Sampling—50 Years After Shannon

MICHAEL UNSER, FELLOW, IEEE

This paper presents an account of the current state of sampling, 50 years after Shannon's formulation of the sampling theorem. The emphasis is on regular sampling, where the grid is uniform. This topic has benefited from a strong research revival during the past few years, thanks in part to the mathematical connections that were made with wavelet theory. To introduce the reader to the modern, Hilbert-space formulation, we reinterpret Shannon's sampling procedure as an orthogonal projection onto the subspace of band-limited functions. We then extend the standard sampling paradigm for a representation of functions in the more general class of "shift-invariant" functions spaces, including splines and wavelets. Practically, this allows for simpler—and possibly more realistic—interpolation models, which can be used in conjunction with a much wider class of (anti-aliasing) prefilters that are not necessarily ideal low-pass. We summarize and discuss the results available for the determination of the approximation error and of the sampling rate when the input of the system is essentially arbitrary; e.g., nonbandlimited. We also review variations of sampling that can be understood from the same unifying perspective. These include wavelets, multiwavelets, Papoulis generalized sampling, finite elements, and frames. Irregular sampling and radial basis functions are briefly mentioned.

Keywords*—*Band-limited functions, Hilbert spaces, interpolation, least squares approximation, projection operators, sampling, sampling theorem, Shannon, splines, wavelets.

I. INTRODUCTION

In 1949, Shannon published the paper "Communication in the Presence of Noise," which set the foundation of information theory [105], [106]. This paper is a masterpiece; both in terms of achievement and conciseness. It is undoubtedly one of the theoretical works that has had the greatest impact on modern electrical engineering [145]. In order to formulate his rate/distortion theory, Shannon needed a general mechanism for converting an analog signal into a sequence of numbers. This led him to state the classical sampling theorem at the very beginning of his paper in the following terms.

Theorem 1 [Shannon]: If a function $f(x)$ contains no frequencies higher than $\omega_{\max}$ (in radians per second), it is completely determined by giving its ordinates at a series of points spaced $T = \pi/\omega_{\max}$ seconds apart.

Manuscript received September 17, 1999; revised January 4, 2000.

The author is with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne CH-1015 Lausanne EPFL, Switzerland (e-mail: Michael.Unser@epfl.ch).

Publisher Item Identifier S 0018-9219(00)02874-7.

While Shannon must get full credit for formalizing this result and for realizing its potential for communication theory and for signal processing, he did not claim it as his own. In fact, just below the theorem, he wrote: "this is a fact which is common knowledge in the communication art." He was also well aware of equivalent forms of the theorem that had appeared in the mathematical literature; in particular, the work of Whittaker [144]. In the Russian literature, this theorem was introduced to communication theory by Kotel'nikov [67], [68].

The reconstruction formula that complements the sampling theorem is

$$f(x) = \sum_{k \in Z} f(kT)\,\mathrm{sinc}(x/T - k) \tag{1}$$

in which the equidistant samples of $f(x)$ may be interpreted as coefficients of some basis functions obtained by appropriate shifting and rescaling of the sinc-function: $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. Formula (1) is exact if $f(x)$ is bandlimited to $\omega_{\max} \leq \pi/T$; this upper limit is the Nyquist frequency, a term that was coined by Shannon in recognition of Nyquist's important contributions in communication theory [88]. In the mathematical literature, (1) is known as the cardinal series expansion; it is often attributed to Whittaker in 1915 [26], [143] but has also been traced back much further [14], [58].

Shannon's sampling theorem and its corresponding reconstruction formula are best understood in the frequency domain, as illustrated in Fig. 1. A short reminder of the key sampling formulas is provided in Appendix A to make the presentation self-contained.

Nowadays the sampling theorem plays a crucial role in signal processing and communications: it tells us how to convert an analog signal into a sequence of numbers, which can then be processed digitally—or coded—on a computer. While Shannon's result is very elegant and has proven to be extremely fruitful, there are several problems associated with it. First, it is an idealization: real world signals or images are never exactly bandlimited [108]. Second, there is no such device as an ideal (anti-aliasing or reconstruction) low-pass filter. Third, Shannon's reconstruction formula is rarely used in practice (especially with images) because of the slow decay of the sinc function [91]. Instead, practitioners typically rely on much simpler techniques such as

0018–9219/00$10.00 © 2000 IEEE

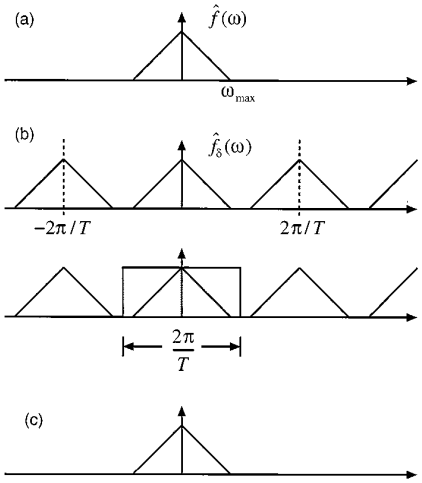

Fig. 1. Frequency interpretation of the sampling theorem: (a) Fourier transform of the analog input signal $f(x)$, (b) the sampling process results in a periodization of the Fourier transform, and (c) the analog signal is reconstructed by ideal low-pass filtering; a perfect recovery is possible provided that $\omega_{\max} \leq \pi/T$.

linear interpolation. Despite these apparent mismatches with the physical world, we will show that a reconciliation is possible and that Shannon's sampling theory, in its modern and extended versions, can perfectly handle such "nonideal" situations.

