NUMER. FUNCT. ANAL. AND OPTIMIZ., 15(1&2), 1–21 (1994)

SAMPLING PROCEDURES IN FUNCTION SPACES AND ASYMPTOTIC EQUIVALENCE WITH SHANNON'S SAMPLING THEORY

Akram Aldroubi and Michael Unser

Biomedical Engineering and Instrumentation Program
National Institutes of Health
Bethesda, Maryland

ABSTRACT

We view Shannon's sampling procedure as a problem of approximation in the space $S = \{s:s(x) = (c * \text{sinc})(x),\ c \in l_2\}$. We show that under suitable conditions on a generating function $\lambda \in L_2$, the approximation problem onto the space $V = \{v:v(x) = (c * \lambda)(x),\ c \in l_2\}$ produces a sampling procedure similar to the classical one. It consists of an optimal prefiltering, a pure jitter-stable sampling, and a postfiltering for the reconstruction. We describe equivalent signal representations using generic, dual, cardinal, and orthogonal basis functions and give the expression of the corresponding filters. We then consider sequences λ^n, where λ^n denotes the n-fold convolution of λ. They provide a sequence of increasingly regular sampling schemes as the value of n increases. We show that the cardinal and orthogonal pre- and postfilters associated with these sequences asymptotically converge to the ideal lowpass filter of Shannon. The theory is illustrated using several examples.

1. INTRODUCTION

A fundamental question in signal processing is how to represent a function defined on $\mathbb{R}$ in terms of a discrete sequence. One way is to sample the function on a uniform grid. However, there are infinitely many continuous functions having the same sample values. A class of functions completely characterized by their sample values is the bandlimited functions described by the Shannon-Whittaker Repre-

Correspondence to: Akram Aldroubi, Rm. 3W13, Bldg. 13, BEIP, National Institutes of Health, Bethesda, MD 20892.

sentation Theorem [23, 25]. Specifically, if $S(f) \in L_2$ is the Fourier transform of $s(x)$ and if $s(x) \in B_{1/2}$, i.e., if Support $(S) \subset [-\frac{1}{2}, \frac{1}{2}]$, then this representation is given by

$$s(x) = \sum_{i=-\infty}^{i=+\infty} s(i)\operatorname{sinc}(x - i) \tag{1}$$

where $\operatorname{sinc}(x) = \sin(\pi x)/\pi x$.

In the mathematical and signal processing literature, this result is known as Shannon's Sampling Theorem. It can be interpreted as an interpolation formula for the class of bandlimited signals. This point of view has been the basis for many generalizations of the classical sampling theory, mainly for nonuniform sampling [5, 13, 15]. We will call these type of extensions *interpolation-sampling* (see the survey papers [7, 11, 14]). A unified view is obtained from representation theorems in *Reproducing Kernel Hilbert Spaces* (RKHS) [19, 26]. Other extensions use the relations between various sampling expansions and Sturm-Liouville boundary value problems [27]. The converse of the Shannon-Whittaker representation is also true, i.e., a well defined series (1) is a bandlimited function.

Using Poisson's summation formula, we can evaluate the Fourier transform of (1)

$$S(f) = \operatorname{rect}(f) \sum_{i=-\infty}^{i=+\infty} S(f - i) \tag{2}$$

where the rect function is given by

$$\operatorname{rect}(f) = \begin{cases} 1 & |f| \leq 1/2 \\ 0 & \text{elsewhere} \end{cases} \tag{3}$$

This function is the ideal lowpass filter used in Shannon's sampling theory. Its effect in (1) is to suppress all the frequency components that are not in the bandpass region (i.e., the interval $[-\frac{1}{2}, \frac{1}{2}]$).

When a signal is not bandlimited, it is first preprocessed and forced to be bandlimited. This is done by prefiltering the signal with the ideal lowpass filter before sampling, as schematized in Figure 1. In effect, this is equivalent to finding the L_2 least square approximation of the signal in the space $B_{1/2}$. This procedure, which prevents aliasing errors, is widely used in signal processing in connection with the Sampling Theorem [23]. It has been extended to signals in finite dimensional spaces [20] and to the representation of functions using polynomial splines [12, 24]. We will call these types of extensions *approximation-sampling procedures*

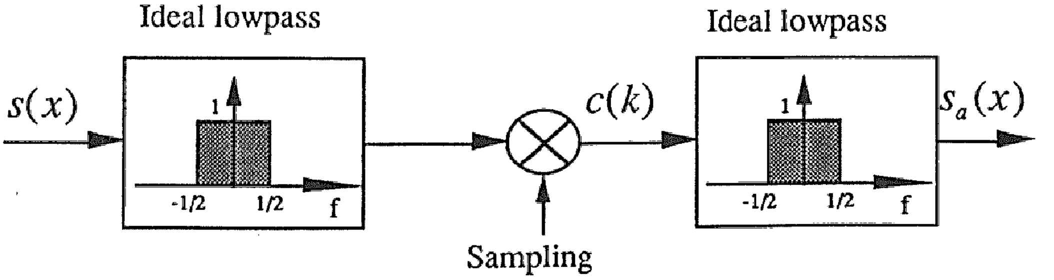

Figure 1 Block diagram of Shannon's sampling procedure for non-bandlimited signals. The samples $c(k) = s_a(k)$ are the coefficients of $s_a(x)$ in the basis $\{\operatorname{sinc}(x - k)\}_{k\in\mathbb{Z}}$ of $B_{1/2}$.

to distinguish them from the interpolation-sampling techniques discussed in the previous paragraph.

In this paper, we adopt the approximation-sampling point of view and propose some generalizations. First, we will consider the extension of Shannon's sampling procedure for the class of "shift-invariant" function spaces of the form $V(\lambda) = \{v: v(x) = (c * \lambda)(x)\ c \in l_2\}$ where λ is an appropriate function of L_2. An important point will be to find conditions on λ that insure that $V(\lambda)$ is a closed subspace of L_2 with $\{\lambda(x - k)\}_{k\in\mathbb{Z}}$ as its Riesz-Schauder basis. We will also show that the orthogonal projection of a function in $V(\lambda)$ can be computed using a procedure similar to the one described in Figure 1 with an appropriate set of filters. Second, we will investigate the properties of sequences of increasingly regular approximation spaces obtained by using the n-fold convolution of a regular function λ. This construction provides a natural extension of the sampling theory for polynomial splines described in [12, 24]. It is also motivated by a convergence result established by de Boor, Höllig, and Riemenschneider [8]. We will show that, under mild conditions, some of the corresponding approximation and interpolation filters tend to the ideal lowpass filter of Shannon as n goes to infinity. The connection between the present results and the theory of multiresolutions and wavelets is discussed in [2]. The results also appear, in a simplified version without proofs, in our chapter [3].

The paper is organized as follows: We start by introducing some definitions and notation in Section 2. In Section 3, we consider the problem of the L_2 approximation of a function in $V(\lambda)$ and derive a general sampling procedure; we also provide extensions for the approximation in the norm associated with the Sobolev spaces $W^{p,2}$. In Section 4, we describe equivalent representations of signals using different sets of basis functions with some special properties (orthogonality, interpolation, duality). In Section 5, we look at the families of n-fold convolution sequences and study their asymptotic properties. Finally, in Section 6, we illustrate the theory with several examples.

2. DEFINITIONS AND NOTATION

The signals considered here are real valued functions defined on $\mathbb{R}$. They belong to the space of measurable, square-integrable functions: L_2. We also consider the Sobolev spaces $W^{m,p}$, which consist of L_p functions with m distributional derivatives in L_p [1]. In particular, a function g belongs to $W^{m,2}$ if and only if its Fourier transform has the property that $(1 + |f|^m)\hat{g}(f)$ is in L_2. The space of square summable real valued sequences (discrete signals) is denoted by l_2 and its associated norm by $\|\cdot\|_{l_2}$.

The symbol $*$ will be used for three slightly different binary operations: the convolution, the mixed convolution, and the discrete convolution; these are defined below. The ambiguity should be easily resolved from the context. For two functions f and g defined on $\mathbb{R}$, $*$ denotes the usual convolution

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(\xi)g(x - \xi)\, d\xi \qquad x \in \mathbb{R} \tag{4}$$

The mixed convolution between a sequence $b(k)$, $k \in \mathbb{Z}$, and a function f on $\mathbb{R}$ is the function $b * f$ on $\mathbb{R}$, given by

$$(b * f)(x) = \sum_{k=-\infty}^{k=+\infty} b(k)f(x-k) \qquad x \in \mathbb{R} \tag{5}$$

The discrete convolution between two sequences a and b is the sequence $a * b$

$$(a * b)(l) = \sum_{k=-\infty}^{k=+\infty} a(k)b(l-k) \qquad l \in \mathbb{Z} \tag{6}$$

A filter $\hat{\lambda}(f)$ is the Fourier transform of a function λ that defines a bounded convolution operator on L_2:

$$\lambda : g \in L_2 \rightarrow \lambda * g \in L_2 \tag{7}$$

Since the convolution product $\lambda * g$ becomes a multiplication product $\hat{\lambda}\hat{g}$ in Fourier space, the filter $\hat{\lambda}$ selectively alters the frequency components of $\hat{g}$.

