(x) \right|.$$

We denote the space of distributions with compact support by $\mathcal{E}'(\mathbb{R})$. Integrable functions and measures with compact supports belong to $\mathcal{E}'(\mathbb{R})$. If f is a compactly supported integrable function then it is associated with the distribution, which we still denote by f, defined by

$$\langle f, v \rangle := \int_{\mathbb{R}} v(x) f(x) dx, \quad v \in C^\infty(\mathbb{R}).$$

If m is a compactly supported measure on $\mathbb{R}$, then it is associated with the distribution, which we still denote by m, defined by

$$\langle m, v \rangle := \int_{\mathbb{R}} v(x) dm(x), \quad v \in C^\infty(\mathbb{R}).$$

Any $\phi \in \mathcal{E}'(\mathbb{R})$ has derivatives $\phi^{(n)}$ of any order n and they are defined by

$$\langle \phi^{(n)}, v \rangle = (-1)^n \langle \phi, v^{(n)} \rangle, \quad n = 0, 1, \ldots.$$

Take a compactly supported distribution $\phi \in \mathcal{E}'(\mathbb{R})$, not necessarily a scaling function. Then for any integer $n \geq 0$,

$$\langle \phi^{(n)}, e^{(\cdot)z} \rangle = (-1)^n \langle \phi, z^n e^{(\cdot)z} \rangle = (-1)^n z^n \widehat{\phi}(iz).$$

If $\widehat{\phi}(0) \neq 0$,

$$\left\langle (-1)^n \phi^{(n)}, \frac{e^{(\cdot)z}}{\widehat{\phi}(iz)} \right\rangle = z^n \tag{3.2}$$

in a neighbourhood of 0. Since ϕ is compactly supported, $\widehat{\phi}$ is analytic. So we can define a sequence of polynomials, P_m, by the generating function

$$\frac{e^{xz}}{\widehat{\phi}(iz)} = \sum_{m=0}^{\infty} \frac{P_m(x)}{m!} z^m \ . \tag{3.3}$$

It follows from (3.2) and (3.3) that for any integer $n \geq 0$,

$$z^n = \sum_{m=0}^{\infty} \left\langle (-1)^n \phi^{(n)}, \frac{P_m(x)}{m!} \right\rangle z^m \ ,$$

which gives the biorthogonal relation

$$\left\langle (-1)^n \phi^{(n)}, \frac{P_m(x)}{m!} \right\rangle = \delta_{m,n} \ . \tag{3.4}$$

Now, let ϕ be the compactly supported solution (a probability measure) of the refinement equation,

$$\phi(x) = \int_{\mathbb{R}} \alpha\phi(\alpha x - y) d\psi(y) \tag{3.5}$$

where ψ is a compactly supported probability measure on $\mathbb{R}$ and $\alpha > 1$ a fixed number. Since ϕ and ψ are compactly supported, they have finite moments of any order. Equation (3.5) can be equivalently expressed in terms of Fourier transforms in the form:

$$\widehat{\phi}(u) = \widehat{\psi}(u/2)\widehat{\phi}(u/2), \quad u \in \mathbb{R}. \tag{3.6}$$

Then ϕ is the unique (up to a multiple) fixed point of the *scaling operator*,

$$(T_{\psi,\alpha} f)(x) := \int_{\mathbb{R}} \alpha f(\alpha x - t) d\psi(t), \quad f \in \mathcal{E}'(\mathbb{R}). \tag{3.7}$$

The Fourier transform of ϕ is given by

$$\widehat{\phi}(u) = \prod_{j=1}^{\infty} \widehat{\psi}(u/\alpha^j), \tag{3.8}$$

where the infinite product converges locally uniformly (see [5], [6] or [9]). Taking derivatives (in the sense of distribution) on both sides of (3.5) gives

$$T_{\psi,\alpha}\phi^{(n)} = \alpha^{-n}\phi^{(n)}, \quad n = 0, 1, 2, \ldots, \tag{3.9}$$

which shows that the derivatives $\phi^{(n)}$ are eigenfunctions of $T_{\psi,\alpha}$ with eigenvalues α^{-n}.