Ten to 15 years ago, the subject of sampling had reached what seemed to be a very mature state [26], [62]. The research in this area had become very mathematically oriented, with less and less immediate relevance to signal processing and communications. Recently, there has been strong revival of the subject, which was motivated by the intense activity taking place around wavelets (see [7], [35], [80], and [85]). It soon became clear that the mathematics of wavelets were also applicable to sampling, but with more freedom because no multiresolution is required. This led researchers to reexamine some of the foundations of Shannon's theory and develop more general formulations, many of which turn out to be quite practical from the point of view of implementation.

Our goal in this paper is to give an up-to-date account of the recent advances that have occurred in *regular* sampling. Here, the term "regular" refers to the fact that the samples are taken on a uniform grid—the situation most commonly encountered in practice. While the paper is primarily conceived as a tutorial, it contains a fair amount of review material—mostly recent work: This should make it a useful complement to the excellent survey article of Jerri, which gives a comprehensive overview of sampling up to the mid-1970's [62].

The outline of this paper is as follows. In Section II, we will argue that the requirement of a perfect reconstruction is an unnecessarily strong constraint. We will reinterpret the standard sampling system, which includes an anti-aliasing prefilter, as an orthogonal projection operator that computes the minimum error band-limited approximation of a not-necessarily band-limited input signal. This is a crucial observation that changes our perspective: instead of insisting that the reconstruction be exact, we want it to be as close as possible to the original; the global system, however, remains unchanged, except that the input can now be arbitrary. (We can obviously not force it to be bandlimited.)

In Section III, we will show that the concept extends nicely to the whole class of spline-like (or wavelet-like) spaces generated from the integer shifts of a generating function. We will describe several approximation algorithms, all based on the standard three-step paradigm: prefiltering, sampling, and postfiltering—the only difference being that the filters are not necessarily ideal. Mathematically, these algorithms can all be described as projectors. A direct consequence is that they reconstruct all signals included within the reconstruction space perfectly—this is the more abstract formulation of Shannon's theorem.

In Section IV, we will investigate the issue of approximation error, which becomes relevant once we have given up the goal of a perfect reconstruction. We will present recent results in approximation theory, making them accessible to an engineering audience. This will give us the tools to select the appropriate sampling rate and to understand the effect of different approximation or sampling procedures.

Last, in Section V, we will review additional extensions and variations of sampling such as (multi)wavelets, finite elements, derivative and interlaced sampling, and frames. Irregular sampling will also be mentioned, but only briefly, because it is not the main focus of this paper. Our list of sampling topics is not exhaustive—for instance, we have completely left out the sampling of discrete sequences and of stochastic processes—but we believe that the present paper covers a good portion of the current state of research on regular sampling. We apologize in advance to those authors whose work was left out of the discussion.

II. Shannon's Sampling Theorem Revisited

Shannon's sampling theory is applicable whenever the input function is bandlimited. When this is not the case, the standard signal-processing practice is to apply a low-pass filter prior to sampling in order to suppress aliasing. The optimal choice is the ideal filter $H_T(\omega) = \mathrm{rect}(T\omega)$, which suppresses aliasing completely without introducing any distortion in the bandpass region. Its impulse response is $T^{-1}\mathrm{sinc}(x/T)$. The corresponding block diagram is shown in Fig. 2. In this section, we provide a geometrical Hilbert space interpretation of the standard sampling paradigm. For notational simplicity, we will set $T = 1$ and rescale the time dimension accordingly.

In 1941, the English mathematician Hardy, who was referring to the basis functions in Whittaker's cardinal series (1), wrote: "It is odd that, although these functions occur repeatedly in analysis, especially in the theory of interpolation, it does not seem to have been remarked explicitly that they form an orthogonal system" [55]. Orthonormality is a fundamental property of the sinc-function that has been revived recently.