Whenever it exits, the convolution inverse $(b)^{-1}$ of a sequence b is defined to be

$$((b)^{-1} * b)(l) = \delta_0(l) \tag{8}$$

where δ_0 is the unit impulse; i.e., $\delta_0(0) = 1$ and $\delta_1(l) = 0$ for $l \neq 0$.

The reflection of a function f (resp., a sequence b) is the function $f^{\vee}$ (resp., the sequence $b^{\vee}$), is given by

$$f^{\vee}(x) = f(-x) \qquad \forall x \in \mathbb{R} \tag{9}$$

$$b^{\vee}(l) = b(-l) \qquad \forall l \in \mathbb{Z} \tag{10}$$

3. GENERALIZED SAMPLING THEORY

If we consider the subspace of L_2 generated by the span of all the shifted sinc functions

$$V = \left\{ v : v(x) = (c * \text{sinc})(x) = \sum_{k=-\infty}^{k=+\infty} c(k)\text{sinc}(x-k) \qquad c \in l_2 \right\} \tag{11}$$

then Shannon's sampling procedure [6, 23], which is depicted in Figure 1, is equivalent to computing the L_2 least square approximation in V. Taking this point of view, the whole scheme applied to a function s can be expressed as

$$s_a(x) = \sum_{k=-\infty}^{k=+\infty} (\text{sinc} * s)(k)\text{sinc}(x-k) \tag{12}$$

where s_a is the orthogonal projection of s in V. The function $(\text{sinc} * s)(x)$ is an ideal lowpass filtered- or bandlimited-signal. Its Fourier transform has a compact support which lies in the interval $[-\frac{1}{2}, \frac{1}{2}]$. By the Paley-Wiener Theorem, it must be an entire function of exponential type. Thus, the operation of taking the samples $c(k) = (\text{sinc} * s)(k)$ of $(\text{sinc} * s)(x)$ makes sense. These samples are then used as coefficients in the series (12) to obtain the approximation s_a. A sampling $c_h(k)$ of a translation $s(x+h)$ of s has the property that $\|c_h - c\|_{l_2}$ goes to zero as $|h| \rightarrow 0$. Because of this property, we say that the sampling scheme is *strongly jitter-stable*.

3.1. Generalized Sampling in L_2

We use the previous interpretation to generalize Shannon's sampling procedure. We start with an "appropriate" real valued function $\lambda \in L_2$, and define the space $V(\lambda)$ by

$$V(\lambda) = \left\{ v : v(x) = (c * \lambda)(x) = \sum_{k=-\infty}^{k=+\infty} c(k)\lambda(x-k) \qquad c \in l_2 \right\} \tag{13}$$

We view sampling as a problem of least square approximation in the space $V(\lambda)$. The whole procedure is then reduced to finding the coefficients $c(k)$ of the orthogonal projection in $V(\lambda)$ of a function $g \in L_2$. The reconstruction is then obtained by $g_a = c * \lambda$. To do this, we must choose the function λ appropriately. First, we must insure that $V(\lambda)$ is well defined and is a subspace of L_2. Second, the subspace $V(\lambda)$ must be closed. Finally, the set $\{\lambda(x-k)\}_{k\in\mathbb{Z}}$ must form a Riesz-Schauder basis of $V(\lambda)$.

Since we seek a class of sampling procedures that extends the classical scheme (see Figure 1), we require that the orthogonal projection on $V(\lambda)$ be implementable using the following three-step procedure:

(i) a shift-invariant prefilter
(ii) a strongly jitter-stable sampling
(iii) a shift-invariant postfilter

where filter is defined in Section 2. The sequence obtained from the sampling is the discrete representation that we seek.

If we start from a function $\lambda \in L_2$ satisfying $\|c * \lambda\|_{L_2}^2 \le M\|c\|_{l_2}^2$ for all $c \in l_2$, then $V(\lambda)$ is a well-defined subspace of L_2. Obviously, this does not guarantee that $V(\lambda)$ is closed. However, if we further assume that $m\|c\|_{l_2}^2 \le \|c * \lambda\|_{L_2}^2$ for some number $m > 0$, then we insure that $V(\lambda)$ is closed and that it is generated by the Riesz-Schauder basis $\{\lambda(x-k)\}_{k\in\mathbb{Z}}$. To see this, we note that if the sequence $c_n * \lambda$ of elements in $V(\lambda)$ is convergent in L_2, then the inequality $\|c\|_{l_2}^2 \le m^{-1}\|c * \lambda\|_{L_2}^2$ implies that c_n is a Cauchy sequence in l_2. Thus, it converges to an element $c_\infty \in l_2$. The inequality $\|c * \lambda\|_{L_2}^2 \le M\|c\|_{l_2}^2$ then implies that $c_n * \lambda$ converges to $(c_\infty * \lambda) \in V(\lambda)$ as n tends to infinity. Finally, if $c * \lambda = 0$, then $\|c\|_{l_2}^2 \le m^{-1}\|c * \lambda\|_{L_2}^2$ implies that $c = 0$. Hence, $\{\lambda(x-k)\}_{k\in\mathbb{Z}}$ is a basis of $V(\lambda)$. Thus we have:

Proposition 1: If for some $M \ge m > 0$

$$m\|c\|_{l_2}^2 \le \|c * \lambda\|_{L_2}^2 \le M\|c\|_{l_2}^2 \qquad \forall c \in l_2 \tag{14}$$

then $V(\lambda)$ is a closed subspace of L_2. Moreover, $\{\lambda(x-k)\}_{k\in\mathbb{Z}}$ is a Riesz-Schauder basis of $V(\lambda)$.

Remark 3.1: By definition, a set $\{e_i\}_{i\in\mathbb{Z}} \subset \mathcal{H}$ of a Hilbert space $\mathcal{H}$ is a Riesz-Schauder basis if there exist two constants $M \ge m > 0$ such that $m\|c\|_{l_2}^2 \le \left\|\sum_i c(i)e_i\right\|_{\mathcal{H}}^2 \le M\|c\|_{l_2}^2$. Thus, Condition (14) is a necessary and sufficient condition for $V(\lambda)$ to be a closed subspace of L_2 with the Riesz-Schauder basis $\{\lambda(x-k)\}_{k\in\mathbb{Z}}$.

If we take the Fourier transform of an element $(c * \lambda) \in V(\lambda)$, and use Plancherel's Theorem and the fact that $C(f+i) = C(f)$, $\forall i \in \mathbb{Z}$, we obtain

$$\|c * \lambda\|_{L_2}^2 = \int_{\mathbb{R}} |C(f)L(f)|^2 \, df = \int_0^1 |C(f)|^2 \sum_i |L(f+i)|^2 \, df \tag{15}$$

where $C(f)$ and $L(f)$ are the Fourier transforms of c and λ, respectively. Clearly, if $m \le \Lambda(f) = \sum_i |L(f-i)|^2 \le M$, then it follows from (15) that (14) is satisfied. The converse is also true; see, for example, [16]. These facts, together with a way of obtaining the orthogonal projection on $V(\lambda)$, are stated in the following theorem:

Theorem 2: Let λ be a function in L_2 and let $M \ge m > 0$ be two positive constants. Then the following two conditions are equivalent:

(i) $m\|c\|_{l_2}^2 \le \|c * \lambda\|_{L_2}^2 \le M\|c\|_{l_2}^2$

(ii) $m \le \Lambda(f) = \sum_i |L(f-i)|^2 \le M$ a.e.

Moreover, the orthogonal projection g_a of a function $g \in L_2$ on $V(\lambda)$ is given by

$$g_a(x) = \sum_{k \in \mathbb{Z}} \langle g(\zeta), \mathring{\lambda}(\zeta - k)\rangle_{L_2} \lambda(x-k) \tag{16}$$

where the dual function $\mathring{\lambda}$, which belongs to $V(\lambda)$, is given by

$$\mathring{\lambda}(x) = ((a)^{-1} * \lambda)(x) \qquad \forall x \in \mathbb{R} \tag{17}$$

and where $(a)^{-1}$ (which is the inverse Fourier transform of $1/\Lambda(f)$) is the convolution inverse of the autocorrelation function $a(k) = (\lambda * \lambda^{\vee})(k)$, $\forall k \in \mathbb{Z}$.

