REFINEMENT DIFFERENTIAL EQUATIONS AND WAVELETS

Jianhong Shen

ABSTRACT. We study the following type of refinement differential equations (RDE):
$$P(D)\phi(x) = 2[H(E)\phi](2x)$$
where $P(\lambda)$ is a real polynomial, $H(z)$ a real Laurent polynomial, $D = d/dx$, and E is the backward translation operator $Ef(x) = f(x + 1)$. If the differential part $P(\lambda) \equiv 1$, the equation is the famous refinement equation for designing scaling functions in wavelet theory. In this paper, we reveal the general structure of solutions to RDEs and establish the relation between RDEs and certain types of refinement functional equations (RFE). This makes it possible to solve RDEs using the generalized subdivision scheme. The probability idea of Rvachev and Derfel is explored in a more systematic way. Finally, our results are applied to the construction of *smoothed wavelets* and *quasi-multiresolution*.

1. Introduction

Let E and D denote the translation operator and derivative operator

$$f(x) \rightarrow E(f) = f(x + 1), \qquad f(x) \rightarrow D(f) = f'(x).$$

Let $P(\lambda)$ be a polynomial in λ and $H(z)$ a Laurent polynomial in z, both with real coefficients:

$$P(\lambda) = c_N \lambda^N + c_{N-1} \lambda^{N-1} + \cdots + c_0, \qquad N \geq 0, \quad c_N \neq 0,$$

$$H(z) = h_m z^{-m} + h_{m-1} z^{-m+1} + \cdots + h_{m-L} z^{-m+L}, \qquad h_m h_{m-L} \neq 0.$$

In this paper, we consider equations of the form

$$P(D)\phi(x) = 2[H(E)\phi](2x). \tag{1.1}$$

We call it a *refinement differential equation* (RDE) of type (P, H), order N and length L.

If $H(z) = 0$, the RDE becomes an ordinary differential equation (linear homogeneous equation of order N). If, on the other hand, $P(\lambda) \equiv 1$, the RDE is the famous refinement equation (for designing scaling functions with compact supports) in wavelet theory. Therefore, in this paper, we shall assume $H \neq 0$ and $\deg P = N \geq 1$. The dilation factor 2 in the equation is not essential and can be replaced by any general integer $k > 1$. It is kept here to follow the literature in wavelet theory and digital signal processing (see Daubechies [4] and Strang and Nguyen [35]).

RDEs of form (1.1) arise in many contexts. The first work that should be mentioned is Mahler's remarkable paper [24] in 1939. Initiated by an integer partition problem in combinatorics, he studied the functional equation

$$\frac{\phi(x + a) - \phi(x)}{a} = \phi(qx), \qquad a \neq 0, \quad 0 < q < 1.$$

As the difference parameter a tends to zero, the equation evolves to an RDE: $\phi'(x) = \phi(qx)$ ($q \neq 2$, however). He constructed a special solution through a tricky integral transform. De Bruijn's work [7] on equations of the following type

$$\Phi'(s) = e^{\alpha s + \beta} \Phi(s - 1)$$

gives a complete account of equations like $\phi'(x) = a\phi(qx)$ with $q \in (0, 1)$. The connection is realized by the change of variables (Kato and McLeod [22])

$$x = e^s, \quad q = e^c, \quad \Phi(s) = \phi(x).$$

In 1971, many authors (Fox and Mayers [16], Kato and McLeod [22], Frederickson [17]) studied the special *functional-differential equation*

$$\phi'(x) = a\phi(qx) + b\phi(x),$$

which had arisen from the mathematical modeling of an industrial problem involving wave motion in the overhead supply line for a high speed train. It was Kato and McLeod who gave the complete investigation on this equation for all types of parameters. As an initial value problem, they showed the equation is well-posed if $q < 1$ and ill-posed if $q > 1$. Moreover, when $q > 1$, the solution to the IVP is not unique. The non-uniqueness had made the equation less interesting to people working on ODEs, to whom existence and uniqueness are two of the most fundamental elements. Interest in this classically "singular" case has arisen only after the ODE sunglasses were abandoned and wavelet theory appeared.

The early 70s saw many papers (see Myshkis' survey paper [26]) on functional-differential equations, most of which were motivated by applications in circuit and control theory. It was probably Rvachev who first deviated from the ODE point of view of the main stream. His interests in the functional properties of the solutions led him to the study of the interesting functional-differential equation (see Rvachev [30] and Rvachev and Rvachev [31]),

$$\phi'(x) = 2\phi(2x + 1) - 2\phi(2x - 1). \tag{1.2}$$

In this paper, we shall call it *the Rvachev equation*. The unique solution with a unit total integral was named up(x) by Rvachev. It is an even and non-negative C^∞ function supported on $[-1, 1]$. The significance of up(x) in the function-theoretic sense can be seen from its "atomic" role in certain spaces consisting of C^∞ functions (Rvachev [30]), parallel to what a wavelet does in $L^2(R)$. The connection between Rvachev's up function and wavelet theory was explained by the work of Derfel et al. [10] on the Stieltjes subdivision scheme and the non-stationary subdivision process.

Their work on this can be summarized as follows. The classical subdivision process for the Bernoulli (two-point) random distribution leads to the Haar scaling function. Then the continuous subdivision process associated with the Haar scaling function generates Rvachev's up function. Part of our current work is to reveal the general principle hidden in this intriguing example (Section 5). We show that, generally, for

each refinement differential equation of form (1.1), one can associate to it a *refinement functional equation* (RFE)

$$\phi(x) = \langle T, \phi(2x - \cdot) \rangle,$$

where T is a suitable distribution. If the distribution allows a "density" function, then the equation can be solved using the continuous subdivision process (Section 6.5). This opens an entirely new window (contrast to the classical ODE point of view for functional-differential equations) to RDEs. It is one of our initial goals to find the bridge connecting wavelet theory and functional-differential equations.

If, in certain circumstances, the associated distribution T is a probability distribution, then a probability method applies to the underlying RDE. For the up function, this already has been noticed by Rvachev [30] and Derfel [9] and can be traced back further to Jessen and Wintner's work in 1935 on infinite convolutions of symmetric Bernoulli distributions [21]. Related work also can be found in Erdös [14, 15], Garsia [18], and Brown and Moran [2] (as initially referred to in Daubechies and Lagarias [5]). In Section 6, we develop the probability interpretation of certain basic RDEs, particularly for the Rvachev equation and the kam equation (see Section 2 for definition). The connection between the exponential distribution and the kam equation is entirely new and provides an efficient approach to construct a uniform approximation to the kam function. We also explain why the normal distribution cannot produce new functions along this probability line.

Our paper is organized as follows. In Section 2, we introduce some necessary concepts and make a few assumptions for the rest of the paper. The main result of the section is written in the Structure Theorem (Theorem 1), which is not difficult to prove but plays a crucial role in the paper. Section 3 presents the main results on RDEs of type $(P(\lambda), 1)$. New functions $\mathrm{kam}(x)$ and $\Phi_\theta(x)$ are introduced as the atomic solutions to RDEs of type $(P(\lambda), 1)$ when $P(\lambda)$ contains no purely imaginary roots. When $P(\lambda)$ has at least one purely imaginary root, we construct a set of linearly independent, periodic and C^∞ solutions based on the well-adapted structure of Dirichlet series. As a by-product, we also show that RDEs of type $(P(\lambda), 1)$ always carry a single-parameter family of C^∞ and *almost periodic* solutions. In Section 4, we construct the solutions to general RDEs. The main result is that a scaling function designed by the $(1, H(z))$ refinement equation is smoothed by the $(P(\lambda), 1)$ RDE to yield a C^∞ solution to the original $(P(\lambda), H(\lambda))$ RDE. Sections 5 and 6 are devoted to developing the connections among RDE, RFE, the probabilistic method, and the generalized subdivision process. In the last section, we demonstrate one possible application of our results in wavelet theory, namely, the construction of smoothed wavelets and quasi-multiresolution.

2. Regular equations and the structure theorem

To formulate a well-posed problem, let us first understand the major difference between RDEs of form (1.1) and linear homogeneous ODEs. This is done by considering the example of an initial-value problem,

$$\phi'(t) + \phi(t) = q\phi(qt), \qquad \phi(0) = 1.$$

Assume first that $q \in (0, 1)$. Integration yields

$$\phi(t) = 1 - \int_{qt}^{t} \phi(\tau) \, d\tau = R_q \phi(t).$$

It is clear that the affine operator R_q is contracting if restricted on $C[0, \rho]$ for any $\rho < 1$. Hence, the local existence and uniqueness follow immediately from the Contracting Mapping Theorem, and the local solution can be obtained from the iterations of R_q on any initial continuous function. Therefore, if $q < 1$, the equation is essentially the same as an ordinary differential equation.

This is not the case when $q > 1$. The integration now gives

$$\phi(t) = 1 + \int_{t}^{qt} \phi(\tau) \, d\tau = R_q \phi(t).$$

R_q is not contracting any more since the domain expands! The system is *non-causal* if we borrow the terminology from signal processing. What occurs at time t is influenced by "future" events up to time qt since $q > 1$. This causes the difficulty. However, the equation is *backward* well-posed in the following sense. Suppose we already know a "future" segment of the solution

$$\phi(t), \qquad t > q^n, \qquad \text{for certain integer } n.$$

Then, by solving the non-homogeneous equations iteratively

$$\phi'(t) + \phi(t) = f_m(t) = q\phi(qt), \qquad q^{m-1} < t \leq q^m, \qquad m \leq n,$$

the "history" of $\phi(t) : 0 < t \leq q^n$ can be recovered uniquely!

Therefore, for $q > 1$, it is easy for us to construct infinitely many solutions on $(0, \infty)$. Start with any function $k(x) \in C^{\infty}[1, q]$ satisfying the compatibility condition,

$$k^{(j+1)}(1) + k^{(j)}(1) = q^{j+1} k^{(j)}(q), \quad j = 0, 1, 2, \ldots$$

(for example, any function in $C_0^{\infty}[1, q]$). Define $\phi(t) = k(t)$ for $t \in [1, q]$. By the above discussion, $\phi(t)$ on $(0, 1)$ can be determined uniquely. For $t > q$, $\phi(t)$ is solved by the following explicit iteration:

$$\phi(qt) = \frac{1}{q}\big[\phi'(t) + \phi(t)\big], \qquad q^n \leq t \leq q^{n+1}, n = 0, 1, \ldots.$$

The example makes it clear that the ODE point of view for RDEs of form (1.1) is inappropriate. In fact, as shown in Section 6, the subdivision scheme S_c (a linear operator) works much better than R_q, which is a major contribution from wavelet theory.

Just as the refinement equation has to satisfy certain compatibility conditions in wavelet theory (Daubechies and Lagarias [5]), RDEs must satisfy certain solvability conditions before we can start to discuss their solutions.

Definition 2.1. An RDE of type $(P(\lambda), H(z))$ is said to be *regular* if

$$P(\lambda) = \left(\frac{\lambda}{2}\right)^r p(\lambda), \qquad H(z) = (1 - z^{-1})^r h(z)$$

for some non-negative integer r and

$$p(0) = h(1) \neq 0;$$

r is called the *index* of the equation.

We shall restrict ourselves to regular equations. The class already includes most functionally interesting examples such as the Rvachev equation and kam equation.

It is easy to see that $\phi(x) = 0$ is a solution. But it is a trivial one. By a non-trivial solution, we mean, in the most relaxed form, a non-zero C^N function $\phi(x)$ such that (1.1) is satisfied pointwise. Here N is the order of the equation. Such a solution is necessarily to be C^∞. However, unlike the homogeneous ODE with constant coefficients, it is not C^ω.

In addition, we also are interested in the following two properties: (1) L^1 and (2) periodicity or almost periodicity (see Section 3.3 for the definition). If the solution is in $L^1(R)$, we secure the uniqueness by imposing the integral normalization condition

$$\langle 1, \phi \rangle = \int_R \phi(x) \, dx = C,$$

for some convenient non-zero constant C. The existence of (almost) periodic solutions to RDEs is shown through the construction of Dirichlet series.

There are two alternative methods for studying regular RDEs, namely, the frequency-domain method and the time-domain method (Daubechies [4]). The first technique is particularly powerful for regularity analysis of the solution, now widely known and practiced in wavelet theory. In this paper, however, we prefer the time-domain method because (1) unlike wavelets and scaling functions in wavelet theory, solutions to regular RDEs are always C^∞ and, hence, analysis on regularity is redundant, and (2) the solutions to RDEs carry very rich structures and contents in the time-domain. The time-domain method is based on the convolution operator "$*$".

Here we list some properties of the convolution operator that will be useful in the paper.