The scaling operator $T_{\psi,\alpha}$ is defined on the space of compactly supported distributions. Its adjoint, $T^*_{\psi,\alpha}$, is defined on the space of test functions $C^\infty(\mathbb{R})$ by

$$(T^*_{\psi,\alpha} g)(x) = \int_{\mathbb{R}} g\left(\frac{x+t}{\alpha}\right) d\psi(t). \tag{3.10}$$

The adjoint relationship is realized by the usual action of a distribution on a test function, i.e. $\langle T_{\psi,\alpha} f, g\rangle = \langle f, T^*_{\psi,\alpha} g\rangle$. The polynomials P_m that satisfy the biorthogonal relation (3.4) are eigenfunctions of $T^*_{\psi,\alpha}$ with eigenvalues α^{-m}. Some stochastic properties of the biorthogonal sequences are given in [12].

Proposition 3.1. *Let ψ be a compactly supported probability measure on $\mathbb{R}$, ϕ the scaling function with mask ψ and scale $\alpha > 1$, P_m the sequence of polynomials generated by (3.3) and*

$$\widehat{\psi}(iz) =: \sum_{k=0}^{\infty} b_k z^k.$$

Then

$$P_m(\alpha x) = m! \sum_{k=0}^{m} b_{m-k} \alpha^k \frac{P_k(x)}{k!}, \quad m = 0, 1, \ldots, \tag{3.11}$$

and

$$P_m(x) = \sum_{k=0}^{m} P_{m-k}(0) \binom{m}{k} x^k, \quad m = 0, 1, \ldots\ , \tag{3.12}$$

where $P_m(0)$ are generated recursively by

$$P_m(0) = \frac{m!}{1-\alpha^m} \sum_{k=0}^{m-1} b_{m-k} \alpha^k \frac{P_k(0)}{k!}, \quad m = 1, 2 \ldots, \tag{3.13}$$

with $P_0(0) = 1$.

Proof. By (3.3) and (3.6), we have

$$\frac{e^{xz}}{\widehat{\psi}(iz/\alpha)\widehat{\phi}(iz/\alpha)} = \sum_{m=0}^{\infty} \frac{P_m(x)}{m!} z^m\ .$$

Then

$$\begin{aligned}
\sum_{m=0}^{\infty} \frac{P_m(\alpha x)}{m!} \frac{z^m}{\alpha^m} &= \frac{e^{xz}}{\widehat{\phi}(iz/\alpha)} \\
&= \widehat{\psi}(iz/\alpha) \sum_{m=0}^{\infty} \frac{P_m(x)}{m!} z^m \\
&= \left(\sum_{k=0}^{\infty} b_k \frac{z^k}{\alpha^k} \right) \left(\sum_{m=0}^{\infty} \frac{P_m(x)}{m!} z^m \right) \\
&= \sum_{m=0}^{\infty} \left\{ \sum_{k=0}^{m} \frac{b_k}{\alpha^k} \frac{P_{m-k}(x)}{(m-k)!} \right\} z^m\ .
\end{aligned}$$

Hence,

$$\frac{P_m(\alpha x)}{\alpha^m m!} = \sum_{k=0}^{m} \frac{b_k}{\alpha^k} \frac{P_{m-k}(x)}{(m-k)!}\ , \quad m = 0, 1, \ldots\ ,$$

which gives (3.11).

Differentiating (3.3) with respect to x and equating coefficients of z^m in the resulting equation gives

$$P_m'(x) = m P_{m-1}(x), \quad m = 1, 2, \ldots\ , \tag{3.14}$$

and hence

$$P_m(x) = m\int_0^x P_{m-1}(t)dt + P_m(0),$$

which leads to (3.12) by induction. The relation (3.13) follows from (3.11) . ∎

Remark 3.1. The relation (3.11) corresponds to the refinement equation (3.5) for the scaling function. We shall call the sequence of biorthogonal polynomials P_m that satisfy (3.11) *scaling biorthogonal polynomials.*

Remark 3.2. The relation (3.14) is referred to by saying that the sequence $P_m/m!$, $m = 0, 1, \dots$, is an *Appell sequence.*

Remark 3.3. The generating function (3.3) cannot be used directly to generate the polynomials P_m when $\widehat{\phi}(iz)$ does not have a closed form, which is generally the case. The relations (3.12) and (3.13) provide an algorithm for generating P_m.