To understand the modern point of view, we have to consider the Hilbert space L_2, which consists of all functions that

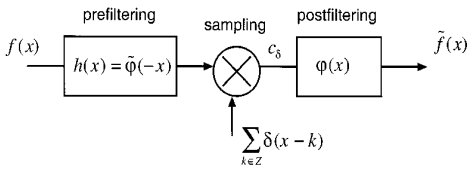

Fig. 2. Schematic representation of the standard three-step sampling paradigm with $T = 1$: 1) the analog input signal is prefiltered with $h(x)$ (anti-aliasing step), 2) the sampling process yields the sampled representation $c_\delta(x) = \sum_{k\in Z} c(k)\delta(x-k)$, and 3) the reconstructed output $\tilde{f}(x) = \sum_{k\in Z} c(k)\varphi(x-k)$ is obtained by analog filtering of c_δ with φ. In the traditional approach, the pre- and postfilters are both ideal low-pass: $h(x) = \varphi(x) = \mathrm{sinc}(x)$. In the more modern schemes, the filters can be selected more freely under the constraint that they remain biorthogonal: $\langle \varphi(x-k),\, \tilde{\varphi}(x-l)\rangle = \delta_{k-l}$.

are squareintegrable in Lebesgue's sense. The corresponding L_2-norm is

$$\|f\| = \left(\int_{-\infty}^{+\infty} |f(x)|^2\, dx\right)^{1/2} = \sqrt{\langle f,\, f\rangle}. \tag{2}$$

It is induced by the conventional L_2-inner product

$$\langle f,\, g\rangle = \int_{-\infty}^{+\infty} f^*(x) g(x)\, dx. \tag{3}$$

We now assume that the input function that we want to sample is in L_2, a space that is considerably larger than the usual subspace of band-limited functions, which we will refer to as V. This means that we will need to make an approximation if we want to represent a non-band-limited signal in the band-limited space V. To make an analogy, L_2 is to V what R (the real numbers) is to Z (the integers).

The countable nature of V is apparent if we rewrite the normalized form of (1) with $T = 1$ as

$$f(x) = \sum_{k\in Z} c(k)\varphi_k \quad \text{with} \quad \varphi_k = \mathrm{sinc}(x-k) \tag{4}$$

where the $c(k)$'s are some coefficients, and where the φ_k's are the basis functions to which Hardy was referring. It is not difficult to show that they are orthonormal

$$\begin{aligned}\langle \varphi_k,\, \varphi_l\rangle &= \int_{-\infty}^{+\infty} \mathrm{sinc}(x-k)\,\mathrm{sinc}(x-l)\, dx \\ &= \mathrm{sinc}(k-l) \\ &= \delta_{k-l}\end{aligned} \tag{5}$$

where the autocorrelation function is evaluated as follows:

$$\begin{aligned}(\mathrm{sinc} * \mathrm{sinc})(x) = \mathrm{sinc}(x) \longleftrightarrow \\ \mathrm{rect}(\omega/2\pi)\cdot \mathrm{rect}(\omega/2\pi) = \mathrm{rect}(\omega/2\pi).\end{aligned}$$

This orthonormality property greatly simplifies the implementation of the approximation process by which a function $f \in L_2$ is projected onto V. Specifically, the orthogonal projection operator P_V: $L_2 \to V$ can be written as

$$P_V f = \sum_{k\in Z} \langle f,\, \varphi_k\rangle \varphi_k \tag{6}$$

where the inner product $c(k) = \langle f,\, \varphi_k\rangle$ represents the signal contribution along the direction specified by φ_k—the approximation problem is decoupled component-wise because of the orthogonality of the basis functions. The projection theorem (see [69]) ensures that this projection operation is well defined and that it yields the minimum-error approximation of f into V.

$$\tilde{f} = P_V f = \arg\min_{g\in V} \|f-g\|^2. \tag{7}$$

By a lucky coincidence, this inner product computation is equivalent to first filtering the input function f with the ideal low-pass filter and sampling thereafter. More generally, we observe that any combination of prefiltering and sampling can be rewritten in terms of inner products

$$\begin{aligned}(h * f)(x)|_{x=k} &= \int f(x) h(k-x)\, dx \\ &= \langle f,\, \tilde{\varphi}_k\rangle \text{ with } \tilde{\varphi}_k = h^T(x-k).\end{aligned} \tag{8}$$

That is, the underlying *analysis* functions correspond to the integer shifts of $h^T(x) = h(-x)$, the time-reversed impulse response of the prefilter h (which can be arbitrary). In the present case, $h(x) = h^T(x) = \mathrm{sinc}(x)$ and $(h * f)(x)$ is the ideal low-pass filtered version of $f(x)$.

The conclusion of this section is that the traditional sampling paradigm with ideal prefiltering yields an approximation $\tilde{f} = P_V f$, which is the orthogonal projection of the input function onto V (the space of band-limited functions). In other words, $\tilde{f}$ is the approximation of f in V with minimum error. In light of this geometrical interpretation, it is obvious that $f \in V \Leftrightarrow f = P_V f$ (since P_V is a projection operator), which is a more concise statement of Shannon's theorem.

III. Sampling in Shift-Invariant (or Spline-Like) Spaces

Having reinterpreted the sampling theorem from the more abstract perspective of Hilbert spaces and of projection operators, we can take the next logical step and generalize the approach to other classes of functions.