(a) Suppose $f(x) \in C^\infty$, and $g(x) \in L^1$ and is compactly supported. Then $f * g \in C^\infty$.

(b) Suppose $f(x)$ is absolutely continuous and has a finite total variation, and $g(x) \in L^1$. Then at least in L^1,

$$D(f * g) = (Df) * g.$$

(c) Suppose $f * g$ is well-defined (in the ordinary sense or L^1). Then

$$2(f * g)(2x) = [2f(2x)] * [2g(2x)],$$
$$E(f * g) = E(f) * g = f * E(g).$$

Theorem 2.1 (Structure Theorem). *Consider an RDE of type $(P(\lambda), H(z))$. Suppose*

$$P(\lambda) = P_1(\lambda) \cdot P_2(\lambda), \qquad H(z) = H_1(z) \cdot H_2(z)$$

where $P_i(\lambda)$ are polynomials in λ and $H_i(z)$ Laurent polynomials in z. Assume $\phi_i(x)$ is the solution to the RDE of type (P_i, H_i), $i = 1, 2$. If $\phi = \phi_1 * \phi_2$ is well defined, then it is a solution to the RDE of type $(P(\lambda), H(z))$.

Proof. Let us check this directly.

$$P(D)\phi = P_1(D) \cdot P_2(D)[\phi_1 * \phi_2]$$
$$= [P_1(D)\phi_1] * [P_2(D)\phi_2]$$
$$= 2[H_1(E)\phi_1](2x) * 2[H_2(E)\phi_2](2x)$$
$$= 2[H_1(E)\phi_1 * H_2(E)\phi_2](2x)$$
$$= 2[H_1 \cdot H_2(E)\phi_1 * \phi_2](2x)$$
$$= 2[H(E)\phi](2x).$$

Hence, $\phi(x)$ is the solution to the RDE of type $(P(\lambda), H(z))$. $\qquad\square$

Corollary 2.1 (Smoothed Scaling Function). *Suppose a regular equation of type* $(P(\lambda), H(z))$ *has index* $r = 0$. *If* $\phi_p(x)$ *is the solution to* $(P(\lambda), 1)$, $\phi_h(x)$ *(the scaling function) to* $(1, H(z))$ *(the refinement equation), and* $\phi(x) = \phi_p * \phi_s$ *is well defined, then* $\phi(x)$ *is the solution to* $(P(\lambda), H(z))$.

This corollary is applied to the construction of smoothed wavelets and quasi-multiresolution in Section 7.

In the following, we first consider regular equations of type $(P(\lambda), 1)$. The complete discussion on general regular equations is in Section 4.

3. Solutions to regular RDEs of type $(P(\lambda), 1)$

Any real polynomial $P(\lambda)$ with $P(0) = 1$ can be factored to

$$\prod_i (1 + a_i\lambda) \prod_j d_j(\lambda)$$

where a_i are non-zero real scalars, and $d_j(\lambda)$ irreducible quadratic polynomials with $d_j(0) = 1$. By the preceding proposition, to solve RDE of type $(P(\lambda), 1)$, it suffices to consider the two types of equations

$$a\phi'(x) + \phi(x) = 2\phi(2x), \tag{3.1}$$

$$\alpha\phi''(x) + b\phi'(x) + \phi(x) = 2\phi(2x). \tag{3.2}$$

3.1. $a\phi'(x) + \phi(x) = 2\phi(2x)$ **and the kam function.** By a change of variables, $x \to at$, we can assume $a = 1$.

In 1971, Kato and McLeod [22] studied the initial value problem

$$y'(t) = ay(qt) + by(t), \qquad t > 0,$$

in full details. Since wavelet theory was not formally on the stage of mathematical analysis at that time, they treated it as a regular initial value problem. The existence, uniqueness, and asymptotic behavior of the solution were of the most concern.

Their results lead to the following:

(i) The only solution to ($a = q = 2$ and $b = -1$)

$$y'(t) + y(t) = 2y(2t), \qquad t > 0,$$

that decays faster than $O(t^{-1})$ as $t \to +\infty$ is a constant multiple of the following function (which in fact decays exponentially fast, and is named after the two authors

here: "ka" for Kato and "m" for McLeod)

$$\mathrm{kam}(t) = e^{-t}\Big\{1 + \sum_{n=1}^{\infty}(-1)^n \frac{2^n \exp[-(2^n-1)t]}{(2-1)(2^2-1)\cdots(2^n-1)}\Big\}, \qquad t > 0. \qquad (3.3)$$

(ii) Any other solution is not in $L^1(0,\infty)$. The two authors constructed a family of solutions that have the exact order $O(t^{-1})$ near $+\infty$.

The Dirichlet series solution in (3.3) also was studied by Frederickson [17] at the same time. He considered general solutions of the form,

$$\phi_\beta(t) = \sum_{n=-\infty}^{\infty} C_{n,\beta}\exp(\beta q^n t)$$

parameterized by β for equations with $q > 1$. It can be verified easily that $\beta = -1$ is the only parameter that allows $C_{n,\beta} = 0$ for all $n < 0$ and hence $\phi_\beta(t) = o(t^{-1})$ near $+\infty$.

We, therefore, conclude that if there is a solution $\phi(x) \in C^1(R) \cap L^1(R)$ to

$$\phi'(x) + \phi(x) = 2\phi(2x), \qquad \int_R \phi(x)\,dx = 1,$$

then $\phi(x)|_{x>0}$ must be a constant multiple of $\mathrm{kam}(x)$. For future reference, we call this *the* kam *equation*.

For convenience, we introduce the combinatorial notation $(n)_q$ for a q-analog number (see Goldman and Rota [19, 20]) defined by

$$(n)_q = \frac{q^n - 1}{q - 1} = 1 + q + \cdots q^{n-1},$$

and $(n)_q!$ for the q-analog factorial given by

$$(n)_q! = (1)_q (2)_q \cdots (n)_q.$$

Then the expression of $\mathrm{kam}(x)$ simplifies to

$$\mathrm{kam}(x) = \sum_{n\geq 0} \frac{(-2)^n}{(n)_2!}\exp(-2^n x)$$

where $(0)_q!$ is defined to be 1. Series in this form are called *Euler Series* in combinatorics, if x is treated as a parameter.

Theorem 3.1 (Properties of kam(x)). *Define* $\mathrm{kam}(x) = 0$ *for all* $x \leq 0$. *Then*

(a) *The alternating infinite series converges uniformly to* $\mathrm{kam}(x)$ *at a rate of* $O(2^{-\binom{n}{2}})$ *for* $x > 0$.

(b) $\mathrm{kam}(x) \in C^\infty(R)$.

(c) $\mathrm{kam}(x) > 0$ *for all* $x > 0$. *Hence* $\mathrm{supp}[\mathrm{kam}(x)] = [0,\infty)$.

(d) $\int_R \mathrm{kam}(x)\,dx = \exp_2(-1) \in (\frac{2}{7},\frac{1}{3})$ *where* $\exp_q(x)$ *is the* q-analog *exponential function defined by*

$$\exp_q(x) = \sum_{n=0}^{\infty} \frac{x^n}{(n)_q!}.$$

Proof. Set

$$a_n(x) = \frac{2^n}{(n)_2!}\exp(-2^n x) = \frac{2^n \exp(-2^n x)}{(2-1)(2^2-1)\cdots(2^n-1)}, \qquad n = 1, 2, \ldots.$$

Then

$$\frac{a_{n+1}(x)}{a_n(x)} = \frac{2}{(2^{n+1}-1)\exp(2^n x)} < 1,$$

for all $n \geq 1$ and $x \geq 0$. Hence, the alternating infinite series converges to kam(x) at a rate of

$$a_n(0) = O(2^{-\binom{n}{2}})$$

for all $x \geq 0$ (and even much faster away from $x = 0$).

Fix any positive integer k,

$$|a_n^{(k)}(x)| = \frac{2^{(k+1)n}\exp(-2^n x)}{(2-1)(2^2-1)\cdots(2^n-1)}$$

$$\leq \frac{2^{(k+1)n}}{(2-1)(2^2-1)\cdots(2^n-1)}$$

$$= O(2^{-\binom{n}{2}+kn}),$$

which is uniformly exponentially small as $n \to +\infty$. Therefore, for $x > 0$,

$$\frac{d^k}{dx^k}\text{kam}(x) = \sum_{n=0}^{\infty}(-1)^{n+k}\frac{2^{(k+1)n}\exp(-2^n x)}{(2-1)(2^2-1)\cdots(2^n-1)},$$

which implies kam$(x) \in C^\infty(0, \infty)$. To prove that kam$(x) \in C^\infty(R)$, it suffices to show that

$$\text{kam}^{(k)}(0^+) = 0, \qquad \text{for all } k = 0, 1, \ldots.$$

In fact,

$$(-1)^k\frac{d^k}{dx^k}\text{kam}(0^+) = 1 + \sum_{n=1}^{\infty}(-1)^n\frac{2^{(k+1)n}}{(2-1)(2^2-1)\cdots(2^n-1)}$$

$$= 1 + \sum_{n=1}^{\infty}(-1)^n\frac{2^{kn}(2^n-1)+2^{kn}}{(2-1)(2^2-1)\cdots(2^n-1)}$$

$$= -2^k(-1)^{k-1}\frac{d^{k-1}}{dx^{k-1}}\text{kam}(0^+) + (-1)^{k-1}\frac{d^{k-1}}{dx^{k-1}}\text{kam}(0^+)$$

$$= (1-2^k)(-1)^{k-1}\frac{d^{k-1}}{dx^{k-1}}\text{kam}(0^+).$$

Thus, an induction proves (b).

Since $(a_n(x))$ is a strictly decreasing positive sequence, (c) is obvious. The equality part of (d) can be obtained by integrating the infinite series term by term; and the rest is due to

$$\exp_2(-1) = 1 - 1 + \frac{1}{3} - \frac{1}{3\cdot 7} + \cdots.$$

The proof is complete.   □

From now on, kam(x) denotes its zero-extended version.

Corollary 3.1. *When $x \gg 1$, the leading term of* $\mathrm{kam}(x)$ *is* $\exp(-x)$.

Corollary 3.2. *For any non-zero real scalar a, the first-order RDE*

$$a\phi'(x) + \phi(x) = 2\phi(2x), \qquad \int_R \phi(x) = \exp_2(-1)$$

has the unique $C^1(R)$ solution

$$\phi_a(x) = \frac{1}{|a|}\,\mathrm{kam}(\frac{x}{a}).$$

In particular, $\mathrm{supp}\phi_a = \mathrm{sign}(a) \cdot [0, \infty)$.

Proposition 3.1 (Integration versus Summation)**.** *Define*

$$K_+(x) = \int_0^x \mathrm{kam}(t)\,dt, \qquad K_-(x) = \int_x^\infty \mathrm{kam}(t)\,dt, \qquad x > 0.$$

Then

$$K_+(x) = \sum_{n=1}^\infty \mathrm{kam}(\frac{x}{2^n}), \qquad K_-(x) = \sum_{n=0}^\infty \mathrm{kam}(2^n x).$$

Proof. Take the second equation for example. Call the function on the right side $L(x)$. Then

$$L'(x) = \sum_{n=0}^\infty 2^n \,\mathrm{kam}'(2^n x)$$

$$= \sum_{n=0}^\infty 2^n [2\,\mathrm{kam}(2^{n+1}x) - \mathrm{kam}(2^n x)]$$

$$= \sum_{n=1}^\infty 2^n \,\mathrm{kam}(2^n x) - \sum_{n=0}^\infty 2^n \,\mathrm{kam}(2^n x)$$

$$= -\,\mathrm{kam}(x).$$

Obviously, $L(x) \to 0$ as $x \to \infty$. Hence, $L(x) = K_-(x)$. $\qquad\square$

Notice that $L(0^+) \neq L(0)$.

3.2. $\alpha\phi''(x) + 2b\phi'(x) + \phi(x) = 2\phi(2x)$, $\alpha > b^2$. By a suitable change of variables $x = \pm\sqrt{\alpha}t$, we can assume $\alpha = 1$ and $b \geq 0$. Hence, it suffices to consider the standard equation,

$$\phi''(x) + 2\cos\theta\,\phi'(x) + \phi(x) = 2\phi(2x), \qquad \theta \in \left(0, \frac{\pi}{2}\right]. \tag{3.4}$$

To make the solution unique, the following integral normalization condition is applied:

$$\int_R \phi(t)\,dt = [\exp_2(-1)]^2. \tag{3.5}$$

The solution is denoted by $\Phi_\theta(x)$.