3.2. Examples of Appell sequences of biorthogonal polynomials

We use the algorithm defined by (3.12) and (3.13) to compute the biorthogonal polynomials for some simple scaling functions.

Example 3.1. Let ψ_p be the discrete probability measure supported on $\{0, 1\}$ with weights q at 0 and p at 1, where $0 < p < 1$ and $p + q = 1$. Then $\widehat{\psi}_p(u) = q + pe^{-iu}$ and the Fourier transform of the refinable function ϕ_p with mask ψ_p and scale 2,

$$\widehat{\phi}_p(u) = \prod_{j=1}^{\infty}(q + pe^{-iu/2^j}),$$

which does not have a close form except when $p = q = 1/2$. The scaling function ϕ_p generates an Appell sequence $\{P_{p,m} : m = 0, 1, \dots\}$ by (3.3). Since

$$\psi(iz) = q + pe^z = 1 + \sum_{k=1}^{\infty}\frac{p}{k!}z^k,$$

$b_0 = 1$ and $b_k = \frac{p}{k!}$, $k = 1, 2, \dots$, so that

$$P_{p,m}(x) = \sum_{k=0}^{m} P_{p,m-k}(0)\binom{m}{k}x^k, \quad m = 0, 1, \dots\ , \tag{3.15}$$

where

$$P_{p,m}(0) = \frac{p}{1-2^m}\sum_{k=0}^{m-1}\binom{m}{k}2^k P_{p,k}(0), \quad m = 1, 2\dots, \tag{3.16}$$

with $P_{p,0}(0) = 1$.

When $p = q = 1/2$,

$$\widehat{\phi}_{1/2}(u) = \prod_{j=1}^{\infty}\frac{1+e^{-iu/2^j}}{2}$$
$$= \lim_{n\to\infty}\prod_{j=1}^{n}\frac{1+e^{-iu/2^j}}{2}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \prod_{j=1}^{n} \frac{1-e^{-iu/2^{j-1}}}{2(1-e^{-iu/2^j})} \\
&= \lim_{n\to\infty} \frac{1-e^{-iu}}{2^n(1-e^{-iu/2^n})} \\
&= \frac{1-e^{-iu}}{iu},
\end{aligned}
$$

which shows that $\phi_{1/2} = B_1$, the uniform B-spline of order 1. Since $\widehat{\phi}_{1/2}(iz) = \frac{e^z-1}{z}$, (3.3) becomes

$$
\frac{ze^{xz}}{e^z-1} = \sum_{m=0}^{\infty} \frac{P_{1/2,m}(x)}{m!} z^m \ , \tag{3.17}
$$

which shows that when $p = 1/2$, $P_{1/2,m}$ are Bernoulli polynomials.

Example 3.2. Up-function and Kabaya-Iri polynomials.
Let ψ be the probability measure for the probability distribution with density function B_1, the uniform B-spline of order 1. The refinable function ϕ with mask ψ and scale 2 is called the *up-function* by its discoverer, Rvachev [28]. It was independently discovered by Kabaya and Iri ([20], [21]), who also studied the corresponding biorthogonal polynomials. The up-function is infinitely differentiable with compact support. It is the solution of a continuous refinement equation and so is the Gaussian. The up-function and the Gaussian belong to a large class of refinable functions studied in [18].

We shall compute the Kabaya-Iri scaling polynomials that are biorthogonal to the derivatives of the up-function, using the algorithm in (3.12) and (3.13). Now, $\widehat{\psi}(u) = \frac{1-e^{-iu}}{iu}$ gives

$$
\widehat{\psi}(iz) = \frac{e^z-1}{z} = \sum_{k=0}^{\infty} \frac{z^k}{(k+1)!},
$$

so that

$$
b_k = \frac{1}{(k+1)!}, \quad k = 0, 1, \ldots
$$

and $P_m(x)$ are given in (3.12), where

$$
P_m(0) = \frac{1}{(1-2^m)(m+1)} \sum_{k=0}^{m-1} \binom{m+1}{k} 2^k P_k(0), \quad m = 1, 2 \ldots, \tag{3.18}
$$

with $P_0(0) = 1$.