We will call a function *generating* if it satisfies (i) (equivalently (ii)) of Theorem 2.

Proof: We now assume (i) and prove (ii) with an argument that uses contradiction. First, if the periodic function $\Lambda(f) := \sum_i |L(f-i)|^2$ is not bounded above by M, then the measure meas(E_M) of the set $E_M := \{f \in (0,1): \Lambda(f) > M\}$ is strictly positive. We consider the periodic function $C_M(f+1) = C_M(f)$ defined by $C_M(f) := (\text{meas}(E_M))^{-1/2}\chi_{E_M}$ for all $f \in (0,1)$, where χ_{E_M} is the characteristic (or indicator) function of the set E_M. From its construction, $C_M(f)$ is the Fourier transform of a sequence $c_M \in l_2$ with $\|c_M\|_{l_2} = 1$. Using Identity (15), and the hypothesis that $\|c * \lambda\|_{L_2}^2 \le M\|c\|_{l_2}^2$, we obtain $M < \int_{E_M} |C_M(f)|^2 \Lambda(f)\, df \le M$, which is a contradiction. Thus, $\Lambda(f) \le M$ almost everywhere. Similarly, if we assume that $\Lambda(f)$ is not bounded below by the positive constant m, then we can construct the periodic function $C_m(f) := (\text{meas}(E_m))^{-1/2}\chi_{E_M}$ for all $f \in (0,1)$, where the measure meas(E_m) of the set $E_m := \{f \in (0,1): \Lambda(f) < m\}$ is strictly positive. We immediately obtain that $m \le \int_{E_m} |C_m(f)|^2 \Lambda(f)\, df < m$, which is a contradiction.

Since $V(\lambda)$ is closed, the least square approximation g_a of g exists and is equal to the orthogonal projection of g in $V(\lambda)$. Let $g_a = \sum_i c(i)\lambda(x-i)$. To find the coefficient $c(i)$, we use the fact that $g - g_a$ is orthogonal to $V(\lambda)$:

$$\langle (g - g_a)(x), \lambda(x-l)\rangle_{L_2} = 0 \qquad \forall l \in \mathbb{Z} \tag{18}$$

By simple substitution, we get

$$(c * a)(l) = \langle g(x), \lambda(x - l)\rangle_{L_2} \qquad \forall l \in \mathbb{Z} \tag{19}$$

where the sequence a is the sampled autocorrelation function, given by

$$a(l) = (\lambda * \lambda^\vee)(l) \qquad \forall l \in \mathbb{Z} \tag{20}$$

To see that the sampling $(\lambda * \lambda^\vee)(l)$, $\forall l \in \mathbb{Z}$, makes sense, we note that because $\lambda \in L_2$, the Fourier transform $|L(f)|^2$ of $\lambda * \lambda^\vee$ belongs to L_1. Thus, $(\lambda * \lambda^\vee)(x)$ is a continuous function. Clearly, by Poisson's formula, the sequence a is the inverse Fourier transform of $\Lambda(f) = \sum_i |L(f - i)|^2 \neq 0$. The coersivity property $\Lambda(f) \geq m > 0$ in (ii) implies that $1/\Lambda(f)$ is bounded above a.e. by m^{-1}. Thus, a has an l_2 convolution inverse $(a)^{-1}$ with Fourier transform $1/\Lambda(f)$. Hence, we can solve Equation (19) for c to get

$$c(l) = \langle g(x), \mathring{\lambda}(x - l)\rangle_{L_2} \tag{21}$$

where $\mathring{\lambda}$ is given by Equation (17).

Theorem 2 may be viewed as a generalization of Shannon's sampling procedure. To make this link explicit, we note that the expansion coefficients in (16) can be obtained by sampling a prefiltered signal at the integers:

$$c(k) = \langle g(x), \mathring{\lambda}(x - k)\rangle = (\mathring{\lambda}^\vee * g)(x)|_{x=k} \tag{22}$$

The impulse response of the optimal prefilter therefore corresponds to the reflection $\mathring{\lambda}^\vee$ of the function $\mathring{\lambda}$ (see Equation (9)). This operator has a role that is analogous to the antialiasing lowpass filter required in conventional sampling theory. The least squares approximation of g (or reconstruction) in $V(\lambda)$ is then obtained from the mixed convolution between the coefficient sequence $c(k)$ and the function λ:

$$g_a(x) = (c * \lambda)(x) \tag{23}$$

a step that can be interpreted as a postfiltering. The full approximation procedure can therefore be described by a block diagram that is essentially equivalent to the one shown in Figure 1. The main difference is that the ideal lowpass filters must be replaced by the appropriate pre- and postfilters: $(\mathring{\lambda}^\vee)^\wedge$ and $\hat{\lambda}$, respectively. Moreover, the sampling defined by (22) is strongly jitter-stable. To see this, let $c_g * \lambda$ denote the orthogonal projection $P_V g$ of g on $V(\lambda)$. We immediately see that

$$m\|c_g\|_{l_2}^2 \leq \|c_g * \lambda\|_{L_2}^2 = \|P_V g\|_{L_2}^2 \leq \|g\|_{L_2}^2 \tag{24}$$

This, together with the fact that $m > 0$ and Theorem 2, allows us to state:

Theorem 3: If $\lambda \in L_2$ is a generating function, then the orthogonal projection in $V(\lambda)$ is a well defined sampling procedure that consists of a prefiltering followed by a strongly jitter-stable sampling and, finally, a postfiltering for the reconstruction.

To obtain an interpretation of the present sampling theory within the general framework of Reproducing Kernel Hilbert Spaces [19, 26], we consider a function $h \in V(\lambda)$ that we reproject into $V(\lambda)$ according to (16). Formally, we have that

$$h(x) = \sum_{k\in\mathbb{Z}} \langle h(\zeta), \mathring{\lambda}(\zeta - k)\rangle_{L_2}\lambda(x - k)$$
$$= \left\langle h(\zeta), \sum_{k\in\mathbb{Z}} \mathring{\lambda}(\zeta - k)\lambda(x - k)\right\rangle_{L_2} \tag{25}$$

Thus, if $\lambda \in L_2$ is smooth and has appropriate decay, then $V(\lambda)$ is a Reproducing Kernel Hilbert Space with reproducing kernel

$$K(x, \zeta) = \sum_{k\in\mathbb{Z}} \mathring{\lambda}(\zeta - k)\lambda(x - k) \tag{26}$$

In particular, we note that the orthogonal projection of a function $g \in L_2$ can be represented in the more compact form

$$g_a(x) = \langle g(\zeta), K(x, \zeta)\rangle_{L_2} \tag{27}$$

There are two problems with the necessary and sufficient Conditions of Theorem 2. First, it is not easy in general to verify that Condition (ii) is satisfied for an arbitrary function λ, except for some special cases like λ = sinc. The difficulty comes from the fact that $\Lambda(f)$ is an infinite sum. Second, the condition is not preserved under convolution. In fact, the convolution $\lambda_1 * \lambda_2$ of two functions $\lambda_1 \in L_2$ and $\lambda_2 \in L_2$ is not necessarily in L_2.

Our goal for the rest of this section will be to find conditions on λ that are more directly verifiable and yet sufficiently general for most applications. Another motivation is to find a set of sufficient conditions that is preserved under convolution. To do this, we start by grouping the zeros of $L(f)$ into sets A_i, $i \in \mathbb{Z}$ as follows:

$$A_i := \{f \in [0, 1) : L(f - i) = 0\} \tag{28}$$

We now have all the elements to state:

Theorem 4: If the Fourier transform $L(f)$ of a real function $\lambda \in L_1$ is such that $L(f) = O(|f|^{-r})$ for some $r > \frac{1}{2}$, and if the set $\bigcap_i A_i = \emptyset$, then the periodic functions $\sum_i |L(f - i)|^2$ is continuous. Moreover, λ is a generating function.