For a given angle θ, define $\omega = \exp(i\theta)$, and

$$J^\theta(t) = 2\sum_{n=0}^\infty \frac{(-1)^m}{(m)_2!}\,\frac{1}{\bar\omega - t2^{-m}\omega} = J_r^\theta(t) + iJ_i^\theta(t),$$

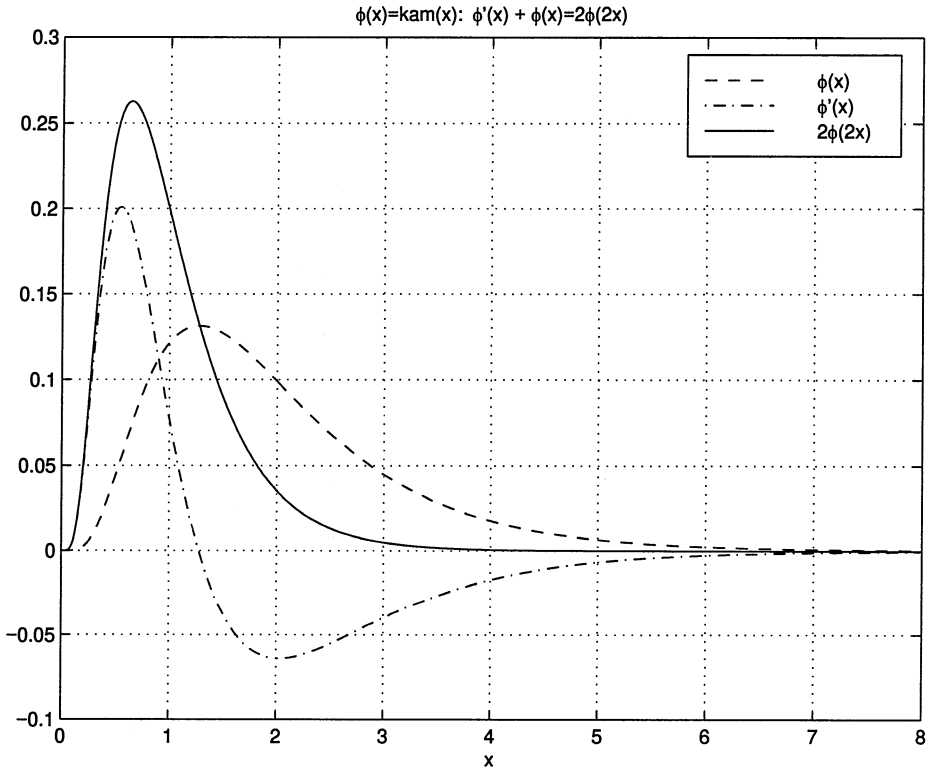for all $t \geq 0$. Here J_r^θ and J_i^θ denote the real and imaginary parts of J^θ.

FIGURE 1. $\mathrm{kam}(x), \mathrm{kam}'(x)$, and $\mathrm{kam}(2x)$

Lemma 3.1. *Suppose* $0 < \theta \leq \pi/2$. *Then,*

$$\sup_{t \geq 0} |J(t)| \leq 2 \exp_2(1) \csc \theta.$$

Proof. For any $t \geq 0$,

$$|J(t)| \leq 2 \sum_{m \geq 0} \frac{1}{(m)_2!} \sup_{t \geq 0, m \geq 0} \frac{1}{|\bar{\omega} - t2^{-m}\omega|}$$

$$= 2 \exp_2(1) \sup_{t \geq 0} \frac{1}{|\bar{\omega} - t\omega|}$$

$$\leq \frac{2 \exp_2(1)}{\sin \theta}.$$

$\square$

Theorem 3.2 $(\Phi_\theta(x))$. *Suppose* $0 < \theta < \frac{\pi}{2}$. *Subject to the normalization condition* (3.5), (3.4) *has the unique solution,*

$$\Phi_\theta(x) = \begin{cases} \sum_{n=0}^{\infty} \frac{(-2)^n}{(n)_2!} \exp(-2^n bx) \left[J_r^\theta(2^n) \cos(2^n ax) + J_i^\theta(2^n) \sin(2^n ax) \right] & x > 0, \\ 0 & x \leq 0. \end{cases}$$

$$(3.6)$$

Here $b = \cos\theta$ and $a = \sin\theta$. $\Phi_\theta(x) \in C^\infty(R)$, and for $x \geq 0$, its derivatives can be obtained by differentiating the infinite series in (3.6) term by term, whose converging rate is $O(2^{-\binom{n}{2}+kn})$ for the k-th derivative.

Proof. We only sketch the proof. Uniqueness can be proved by the technique of Kato and McLeod [22] or the frequency-domain method. The rest of the proof is done in the following three steps.

First, with the help of Lemma 3.1, we can show that $\Phi_\theta(x)$ is C^∞ for $x > 0$. The derivatives can be obtained by taking differentiation on the infinite series term by term and have the predicted converging rate (see the proof in the preceding theorem).

Secondly, we show $\Phi_\theta(0^+) = \Phi'_\theta(0^+) = 0$. Take $\Phi'_\theta(0^+) = 0$, for example; then

$$
\begin{aligned}
\Phi'_\theta(0^+) &= -\mathrm{Re} \sum_{n \geq 0} \frac{(-2)^n}{(n)_2!} 2^n \omega J^\theta(2^n) \\
&= -2\mathrm{Re} \sum_{m,n \geq 0} \frac{(-2)^{n+m}}{(n)_2!(m)_2!} \frac{2^n \omega}{2^m \bar\omega - 2^n \omega} \\
&= 2 \sum_{m,n \geq 0} \frac{(-2)^{n+m}}{(n)_2!(m)_2!} - 2\mathrm{Re} \sum_{m,n \geq 0} \frac{(-2)^{n+m}}{(n)_2!(m)_2!} \frac{2^m \bar\omega}{2^m \bar\omega - 2^n \omega} \\
&= -2\mathrm{Re} \sum_{m,n \geq 0} \frac{(-2)^{n+m}}{(n)_2!(m)_2!} \frac{2^n \bar\omega}{2^n \bar\omega - 2^m \omega} \\
&= 2\mathrm{Re} \sum_{m,n \geq 0} \frac{(-2)^{n+m}}{(n)_2!(m)_2!} \frac{2^n \omega}{2^m \bar\omega - 2^n \omega}.
\end{aligned}
$$

Comparison of the second line with the last line verifies that $\Phi'_\theta(0^+) = 0$.

In the final step, we show that for $x > 0$, $\Phi_\theta(x)$ is the solution to the given RDE. We ask readers to fill in the proof.

The combination of the last two steps and the original RDE implies that $\Phi_\theta^{(k)}(0^+) = 0$ for all non-negative integer k. Hence, $\Phi_\theta(x) \in C^\infty(R)$ and satisfies the given RDE. Finally, a direct computation shows that Φ_θ satisfies the prescribed integral normalization condition. This completes the proof. □

Corollary 3.3 (Damped Oscillation). *If $0 < \theta < \pi/2$, or equivalently $b > 0$, the leading term of the solution to (3.4) and (3.5) for $x \gg 1$ is the damped oscillation,*

$$\Phi_\theta(x) \simeq e^{-bx}[A_\theta \cos(ax) + B_\theta \sin(ax)]$$

where, $b = \cos\theta$, $a = \sin\theta$, and the real constants A_θ and B_θ are given by

$$A_\theta + iB_\theta = 2 \sum_{m \geq 0} \frac{(-1)^m}{(m)_2!} \frac{1}{\bar\omega - 2^{-m}\omega}.$$

Notice the leading term is a special solution to the ODE defined by the left side of the RDE.

Corollary 3.4 (Periodic Solution for $\theta = \frac{\pi}{2}$). *Suppose $\theta = \frac{\pi}{2}$. Then $\Phi_{\frac{\pi}{2}}(x)$ defined in (3.6) is forward periodic with period 2π when $x > 0$; that is,*

$$\Phi_\theta(x) = \Phi_\theta(x + 2\pi), \qquad x > 0.$$

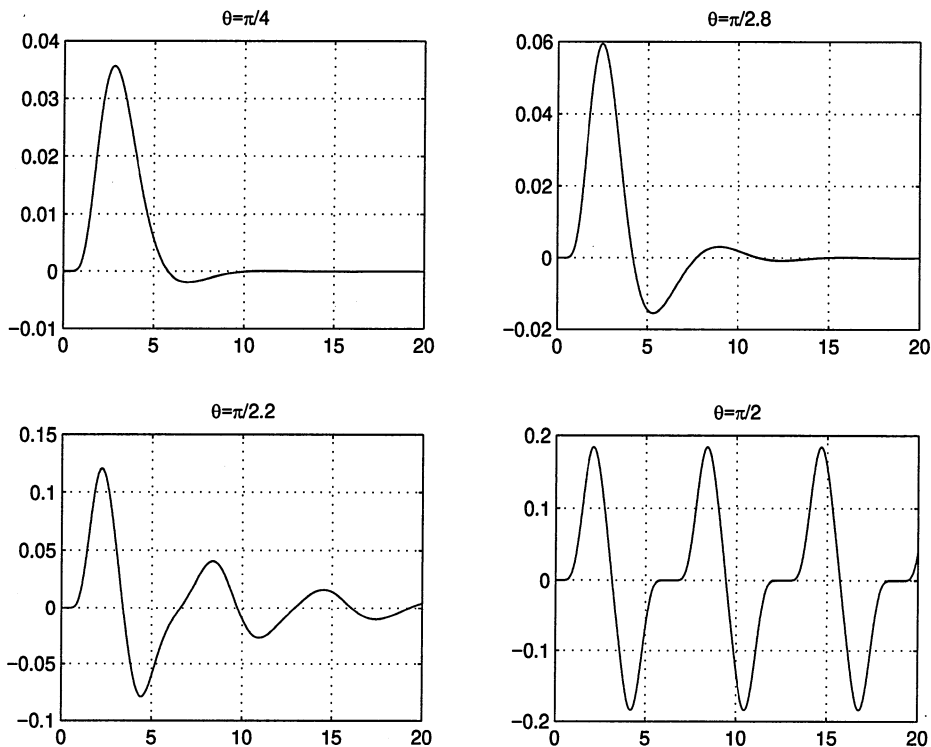$\Phi_{\frac{\pi}{2}}(x)$ *is a special solution to (3.4).*

FIGURE 2. $\Phi_\theta(x)$ for different angles: Damped oscillation ($\theta < \frac{\pi}{2}$) and periodicity ($\theta = \frac{\pi}{2}$).

Remark. It is not difficult to see that both $\Phi_{\frac{\pi}{2}}(x)$ and $\Phi_{\frac{\pi}{2}}(-x)$ are solutions to (3.4). Hence, the solution space is at least 2-dimensional. It can be shown further that in this extremal case, (3.4) has no solution in $C^1(R) \cap L^1(R)$. We will discuss this case in more detail in the coming subsection.

For convenience, we extend the range of the parameter θ in Φ_θ to include $\theta \in (\pi/2, \pi)$ by defining

$$\Phi_\theta(x) = \Phi_{\pi-\theta}(-x), \qquad \theta \in (\pi/2, \pi).$$

Corollary 3.5. *Suppose the real quadratic polynomial $\alpha\lambda^2 + b\lambda + 1$ has two complex roots $re^{\pm i\theta}$ for some $r > 0$ and $\theta \in (0, \pi)\backslash\frac{\pi}{2}$. Then, the unique $C^2 \cap L^1$ solution to*

$$\alpha\phi''(x) + b\phi(x) + \phi(x) = 2\phi(2x), \qquad \int_R \phi(x)dx = 1$$

is $\Phi_\theta(-rx)$ up to a multiplicative constant.

Finally, by the Structure Theorem, we achieve the main result of this section.

Theorem 3.3 (Main Theorem, Part I). *Consider an RDE of type $(P(\lambda), 1)$ with $P(0) = 1$. Suppose $P(\lambda)$ has no imaginary roots. Then, (3.4) has a $C^\infty(R)$ solution which decays exponentially fast near $\pm\infty$. In fact, up to a constant multiple, the*

solution takes the form

$$\phi(x) = \text{kam}(-\lambda_1 x) * \cdots * \text{kam}(-\lambda_k x) * \Phi_{\theta_1}(-r_1 x) * \cdots * \Phi_{\theta_m}(-r_m x).$$

Here, $\lambda_1, \ldots, \lambda_k$ *are all real roots of* $P(\lambda)$, *and* $r_j \exp(\pm\theta_j)$, $r_j > 0$, $\theta_j \in (0, \pi)$, $j = 1, \ldots, m$, *are all the complex roots.*

3.3. Dirichlet series solution and periodicity.
Generally, for a given regular RDE of type $(P(\lambda), 1)$, $P(0) = 1$,

$$P(D)\phi(x) = 2\phi(2x),$$

one always can try the Dirichlet series solution,

$$\phi_\beta(x) = \sum_n C_{n,\beta} e^{-2^n \beta x}$$

where the index n runs over all integers and β is a parameter to be discussed. The necessary condition for such a function $\phi(x)$ to be a solution is the following recursion formula for the coefficients:

$$C_{n,\beta} P(-2^n \beta) = 2 C_{n-1,\beta}. \tag{3.7}$$

First, let us assume the parameter β is chosen so that $\lambda = -2^n\beta$, $n = 0, \pm 1, \ldots,$ are all not roots of $P(\lambda)$. Then, for any $n > 0$,

$$C_{n,\beta} = \frac{2}{P(-2^n\beta)} C_{n-1,\beta},$$

and for any $n < 0$,

$$C_{n,\beta} = \frac{1}{2} P(-2^{n+1}\beta) C_{n+1,\beta}.$$

Therefore, there exists a unique Dirichlet series "solution" subject to $C_{0,\beta} = 1$.