The first 6 polynomials are

$$
\begin{aligned}
P_0(x) &\equiv 1 \\
P_1(x) &= -\frac{1}{2} + x \\
P_2(x) &= \frac{2}{9} - x + x^2 \\
P_3(x) &= -\frac{1}{12} + \frac{2}{3}x - \frac{3}{2}x^2 + x^3
\end{aligned}
$$

$$P_4(x) = \frac{16}{675} - \frac{1}{3}\,x + \frac{4}{3}\,x^2 - 2\,x^3 + x^4$$
$$P_5(x) = -\frac{1}{270} + \frac{16}{135}\,x - \frac{5}{6}\,x^2 + \frac{20}{9}\,x^3 - \frac{5}{2}\,x^4 + x^5.$$

4. Hermite polynomials as limits of scaling biorthogonal polynomials

In this section we consider a family of sequences of polynomials $\{P_{N,m} : m = 0, 1, \ldots, \}$, $N = 1, 2, \ldots$, that are generated by a sequence of scaling functions, ϕ_N, which converges to the Gaussian function G. We assume that the scaling functions satisfy the sequence of refinement equations,

$$\phi_N(x) = \sum_{k=0}^{N} a_{N,k}\phi_N(\alpha x - k) \tag{4.1}$$

with mask $\{a_{N,k} : k = 0, 1, \ldots, N\}$ whose Z-transforms, $A_N(z) := \sum_{k=0}^{N} a_{N,k} z^k$, satisfy the conditions of Theorem 2.3 so that the standardized scaling functions converge to the standard Gaussian. Let $P_{N,m}$, $m = 0, 1, \ldots,$ be the sequence of scaling biorthogonal polynomials generated by ϕ_N as in (3.3), i.e.

$$\frac{e^{xz}}{\widehat{\phi}_N(iz)} = \sum_{m=0}^{\infty} \frac{P_{N,m}(x)}{m!} z^m \ . \tag{4.2}$$

Let $\widetilde{\phi}_N$ be the standardized form of ϕ_N, i.e.

$$\widetilde{\phi}_N(x) = \sigma_N \phi_N(\sigma_N x + \mu_N),$$

where μ_N and σ_N^2 are the mean and variance of ϕ_N. Define the standardized form of the scaling biorthonormal polynomials, $\widetilde{P}_{N,m}$, of $P_{N,m}$, $m = 0, 1, \ldots,$ by

$$\widetilde{P}_{N,m}(x) = \sigma_N^{-m} P_{N,m}(\sigma_N x + \mu_N). \tag{4.3}$$

Then the following biorthogonal relations for the standardized scaling functions and the scaling biorthogonal polynomials follow from (3.4):

$$\langle (-1)^n \widetilde{\phi}_N^{(n)}, \widetilde{P}_{N,m} \rangle = \delta_{m,n} \quad \forall\ m, n \geq 0.$$

Further, the generating functions of $\widetilde{P}_{N,m}$ are given by

Lemma 4.1.

$$\frac{e^{xz}}{\widehat{\widetilde{\phi}}_N(iz)} = \sum_{m=0}^{\infty} \frac{\widetilde{P}_{N,m}(x)}{m!} z^m. \tag{4.4}$$

Proof. By (4.3),

$$\begin{aligned} \sum_{m=0}^{\infty} \frac{\widetilde{P}_{N,m}(x)}{m!} z^m &= \sum_{m=0}^{\infty} \frac{P_{N,m}(\sigma_N x + \mu_N)}{m!} \left(\frac{z}{\sigma_N}\right)^m \\ &= \frac{e^{\frac{z}{\sigma_N}(\sigma_N x + \mu_N)}}{\widehat{\phi}_N(iz/\sigma_N)} \ . \end{aligned} \tag{4.5}$$

Noting that

$$\widehat{\widetilde{\phi}}_N(iz) = e^{-z\frac{\mu_N}{\sigma_N}} \widehat{\phi}_N(iz/\sigma_N),$$

we see that (4.5) gives (4.4) . ∎

The Hermite polynomials, H_m, are generated by

$$\frac{e^{xz}}{e^{z^2/2}} = \sum_{m=0}^{\infty} \frac{H_m(x)}{m!} z^m. \tag{4.6}$$