The proof of this theorem is postponed to the end of this section. Clearly, the n-fold convolution of a generating function satisfying the conditions of Theorem 4 is also a generating function satisfying the same conditions. The growth condition on $L(f)$ is essentially a smoothness condition on λ. In fact, it implies that λ belongs to the Sobolev space $W^{s,2}$ for any s for which $s < r - \frac{1}{2}$ (see Section 2). However, $\lambda \in W^{s,2}$ for all $s < r - \frac{1}{2}$ does not imply that $L(f) = O(|f|^{-r})$. A slightly stronger condition must be imposed. For instance, if the derivative of λ is $C^1(\mathbb{R}) \cap L_1$, then $L(f)$ is $O(|f|^{-1})$. In fact, if $\lambda \in W^{s,1}$ with $s \geq 2$, then the Sobolev Embedding Theorems imply that $\lambda \in L_2$ and that $\lambda \in C^1(\mathbb{R})$. Thus, in this case, $\lambda \in L_1 \cap L_2$, and $L(f) \in L_2$ with the decay $L(f) = O(|f|^{-1})$. As a corollary of Theorem 4, we immediately obtain:

Corollary 5: If the Fourier transform $L(f)$ of a real function $\lambda \in L_1$ is such that $L(f) \neq 0$ for all $f \in [-\frac{1}{2}, \frac{1}{2}]$ and $L(f) = O(|f|^{-r})$ with $r > \frac{1}{2}$, then λ is a generating function.

Clearly, if λ_1 and λ_2 satisfy the assumptions of Corollary 5, then so does the convolution product $\lambda_1 * \lambda_2$.

3.1.1. Proof of Theorem

Proof: By the Riemann-Lebesgue Lemma, $L(f)$ is continuous. This and the fact that $L(f) = O(|f|^{-r})$ for some $r > \frac{1}{2}$ imply that the series

$$\sum_i |L(f - i)|^2 \leq \text{const} \sum_{i=1}^{\infty} |i|^{-2r} \tag{29}$$

is absolutely convergent, independent of f. Thus, it is continuous on $[0, 1]$. From its expression, this series can be seen to be periodic with period 1.

Since $\bigcap_i A_i = \emptyset$, we get that $\Lambda(f) = \sum_i |L(f - i)|^2 \neq 0$, $\forall f \in [0, 1]$. Therefore, because the series Λ is continuous on $[0, 1]$, it follows that it is uniformly bounded below by a positive constant $m > 0$. This fact, combined with (29), yields the desired result.

3.2. Sampling in Sobolev Spaces

The sampling procedures considered so far provide the best representation of a signal in the least squares sense (minimum L_2-norm). However, one may conceive of applications in which a good rendition of the derivatives of signals is also of interest. For this purpose, we will now consider approximation-sampling schemes using the norm associated with the Sobolev spaces $W^{s,2}$.

Specifically, if the Fourier transform of $\lambda \in L_1$ is such that $L(f) = O(|f|^{-r})$ for some $r > n + \frac{1}{2}$, $n \geq 0$, and $n \in \mathbb{Z}$, then an estimate similar to Inequality (29) can be used to show that $c * \lambda$ is in $W^{n,2}$. Moreover, if the $\bigcap_i A_i = \emptyset$ (See Definition (28)), then we have the following estimate:

$$m\|c\|_2^2 \leq \|c * \lambda\|_{L_2}^2 \leq \|c * \lambda\|_{W^{n,2}}^2 \tag{30}$$

where the constant m is the same as in Theorem 2. In this case, the best approximation in $(V(\lambda), \|\cdot\|_{W^{n,2}})$ of a function $g \in W^{n,2}$ is obtained by a series of parallel prefilters followed by a pure jitter-stable sampling and, finally, a postfiltering with the function λ. The samples $c(l)$ are given by

$$c(l) = (\mathring{\lambda}_1^\vee * g + \mathring{\lambda}_2^\vee * (D^{(1)}g) + \cdots + \mathring{\lambda}_{n+1}^\vee * (D^{(n)}g))(l) \qquad \forall l \in \mathbb{Z} \tag{31}$$

where $D^{(j)}g$ is the jth distributional derivative of g and where

$$\mathring{\lambda}_i = (a_1 + a_2 + \cdots + a_{n+1})^{-1} * (D^{(i-1)}\lambda) \tag{32}$$

$$a_i(l) = ((D^{(i-1)}\lambda) * (D^{(i-1)}\lambda)^\vee)(l) \qquad \forall l \in \mathbb{Z} \tag{33}$$

4. EQUIVALENT GENERATING FUNCTIONS AND SAMPLING STRATEGIES

Two nonidentical functions λ_1 and λ_2 may generate the same subspace V. Depending on the signal processing application, a particular basis for V may have advantages over other choices. If phase distortion is not acceptable, then a generating function with an axis of symmetry is desirable. If the sequence of the expansion coefficients is required to be a faithful representation of the underlying continuous signal, then an interpolating function is the most appropriate. In some applications, the generating function is chosen so that it is localized in time or frequency. In others, the only requirement is to have an orthonormal basis. Given a function λ that generates the subspace $V(\lambda)$, we can use the mixed convolution to construct another function $\lambda_{=}$ generating the same space $V(\lambda_{=}) = V(\lambda)$. This is done by convolving λ with a sequence p (mixed convolution as in Definition (5)):

$$\lambda_{=}(x) = (p * \lambda)(x) \tag{34}$$

where p is an invertible convolution operator from l_2 into itself. A necessary and sufficient condition for p to define such an operator is provided by the following proposition:

Proposition 6: Let p be a sequence and $P(f)$ its Fourier transform. The sequence p defines a bounded invertible convolution operator onto l_2 if and only if ess sup$|P(f)| < \infty$ and ess inf$|P(f)| > 0$.

Proof: We first assume that ess sup$|P(f)| < \infty$, and ess inf$|P(f)| > 0$. Since ess sup$|P(f)| < \infty$, $P(f) \in L_2[0, 1]$. Hence, the sequence p is in l_2. Similarly, since ess inf$|P(f)| > 0$, we also have that $1/P(f)$ is in $L_2[0, 1]$. Therefore it is the Fourier transform of $(p)^{-1} \in l_2$, the convolution inverse of p (see Definition (8)). If $c \in l_2$, then using Plancherel's Theorem, we have

$$\|P^{-1}\|_\infty^2 \|c\|_{l_2}^2 \leq \|p * c\|_{l_2}^2 = \int_0^1 |P(f)C(f)|^2 \, df \leq \|P\|_\infty^2 \|c\|_{l_2}^2 \tag{35}$$

It follows that p defines a bounded convolution operator from l_2 into itself, and that its inverse $(p)^{-1}$ is also a bounded convolution operator from l_2 into itself. This proves the first part of the proposition.

Next, assume that p defines a bounded invertible convolution operator onto l_2. Thus, by assumption, there exist two positive constants $0 < m \leq M$ such that $m\|c\|_{l_2}^2 \leq \|p * c\|_{l_2}^2 \leq M\|c\|_{l_2}^2$. By using (35) and the same argument as in the proof of Theorem 2, in which we replace $\Lambda(f)$ by $P(f)$, we then obtain the desired result.

Remark 4.1: An example in which the conditions of the proposition are satisfied occurs when $P(f)$ is continuous and $P(f) \neq 0$ in $[0, 1]$.

The set of generating functions can be partitioned into equivalence classes by the relation that associates two functions whenever they generate the same space. Moreover, two generating functions belonging to the same equivalence class are always related by Eq. (34). By appropriately choosing the sequence p, we can construct generating functions with some desired properties. Next, we will show how to construct biorthogonal, interpolating, and orthogonal basis functions. In all cases, the orthogonal projection in V does not change. However, the prefilter and postfilter in the sampling procedure depend on the particular choice of basis.

4.1. Dual Generating Function

The inner product between λ and the generating function $\lambda_{=}$ defined by (34) is simply

$$\langle \lambda(x), \lambda_{=}(x - k)\rangle_{L_2} = (p^{\vee} * a)(k) \tag{36}$$

where the symmetrical autocorrelation sequence a is defined in Theorem 2. It follows that the corresponding sets of basis functions are biorthogonal if and only if $p = p^\vee = (a)^{-1}$. For this particular choice of p, λ is precisely the dual generating function $\mathring{\lambda}$ defined by (17) in Theorem 2. Since λ and $\mathring{\lambda}$ generate the same space $V(\lambda)$, we conclude that their role in Equation (16) can be simply interchanged.

4.2. Orthogonal Generating Function

We will say that a generating function ϕ is orthogonal if $\phi(x - k)$, $k \in \mathbb{Z}$ constitutes an orthonormal basis of $V(\phi)$:

$$\langle \phi(x), \phi(x - k) \rangle_{L_2} = (\phi * \phi^\vee)(k) = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \ldots \end{cases} \tag{37}$$

In this case, the function ϕ and its dual $\mathring{\phi}$ are identical. An example of such an orthogonal generating function is $\mathrm{sinc}(x)$.