A. Extending Shannon's Model

While it would be possible to consider arbitrary basis functions, we want a sampling scheme that is practical and retains the basic, shift-invariant flavor of the classical theory. This is achieved by simply replacing $\mathrm{sinc}(x)$ by a more general template: the *generating function* $\varphi(x)$. Accordingly, we specify our basic approximation space V as

$$V(\varphi) = \left\{ s(x) = \sum_{k\in Z} c(k)\varphi(x-k) \colon c \in l_2 \right\}. \tag{9}$$

This means that any function $s(x) \in V(\varphi)$, which is continuously defined, is characterized by a sequence of coefficients $c(k)$; this is the *discrete* signal representation that will be used to do signal processing calculations or to perform coding. Note that the $c(k)$'s are not necessarily the samples

of the signal, and that $\varphi(x)$ can be something quite different from $\mathrm{sinc}(x)$. Indeed, one of our motivations is to discover functions that are simpler to handle numerically and have a much faster decay.

For such a continuous/discrete model to make sense, we need to put a few mathematical safeguards. First, the sequence of coefficients must be square-summable: $c(k) \in l_2$. Second, the representation should be stable[1] and unambiguously defined. In other words, the family of functions $\{\varphi_k = \varphi(x-k)\}_{k\in Z}$ should form a Riesz basis of $V(\varphi)$, which is the next best thing after an orthogonal one [35]. The definition of a Riesz basis is that there must exist two strictly positive constants $0 < A$ and $B < +\infty$ such that

$$\forall\, c(k) \in l_2,\ A \cdot \|c\|_{l_2}^2 \le \left\| \sum_{k\in Z} c(k)\varphi_k \right\|^2 \le B \cdot \|c\|_{l_2}^2 \quad (10)$$

where $\|c\|_{l_2}^2 = \sum_k |c(k)|^2$ is the squared l_2-norm (or energy) of $c(k)$. A direct consequence of the lower inequality is that $s = \sum_k c(k)\varphi_k = 0$ implies $c = 0$. Thus, the basis functions are linearly independent, which also means that every signal $s(x) \in V(\varphi)$ is uniquely specified by its coefficients $c(k) \in l_2$. The upper bound in (10) implies that the L_2-norm of the signal s is finite, so that $V(\varphi)$ is a valid subspace of L_2. Note that the basis is orthonormal if and only if $A = B = 1$, in which case we have a perfect norm equivalence between the continuous and the discrete domains (Parseval's relation). Because of the translation-invariant structure of the construction, the Riesz basis requirement has an equivalent expression in the Fourier domain [9]

$$A \le \sum_{k\in Z} |\hat{\varphi}(\omega + 2\pi k)|^2 \le B \quad (11)$$

where $\hat{\varphi}(\omega) = \int \varphi(x)\, e^{-j\omega x}\, dx$ is the Fourier transform of $\varphi(x)$. Note that the central term in (11) is the Fourier transform of the sampled autocorrelation function

$$a_\varphi(k) = \langle \varphi(x-k),\, \varphi(x) \rangle. \quad (12)$$

It can therefore also be written $A_\varphi(e^{j\omega}) = \sum_k a_\varphi(k) e^{-j\omega k}$ [see (A.7) in Appendix A].

The final requirement is that the model should have the capability of approximating any input function as closely as desired by selecting a sampling step that is sufficiently small (similar to the Nyquist criterion $T \le \pi/\omega_{\max}$). As shown in Appendix B, this is equivalent to the partition of unity condition

$$\forall\, x \in R,\ \sum_{k\in Z} \varphi(x+k) = 1. \quad (13)$$

In practice, it is this last condition that puts the strongest constraint of the selection on an admissible generating function $\varphi(x)$.

Let us now look at some examples. The first one, which has already been discussed in great length, is the classical choice $\varphi(x) = \mathrm{sinc}(x)$. It is easy to verify that the corresponding Riesz bounds in (11) are $A = B = 1$, which is consistent with the orthonormality property (5). We now show that the sinc function satisfies the partition of unity: using Poisson's summation formula (cf. Appendix A), we derive an equivalent formulation of (13) in the Fourier domain[2]

$$\sum_{k\in Z} \varphi(x+k) = \sum_{n\in Z} \hat{\varphi}(2\pi n)\, e^{j2\pi n x} = 1 \Leftrightarrow \hat{\varphi}(2\pi n) = \delta_n \quad (14)$$

a relation that is obviously satisfied by $\mathrm{rect}(\omega/2\pi)$, the Fourier transform of $\mathrm{sinc}(x)$.

The sinc function is well localized in the frequency domain but has very poor time decay. At the other extreme, we can look for the simplest and shortest function that satisfies (13). It is the box function (or B-spline of degree 0)

$$\beta^0(x) = \begin{cases} 1, & |x| < \frac{1}{2} \\ 0, & \frac{1}{2} < |x|. \end{cases} \quad (15)$$

The corresponding basis functions are clearly orthogonal for they do not overlap.