However, we have to check the convergence properties of the resulting series and justify the above differentiation term by term. This is done by the following estimation. Suppose $P(\lambda)$ is of order $N \geq 1$. The above recurrence formula implies that

$$\frac{C_{n,\beta}}{C_{n-1,\beta}} = O(2^{-nN}), \qquad n \gg 1,$$

$$\frac{C_{n,\beta}}{C_{n+1,\beta}} \simeq \frac{1}{2}, \qquad n \ll -1.$$

Hence,

$$C_{n,\beta} = O(2^{-\binom{n}{2}N}), \quad n > 0, \qquad \text{and} \qquad C_{n,\beta} = O(2^n), \quad n < 0.$$

Definition 3.1 (Almost Periodicity). A function $f(x)$ is said to be *almost periodic* if for any $\epsilon > 0$, there exists a periodic function $f_\epsilon(x)$ such that the uniform norm $\|f - f_\epsilon\| < \epsilon$ holds.

The above estimation leads to:

Proposition 3.2 (Almost Periodic Solution). *Consider an RDE of type* $(P(\lambda), 1)$ *and a given purely imaginary parameter* β. *If* $P(\lambda)$ *has no root in the form of* $-2^n\beta$ *for any integer* n, *then the real and imaginary parts of* $\phi_\beta(x)$ *with* $C_{0,\beta} = 1$ *are two linearly independent and* $C^\infty(R)$ *real solutions to the RDE. Furthermore, they are both almost periodic.*

Please note that the proposition makes no sense if β is not purely imaginary, in which case, the Dirichlet series does not converge for any $x \neq 0$.

Now we consider the case when there exists an integer n such that $P(-2^n\beta) = 0$. Let n_0 denote the largest integer satisfying this condition. By choosing $\beta' = 2^{n_0}\beta$, we can assume $n_0 = 0$. From the recursion formula,

$$C_{n,\beta} = 0, \ n < 0, \qquad C_{n,\beta} = \frac{2}{P(-2^n\beta)} C_{n-1,\beta}, \ n > 0.$$

Hence, by choosing $C_{0,\beta} = 1$, we obtain the unique (complex) Dirichlet series solution:

$$\phi_\beta(x) = \sum_{n \geq 0} C_{n,\beta} e^{-2^n \beta x}.$$

If the real part of β is not zero, the above expression cannot give a global solution since the series diverges on one of the half-axes. It can be fixed, however, if there exists some x_0 in the convergent half-axis such that $\phi_\beta(x)$ (or one of its real and imaginary parts) and all of its derivatives are zero at this point: $\phi_\beta^{(k)}(x_0) = 0$ for $k = 0, 1, \ldots$. This is indeed the case for $\mathrm{kam}(x)$ and $\Phi_\theta(x)$ ($x_0 = 0$ for both of them). Otherwise, we cannot obtain a global C^1 solution from the Dirichlet series.

If β is purely imaginary, we have:

Proposition 3.3 (Periodic Solution). *Suppose $P(\lambda)$ $(P(0) = 1)$ contains a purely imaginary root $\beta = i\omega$, for which no $-2^n\beta$, $n > 0$, is a root of $P(\lambda)$. Then, the real and imaginary parts of $\phi_\beta(x)$ with $C_{0,\beta} = 1$ are two linearly independent and periodic C^∞ real solutions to the RDE of type $(P(\lambda), 1)$.*

Definition 3.2 (Binary Degree). Two complex numbers a and b are said to be *binary dependent* if a/b is an integer power of 2. A set of complex numbers are said to be *binary independent* if no two distinct numbers are binary dependent. The *binary degree* of a finite set is the cardinality of the maximal subset that is binary independent.

This concept and the above discussion lead to the second part of our Main Theorem.

Theorem 3.4 (Main Theorem, Part II). *Given a real polynomial $P(\lambda)$, $P(0) = 1$, let Z denote all its purely imaginary roots with positive imaginary parts. Suppose the binary degree of Z is $d > 0$. Then the RDE of type $(P(\lambda), 1)$ has at least $2d$ linearly independent and periodic C^∞ solutions.*

3.4. Analytic domain. In the above, we have shown that both $\mathrm{kam}(x)$ and $\Phi_\theta(x)$ ($\theta \in (0, \frac{\pi}{2})$) belong to $C^\infty(R)$. Since both have support on the positive half axis, they cannot be analytic. However, we have the following results.

Proposition 3.4 (Analytic Extension of $\mathrm{kam}(x)$). *There is a unique analytic function $K(z)$ that is defined on the right-half plane: $\mathrm{Re}\ z > 0$ and continuous to the imaginary axis, such that its restriction on the positive half-axis is $\mathrm{kam}(x)$.*

Proof. In fact, $K(z)$ must be the function,

$$K(z) = \sum_{n=0}^{\infty} \frac{(-2)^n}{(n)_2!} \exp(-2^n z), \qquad \mathrm{Re}\ z > 0.$$

It also is not difficult to see that $K(z)$ is continuous to the imaginary axis. $\qquad\square$

Proposition 3.5 (Analytic Extension of $\Phi_\theta(x)$). *Suppose $\theta \in (0, \frac{\pi}{2})$. There is a unique analytic function that is analytic inside the angular domain defined by*

$$-\frac{\pi}{2} + \theta < \operatorname{Arg} z < \frac{\pi}{2} - \theta,$$

and continuous to the boundary, such that its restriction on the positive half-axis is $\Phi_\theta(t)$.

Proof. In fact, this function must be given by the infinite series (see Eq.(3.6));

$$\sum_{n=0}^{\infty} \frac{(-2)^n}{(n)_2!} \exp(-2^n b z) \left[J_r^\theta(2^n) \cos(2^n a z) + J_i^\theta(2^n) \sin(2^n a z) \right].$$

$\square$

4. General regular RDEs

In this section, we consider general regular RDEs of type $(P(\lambda), H(z))$ and with index r. Since the equation is regular, we can assume

$$P(\lambda) = \left(\frac{\lambda}{2}\right)^r p(\lambda), \qquad H(z) = (1 - z^{-1})^r h(z),$$

with $p(0) = h(1) \neq 0$. Without loss of generality, assume $p(0) = h(1) = 1$.

4.1. Type $(\frac{\lambda}{2}, 1 - z^{-1})$ and Rvachev's up function. An RDE of type $(\frac{\lambda}{2}, 1 - z^{-1})$ has the form,

$$\frac{1}{2}\phi'(x) = 2\phi(2x) - 2\phi(2x - 1), \qquad \int_R \phi(x)\,dx = 1. \tag{4.1}$$

By the change of variables:

$$y = 2x - 1, \qquad \bar{\phi}(y) = \frac{1}{2}\phi(x) = \frac{1}{2}\phi\left(\frac{y+1}{2}\right),$$

the equation is transformed to

$$\bar{\phi}'(y) = 2\bar{\phi}(2x + 1) - 2\bar{\phi}(2x - 1),$$

which is exactly the Rvachev equation! Therefore,

Proposition 4.1. *The unique solution to* (4.1) *is given by*

$$\mathrm{up}_+(x) = 2\,\mathrm{up}(2x - 1),$$

which is a non-negative C^∞ function supported on $[0, 1]$.

The Fourier transform of the up function first appeared in Jessen and Wintner's paper [21] in 1935 as an example of infinite convolutions of symmetric Bernoulli distributions. Later in 1971, Rvachev [31] studied (1.2) and obtained up(x) as a solution. Since then, it was rediscovered by many other authors in different contexts (see Kirov and Totkov [23] and de Reina Martinez [8], for examples), and its roles in approximation theory and in the representation of smooth functions have been studied extensively. More references can be found in Myshkis' survey paper [26] in 1977 and Rvachev [30] in 1990.

Integrating both sides of the Rvachev equation

$$\phi'(x) = 2\phi(2x + 1) - 2\phi(2x - 1),$$

with the assumption that $\phi(x) \in C^1(R) \cap L^1(R)$, we obtain

$$\phi(x) = \int_{2x-1}^{2x+1} \phi(t)\, dt. \tag{4.2}$$

Define the Rvachev operator R to be

$$Rf = \int_{2x-1}^{2x+1} f(t)\, dt.$$

Restrict Dom $R = L^1(R)$. Then R has the following three properties: (1) the subspace of all $L^1(R)$ functions that are supported on $[-1, 1]$ is invariant; (2) the total integral $\int_{-\infty}^{\infty} f(t)dt$ is conserved; and finally (3) R improves smoothness by one order.

Rvachev obtained the up function by iterating R on the initial candidate $\phi_0(x) = \frac{1}{2}1_{[-1,1]}(x)$. That is, define

$$\phi_n = R^n \phi_0,$$

for all $n = 0, 1, \ldots$. He showed that the (spline) sequence $\phi_n(x)$ converges uniformly. The limit is called the up function, which is a non-negative C^∞ function supported on $[-1, 1]$.

We list some other functional properties of $\mathrm{up}(x)$:

(1) $\mathrm{up}(x) + \mathrm{up}(x-1) = 1$ for $x \in [0, 1]$. This follows directly from taking the first derivative. In particular,

$$\sum_{n=-\infty}^{\infty} \mathrm{up}(x-n) = 1.$$

(2) For any non-negative integer j, $\mathrm{up}^{(j)}(x)$ is a linear combination of the translated and dilated copies $\mathrm{up}(2^j x + k)$. This is proved by induction.

From the multiresolution point of view, $\mathrm{up}^{(j)}(x)$ is inside $V_j(\mathrm{up})$, the space spanned (and closed) by all functions $\mathrm{up}(2^j x - k)$ for $k = 0, \pm 1, \ldots$.

4.2. Type $(1, H(z))$ and the scaling function. A regular RDE of type $(1, H(z))$ $(H(1) = 1)$

$$\phi(x) = 2 \sum_{k=m-L}^{m} h_k \phi(2x - k)$$

is the famous refinement equation in wavelet theory and computer aided design. Many authors have contributed to its development. Complete references can be found in Daubechies and Lagarias [5, 6], and Daubechies [4]. In the following, we summarize some main results.

Deslauriers and Dubuc [12] showed that the equation is always solvable in the distributional sense (the solution therefore is called *the scaling distribution* in the following). The non-trivial distributional solution is supported on $[m - L, m]$. Mallat [25] considered the equation which satisfies the "orthogonal condition"

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 1.$$

He showed that a refinement equation with such an orthogonal real filter $H(z)$ has a non-trivial $L^2(R)$ solution. Daubechies [3] studied the equation when the filter $H(z)$ satisfies the strong lowpass condition

$$H(z) = \left(\frac{1 + z^{-1}}{2}\right)^p L(z).$$

She established the following regularity result: if $\|L(z)\|_{S^1} < 2^{p-q}$ for some non-negative number q, then $\phi(x) \in C^q$. Here the norm refers to the supremum norm restricted on the unit circle $S^1 : |z| = 1$. More results on the existence and regularity properties of the solution can be found in Daubechies and Lagarias [5, 6].

4.3. Solutions to general regular equations. Combining all the results so far, we achieve the last part of the main theorem.