As is well-known [16, 18], to obtain an orthogonal function from an arbitrary λ, we can choose p in (34) to be the square-root convolution inverse of the sampled autocorrelation function: $p_o = (a)^{-1/2}$. This is the inverse Fourier transform of

$$P_o(f) = \left(\sum_i |L(f - i)|^2 \right)^{-1/2} \tag{38}$$

Clearly, because of Theorem 2, $P_o(f)$ satisfies the conditions of Proposition 6. Of course, there are many other orthogonal generating functions associated with $V(\lambda)$.

4.3. Interpolating Generating Function

Another generating function of interest is the interpolating function η, which is continuous and vanishes at all the integers except the origin where it takes the value 1:

$$\eta(k) = \delta_0(k) = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \ldots \end{cases} \tag{39}$$

Because of this property, it can be used to interpolate between samples on $\mathbb{Z}$. An example of such a function is $\mathrm{sinc}(x)$. An interpolating function can be obtained from λ as described in the following proposition:

Proposition 7: If λ is a generating function satisfying the conditions of Theorem 4 with $r > 1$, and if $S(f) = \left| \sum_i L(f - i) \right| \neq 0$, $\forall f \in \mathbb{R}$, then we can choose a sequence p_I so that $\eta = p_I * \lambda$ is an interpolating generating function. The sequence p_I is given by the convolution inverse of $\lambda(k)$, $k \in \mathbb{Z}$; it can be specified by its Fourier transform,

$$P_I(f) = \left(\sum_i L(f - i) \right)^{-1} \tag{40}$$

Proof: Since $r > 1$, the series $\left| \sum_i L(f - i) \right|$ converges uniformly and is therefore continuous. Moreover, since the series is never zero, $P_l(f)$ defined by (40) exists, and it is continuous. Thus, η is a generating function. Its Fourier transform $H(f)$ is continuous, and $|H(f)| = O(|f|^{-r})$ with $r > 1$. This implies that η is continuous and the sequence $\eta(l), l \in \mathbb{Z}$ is well defined. If we use Poisson's formula to get the Fourier transform $E(f) = \sum_i H(f - i)$ of the discrete sequence $\eta(l)$, $l \in \mathbb{Z}$, we immediately get that $E(f) = 1$. Hence, η is interpolating.

5. CONVOLUTION SEQUENCES OF GENERATING FUNCTIONS

The n-fold convolution of a generating function that satisfies the conditions of Theorem 4, is also a generating function. Thus, we can start with a single function and construct, by repeated convolution, an infinite sequence of generating functions. In fact, the well known B-splines can be obtained precisely in this fashion. These functions turn out to be very useful for generating certain polynomial spline function spaces [21, 22]. In this sense, the constructions that follow can be viewed as a generalization of polynomial splines.

5.1. The Basic Sequence

We start with a function λ that satisfies the conditions in Theorem 4 and construct, by convolution, a sequence of increasingly regular functions λ^n, as follows:

$$\lambda^n(x) = \lambda * \lambda * \lambda * \cdots * \lambda \qquad (n - 1 \text{ convolution}) \tag{41}$$

These functions will be called the basic generating functions of order n. The fact that λ^n is also a generating function satisfying the conditions in Theorem 4 can be deduced by taking the Fourier transform of (41). In fact, if the Fourier transform $L(f)$ of λ is $O(|f|^{-r})$, then the Fourier transform $(L(f))^n$ of λ^n is $O(|f|^{-nr})$. It follows that the regularity of λ^n improves with increasing n. Moreover, if λ is "linear phase," then λ^n will also be "linear phase" (i.e., having a vertical axis of symmetry). This property is relevant in signal processing for obtaining representations that have no phase distortions.

Another result that motivates such a construction is a general convergence theorem by de Boor, Höllig, and Riemenschneider [8], which we give here in one dimension using our notation.

These authors start with a compactly supported function $\mu \in L_2$ with Fourier transform $|M(f)| = O(|f|^{-1})$, and introduce the set

$$\Omega = \{f: |M(f + j)| < |M(f)|, j \in \mathbb{Z}\backslash 0\} \tag{42}$$

Next, they define the class of functions

$$V_\infty = \left\{ g \in L_2: \lim_{n \to +\infty} \operatorname{dist}(g, V(\mu^n)) = 0 \right\} \tag{43}$$

They then prove the following theorem:

Theorem 8: $g \in V_\infty$ if and only if the support of the Fourier transform of g is contained in $\bar{\Omega}$.

In the special case of polynomial splines, stronger results can be found in [17]. In particular, Marsden, Richards, and Riemenschneider show that the spline interpolant $\Sigma\, g(j)\eta^{2n}(x - j)$ of a bandlimited function $g \in B_{1/2} \cap L_p$ converges to g in the L_p-norms as $n \to \infty$.

Theorem 8 applies directly to our construction, provided that the function λ in (41) is compactly supported. A case of special interest is $\bar{\Omega} = [-\frac{1}{2}, \frac{1}{2}]$, in which V_∞ is precisely the class of bandlimited functions considered in Shannon's sampling theorem. This observation suggests a close connection between the sampling in the subspaces $V(\lambda^n)$ and the classical sampling procedure. Next, we make this connection explicit and extend the results above to noncompactly supported functions, and to the case of L_p spaces (see Remark 5.2 below). Specifically, we use our results in Sections 3 and 4 to obtain approximations of signals in $V(\lambda^n)$. We will then prove that, under mild conditions, the filters in the interpolating and orthogonal representations tend to the ideal lowpass filter. However, filters associated with other basis functions (e.g., λ^n) do not necessarily converge to the ideal filter.

5.2. The Interpolating and Orthogonal Sequences

Our previous results provide us with a number of possibilities for characterizing the sequence of spaces $V(\lambda^n)$. A natural approach is to use an orthogonal generating function. We can use (34) and (38) to obtain a generating function ϕ^n that satisfies the orthogonality property (37). It is obtained from λ^n as follows:

$$\phi^n(x) = ((a^n)^{-1/2} * \lambda^n)(x) \tag{44}$$

$$a^n(k) = (\lambda^n * \lambda^{n\vee})(k) \qquad \forall k \in \mathbb{Z} \tag{45}$$

Remark 5.1: It should be noted that the sampled autocorrelation function a^n is not equal to the $(n - 1)$-fold convolution of $a(k) = (\lambda * \lambda^\vee)(k)$.

Similarly, we can use the interpolating representation in which the expansion coefficients are precisely the sampled values $s(k)$, $k \in \mathbb{Z}$, of the underlying signal $s(x)$. This implies that the expansion coefficients can themselves be viewed as a faithful representation of the signal $s(x)$. For this purpose, if λ is as in Proposition 7, we can define

$$b^n(k) = \lambda^n(k) \qquad \forall k \in \mathbb{Z} \tag{46}$$

From Proposition 7, we know that an interpolating generating function exists provided that the sequence b^n is invertible. It is given by

$$\eta^n(x) = ((b^n)^{-1} * \lambda^n)(x) \tag{47}$$

A case of special interest occurs when λ is symmetrical. In this case, $a^n(k) = b^{2n}(k)$, and there is a direct relation between the orthogonal and interpolating functions: $\eta^{2n} = \phi^n * \phi^n$. This relation also implies the existence of the interpolating representation for n even. Note that symmetrical functions are easily constructed; starting from any generating function τ that satisfies the conditions of theorem 4, it is always possible to obtain a symmetrical one: $\lambda = \tau * \tau^\vee$.

5.3. Asymptotic Convergence Results

As n increases, the sequence of basic functions λ^n converges to a Gaussian. This result is a consequence of the well-known Central Limit Theorem [9]. If, instead, we consider the interpolating representation, then the underlying filters tend to the ideal lowpass filter of Shannon, as expressed by the following theorem:

Theorem 9: If $L(f)$ is the Fourier transform of a symmetrical generating function λ satisfying the conditions of Theorem 4, and if $L(f)$ satisfies

$$\min_{f \in I} |L(f)| > |L(f)| \qquad \forall f \notin I = \left[-\frac{1}{2}, \frac{1}{2}\right] \tag{48}$$

then the Fourier transforms $H^{2n}(f)$ of the interpolating functions η^{2n} generated by (47) and the Fourier transforms $\mathring{H}^{2n}(f)$ of their duals $\mathring{\eta}^{2n}$ converge pointwise a.e. and in L_p-norms, $p \in [1, \infty)$, to the ideal lowpass filter as n tends to infinity:

$$L_p - \lim_{n \to \infty} H^{2n}(f) = \text{rect}(f) \tag{49}$$

$$L_p - \lim_{n \to \infty} \mathring{H}^{2n}(f) = \text{rect}(f) \tag{50}$$

Corollary 10: The interpolating functions η^{2n} and their duals $\mathring{\eta}^{2n}$ converge, in L_q-norms, $q \in [2, \infty]$, to the ideal sinc interpolator of Shannon as n tends to infinity.