By convolving this function with itself repeatedly, we construct the B-splines of degree n, which are defined recursively as

$$\beta^n(x) = \beta^0 * \beta^{n-1}(x), \qquad n \ge 1. \quad (16)$$

These functions are known to generate the polynomial splines with equally spaced knots [98], [99]. Specifically, if $\varphi(x) = \beta^n(x)$, then the signals defined by (9) are polynomials of degree n within each interval $(k, k+1)$ (n odd), respectively, $(k - (1/2),\ k + (1/2))$ when n is even, with pieces that are patched together such as to guarantee the continuity of the function and its derivatives up to order $n-1$ (i.e., $s(x) \in C^{n-1}$). The B-spline basis functions up to degree 4 are shown in Fig. 3. They are symmetric and well localized, but not orthogonal—except for $n = 0$. Yet, they all satisfy the Riesz basis condition and the partition of unity. This last property is easily verified in the Fourier domain using (14). The B-splines are frequently used in practice (especially for image processing) because of their short support and excellent approximation properties [126]. The B-spline of degree 1 is the tent function, and the corresponding signal model is piecewise linear. This representation is quite relevant because linear interpolation is one of the most commonly used algorithm for interpolating signal values.

As additional examples of admissible generating functions, we may consider any *scaling* function (to be defined below) that appears in the theory of the wavelet transform [35], [79], [115], [139]. It is important to note, however, that scaling functions, which are often also denoted by $\varphi(x)$, satisfy an additional two-scale relation, which is not required here but not detrimental either. Specifically, a scaling function is valid (in the sense defined by Mallat

[1]By stable, we mean that a small variation of the coefficients must result in a small variation of the function. Here, the upper bound of the Riesz condition (10) ensures L_2-stability.

[2]The equality on the right hand side of (14) holds in the distributional sense provided that $\varphi(x) \in L_1$, or by extension, when $\varphi(x) + \varphi(x+1) \in L_1$, which happens to be the case for $\varphi(x) = \mathrm{sinc}(x)$

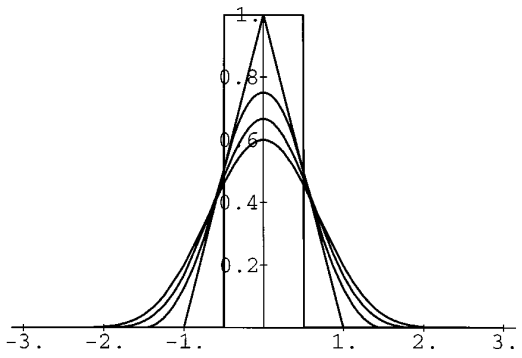

Fig. 3. The centered B-splines for $n = 0$ to 4. The B-splines of degree n are supported in the interval $[-((n+1)/2), ((n+1)/2)]$; as n increases, they flatten out and get more and more Gaussian-like.

[81]) if and only if 1) it is an admissible generating function (Riesz basis condition + partition of unity) and 2) it satisfies the two-scale relation

$$\varphi(x/2) = \sqrt{2} \sum_{k \in Z} h(k)\varphi(x-k) \tag{17}$$

where $h(k)$ is the so-called refinement filter. In other words, the dilated version of φ must live in the space $V(\varphi)$, a property that is much more constraining than the conditions imposed here.

B. Minimum Error Sampling

Having defined our signal space, the next natural question is how to obtain the $c(k)$'s in (9) such that the signal model is a faithful approximation of some input function $f(x) \in L_2$. The optimal solution in the least squares sense is the orthogonal projection, which can be specified as

$$P_{V(\varphi)} f = \sum_{k \in Z} \langle f, \overset{\circ}{\varphi}_k \rangle \varphi_k \tag{18}$$

where the $\overset{\circ}{\varphi}_k$'s are the dual basis functions of $\varphi_k = \varphi(x-k)$. This is very similar to (6), except that the *analysis* and *synthesis* functions ($\overset{\circ}{\varphi}$ and φ, respectively) are not identical—in general, the approximation problem is not decoupled. The dual basis $\{\overset{\circ}{\varphi}_k\}_{k \in Z}$ with $\overset{\circ}{\varphi}_k \in V(\varphi)$ is unique and is determined by the biorthogonality condition

$$\langle \overset{\circ}{\varphi}_k, \varphi_l \rangle = \delta_{k-l}. \tag{19}$$

It also inherits the translation-invariant structure of the basis functions: $\overset{\circ}{\varphi}_k = \overset{\circ}{\varphi}(x-k)$.