Theorem 4.1 (Main Theorem, Part III). *Suppose a regular equation $(P(\lambda), H(z))$ has order $N = r + n \geq 1$, length $L = r + l$, and index r.*

(a) *If $P(\lambda)$ has no purely imaginary roots, then the equation has a unique (up to a multiplicative constant) $C^N(R) \cap L^1(R)$ solution in the form,*

$$\phi(x) = \mathrm{up}_+^{*r}(x) * K(x) * \Phi(x) * \phi_h(x)$$

where

$\quad \mathrm{up}_+^{*r}(x) \quad$ *is the r-th convolutional power of $\mathrm{up}_+(x)$,*
$\quad K(x) \quad$ *is a convolutional product of some scaled $\mathrm{kam}(x)$ functions,*
$\quad \Phi(x) \quad$ *is a convolutional product of some scaled $\Phi_\theta(x)$ functions,*
$\quad \phi_h(x) \quad$ *is a certain scaling distribution with a compact support of length l.*

(b) *If the set of all purely imaginary roots of $P(\lambda)$ with positive parts has a binary degree d (Section 3.3), then the equation has at least $2d$ linearly independent, periodic, and C^N real solutions in the form,*

$$\phi(x) = \mathrm{up}_+^{*r}(x) * \Phi_p(x) * \phi_h(x)$$

where $\Phi_p(x)$ is some periodic solution of trigonometric Dirichlet series (Section 3.3).

5. Distributions and refinement functional equations

The probability explanation of the up function can be found in Rvachev [30]. Derfel [9] generalized the refinement equation in wavelet theory by allowing arbitrary probability masks (filters). Derfel et al. [10] studied the convergence properties of the continuous and non-stationary subdivision process with more general Stieltjes masks, which gave further probability explanations to Rvachev's up function. Following this line, we will discuss the probability side of some typical RDEs in the next section. In this section, we first develop a more general framework based on distribution theory, by which a close connection is found between regular RDEs and a class of refinement equations with distributional masks, or the refinement functional equations (RFE) as called in this paper.

In what follows, we consider only the class $\mathcal{S}$ of Schwartz functions and the space $\mathcal{S}'$ of Schwartz distributions, though some of the theory applies to a much larger class of distributions.

Recall that a C^∞ function $\phi(x)$ belongs to $\mathcal{S}$ if and only if, for any non-negative integer k and N,

$$\sup_{x \in R} |\phi^{(k)}(x)| \, (1 + |x|)^N < \infty.$$

Hence $\mathcal{S}$ is closed under the differentiating operator D. The space of all linear functionals on $\mathcal{S}$ is denoted by $\mathcal{S}'$. Elements in $\mathcal{S}'$ are usually denoted by capital letters $T, F, \ldots$. and called Schwartz distributions (or *tempered distributions*). The value $T(\phi)$ is conventionally denoted by the scalar product $\langle T, \phi \rangle$.

Given a Schwartz distribution T, let us consider the refinement functional equation(RFE),

$$\phi(x) = \langle T, \phi(2x - \cdot) \rangle, \qquad \phi(x) \in \mathcal{S}. \tag{5.1}$$

Here, T acts on the variable in the position of $\cdot$ and x plays the role of a parameter. Since $\mathcal{S}$ is invariant under translation $(\phi(t) \to \phi(t - a))$ and dilation $(\phi(t) \to \phi(at))$ for any non-zero constant a, the equation is well defined.

All solutions to a given RFE constitute a linear subspace of $\mathcal{S}$. In most interesting cases, it is a line. Therefore, to make the solution unique, we usually add another scalar character for the solution, such as

$$\langle 1, \phi \rangle = \int_R \phi(t) \, dt = c,$$

for a specified constant $c \neq 0$.

RFEs and RDEs are connected through the following concepts.

Definition 5.1 (δ-train). A *δ-train* is the functional

$$F = \sum_{n=-\infty}^{\infty} c_n \delta(x - n),$$

with the coefficients c_n satisfying the *temper growth* condition: there exists an integer K, such that

$$\sup_n \frac{|c_n|}{(1 + |n|)^K} < \infty.$$

This condition makes the δ-train a Schwartz distribution. If there are only finitely many non-zero c_ns, the δ-train is said to be compactly supported. The maximal non-negative integer L such that there exists m, $c_m c_{m-L} \neq 0$, is called the *length* of the train.

The most famous δ-train is the uniform train:

$$P = \sum_{n=-\infty}^{\infty} \delta(x - n),$$

which is closely related to the famous Poisson Summation Formula (see Strichartz [36]), and also is directly connected to Shannon's Sampling Theorem (see Oppenheim and Schafer [28]). The interesting class of δ-trains in wavelet theory are those with compact supports.

Definition 5.2 (δ-simple). A Schwartz distribution T is said to be δ-*simple* if it solves the distributional differential equation,

$$a_N T^{(N)} + a_{N-1} T^{(N-1)} + \cdots + a_1 T' + a_0 T = F, \qquad (5.2)$$

for some scalars a_k, $k = 0, 1, \ldots, N$, $a_N \neq 0$, and some δ-train F. N is called the *order* of T. If F is compactly supported and with length L, then so is T.

The following properties are obvious:

(a) T is δ-simple if and only if $T(-x)$ is.

(b) If T_1 and T_2 are δ-simple, and $T = T_1 * T_2$ is well-defined, then T is δ-simple.

(c) Suppose T is δ-simple. Then any finite sum of the following form is also δ-simple,

$$\sum_k c_k T(x - k).$$

It is easily seen that a δ-simple distribution is a linear combination of integer translates of the fundamental solution (to the associated ODE).

Theorem 5.1 (Second Main Theorem). *Suppose T is a δ-simple Schwartz distribution of order N and length L, and $\phi(x) \in \mathcal{S}$ is the solution to RFE (5.1). Then $\phi(x)$ is the solution to an RDE with the same order and length. Conversely, suppose a regular RDE of order N and length L has a non-trivial solution ϕ that belongs to the Schwartz class $\mathcal{S}$. Then $\phi(x)$ must solve an RFE with a δ-simple distribution mask T of the same order and length.*

Proof. Suppose T satisfies

$$a_N T^{(N)} + \cdots + a_1 T' + a_0 T = c_m \delta(t - m) + \cdots + c_{m-L} \delta(t - m + L).$$

Applying both sides to the t-function $\phi(2x - t)$ parameterized by x, we have

$$\sum_{k=0}^{N} a_k \langle T^{(k)}, \phi(2x - t) \rangle = \sum_{l=m-L}^{m} c_l \phi(2x - l).$$

On the other hand, the left side of the above equation is

$$\text{l.h.s} = \sum_{k=0}^{N} a_k \langle T, (-1)^k \partial_t^{(k)} \phi(2x - t) \rangle$$

$$= \sum_{k=0}^{N} a_k \left\langle T, \frac{\partial_x^{(k)}}{2^k} \phi(2x - t) \right\rangle$$

$$= \sum_{k=0}^{N} \frac{a_k}{2^k} D^k \langle T, \phi(2x - t) \rangle$$

where $D = d/dx$. Suppose $\phi(x)$ solves the RFE with mask T. Then

$$\text{l.h.s} = \sum_{k=0}^{N} \frac{a_k}{2^k} D^k \phi(x).$$

Hence, $\phi(x)$ solves the RDE of type $(P(\lambda), H(\lambda))$

$$P(\lambda) = \sum_{k=0}^{N} a_k \left(\frac{\lambda}{2} \right)^k, \qquad H(z) = \frac{1}{2} \sum_{l=m-L}^{m} c_l z^{-l}, \qquad (5.3)$$

which obviously has order N and length L.

The converse is proved in a similar way.   □

Relation (5.3) is very useful in computations.

Examples:

(i) Suppose T itself is a δ-train:

$$T = 2c_0\delta(x) + 2c_1\delta(x-1) + \cdots + 2c_L\delta(x-L).$$

Then T is δ-simple (of order 0 and length L). The corresponding RFE (5.1) directly gives the refinement equation in wavelet theory:

$$\phi(t) = 2c_0\phi(2t) + 2c_1\phi(2t-1) + \cdots + 2c_L\phi(2t-L).$$

(ii) Define

$$T = \alpha\exp(-\alpha|x|), \qquad \text{for some positive constant } \alpha.$$

Obviously, T is a Schwartz distribution solving

$$-\frac{T''}{\alpha^2} + T = 2\delta.$$

Therefore, T is a second-order δ-simple distribution of length 0. If $\phi(x) \in \mathcal{S}$ satisfies

$$\phi(x) = \int_R \alpha e^{-\alpha|t|}\phi(2x-t)\,dt,$$

then it must solve the second-order RDE of length 0,

$$-\frac{\phi''(x)}{4\alpha^2} + \phi(x) = 2\phi(2x).$$

(iii) Let T be the characteristic of interval $[-1,1]$:

$$\langle T, \phi(x)\rangle = \int_{-1}^{1}\phi(x)\,dx.$$

Then,

$$T' = \delta(x+1) - \delta(x-1).$$

Hence, the associated RDE is

$$\phi'(x) = 2\phi(2x+1) - 2\phi(2x-1),$$

which is exactly the Rvachev equation.

(iv) Set $T = \mathbf{1}_{x>0}(x)2\sin\frac{x}{2}$, which is the solution to the distributional differential equation,

$$4T'' + T = 2\delta.$$

Hence, the associated RDE is given by

$$\phi''(x) + \phi(x) = 2\phi(2x).$$

Remark. As an equation, RFE (5.1) is well defined for any Schwartz distribution. However, it may have no solutions in the Schwartz class $\mathcal{S}$. On the contrary, a regular RDE as a point-wise equation always can be solved as shown in the previous sections. For instance, in the last example, the associated RDE does have at least one non-trivial C^∞ solution which is periodic (see Section 3). But a non-zero Schwartz function cannot be periodic.

The second main theorem connects regular RDEs to a special class of RFEs with δ-simple distributional masks. It provides a new way to interpret and solve generic regular RDEs. In the section below, we show that the two "building block" functions, $\mathrm{up}(x)$ and $\mathrm{kam}(x)$, can be studied successfully in this way.

6. Probability method and continuous subdivision process

Following the discussion in the preceding section, we consider a special class of distribution masks called probability measure masks. The connection between Rvachev's up function and probability has been pointed out in Rvachev [30], Derfel [9], and Derfel et al. [10]. In this section, we develop systematically the probabilistic method for RDEs, especially for the two "building block" equations, the Rvachev equation and the kam equation. Derfel's generalized subdivision process is applied to generic RDEs. By "generic equations," we mean regular RDEs of type $(P(\lambda), H(z))$ such that $P(\lambda)$ contains no imaginary roots.

6.1. Probability interpretation of certain RFEs. When T is a probability distribution of some random variable X, RFE (5.1) can be rewritten as

$$\phi(x) = \mathbf{E}(2\phi(2x - X)). \tag{6.1}$$

Here $\mathbf{E}$ is the expectation operator. The first number 2 indicates that we are taking $2d\mu$ for T, if $d\mu$ stands for the probability measure of X.

Given a random variable X, we define an "X-averaging" operator $\mathcal{A}_X$, which transforms any random variable Y that is independent of X to a new random variable $\mathcal{A}_X(Y)$ defined by

$$\mathcal{A}_X(Y) = \frac{X + Y}{2}.$$

The "fixed point" of this operator is the most interesting. A random variable Y independent of X is called a *fixed point* of $\mathcal{A}_X$ if $\mathcal{A}_X(Y)$ has the same distribution as Y.

By recursion, it is not difficult to show the

Proposition 6.1. *Let* X, X_n, $n = 1, 2, \ldots$ *be a sequence of i.i.d. random variables on some probability space. If the following infinite series converges a.s. to a random variable* Y_X

$$Y_X = \sum_{n=1}^{\infty} \frac{X_n}{2^n},$$

then Y_X *is the unique (in the sense of distribution) fixed point for* $\mathcal{A}_X$.

The following lemma describes a qualitative relation between Y_X and X.

Lemma 6.1. *Suppose* $\mathrm{supp}\, X = [a, b]$ *and* $\mathrm{Prob}(c < X < d) > 0$ *for any* c, d, $a < c < d < b$. *Then,* Y_X *has the same properties.*

Proof. It is not difficult to see supp $Y_X \subset [a, b]$ from the infinite summation. Hence, we only need to show that Y_X shares the second property. Set $A = \max(|a|, |b|)$. Then,

$$\left| \sum_{n > N} \frac{X_n}{2^n} \right| \leq A 2^{-N}, \qquad \text{a.s.}$$

Denote $A 2^{-N}$ by δ. For any $c, d, a < c < d < b$,

$$\text{Prob}\big[c < Y_X < d\big] \geq \text{Prob}\Big[c + \delta < \sum_{n=1}^{N} \frac{X_n}{2^n} < d - \delta\Big]$$

$$\geq \text{Prob}\Big[\frac{c + \delta}{1 - 2^{-N}} < X_n < \frac{d - \delta}{1 - 2^{-N}} \ : \ n = 1, 2, \ldots, N\Big].$$

Choose N large enough so that $a < c' < d' < b$ where $c' = (c + \delta)/(1 - 2^{-N})$ and $d' = (d - \delta)/(1 - 2^{-N})$. Then

$$\text{Prob}(c < Y_X < d) \geq [\text{Prob}(c' < X < d')]^N > 0.$$

$\square$

By the standard truncation technique, a and/or b can be relaxed to ∞ provided that $\mathbf{E}(|X|) < \infty$.