The proof of Theorem 9 is postponed until the end of this section. The assumption (Theorem 4) that λ belongs to L_1 implies that $L(f)$ is a continuous function. Thus, Condition (48) implies that the minimum of $|L(f)| = |L(-f)|$ in $I = [-\frac{1}{2}, \frac{1}{2}]$ is achieved at $f = \pm\frac{1}{2}$. Clearly, this minimum can be achieved elsewhere in I, and $|L(f)|$ can have other local minima in $[-\frac{1}{2}, \frac{1}{2}]$. An equivalent statement is: $|L(f_1)| > |L(f_2)|$ for all $f_1 \in I$ and for all $f_2 \notin I$. Condition (48) essentially means that $L(f)$ is a nonideal lowpass filter in the frequency band $(-\frac{1}{2}, \frac{1}{2})$. The theorem can be viewed as stating that the ideal lowpass filter can be approximated as closely as necessary by the sequences η^{2n} and η^{2n}. These are obtained by repeated convolutions and a simple correction of a single nonideal lowpass filter. If the basic generating function is not symmetrical, we can define λ^{2n} to be $(\lambda * \lambda^{\vee})^n$ and obtain the same convergence results as in Theorem 9.

It should be noted that Condition (48) can be relaxed. In fact, we only need to have

$$|L(f)| > |L(f - i)| \qquad i = \pm 1, \pm 2, \ldots, \forall f \notin \left(-\frac{1}{2}, \frac{1}{2}\right) \tag{51}$$

Similar to the interpolating family, the Fourier transforms of the sequence ϕ^n also tend to the ideal lowpass filter as n goes to infinity. This fact is stated in the following theorem:

Theorem 11: If the Fourier transform $L(f)$ of a generating function λ satisfying the conditions of Theorem 4 is such that

$$\min_{f \in I} |L(f)| > |L(f)| \qquad \forall f \notin I = \left[-\frac{1}{2}, \frac{1}{2}\right] \tag{52}$$

then the modulus $|F^n(f)|$ of the Fourier transform of the orthogonal functions and its dual $\mathring{\phi}^n = (\phi^n)^{\vee}$ converge pointwise a.e. and in L_p-norms, $p \in [1, \infty)$, to the ideal lowpass filter as the order n tends to infinity:

$$L_p - \lim_{n \to \infty} |F^n(f)| = \text{rect}(f) \tag{53}$$

Remark 5.2: The remarkable result in Theorem 8 of de Boor, Höllig, and Riemenschneider contains the essence of the asymptotic equivalence of our generalized sampling procedures with the classical scheme. Our theorems which describe the convergence of specific basis functions make this connnection explicit. They also provide extensions for L_p-norms and noncompactly supported functions. By setting $p = 2$, and choosing a compactly supported function λ, we can use (49) to obtain Theorem 8. The converse is not true. Theorem 8 implies that the L_2-limit of a convergent sequence is a bandlimited function. However, this result is not sufficient to infer that the various sequences of basis functions in Theorems 9 and 11 tend to be the ideal filter. Moreover, a sequence of functions $s_n \in V(\mu^n)$ is not necessarily convergent.

Equation (49) in Theorem 9 can be used to obtain a generalization of the convergence result for polynomial splines in [17]. Specifically, we can show that the interpolants $s_n \in V(\lambda^n)$ of the samples $\{s(k)\}_{k \in \mathbb{Z}}$ of an L_p bandlimited function $s \in B_{1/2}$ (i.e., s is a function of exponential type in L_p) tend to s and n goes to infinity, in L_p-norms for all $p \in [2, \infty]$.

5.4. Proof of Theorems 9 and 11

5.4.1. Proof of Theorem 9

Proof: The Fourier transfrom $H^{2n}(f)$ is given by

$$H^{2n}(f) = \frac{|L(f)|^{2n}}{\sum_i |L(f-i)|^{2n}} \tag{54}$$

It is well-defined, since the denominator $\sum_i |L(f-i)|^{2n}$ is strictly positive. This follows from the fact that $L(f)$ is continuous and $L(f) \neq 0$, for all $f \in I = [-\frac{1}{2}, \frac{1}{2}]$. For $f \in I$, we estimate $|H^{2n}(f) - 1|$ by

$$\begin{aligned} |H^{2n}(f) - 1| &= \left| \frac{|L(f)|^{2n}}{|L(f)|^{2n} + \sum_{i \neq 0} |L(f-i)|^{2n}} - 1 \right| \\ &\leq \sum_{i \neq 0} |L(f-i)/L(f)|^{2n} \end{aligned} \tag{55}$$

We rewrite the series on the right side of the inequality in (55) to get

$$\sum_{i \neq 0} |L(f-i)/L(f)|^{2n} = \max_{i \neq 0} |L(f-i)/L(f)|^{2n} \sum_{i \neq 0} \frac{|L(f-i)/L(f)|^{2n}}{\max_{i \neq 0} |L(f-i)/L(f)|^{2n}} \tag{56}$$

The fact that the term of the series in the right side of (56) is $O(|i|^{-2r})$ with $r > \frac{1}{2}$ implies that the series converges. Moreover, since the terms in the series are no bigger than 1, we can choose $n = 1$ in the series and obtain the upper bound:

$$\sum_{i \neq 0} |L(f-i)/L(f)|^{2n} \leq \max_{i \neq 0} |L(f-i)/L(f)|^{2n-2} \sum_{i \neq 0} |L(f-i)/L(f)|^2 \tag{57}$$

The series on the right side of the Inequality (57) can be bounded above by a constant independent of $f \in I$. We have

$$\sum_{i \neq 0} |L(f - i)/L(f)|^2 \leq \left(\min_{f \in I} |L(f)|^2\right)^{-1} \sum_{i \neq 0} |L(f - i)|^2 \leq C \tag{58}$$

where C is a constant independent of f. Thus, by (55), (57), and (58), we obtain the upper bound:

$$|H^{2n}(f) - 1| \leq C \max_{i \neq 0} |L(f - i)/L(f)|^{2n-2} \tag{59}$$

Since $|L(f - i)/L(f)|$ is $O(|i|^{-r})$ and because of Condition (48) in the theorem, we have that for $f \in (-\frac{1}{2}, \frac{1}{2})$, $\max_{i \neq 0} |L(f - i)/L(f)| < 1$. Thus, we conclude that $|H^{2n}(f) - 1|$ converges pointwise to 0. Moreover, $|H^{2n}(f) - 1|$ is bounded above by the constant C, which is independent of n. Hence, by the Lebesgue Dominated Convergence Theorem, we have that $|H^{2n}(f) - 1|$ converges to zero in $L_p(-\frac{1}{2}, \frac{1}{2})$.

Because $|H^{2n}(f)| = |H^{2n}(-f)|$, it only remains to look at the case for which $f \in (\frac{1}{2}, \infty)$. For $L(f) \neq 0$, we rewrite (54) to obtain

$$|H^{2n}(f)| = \frac{|L(f)|^{2n}}{\sum_i |L(f - i)|^{2n}} = \frac{1}{1 + \sum_{i \neq 0} |L(f - i)/L(f)|^{2n}} \tag{60}$$

Let $f \in (k - \frac{1}{2}, k + \frac{1}{2}]$, $k = 1, 2, \ldots$. Since Condition (48) implies that $L(f - k)$ is the largest term, we can estimate $|H^{2n}(f)|$ by

$$|H^{2n}(f)| \leq |L(f)/L(f - k)|^{2n} < 1 \tag{61}$$

Together, the first inequality of (61) and Condition (48) imply that $N^{2n}(f)$ tends to zero as n tends to infinity. For large values of f, we use the first inequality of (61) to obtain

$$\begin{aligned} |H^{2n}(f)| &\leq |L(f)/L(f - k)|^{2n} \\ &\leq |L(f)/L(f - k)|^2 \\ &\leq \left(\min_{f \in I} |L(f)|^2\right)^{-1} |L(f)|^2 \\ &\leq C|f|^{-2r} \end{aligned} \tag{62}$$

where C is a constant independent of f and n. Because of the last inequality in (62), and because $|H^{2n}(f)|$ is bounded above by 1 (Eq. (61)), we can use Lebesgue's Dominated Convergence Theorem to conclude that $H^{2n}(f)$ converges to zero in $L_p(0, \infty)$. The proof of part (i) of the theorem is thus completed.