Since $\overset{\circ}{\varphi}(x) \in V(\varphi)$, it is a linear conbination of the φ_k's. Thus, we may represent it as

$$\overset{\circ}{\varphi}(x) = \sum_{k \in Z} p(k)\varphi(x-k) \overset{\text{Fourier}}{\longleftrightarrow} \hat{\varphi}(\omega)P(e^{j\omega})$$

where $p(k)$ is a suitable sequence to be determined next. Let us therefore evaluate the inner product

$$\left\langle \overset{\circ}{\varphi}(x), \varphi(x-k) \right\rangle = \sum_{l \in Z} p(l) \langle \varphi(x-l), \varphi(x-k) \rangle = p * a_\varphi(k)$$

where $a_\varphi(k)$ is the autocorrelation sequence in (12). By imposing the biorthogonality constraint $p * a_\varphi(k) = \delta_k$, and by solving this equation in the Fourier domain [i.e., $P(e^{j\omega})A_\varphi(e^{j\omega}) = 1 \Rightarrow P(e^{j\omega}) = 1/A_\varphi(e^{j\omega})$], we find that

$$\hat{\overset{\circ}{\varphi}}(\omega) = \hat{\varphi}(\omega)P(e^{j\omega}) = \frac{\hat{\varphi}(\omega)}{A_\varphi(e^{j\omega})} = \frac{\hat{\varphi}(\omega)}{\sum_{k \in Z} |\hat{\varphi}(\omega + 2\pi k)|^2}. \tag{20}$$

Note that the Riesz basis condition (11) guarantees that this solution is always well defined [i.e., the numerator on the right-hand side of (20) is bounded and nonvanishing].

Similar to what has been said for the band-limited case [see (4)], the algorithm described by (18) has a straightforward signal-processing interpretation (see Fig. 2). The procedure is exactly the same as the one dictated by Shannon's theory (with anti-aliasing filter), except that the filters are not necessarily ideal anymore. Note that the optimal analysis filter is entirely specified by the choice of the generating function (reconstruction filter); its frequency response is given by (20). If φ is orthonormal, then it is its own analysis function (i.e., $\tilde{\varphi} = \overset{\circ}{\varphi} = \varphi$) and the prefilter is simply a flipped version of the reconstruction filter. For example, this implies that the optimal prefilter for a piecewise constant signal approximation is a box function.

C. Consistent Sampling

We have just seen how to design an optimal sampling system. In practice however, the analog prefilter is often specified *a priori* (acquisition device), and not necessarily optimal or ideal. We will assume that the measurements of a function $f \in H$ are obtained by sampling its prefiltered version, $h * f$, which is equivalent to computing the series of inner products

$$c_1(k) = \langle f, \varphi_1(x-k) \rangle \tag{21}$$

with *analysis* function $\varphi_1(x) = h(-x)$ [see (8)]. Here, it is important to specify the input space H such that the measurements are square-summable: $\forall\ f \in H$, $c_1 \in l_2$. In the most usual cases (typically, $\varphi_1 \in L_2$), we will be able to consider $H = L_2$; otherwise (φ_1 is a Delta Dirac or a differential operator), we may simply switch to a slightly more constrained Sobolev space.[3] Now, given the measurements in (21), we want to construct a meaningful approximation of the form (9) with *synthesis* function $\varphi_2 = \varphi$. The solution is to apply a suitable digital correction filter q, as shown in the block diagram in Fig. 4.

Here, we consider a design based on the idea of *consistency* [127]. Specifically, one seeks a signal approximation that is such that it would yield exactly the same measurements if it was reinjected into the system. This is a reasonable requirement, especially when we have no other way of probing the input signal: if it "looks" the same, we may as well say that it is the same for all practical purposes.

[3]The Sobolev space W_2^r specifies the class of functions that are r times differentiable in the L_2-sense. Specifically, $W_2^r = \{f\colon \int (1+\omega^2)^r |\hat{f}(\omega)|^2\, d\omega < +\infty\}$ where $\hat{f}(\omega)$ is the Fourier transform of $f(x)$.

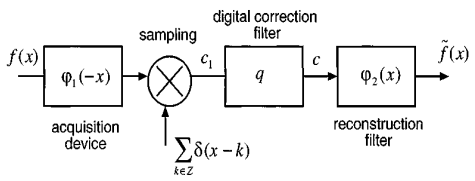

Fig. 4. Sampling for nonideal acquisition devices. The block diagram is essentially the same as in Fig. 2, except for the addition of the digital correction filter q.

Before stating the main theorem, we define the cross-correlation sequence

$$a_{12}(k) = \langle \varphi_1(x-k),\, \varphi_2(x) \rangle \tag{22}$$

where φ_1 is the analysis function and where φ_2 is the generating (or synthesis) function on the reconstruction side.

Theorem 2 [127]: Let $f \in H$ be an unknown input function. Provided there exists $m > 0$ such that $|A_{12}(e^{j\omega})| \geq m$, a.e., then there is a unique signal approximation $\tilde{f}$ in $V(\varphi_2)$ that is consistent with f in the sense that

$$\forall\, f \in H,\, c_1(k) = \langle f,\, \varphi_1(x-k) \rangle = \langle \tilde{f},\, \varphi_1(x-k) \rangle. \tag{23}$$

This signal approximation is given by

$$\tilde{f} = \tilde{P}f(x) = \sum_{k\in Z} (c_1 * q)(k)\varphi_2(x-k) \tag{24}$$

where

$$Q(z) = \frac{1}{\displaystyle\sum_{k\in Z} a_{12}(k) z^{-k}} \tag{25}$$

and the underlying operator $\tilde{P}$ is a projector from L_2 into $V(\varphi_2)$.