For any random variable Y independent of X, define $Z = \mathcal{A}_X(Y)$. Suppose Y has the probability density function (p.d.f.) ρ^Y.

Proposition 6.2. *If the p.d.f. ρ^Z of Z exists, then*

$$\rho^Z(x) = \mathbf{E}(2\rho^Y(2x - X)).$$

Proof. This is because (1) $\mathbf{E}(\rho^Y(x - X))$ is the p.d.f. of $X + Y$ whenever X and Y are independent; and (2) $2\rho(2x)$ is the p.d.f. of $X/2$ if X has p.d.f. $\rho(x)$. $\square$

Corollary 6.1 (Probability Meaning of RFEs). *If the fixed point Y_X of $\mathcal{A}_X$ exists and has p.d.f. $\rho(x)$, then $\phi(x) = \rho(x)$ solves (6.1).*

Remark. In the above argument, we have left out some technical details of the regularity conditions on the random variables involved. For instance, in Proposition 6.1, it is not difficult to show that, by applying Kolmogrov's Three Series Theorem (see Rao [29], for example), if $\mathbf{E}(|X|) < \infty$, then the infinite series of Y_X does converge almost surely. We refer to Derfel et al. [10] for more information on convergence and regularity conditions.

6.2. The uniform distribution and up(x). Let $X_u \sim U[-1, 1]$ be a random variable uniformly distributed on $[-1, 1]$. The corresponding T is denoted by T_u and given by

$$T_u = \mathbf{1}_{[-1,1]}(x),$$

the characteristic of interval $[-1, 1]$. T_u is a δ-simple distribution of order 1 and length 2 since

$$T_u' = \delta(x + 1) - \delta(x - 1).$$

Hence, the solution $\phi(x)$ to (6.1) satisfies the following RDE according to the previous section

$$\phi'(x) = 2\phi(2x + 1) - 2\phi(2x - 1),$$

which is the Rvachev Equation. Thus, up to a multiplicative constant, the solution is $\mathrm{up}(x)$.

Corollary 6.2. $\mathrm{supp}[\mathrm{up}] = [-1, 1]$, and $\mathrm{up}(x) \geq 0$ for all $x \in [-1, 1]$.

Proof. Let X_n, $n = 1, 2, \ldots$. be a sequence of i.i.d. random variables of type X_u. Define

$$Y_u = \sum_{n=1}^{\infty} \frac{X_n}{2^n}.$$

By the second Main Theorem and the integral normalization condition, $\mathrm{up}(x) = \rho^Y(x)$. The proof is complete after applying the preceding lemma to (X_u, Y_u). $\square$

6.3. The exponential distribution and $\mathrm{kam}(x)$. Let X_e be any random variable with mean 2 and exponentially distributed along the positive half-axis. The corresponding T is denoted by T_e:

$$T_e = \exp(-\frac{x}{2})\mathbf{1}_{x \geq 0}.$$

T_e is a δ-simple distribution of first order since it satisfies

$$2T_e' + T_e = 2\delta(x).$$

Therefore, the solution $\phi(x)$ to the associated RFE also must solve

$$\phi'(x) + \phi(x) = 2\phi(2x).$$

Up to a multiplicative constant, the only solution is $\mathrm{kam}(x)$.

Corollary 6.3. $\mathrm{supp}[\mathrm{kam}] = [0, \infty)$, and $\mathrm{kam}(x) \geq 0$ for all $x > 0$.

Proof. Let X_n, $n = 1, 2, \ldots$ be i.i.d. random variables of exponential type X_e on some probability space. Define

$$Y_e = \sum_{n=1}^{\infty} \frac{X_n}{2^n}.$$

Then, Y_e is well defined, and its p.d.f. is given by

$$\rho^Y(x) = \frac{\mathrm{kam}(x)}{\exp_2(-1)}.$$

An application of the preceding lemma completes the proof. $\square$

6.4. The normal distribution and "fixed point". So far, we have discussed two important probability distributions. To be complete, it is natural to ask what is the function $\phi(x)$ that corresponds to the normal distribution. The answer is: the normal distribution does not lead to new functions because of the semi-group property:

$$N(0, t) + N(0, s) \frown N(0, t + s).$$

Here, $N(0, t)$ stands for a normal random variable of mean 0 and variance t, and the first two random variables are assumed to be independent.

The normal distribution serves as a famous "fixed point" in the Central Limit Theorem and Fourier transform. It does here also for the refinement equations.

For the normal distribution, we define the corresponding Schwartz distribution T_n to be

$$T_n = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{x^2}{2}\right).$$

It solves the following first-order equation,

$$T_n' + x T_n = 0,$$

from which it is obvious that T_n is not δ-simple. Hence, one should not expect that the solution $\phi(x)$ to (6.1) also can solve an RDE.

On the other hand, let X_n, $n = 1, 2, \ldots$ be a sequence of i.i.d. random variables of type $N(0,1)$ and define

$$Y = \sum_{n=1}^{\infty} \frac{X_n}{2^n}.$$

Because of the semi-group property, i.e., *the sum of any two independent normal random variables is still of normal type*, Y must be a normal random variable also. Since

$$\mathbf{E}(Y) = \sum_{n \geq 1} \frac{\mathbf{E}(X_n)}{2^n} = 0, \qquad \sigma^2(Y) = \sum_{n \geq 1} \frac{\sigma^2(X_n)}{4^n} = \frac{1}{3},$$

we conclude that the solution to (6.1) for the normal distribution is (subject to the integral normalization condition $\langle 1, \phi \rangle = 1$)

$$\phi(x) = \sqrt{\frac{3}{2\pi}} \exp\left(-\frac{3x^2}{2}\right).$$

Notice the major difference between the solution in this case and those in the previous two cases: the solution here is C^ω!

6.5. Continuous subdivision scheme for generic RDEs.
The connection between RFEs and RDEs makes it possible to solve generic RDEs using the continuous subdivision process (Derfel et al. [10]).

We first review briefly the role of the subdivision scheme in wavelet theory. For a given refinement equation

$$\phi(x) = 2 \sum_n h_n \phi(2x - n),$$

the associated subdivision scheme S is the "refining operator,"

$$(S\mathbf{f})[n] = 2 \sum_k h_{n-2k} \mathbf{f}[k], \qquad n = 0, \pm 1, \ldots,$$

which maps an infinite sequence $\mathbf{f}[n]$ to another (refined) sequence $S\mathbf{f}$ (Daubechies [4]). The subdivision scheme shall be better viewed as a grid transfer function in the Multigrid Method (see Briggs [1]). The subdivision process (SP) refers to the iteration process starting with the "impulse signal" $\delta[n]$

$$\delta, S\delta, S^2\delta, \ldots.$$

It is said to be *convergent uniformly* if there exists a continuous function $\phi(x)$ such that

$$\lim_{j \to \infty} \|S^j \delta - \phi_j\|_\infty = 0$$

where sequence ϕ_j is defined by $\phi_j[k] = \phi(k/2^j)$. Obviously if the SP converges, $\phi(x)$ must be unique. Indeed, in the case of finite length filter, h_n, $\phi(x)$ is the unique solution to the refinement equation subject to the integral normalization condition $\int \phi(x)dx = 1$.

SP is very useful in computer graphics for generating continuous objects from discrete data (see Deslauriers and Dubuc [11], and Dyn and Levin [13]). It also appears in the analysis of the Picard-Lindelöf iteration in numerical computation of ODE systems (see Nevanlinna [27]).

For refinement equations with continuous masks, Derfel et al. [10] generalizes the subdivision process in a natural way. Their results now also apply to RDEs based on the connection we established in the previous section.

Given a refinement differential equation of type $(P(\lambda), H(z))$, suppose T is the δ-simple distribution associated to it. Assume T has a "density" function $\rho(x) \in L^1(R)$ (not necessary to be non-negative), i.e.,

$$\langle T, g(x) \rangle = \int_R \rho(x)g(t)\, dx,$$

for any test function $g(x)$. We require $\int_R \rho(x)\, dx = 2$.

The continuous subdivision scheme is the operator S_c (subscript c stands for "continuous") defined by

$$S_c f(x) = \int_R \rho(x - 2t)f(t)\, dt.$$

It is easy to see that for any p with $1 \le p \le \infty$,

$$S_c(L^p(R)) \subset L^p(R).$$

Let $\delta = \delta(x)$ be the Dirac distribution. The generalized SP is the iteration:

$$\delta, S_c\delta, S_c^2\delta, \ldots.$$

It is said to converge *uniformly* to a continuous function $\phi(x)$ if

$$\lim_{j \to +\infty} \|S_c^j\delta(x) - \phi(2^{-j}x)\|_\infty = 0$$

and converge *weakly* to a distribution R if for any test function $g(x)$,

$$\lim_{j \to +\infty} \int S_c^j\delta(x)2^{-j}g(2^{-j}x)\, dx = \langle R, g(x) \rangle.$$

The result of Derfel et al. [10] leads to the following.

Proposition 6.3 (Derfel et al. [10], Corollary 15 modified). *If $\rho(x)$ is rapidly decreasing, then the generalized SP converges in the weak sense to an infinitely differentiable function in $L^1(R)$, which is the unique solution of the corresponding RFE in $L^1(R)$ subject to the integral normalization condition. Furthermore, if $\rho(x)$ consists of finitely many smooth pieces, then the convergence is uniform in any $C^m(R)$.*

The continuous subdivision process points to an algorithm for generic regular RDEs. We demonstrate it through the following three examples. Notice that the SP also provides an efficient way to compute convolutions like $f * \phi$ for an arbitrary function $f(x)$ if $\phi(x)$ is an solution to certain RFE: one does not need to know the explicit expression of $\phi(x)$ and simply starts the SP with f, instead of δ.

Examples:

(i) Rvachev equation: $\phi'(x) = 2\phi(2x+1) - 2\phi(2x-1)$.

In this case, $\rho_u(x) = \mathbf{1}_{[-1,1]}(x)$. The subdivision scheme S_u is given by

$$S_u f(x) = \int_R \rho_u(x - 2t) f(t)\, dt = \int_{\frac{x-1}{2}}^{\frac{x+1}{2}} f(t)\, dt.$$

Hence,

$$S_u \delta(x) = \rho_u(x) = \mathbf{1}_{[-1,1]}(x),$$

$$S_u^2 \delta = \begin{cases} \frac{x+3}{2} & -3 \leq x < -1, \\ 1 & -1 \leq x < 1, \\ \frac{3-x}{2} & 1 \leq x < 3, \\ 0, & \text{elsewhere} \end{cases}$$

$$\cdots .$$

If we look at the rescaled (spline) functions $S_u \delta(2x)$, $S_u^2 \delta(4x)$, ... it is easy to observe the following properties (those invariant during the SP) of the limiting function $\mathrm{up}(x)$:

(a) $\mathrm{supp}[\mathrm{up}] = [-1, 1]$ and $\mathrm{up}(x) > 0$ for all $x \in (-1, 1)$;

(b) $\mathrm{up}(x)$ is infinitely flat at $x = \pm 1$ and $x = 0$, and $\mathrm{up}(0) = 1$, $\mathrm{up}(\pm 1) = 0$;

(c) $\mathrm{up}\,|_{[0,1]}$ is mirror symmetric around $\frac{1}{2}$: $\mathrm{up}(1 - x) = 1 - \mathrm{up}(x)$.

(ii) The kam equation: $\phi'(x) + \phi(x) = 2\phi(2x)$.