If the conditions of Theorem 2.3 are satisfied, the estimates (a), (b), (c) in the Theorem can be extended to the complex plane. In particular, there are constants, $r > 0$ and $A > 0$, such that for all sufficiently large N,

$$\left|\widehat{\widetilde{\phi}}_N(iz) - e^{z^2/2}\right| \le \frac{A}{\sigma_N}, \quad |z| < r. \tag{4.7}$$

We shall prove the following

Theorem 4.1. *Let ϕ_N be the scaling functions with masks $\{a_{N,k} : k = 0, 1, \ldots, N\}$, $N = 1, 2, \ldots,$ which satisfy the conditions of Theorem 2.3 and let $\{\widetilde{P}_{N,m} : m = 0, 1, \ldots\}$ be the scaling biorthogonal polynomials generated by the standardized scaling functions, $\widetilde{\phi}_N$, as in (4.4). Then for each $m = 0, 1, \ldots,$ $\widetilde{P}_{N,m}(x)$ converges locally uniformly to the Hermite polynomial $H_m(x)$ as $N \to \infty$.*

Proof. By (4.4) and (4.6),

$$\frac{e^{xz}}{\widehat{\widetilde{\phi}}_N(iz)} - \frac{e^{xz}}{e^{z^2/2}} = \sum_{m=0}^{\infty} \frac{(\widetilde{P}_{N,m}(x) - H_m(x))}{m!} z^m. \tag{4.8}$$

Since $\widehat{\widetilde{\phi}}_N(0) = 1$, we can choose a neighbourhood U of the origin so that $|\widehat{\widetilde{\phi}}_N(iz)| \ge 1/2$ and $|e^{z^2/2}| \ge 1/2$ for all $z \in U$. Take a circle C inside U with centre at 0 and radius r so that (4.7) is satisfied. The coefficients of the Taylor series (4.8) are represented by the Cauchy's formula:

$$\widetilde{P}_{N,m}(x) - H_m(x) = \frac{m!}{2\pi i} \oint_C \frac{e^{xz}(e^{z^2/2} - \widehat{\widetilde{\phi}}_N(iz))}{z^{m+1}\widehat{\widetilde{\phi}}_N(iz)e^{z^2/2}} dz.$$

Therefore,

$$\begin{aligned} \left|\widetilde{P}_{N,m}(x) - H_m(x)\right| &\le \frac{m!}{2\pi} \oint_C \frac{|e^{xz}||e^{z^2/2} - \widehat{\widetilde{\phi}}_N(iz)|}{r^{m+1}|\widehat{\widetilde{\phi}}_N(iz)||e^{z^2/2}|} |dz| \\ &\le \frac{m!}{2\pi} \oint_C \frac{e^{x Re(z)} A/\sigma_N}{r^{m+1}|\widehat{\widetilde{\phi}}_N(iz)||e^{z^2/2}|} |dz| \quad \text{by (4.7)} \\ &\le \frac{4(m!)Ae^{rx}}{r^m \sigma_N}. \end{aligned}$$

Since $\sigma_N \to \infty$ as $N \to \infty$, it follows that for each m, $\widetilde{P}_{N,m}(x) \to H_m(x)$ uniformly on compact sets. ∎

4.1. Uniform B-splines and Bernoulli polynomials

Recall that the uniform B-spline, B_N, of order N, is the refinable function with mask $\{a_{N,k} : k = 0, 1, \ldots, N\}$, where $a_{N,k} = \binom{N}{k}$, $k = 0, 1, \ldots, N$, and scale 2. The Z-transform of its mask, $A(z) = \left(\frac{1+z}{2}\right)^N$, and from the Fourier transform of B_N we have

$$\widehat{B_N}(iz) = \left(\frac{e^z - 1}{z}\right)^N .$$

The scaling biorthogonal polynomials, $B_{N,m}$, $m = 0, 1, \ldots$, that are biorthogonal to the derivatives, $B_N^{(n)}$, of the B-splines are generated by

$$\frac{z^N e^{xz}}{(e^z - 1)^N} = \sum_{n=0}^{\infty} \frac{B_{N,m}(x)}{m!} z^m , \tag{4.9}$$

which shows that $B_{N,m}$, $m = 0, 1, \ldots$, are *Bernoulli polynomials of order* N (see [10] ,[26]).