If, in addition, the analysis function φ_1 satisfies the Riesz basis condition (11), then $H = L_2$ and the approximation operator has an interesting geometrical interpretation: it is the projection onto $V(\varphi_2)$ perpendicular to $V(\varphi_1)$ [127]. A simplified graphical illustration of the situation is given in Fig. 5.

It is interesting to compare the performance of this *oblique projection* algorithm to the least squares solution, which is typically not realizable here. Specifically, we can derive the following optimal error bound (see [127, Theorem 3]):

$$\|f - Pf\| \leq \|f - \tilde{P}f\| \leq \frac{1}{\cos\,\theta_{12}} \|f - Pf\| \tag{26}$$

where Pf denotes the orthogonal projection of f into $V(\varphi_2)$ and where $\cos\,\theta_{12}$ is an abstract quantity that can be assimilated to the cosine of the maximum angle between the subspaces $V(\varphi_1)$ and $V(\varphi_2)$. This latter quantity is given by (see [127])

$$\cos(\theta_{12}) = \inf_{\omega\in[0,\,\pi]} |\hat{r}_{12}(\omega)|^{1/2} \tag{27}$$

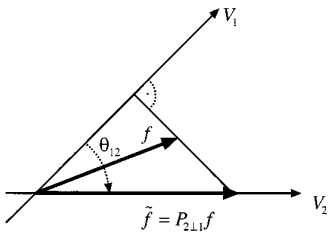

Fig. 5. Principle of an oblique projection onto V_2 perpendicular to V_1 in the simplified case of one-component signal spaces.

where $|\hat{r}_{12}(\omega)|$ is the *spectral coherence* function defined by

$$\hat{r}_{12}(\omega) = \frac{\displaystyle\sum_{k\in Z} \hat{\varphi}_1^*(\omega + 2\pi k)\hat{\varphi}_2(\omega + 2\pi k)}{\sqrt{\displaystyle\sum_{k\in Z} |\hat{\varphi}_1(\omega + 2\pi k)|^2}\,\sqrt{\displaystyle\sum_{k\in Z} |\hat{\varphi}_2(\omega + 2\pi k)|^2}}. \tag{28}$$

Equation (26) indicates that the behavior of both methods is qualitatively the same. The upper bound, which can be quantified, corresponds to the worst possible scenario. The present algorithm provides the least squares solution when $\cos\,\theta_{12} = 1$. In the case where $\varphi_1 \in V(\varphi_2)$, the analysis and synthesis spaces are identical and the oblique projection is equivalent to the orthogonal one. Interestingly, the factor $1/\cos\,\theta_{12}$ also corresponds to the norm of the operator $\tilde{P}$ (see [136, Theorem 4]); it is therefore also a good indicator of the performance of the algorithm in the presence of noise.

Theorem 2 provides a generalization of Shannon's result for nonideal acquisition devices. The projection property implies that the method can reconstruct perfectly any signal that is included in the reconstruction space; i.e., $\forall f \in V(\varphi)$, $f = \tilde{P}f$. This is true not only for band-limited functions but also for functions living in the more general "shift-invariant" subspaces $V(\varphi)$.

The stability of the reconstruction process will obviously depend on the invertibility of the cross-correlation sequence $a_{12}(k)$. Note that, in the special case $\varphi_2 = \mathrm{sinc}$, one recovers the classical inverse filtering solution to the deconvolution problem. Approximate, regularized solutions are also conceivable to improve the robustness in the presence of noise [124].

An attractive feature of the above formulation is the symmetric role assumed by the analysis and synthesis functions φ_1 and φ_2. It is therefore possible to model the distortions that happen at either end of the chain. This allows for the design of a digital filter q that compensates for the nonideal response (φ_2) of some specific ouput device; for example, the sample-and-hold of a digital-to-analog converter or the pixel shape of an image display.

D. Interpolation Revisited

The most basic form of sampling occurs when the signal is specified in terms of its sample values $f(k)$. To make sure that these samples are in l_2 (see [19, Appendix III.A]), we select the input space $H = W_2^1$ (Sobolev space of order one), which simply adds the constraint that the derivative

of f be in L_2 as well. The question is now how to find the coefficients in (9) such that the signal $s(x) \in V(\varphi)$ interpolates the specified samples values exactly. We now show that the solution can be derived directly from Theorem 2. To model the sampling process, we take the analysis function to be $\varphi_1(x) = \delta(x)$ (Dirac delta). In this way, we are able to reformulate the interpolation condition as a consistency requirement: $f(k) = \langle f(x), \delta(x-k) \rangle$.

If we now substitute this choice together with $\varphi_2(x) = \varphi(x)$ in Theorem 2, we get the correction filter

$$Q_{\text{int}}(z) = \frac{1}{\sum_{k \in Z} \varphi(k) z^{-k}} \tag{29}$$

which provides an efficient digital filtering solution to our problem. Note that the filter is the identity if and only if the generating function has the interpolation property: $\varphi(k) = \delta_k$; this is for example the case for the sinc function and the B-splines of degree 0 and 1. For higher order splines a nontrivial filter q is necessary; this kind of inverse filter can be implemented very efficiently using a simple combination of causal and anticausal recursive filters [131], [132].