Here we have $\rho_e(x) = \exp(-\frac{x}{2})\mathbf{1}_{x>0}$. The subdivision scheme S_e is defined by

$$S_e f(x) = \int_R \rho_e(x - 2t) f(t)\, dt = e^{-\frac{x}{2}} \int_{-\infty}^{\frac{x}{2}} e^t f(t)\, dt.$$

In particular, if $\mathrm{supp}\, f \subset [0, \infty)$, then

$$S_e f(x) = \mathbf{1}_{x>0} e^{-\frac{x}{2}} \int_0^{\frac{x}{2}} e^t f(t)\, dt.$$

Hence, the first few steps of the SP are given by

$$S_e \delta = \rho_e = \exp(-\frac{x}{2})\mathbf{1}_{x>0}, \qquad S_e^2 \delta = [2e^{-\frac{x}{4}} - 2e^{-\frac{x}{2}}]\mathbf{1}_{x>0},$$

$$S_e^3 \delta = \left[\frac{8}{3}e^{-\frac{x}{8}} - 4e^{-\frac{x}{4}} + \frac{4}{3}e^{-\frac{x}{2}}\right]\mathbf{1}_{x>0}, \qquad \cdots .$$

Generally, $S_e^k \delta$ can be obtained by solving a linear system in the following way. Assume

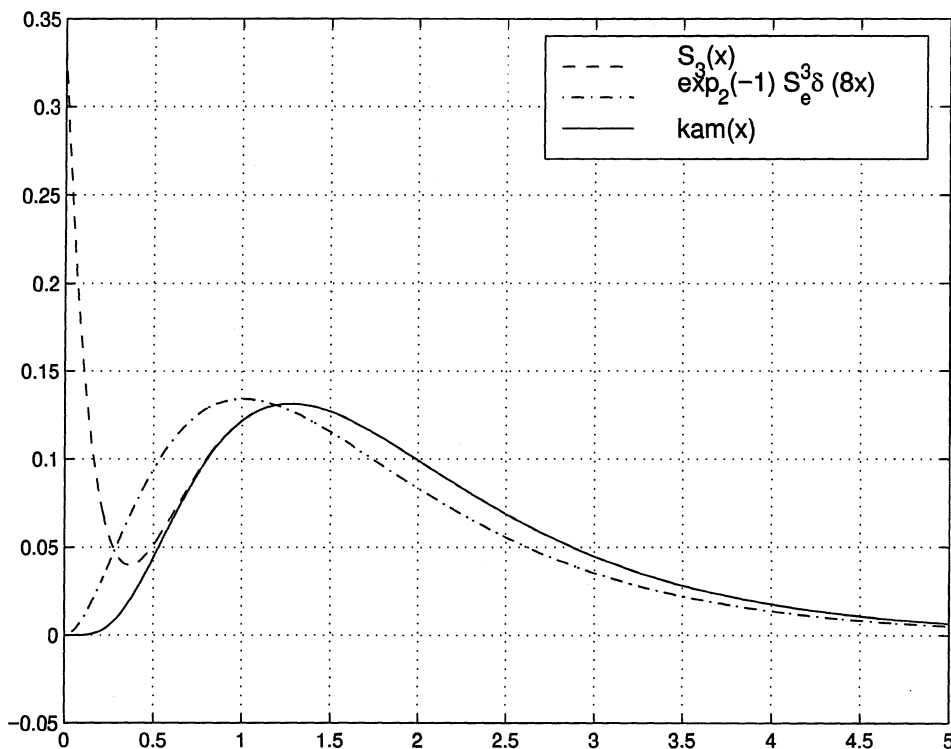$$S_e^k \delta(x) = \mathbf{1}_{x>0} \sum_{j=1}^{k} c_j e^{-x/2^j}.$$

FIGURE 3. $\exp_2(-1)S_e^k\delta(2^kx)$ gives a better uniform approximation to kam(x) compared with the partial sum sequence $S_k(x)$, $k=3$.

To determine the coefficients c_j, we impose the following k conditions:

$$\int S_e^k\delta(2^kx)dx = 1; \quad \frac{d^m}{dx^m}\Big|_{x=0}S_e^k\delta(x) = 0, \quad m = 0,1,\ldots,k-2.$$

This establishes a linear system of Vandermonde type

$$\sum_{j=1}^{k} c_j2^j = 2^k; \quad \sum_{j=1}^{k} c_j2^{-jm} = 0, \quad m = 0,1,\ldots,k-2.$$

It is well defined and has a unique solution.

Now two sequences of functions are available to approximate kam(x): the scaled SP sequence $\exp_2(-1)S_e^k\delta(2^kx)$ and the k-th partial sum in (3.3):

$$S_k(x) = \sum_{m=0}^{k-1} \frac{(-2)^m}{(m)_2!} \exp(-2^mx).$$

From Figure 3, the SP sequence gives a better uniform approximation. $S_k(x)$ is only good away from $x = 0$.

(iii) $\frac{\phi''(x)}{2} + \phi'(x) + \phi(x) = 2\phi(2x)$.

The "density" function for this equation is

$$\rho(x) = \mathbf{1}_{x>0} e^{-\frac{x}{2}} \sin \frac{x}{2}.$$

Hence, the resulting subdivision scheme S for any function f supported in the positive half-axis is

$$Sf(x) = \mathbf{1}_{x>0} e^{-\frac{x}{2}} \left[\sin \frac{x}{2} \int_0^{\frac{x}{2}} e^t \cos t f(t) \, dt - \cos \frac{x}{2} \int_0^{\frac{x}{2}} e^t \sin t f(t) \, dt \right].$$

It is guaranteed by the preceding proposition that $S^k \delta(2^k x)$, $k = 0, 1, \ldots$ converges uniformly to $\sqrt{2}\Phi_{\frac{\pi}{4}}(\sqrt{2}x)$ up to a multiplicative constant.

Remark. The probability method and subdivision algorithm in this section recently have been generalized to combinatorics by the author [32].

7. Application: smoothed wavelets and quasi-multiresolution

In this section, we present one application of the previous results in wavelet theory, namely, the construction of smoothed wavelets and quasi-multiresolution.

The application is controversial because the "smoothed wavelets" are, in fact, no longer wavelets in the classical sense. The multiresolution axioms now only hold approximately!

However, if we go beyond the multiresolution idea and open our minds more widely to the field, the application brings a lot of good news. The new class of functions constructed below are C^∞ and nearly orthogonal. Their essential supports are almost the same as their classical counterparts. Most importantly, their derivatives can be expressed in terms of their translated and dilated copies, a merit so crucial to the numerical computation of differential equations, but pitifully and inevitably missed by classical wavelets.

Therefore, we optimistically anticipate more applications in the future in numerical computations and data processing (in statistics, for example).

7.1. Classical wavelets with compact support. Wavelets with compact supports are of particular interest in applications. The design starts with the refinement equation,

$$\phi(x) = 2[H(E)\phi](2x) = 2 \sum_{n=m-L}^{m} h_n \phi(2x - n). \tag{7.1}$$

Once the scaling function has been worked out, the associated (mother) wavelet $\psi(x)$ is obtained from the wavelet equation,

$$\psi(x) = 2[G(E)\phi](2x) = 2 \sum_{k=p-M}^{p} g_k \phi(2x - k). \tag{7.2}$$

Here

$$G(z) = \sum_{k=p-M}^{p} g_k z^{-k}$$

is the companion *highpass filter* of $H(z)$, which must satisfy the highpass condition: $G(1) = 0$.

The scaling function and wavelet generate a multiresolution (MR) in the following way. For any function $f(x) \in L^2(R)$ and integer j, denote by $V_j(f)$ the closed subspace in $L^2(R)$,

$$\overline{\text{span}\{f(2^j x - k) \mid k = 0, \pm 1, \dots\}}.$$

MR is the iteration (w.r.t. j) of the relation in $L^2(R)$ given by

$$V_j(\phi) = V_{j-1}(\phi) \oplus V_{j-1}(\psi).$$

Here $\oplus$ denotes the direct sum of subspaces. For orthogonal multiresolution, it is an orthonormal sum.

With more restrictions on the filters, the resulting MR is complete in the sense that $\bigcup_{j \in \mathbf{Z}} V_j(\phi)$ is dense in $L^2(R)$. Further conditions on the filters can lead to the orthogonality of the scaling functions and wavelets (Daubechies [4]).

One major defect of classical compact wavelets is that they are not C^∞. In some applications such as applying the wavelet-Galerkin method to solve differential equations numerically, C^∞ (or piecewise C^∞) basis functions are welcomed (classically, trigonometric functions, splines, orthogonal polynomials, and eigenfunctions of a Sturm-Liouville problem). The naive idea to achieve high-order smoothness is to mollify the scaling functions and wavelets in existence. This is indeed what the following perturbed refinement equation (a special RDE) achieves.

7.2. Smoothed wavelets and quasi-multiresolution. For a system of refinement and wavelet equations defined through a given filter pair $(H(z), G(z))$, let us consider the following first-order perturbed system,

$$\epsilon \phi'_\epsilon(x) + \phi_\epsilon(x) = 2[H(E)\phi_\epsilon](2x), \tag{7.3}$$

$$\epsilon \psi'_\epsilon(x) + \psi_\epsilon(x) = 2[G(E)\phi_\epsilon](2x), \tag{7.4}$$

where ϵ is a small perturbation parameter and assumed to be positive harmlessly. We impose again the normalization condition $\int \phi_\epsilon(x)\, dx = 1$.

By the Structure Theorem, if $(\phi(x), \psi(x))$ solves the original unperturbed system, then the solution to the above system is

$$\phi_\epsilon(x) = K_\epsilon(x) * \phi(x), \qquad \psi_\epsilon(x) = K_\epsilon(x) * \psi(x)$$

where $K_\epsilon(x)$ is the scaled kam function

$$K_\epsilon(x) = \frac{\epsilon^{-1} \text{kam}(\epsilon^{-1} x)}{\exp_2(-1)}.$$

It is interesting to see that $K_\epsilon(x)$ plays the exact role of a *mollifier* in functional analysis. Since both $\phi(x)$ and $\psi(x)$ are compactly supported, we have:

(1) ϕ_ϵ and ψ_ϵ both belong to the Schwartz class, and

(2) ϕ_ϵ and ψ_ϵ converge to ϕ and ψ in $C^\alpha(R)$ and $L^p(R)$ whenever $\phi \in C^\alpha(R)$ for some $\alpha \geq 0$.

If ϵ is small enough, $V_j(\phi_\epsilon)$ and $V_j(\psi_\epsilon)$ are two subspaces very "close" to $V_j(\phi)$ and $V_j(\psi)$ in the sense that Shen and Strang gave in [34]. Thus it can be expected that the following relation should still hold approximately:

$$V_j(\phi_\epsilon) \simeq V_{j-1}(\phi_\epsilon) \oplus V_{j-1}(\psi_\epsilon).$$

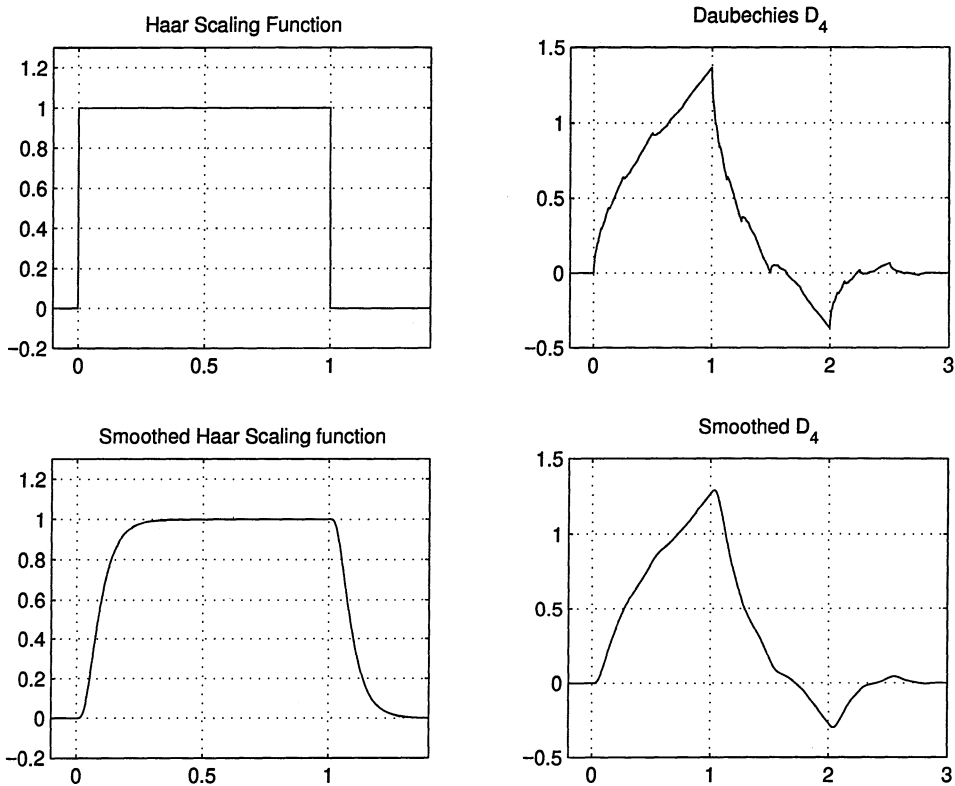The iteration of this approximate relation is called *a quasi-multiresolution* (QMR).

FIGURE 4. First-order perturbed Haar scaling function and D_4 ($\epsilon = 0.04$).