The standardized B-splines,

$$\widetilde{B}_N(x) = \sigma_N B_N(\sigma_N x + N/2),$$

where $\sigma_N = \sqrt{N/12}$, converges uniformly to the Gaussian function G and an estimate of the rate of convergence is given in Theorem 2.4. By Theorem 4.1, the standardized Bernoulli polynomials,

$$\widetilde{B}_{N,m}(x) := \sigma_N^{-m} B_{N,m}(\sigma_N x + N/2), \quad m = 0, 1, \ldots,$$

converges locally uniformly to the Hermite polynomial $H_m(x)$ for each m.

The standardized Bernoulli polynomials of order N are generated by

$$\frac{e^{xz}}{\widehat{\widetilde{B}}_N(iz)} = \sum_{m=0}^{\infty} \frac{\widetilde{B}_{N,m}(x)}{m!} z^m.$$

Acknowledgement. My thanks to the referees for many helpful suggestions.

References

[1] J. Babaub, A. P. Witkin, M. Baudin and R. O. Duda, *Uniqueness of Gaussian kernel for scale-space filtering,* IEEE Trans. Pattern Analysis and Machine Intelligence, #1, **8** (1986), 26–33.

[2] L. Carlitz et al., Asymptotic properties of Eulerian numbers, *Z. Wahrsch. Verw. Gebiete* **23** (1972), 47–54.

[3] G. M. Chaikin, An algorithm for high speed curve generation, *Comp. Graphics and Image Proc. 3* (1974), 346–349.

[4] H. B. Curry and I. J. Schoenberg, On Pólya frequency functions. IV. The fundamental spline functions and their limits, *J. Analyse Math.* **17** (1966), 71–107.

[5] L. H. Y. Chen, T. N. T. Goodman and S. L. Lee, Asymptotic normality of scaling functions, *SIAM J. Math. Anal.* **36** (2004), no. 1, 323–346 (electronic).

[6] C. K. Chui, *An Introduction to Wavelets*, Academic Press, Boston, MA, 1992.

[7] C. K. Chui and J. Wang, On compactly supported spline wavelets and a duality principle, *Trans. Amer. Math. Soc.* **330** (1992), no. 2, 903–915.

[8] C. K. Chui and J. Z. Wang, A study of asymptotically optimal time-frequency localization by scaling functions and wavelets, *Ann. Numer. Math.* **4** (1997), no. 1-4, 193–216.

[9] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[10] A. Di Bucchianico, *Probabilistic and Analytical Aspects of the Umbral Calculus*, Math. Centrum, Centrum Wisk. Inform., Amsterdam, 1997.

[11] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. II*, Second edition, Wiley, New York, 1971.

[12] X. Gao , S. L. Lee and Q. Sun, *Stochastic multiscale operators and Appell sequences,* Proc. Second International Congress of Chinese Mathematicians, Chang-Shou Lin, Lo Yang, Shing Tung Yau (eds.) , International Press, 2001, pp. 569–573.

[13] M. A. Georgeson, K. A. May, T. C. Freeman and G. S. Hesse, *From filters to features: Scale–space analysis of edge and blur coding in human vision* , Journal of Vision #13, **7** (2007), 1–21.

[14] S. S. Goh, T. N. T. Goodman and S. L. Lee, Causality properties of refinable functions and sequences, *Adv. Comput. Math.* **26** (2007), no. 1-3, 231–250.

[15] T. N. T. Goodman and S. L. Lee, Asymptotic optimality in time-frequency localization of scaling functions and wavelets, in *Frontiers in Interpolation and Approximation*, 145–171, Chapman & Hall/CRC, Boca Raton, FL.

[16] T. N. T. Goodman and C. A. Micchelli, On refinement equations determined by Pólya frequency sequences, *SIAM J. Math. Anal.* **23** (1992), no. 3, 766–784.