The expression for the dual filter $\mathring{H}^{2n}(f)$ is given by

$$\mathring{H}^{2n}(f) = \frac{|L(f)|^{2n} \sum_i |L(f - i)|^{2n}}{\sum_i |L(f - i)|^{4n}} = H^{4n}(f) + E^{2n}(f) \tag{63}$$

where $E^{2n}(f)$ is defined by

$$E^{2n}(f) = \frac{|L(f)|^{2n} \sum_{i \neq 0} |L(f - i)|^{2n}}{\sum_i |L(f - i)|^{4n}} \tag{64}$$

The term $H^{4n}(f)$ converges to rect(f) by part (i) of the theorem. Thus, we only need to show that $E^{2n}(f)$ converges pointwise and in L_p to zero as n tends to infinity.

For $f \in (-\frac{1}{2}, \frac{1}{2})$, we have that

$$|E^{2n}(f)| = \frac{\sum_{i \neq 0} |L(f-i)/L(f)|^{2n}}{1 + \sum_{i \neq 0} |L(f-i)/L(f)|^{4n}} \leq \sum_{i \neq 0} |L(f-i)/L(f)|^{2n} \tag{65}$$

This is the same expression as in the last inequality of (55). Thus, $E^{2n}(f)$ converges pointwise and in $L_p(-\frac{1}{2}, \frac{1}{2})$ to 0 as n tends to infinity.

Again because of symmetry, we only need to look at $f \in (0, \infty)$. For $f \in (k - \frac{1}{2}, k + \frac{1}{2}]$, we multiply and divide by the term $L(f-k)$ to rewrite (64) as

$$E^{2n}(f) = \frac{|L(f)|^{2n}|L(f-k)|^{2n} \sum_{i \neq 0} |L(f-i)/L(f-k)|^{2n}}{\sum_{i} |L(f-i)|^{4n}} \tag{66}$$

If we let $z = f - k$, then the series in the numerator can be estimated above as follows:

$$\begin{aligned} \sum_{i \neq 0} |L(f-i)/L(f-k)|^{2n} &\leq \sum_{j} |L(z-j)/L(z)|^{2n} \\ &\leq 1 + \left(\min_{f \in I} |L(z)|^2\right)^{-1} \sum_{i \neq 0} |L(z-i)|^2 \leq C \end{aligned} \tag{67}$$

where C in the last inequality is a constant independent of f and n. Thus, if $L(f) \neq 0$ (otherwise $E^{2n}(f) = 0$), the expression of $E^{2n}(f)$ in Equation (66) can be estimated by

$$\begin{aligned} E^{2n}(f) &\leq C \frac{|L(f)|^{2n}|L(f-k)|^{2n}}{\sum_{i} |L(f-i)|^{4n}} \\ &\leq C \frac{|L(f)|^{2n}}{|L(f-k)|^{2n}} \leq C \end{aligned} \tag{68}$$

Together, Condition (48) and the second inequality of (68) imply the pointwise convergence of $E^{2n}(f)$ to zero. Finally, since $|L(f)/L(f-k)| < 1$, we estimate $E^{2n}(f)$ for large values of f as follows:

$$\begin{aligned} E^{2n}(f) &\leq C \frac{|L(f)|^2}{|L(f-k)|^2} \\ &\leq \left(\min_{f \in I} |L(f)|^2\right)^{-1} |L(f)|^2 \leq C_1 |f|^{-2r} \end{aligned} \tag{69}$$

where C_1 is a constant independent of n and f. Because of this last inequality, and the fact that $|E^{2n}(f)|$ is bounded above by a constant C (Eq. (68)), we can use Lebesgue's Dominated Convergence Theorem to conclude that $E^{2n}(f)$ converges to zero in $L_p(0, \infty)$. The proof of part (ii) of the theorem is thus completed.

5.4.2. Proof of Theorem 11

Proof: The Fourier transform $F^n(f)$ is given by

$$F^n(f) = \frac{(L(f))^n}{\left(\sum_i |L(f-i)|^{2n}\right)^{1/2}} \tag{70}$$

For $f \in (0, \frac{1}{2})$, we get the upper bound estimate,

$$\begin{aligned} ||F^n(f)| - 1| &\leq \left| \frac{1 - \left(1 + \sum_{i\neq 0} |L(f-i)/L(f)|^{2n}\right)^{1/2}}{\left(1 + \sum_{i\neq 0} |L(f-i)/L(f)|^{2n}\right)^{1/2}} \right| \\ &\leq \left| 1 - \left(1 + \sum_{i\neq 0} |L(f-i)/L(f)|^{2n}\right)^{1/2} \right| \\ &\leq \sum_{i\neq 0} |L(f-i)/L(f)|^{2n} \end{aligned} \tag{71}$$

The last inequality is the same as in (55). Thus, the pointwise and L_p convergence follows for $f \in (-\frac{1}{2}, \frac{1}{2})$.

For $f \in (\frac{1}{2}, \infty)$, we have that

$$|F^n(f)| = \left| 1 + \left(\sum_{i\neq 0} |L(f-i)/L(f)|^{2n}\right)\right|^{-1/2} \leq |L(f-k)/L(f)|^{-n} \tag{72}$$

where $f \in (k - \frac{1}{2}, k + \frac{1}{2}]$. Note that if we replace n by $2n$ in the last inequality, we obtain inequality (61) in the proof of Theorem 9. Thus, by following exactly the same arguments, we complete the proof.

6. EXAMPLES

6.1. Polynomial Splines

A particular example of the present theory is provided by polynomial splines. The sampling theory for this class of functions has been investigated by several authors [12, 24].

The B-splines of order m, β^m, are obtained by repeated convolution of a B-spline of order 0 [21]:

$$\beta^m(x) = (\beta^0 * \beta^0 * \cdots * \beta^0)(x) \qquad (m \text{ convolution}) \tag{73}$$

where $\beta^0(x)$ is the characteristic function in the interval $[0, 1)$. Schoenberg has shown that these functions generate the polynomial splines of order m with knot points at the integer. These splines are C^{m-1} functions that are formed by patching together polynomials of degree m at the grid points, which are called the knot points. Accordingly, when the grid is $\mathbb{Z}$, any polynomial spline function $s(x)$ of order m can be represented as

$$s(x) = \sum_{k=-\infty}^{k=+\infty} c(k)\beta^{m}(x - k) \qquad (74)$$

The function β^0 satisfies the conditions of Theorems 9 and 11 if we select $n = m + 1$. Thus, we can conclude that the interpolating and orthogonal spline filters $(\eta^{2n})^\wedge$, $(\mathring{\eta}^{2n})^\wedge$, $(\phi^{2n})^\wedge$, and $(\mathring{\phi}^{2n})^\wedge$ converge to the ideal filter as n tends to infinity. More detailed convergence results can be found in [4, 24]. Related asymptotic properties of spline interpolants are also discussed in [17, 22].

6.2. Gaussian Functions

Another case is provided by the example of Gaussian functions. These functions are optimally localized in the time-frequency plane [10]. Because of this property, they have been widely used in physics and engineering for signal representation.