If the original unperturbed system is orthogonal, that is,

$$\int_R \phi(x)\phi(x-n)\,dx = \delta_n, \qquad \int_R \phi(x)\psi(x-n)\,dx = 0,$$

$$\int_R \psi(x)\psi(x-n)\,dx = \delta_n,$$

then the perturbed system must satisfy

$$\int_R \phi_\epsilon(x)\phi_\epsilon(x-n)\,dx = \delta_n + s_n, \qquad \int_R \phi_\epsilon(x)\psi_\epsilon(x-n)\,dx = r_n, \qquad (7.5)$$

$$\int_R \psi_\epsilon(x)\psi_\epsilon(x-n)\,dx = \delta_n + w_n, \qquad (7.6)$$

where sequences (s_n), (r_n), (w_n) are uniformly (for $\epsilon \leq 1$) exponentially small for large $|n|$ and the sequence supremum norms are $O(\epsilon)$ (in fact, $O(\epsilon^2)$ as we shall show later). This is to say that the resulting QMR is *near orthogonal*.

For example, in Figure 4, we have plotted the smoothed Haar scaling function and Daubechies min-phase orthogonal scaling function D_4 (see Daubechies [4]), both obtained by choosing the perturbation parameter ϵ to be 0.04.

7.3. Smoothing versus small shifting. In this section, we show that the solutions to the perturbed system (7.3) and (7.4) are very close to a small shifting of the original scaling functions and wavelets.

Let us consider another perturbed system,

$$\bar{\phi}_\epsilon(x + \epsilon) = 2[H(E)\bar{\phi}_\epsilon](2x), \tag{7.7}$$

$$\bar{\psi}_\epsilon(x + \epsilon) = 2[G(E)\bar{\phi}_\epsilon](2x), \tag{7.8}$$

with the same integral normalization condition for the scaling function.

Proposition 7.1 (Small Shifting). *If $(\phi(x), \psi(x))$ is the scaling function and wavelet pair for the original system, then (7.7) and (7.8) have solutions*

$$\bar{\phi}_\epsilon(x) = \phi(x - 2\epsilon), \qquad \bar{\psi}_\epsilon(x) = \psi(x - 2\epsilon).$$

Proof. A direct check using the original refinement and wavelet equation. $\qquad\square$

When the lowpass filter $H(z)$ vanishes rapidly at $\omega = \pi$ or $z = -1$, $\phi(x)$ and $\psi(x)$ are differentiable for finitely many times, so are $\bar{\phi}_\epsilon$ and $\bar{\psi}_\epsilon$ according to the proposition. We can apply Taylor expansion to ϵ:

$$\bar{\phi}_\epsilon(x + \epsilon) = \bar{\phi}(x) + \epsilon\bar{\phi}'(x) + O(\epsilon^2), \qquad \bar{\psi}_\epsilon(x + \epsilon) = \bar{\psi}(x) + \epsilon\bar{\psi}'(x) + O(\epsilon^2).$$

Hence, up to the second order,

$$\epsilon\bar{\phi}'_\epsilon(x) + \bar{\phi}_\epsilon(x) \simeq 2[H(E)\bar{\phi}_\epsilon](2x),$$

$$\epsilon\bar{\psi}'_\epsilon(x) + \bar{\psi}_\epsilon(x) \simeq 2[G(E)\bar{\phi}_\epsilon](2x).$$

That is, $(\bar{\phi}_\epsilon, \bar{\psi}_\epsilon)$ solves the first perturbed system (7.3) and (7.4) approximately. In fact,

Corollary 7.1 (Smoothing versus Small Shifting). *Suppose the original scaling function $\phi(x)$ is $C^\alpha(R)$ for some $\alpha \geq 2$. Then, uniformly for all x,*

$$\phi_\epsilon(x) = \phi(x - 2\epsilon) + O(\epsilon^2), \qquad \psi_\epsilon(x) = \psi(x - 2\epsilon) + O(\epsilon^2).$$

Proof. We only sketch the proof. Set $\Delta_\epsilon(x) = \bar{\phi}_\epsilon(x) - \phi_\epsilon(x)$. By the regularity condition, we can assume that

$$\Delta_\epsilon(x) = \epsilon\Delta_1(x) + \frac{\epsilon^2}{2}\Delta_2(x) + \cdots$$

where $\Delta_1(x), \Delta_2(x), \ldots$ are functions independent of ϵ. By the integration normalization condition, $\int_R \Delta_k(x)\,dx = 0$ for $k = 1, 2, \ldots$. On the other hand, since

$$\bar{\phi}_\epsilon(x + \epsilon) = \bar{\phi}_\epsilon(x) + \epsilon\bar{\phi}'_\epsilon(x) + \frac{\epsilon^2}{2}\bar{\phi}''_\epsilon(x) + \cdots,$$

we have

$$2[H(E)\Delta_\epsilon](2x) - \Delta_\epsilon(x) - \epsilon\Delta'_\epsilon(x) = \frac{\epsilon^2}{2}\bar{\phi}_\epsilon(x) + \text{higher order } \epsilon \text{ terms.}$$

Comparison of the first-order ϵ terms yields

$$\Delta_1(x) = 2[H(E)\Delta_1](2x).$$

Hence, Δ_1 is a constant multiple of the unperturbed scaling function $\phi(x)$. It must be 0 since $\int_R \Delta_1(x)\,dx = 0$. Therefore,

$$\bar{\phi}_\epsilon(x) - \phi_\epsilon(x) = \Delta_\epsilon(x) = \frac{\epsilon^2}{2}\Delta_2(x) + \cdots = O(\epsilon^2).$$

The proof for $\psi_\epsilon(x)$ is done in a similar manner. $\square$

Therefore, visually, $\phi_\epsilon(x)$ accomplishes two things simultaneously: infinitely smoothing the original scaling function $\phi(x)$ (hence, also the wavelet) and shifting it (rightward) by a small distance 2ϵ.

Corollary 7.2 (Linear Orthogonality). *Suppose the original scaling function and wavelet lead to an orthogonal MR, and the scaling function is at least C^2, then the sequences (s_n), (r_n), and (w_n) in (7.5) and (7.6) are $O(\epsilon^2)$. In such a case, we say that the QMR is linearly orthogonal.*

Finally, we point out that if one considers the second-order perturbed refinement equation (or wavelet equation),

$$\frac{\epsilon^2}{2}\phi''_{\epsilon,2}(x) + \epsilon\phi'_{\epsilon,2}(x) + \phi_{\epsilon,2}(x) = 2[H(E)\phi_{\epsilon,2}](2x), \qquad \int_R \phi_{\epsilon,2}(x)\,dx = 1,$$

then the following results can be established in the same fashion.

(1) $\phi_{\epsilon,2} = K_{\epsilon,2} * \phi$ where

$$K_{\epsilon,2} = \frac{1}{[\exp_2(-1)]^2}\frac{\sqrt{2}}{\epsilon}\Phi_{\frac{\pi}{4}}\left(\frac{\sqrt{2}}{\epsilon}x\right),$$

(2) $\phi_{\epsilon,2}$ belongs to the Schwartz class, and

(3) (Quadratic Orthogonality) If the original scaling function $\phi(x)$ is C^α for some $\alpha \geq 3$, then

$$\phi_{\epsilon,2}(x) = \phi(x - 2\epsilon) + O(\epsilon^3),$$
$$\psi_{\epsilon,2}(x) = \psi(x - 2\epsilon) + O(\epsilon^3),$$

and the uniform norm of (s_n), (r_n), and (w_n) all are $O(\epsilon^3)$ if the original MR is orthogonal. Hence, the QMR is "more" orthogonal than the previous case.

References

1. W. L. Briggs, *A Multigrid Tutorial*, SIAM, Philadelphia, 1987.
2. G. Brown and W. Moran, *A dichotomy for infinite products of discrete measures*, Proc. Cambridge Philos. Soc. **73** (1973), 307–316.
3. I. Daubechies, *Orthogonal bases of compactly supported wavelets*, Comm. Pure. Appl. Math. **41** (1988), 909–996.
4. ———, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

5. I. Daubechies and J. Lagarias, *Two scale difference equations I*, SIAM J. Math. Anal. **22** (1991), 1388–1410.

6. _____, *Two scale difference equations II*, SIAM J. Math. Anal. **23** (1992), 1031–1079.

7. N. G. de Bruijn, *The difference-differential equation $F'(x) = e^{\alpha x + \beta} F(x-1)$*, I., II. Nederl. Akad. Wetensch. Proc. Ser. A **56**; Indag. Math. **15** (1953), 449–464.

8. J. A. de Reina Martinez, *Definicio y estudio de una funcion indefinidamente diferenciable de soporto compacto*, Rev. Real Cienc. Acad. Exact. Fis. Natur. Madrid **76** (1982), 21–38, 1982.

9. G. Derfel, *Probabilistic method for a class of functional-differential equations*, Ukrain. Math. J. *41* (1989), 1117–1234.

10. G. Derfel, N. Dyn, and D. Levin, *Generalized refinement equations and subdivision processes*, J. Approx. Theory **80** (1995), 272–297.

11. G. Deslauriers and S. Dubuc, *Interpolation dyadique*. In: Fractals, dimensions non entiès et applications, Masson, Paris, pp.44–55, 1987.

12. _____, *Symmetric iterative interpolation processes*, Constr. Approx. **5** (1989), 49–68.

13. N. Dyn and D. Levin, *Interpolating subdivision schemes for the generation of curves and surfaces*, Internat. Ser. Numer. Math. **94** (1990), 91–106.

14. P. Erdös, *On a family of symmetric Bernoulli convolutions*, Amer. J. Math. **61** (1939), 974–976.

15. _____, *On the smoothness properties of a family of symmetric Bernoulli convolutions*, Amer. J. Math. **62** (1940), 180–186.

16. L. Fox and D. F. Mayers, *On a functional differential equation*, J. Inst. Maths. Applics. **8** (1971), 271–307.

17. P. O. Frederickson, *Dirichlet series solutions for certain functional differential equations*, Lecture Notes in Mathematics **243** (1971), 249–254.

18. A. Garsia, *Arithmetic properties of Bernoulli convolutions*, Trans. Amer. Math. Soc. **102** (1962), 409–432.

19. J. Goldman and G.-C. Rota, *On the foundations of combinatorial theory IV: finite vector spaces and Eulerian generating functions*, Stud. Appl. Math. **49**, 239–258.

20. _____, *The number of subspaces of a vector space*. In: Recent Progress in Combinatorics , (W. Tutte ed.), Academic Press, New York, 1969.

21. B. Jessen and A. Wintner, *Distribution functions and the Riemann zeta function*, Trans. Amer. Math. Soc. **38** (1935), 48–88.

22. T. Kato and J. B. McLeod, *The functional-differential equation: $y'(x) = ay(\lambda x) + by(x)$*, Bull. Amer. Math. Soc. **77** (1971), 891–937.

23. G. Kh. Kirov and G. A. Totkov, *Distribution of zeros of derivatives of the function $\lambda(x)$*. In: Differential Equations and Applications, pp.341–344, I. II. tech. univ., Ruse 1982.

24. K. Mahler, *On a special functional equation*, J. London Math. Soc. **15** (1940), 115–123.

25. S. Mallat, *Multiresolution approximation and wavelets*, Trans. Amer. Math. Soc. **315** (1989), 69–88.

26. A. D. Myshkis, *On certain problems in the theory of differential equations with deviating argument*, Russian Math. Surveys **32** (1977), 181–213.

27. O. Nevanlinna, *Power bounded prolongations and Picard-Lindelöf iteration*, Numer. Math. **58** (1990), 479–501.

28. A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing*, Prentice Hall, New Jersey, 1989.

29. B. L. S. Prakasa Rao, *Asymptotic Theory of Statistical Inference*, John Wiley & Sons, 1986.

30. V. A. Rvachev, *Compactly supported solutions of functional-differential equations and their applications*, Russian Math. Surveys *45* (1990), 87–120.

31. V. L. Rvachev and V.A. Rvachev, *On a function with compact support*, Dopov. Akad. Nauk. URSR **8** (1971), 705–707.

32. J. Shen, *Combinatorics for wavelets: the umbral refinement equation*, Studies Appl. Math., to appear.

33. J. Shen and G. Strang, *On wavelet fundamental solutions to the heat equation – heatlets*, Submitted to J. Diff. Eq., 1998.

34. ———, *Asymptotic analysis of Daubechies polynomials*, Proc. Amer. Math. Soc. *124* (1996), 3819–3833.

35. G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley MA, 1996.

36. R. Strichartz, *A Guide to Distribution Theory and Fourier Transform*, CRC Press, Florida, 1993.

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139.
E-mail: `jhshen@math.mit.edu`.

(After Sept. 1, 1998) DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES, CA 90095-1555.
E-mail: `jhshen@math.ucla.edu`.