[17] L. H. Harper, Stirling behavior is asymptotically normal, *Ann. Math. Statist.* **38** (1967), 410–414.

[18] R. Q. Jia, S. L. Lee and A. Sharma, Spectral properties of continuous refinement operators, *Proc. Amer. Math. Soc.* **126** (1998), no. 3, 729–737.

[19] T. Iijima, *Basic theory of pattern observation,* Papers of Technical Group on Automata and Automatic Control, IECE, Japan, Dec. 1999 (in Japanese).

[20] K. Kabaya and M. Iri, Sum of uniformly distributed random variables and a family of nonanalytic C^∞-functions, *Japan J. Appl. Math.* **4** (1987), no. 1, 1–22.

[21] K. Kabaya-Imai and M. Iri, On operators defining a family of nonanalytic C^∞-functions, *Japan J. Appl. Math.* **5** (1988), no. 3, 333–365.

[22] R. Kimmel, N. Sochen and J. Weickert (eds.), *Scale-space and PDE methods in computer vision, 5th International Conference,* in Special Issue of International Journal of Computer Vision, #3, 70(2006), 195–298.

[23] J. J. Koenderink, The structure of images, *Biol. Cybernet.* **50** (1984), no. 5, 363–370.

[24] T. A. Poggio, V. Torre, and C. Koch, *Computational vision and regularization theory,* Nature, 317(1985), 314–319.

[25] V. Protasov, Refinement equations with nonnegative coefficients, *J. Fourier Anal. Appl.* **6** (2000), no. 1, 55–78.

[26] S. Roman, *The Umbral Calculus*, Academic Press, New York, 1984.

[27] Edited by Bart M. ter Haar Romeny, *Geometry-driven Diffusion in Computer Vision*, Kluwer Acad. Publ., Dordrecht, 1994.

[28] V. A. Rvachëv, Uspekhi Mat. Nauk **45** (1990), no. 1(271), 77–103, 222; *Translation in Russian Math. Surveys* **45** (1990), no. 1, 87–120.

[29] I. J. Schoenberg, Contributions to the problem of approximation of equidistant data by analytic functions, *Quart. Appl. Math.* **4** (1946), 45–99; 112–141.

[30] I. J. Schoenberg, *Cardinal spline functions,* SIAM Publication.

[31] W. Schoutens, *Stochastic Processes and Orthogonal Polynomials*, Lecture Notes in Statist., 146, Springer, New York, 2000.

[32] G. J. Székely, On the coefficients of polynomials with negative zeros, *Studia Sci. Math. Hungar.* **8** (1973), 123–124.

[33] M. Unser, *B-splines: A perfect fit for signal processing,* IEEE Signal Processing Magazine, (2000).

[34] M. Unser, A. Aldroubi and M. Eden, On the asymptotic convergence of B-spline wavelets to Gabor functions, *IEEE Trans. Inform. Theory* **38** (1992), no. 2, part 2, 864–872.

[35] M. Unser, A. Aldroubi and M. Eden, *Polynomial spline pyramid,* IEEE Trans. Pattern Analysis and Machine Intelligence, 15(1993), 364–378.

[36] M. Unser, A. Aldroubi and S. J. Schiff, *Fast implementation of continuous wavelet transforms with integer scale,* IEEE Trans. Signal Processing, **42** (1994), 351–3523.

[37] Y. P. Wang and S. L. Lee, *Scale-space derived from B-spline,* IEEE Trans. on Pattern Analysis and Machine Intelligence, #10, 20(1998), 1040–1055.

[38] J. Weickert, S. Ishikawa and A. Imiya, Linear scale-space has first been proposed in Japan, *J. Math. Imaging Vision* **10** (1999), no. 3, 237–252.

[39] A. P. Witkin. *Scale-space filtering,* In Proc. International Joint Conference on Artificial Intelligence, Karlsruhe, Germany, 1983, 1019-1023.

[40] R. A. Young and R. M. Lesperance, *The gaussian derivative model for spatial-temporal vision: II. Cortical data,* Spatial Vision, **14** (2001), 321-389.

[41] A. H. Zemanian, *Distribution Theory and Transform Analysis*, Second edition, Dover, New York, 1987.