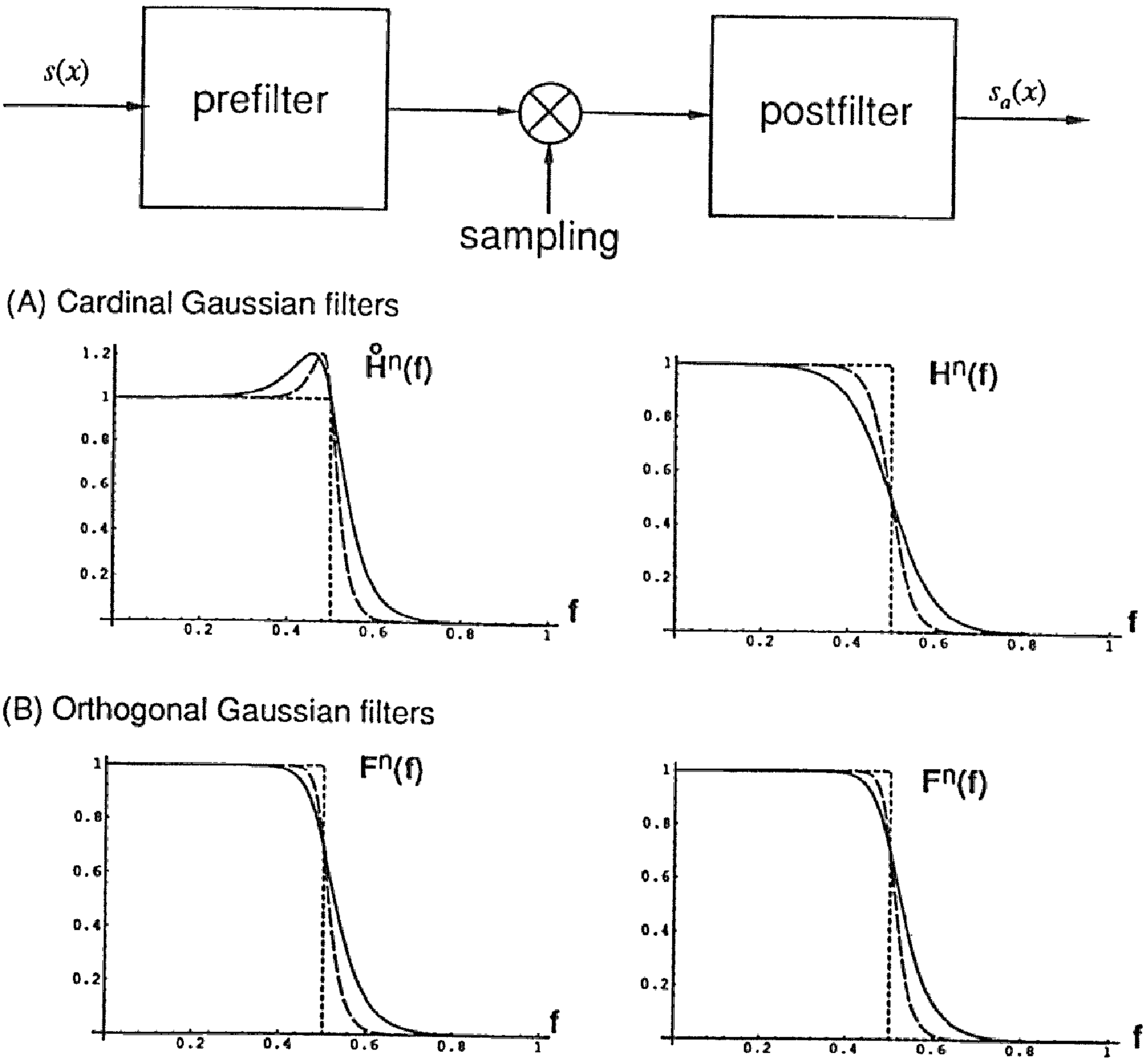

Figure 2 Block diagram of the generalized sampling procedures with the example of the Gaussian functions $\lambda^{2n} = (\pi n)^{-1/2}\exp(-x^2/n)$. (A) Interpolating Gaussian filters, and (B) orthogonal Gaussian filters: continuous lines for $n = 1$, dashed lines for $n = 2$, and dotted lines for the ideal filter.

Any Gaussian function $\lambda(x) = (2\pi\sigma)^{-1/2} \exp(-2^{-1}(x/\sigma)^2)$ satisfies the conditions of Theorems 9 and 11. Thus, we conclude that the interpolating and orthogonal Gaussian filters $(\eta_G^{2n})^\wedge$, $(\mathring{\eta}_G^{2n})^\wedge$, $(\phi_G^{2n})^\wedge$, and $(\mathring{\phi}_G^{2n})^\wedge$ converge to the ideal filter as n tends to infinity. The process of Gaussian sampling with $\lambda^{2n}(x) = (\pi n)^{-1/2} \exp(-x^2/n)$ and the convergence of the corresponding filters are illustrated in Figure 2. The approximation procedure involves a prefiltering followed by a sampling and, finally, a reconstruction using a postfilter. As can be seen from the graphs, the interpolating and orthogonal Gaussian filters of order 4 are already a good approximation to the ideal filter.

6.3. Scaling Functions

Finally, we can obtain similar constructions starting from any scaling function that satisfies the QMF conditions of S. Mallat [16]. These functions generate multiresolution analysis of L_2, and are widely used in the context of wavelet transforms. An interesting property is the fact that the multiresolution structure is preserved under convolution [2].

ACKNOWLEDGMENT

We thank an anonymous referee for his helpful comments. They allowed us to improve the manuscript and strengthen the first part of this paper.

REFERENCES

1. R. Adams, *Sobolev spaces*, Academic Press, New York, 1975.
2. A. Aldroubi and M. Unser, Families of multiresolution and wavelet spaces with optimal properties, *Numer. Funct. Anal. Optimiz.*, *14* (1993), 417–446.
3. A. Aldroubi and M. Unser, *Families of wavelet transforms in connection with Shannon's sampling theory and the Gabor transform*, in Wavelets—A tutorial in theory and applications, C.K. Chui, ed., Academic Press, San Diego, 1992, 509–528.
4. A. Aldroubi, M. Unser and M. Eden, Cardinal spline filters: stability and convergence to the ideal sinc interpolator, *Signal Processing 28* (1992), 127–138.
5. J.J. Benedetto, *Irregular sampling and frames*, in Wavelets—A tutorial in theory and applications, C.K. Chui, ed., Academic Press, Orlando, 1992, 445–507.
6. W.M. Brown, Optimal prefiltering of sampled data, *IRE Trans. Inform. Theory 7* (1961), 269–270.
7. P.L. Butzer, A survey of the Whittaker-Shannon sampling theorem and some of its extensions, *J. Math. Res. Exposition 3* (1983), 185–212.
8. C. de Boor, K. Höllig, and S. Riemenschneider, Convergence of cardinal series, *Proc. Amer. Math. Soc. 98* (1986), 457–460.
9. W. Feller, *An introduction to probability theory and its application. Vol. II.* Wiley, New York, 1971.
10. D. Gabor, Theory of communication, *J. IEE 93* (1945), 429–441.
11. J.R. Higgins, Five short stories about the cardinal series, *Bull. Amer. Math. Soc. 121*, (1985), 45–89.
12. R. Hummel, Sampling for spline reconstruction, *SIAM J. Appl. Math. 43* (1983), 278–288.
13. A. Jerri, Sampling expansion for Laguerre transforms, *J. Res. Nat. Bur. Standards Sec. B (80), (B) 3* (1976), 415–418.

14. A.J. Jerri, The Shannon sampling theorem—its various extensions and applications: A tutorial review, *Proc. IEEE 65* (1977), 1565–1596.
15. H.P. Kramer, A generalized sampling theorem, *J. Math. Phys. 38* (1959), 68–72.
16. S.G. Mallat, Multiresolution approximations and wavelet orthogonal bases of $L_2(\mathbb{R})$, *Trans. Amer. Math. Soc. 315* (1989), 69–87.
17. M.J. Marsden, F.B. Richards and S.D. Riemenschneider, Cardinal sline interpolation operators on l_p data, *Indiana Univ. Math. J. 24* (1975), 677–689.
18. Y. Meyer, *Ondelettes*, Hermann, Paris, Editeurs des Sciences et des Arts, 1990.
19. M.Z. Nashed and G.G. Walter, General sampling theorems for functions, in Reproducing kernel Hilbert spaces, *Math. of Control Signals Sys. 4* (1991), 373–412.
20. H. Ogawa, A generalized sampling theorem, *Electronics and Communications in Japan 72* (1989), 97–105.
21. I.J. Schoenberg, Contribution to the problem of approximation of equidistant data by analytic functions, *Quart. Appl. Math. 4* (1946), 45–99, 112–141.
22. I.J. Schoenberg, Notes on spline functions III: on the convergence of the interpolating cardinal splines as their degree tends to infinity, *Israel J. Math. 16* (1973), 87–92.
23. C.E. Shannon, Communications in the presence of noise, *Proc. IRE 37* (1949), 10–21.
24. M. Unser, A. Aldroubi, and M. Eden, *A sampling theory for polynomial splines*, in Proc. International Symposium on Information Theory and its Applications, Waikiki, Hawaii (1990), 279–282.
25. J.M. Whittaker, Interpolation function theory, in Cambridge tracts in mathematics and mathematical physics, Cambridge University Press, Cambridge, 1935.
26. K. Yao, Application of reproducing kernel Hilbert Spaces-bandlimited signal models, *Information and Control 11* (1967), 429–444.
27. A.I. Zayed, On Kramer's sampling theorem associated with general Sturm-Liouville problems and Lagrange interpolation, *SIAM J. Appl. Math. 51* (1991), 575–604.

Received: January 27, 1992
Revised: November 25, 1992
Accepted: May 3, 1993