The interpolating signal $f \in V(\varphi)$ that results from this process is

$$f(x) = \sum_{k \in Z} \underbrace{f(k) * q_{\text{int}}(k)}_{c(k)} \varphi(x-k).$$

This solution can also be presented in a form closer to the traditional one [see (1)]

$$f(x) = \sum_{k \in Z} f(k) \varphi_{\text{int}}(x-k) \tag{30}$$

where φ_{int} is the interpolation function given by

$$\varphi_{\text{int}}(x) = \sum_{k \in Z} q_{\text{int}}(k) \varphi(x-k). \tag{31}$$

Assuming that q_{int} is well defined, the interpolating function within the space $V(\varphi)$ is uniquely defined. $\varphi_{\text{int}}(x)$ has the remarkable property that it is one at the origin and zero at all other integers. The interpolator for the space of cubic splines is shown in Fig. 6. It is rather similar to the sinc function, which is shown by the dotted line; for more details, refer to [10].

A slight modification of the scheme allows for the representation of signals in terms of samples shifted by some fixed amount Δ [60]. Walter has developed a similar representation in the more constrained setting of the wavelet transform [141].

E. Equivalent Basis Functions

So far, we have encountered three types of basis functions: the generic ones (φ), the duals ($\overset{0}{\varphi}$), and the interpolating ones (φ_{int}). In fact, it is possible to construct many others using equivalent generating functions of the form

$$\varphi_{\text{eq}}(x) = \sum_{k \in Z} p(k) \varphi(x-k) \tag{32}$$

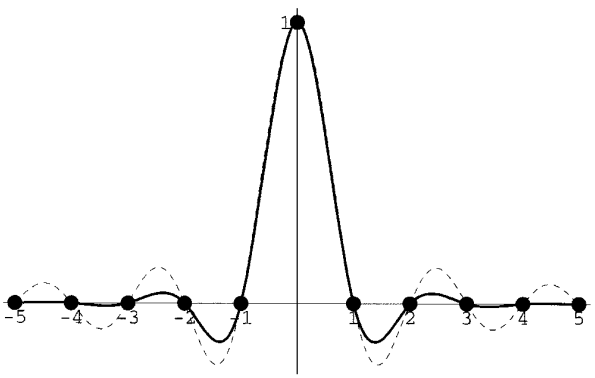

Fig. 6. Comparison of the cubic spline (solid line) and sinc (dashed line) interpolators.

where $p(k)$ is an appropriate sequence of weights. The necessary and sufficient conditions for $\{\varphi_{\text{eq}}(x-k)\}_{k \in Z}$ to yield an equivalent Riesz basis of $V(\varphi)$ are that there exist two strictly positive constants m and M such that $m \leq |P(e^{j\omega})|^2 \leq M$, where $P(e^{j\omega})$ is the Fourier transform of $p(k)$[9]. The most important types of basis functions are summarized in Table 1. Note that the orthogonalized version φ_{ortho} plays a special role in the theory of the wavelet transform [81]; it is commonly represented by the symbol ϕ.

Interestingly, it has been shown that the interpolator φ_{int} and the orthonormal function φ_{ortho} associated with many families of function spaces $V(\varphi)$ (splines, Dubuc–Deslaurier [42], etc.) both converge to Shannon's sinc interpolator as the order (to be defined in Section IV-B) tends to infinity [9], [10], [130]. The sinc function is rather special in the sense that it is both interpolating and orthonormal, which also means that it is its own dual.

A few remarks are in order concerning the usefulness of the various representations; these are applicable not only for single resolution schemes but also for wavelet-like representations, which offer the same kind of freedom [8]. If one is only concerned with the signal values at the integers, then the so-called cardinal representation is the most adequate; here, the knowledge of the underlying basis function is only necessary when one wishes to discretize some continuous signal-processing operator (an example being the derivative) [122]. If, on the other hand, one wishes to perform computation involving values in between samples, one has advantage of using an expansion formula as in (9) where the basis functions have minimal support (e.g., B-splines). This is especially true in higher dimensions where the cost of prefiltering is negligible in comparison to the computation of the expansion formula. More details on the computational issues can be found in [120]. From the point of view of computation, the Shannon model has a serious handicap because there are no band-limited functions that are compactly supported—a consequence of the Paley–Wiener theorem [117]. The use of windowed or truncated sinc functions is not recommended because these fail to satisfy the partition of unity; this has the disturbing consequence that the reconstruction error will

Table 1

Primary Types of Equivalent
Generating Functions with Their Specific Properties

Type	Symbol	$P(e^{j\omega}) = \sum_{k\in Z} p(k)e^{-j\omega k}$	Property
Basic or generic	φ	1	Admissibility $A \le A_0(e^{j\omega}) \le B$
Cardinal or interpolating	φ_{int}	$\frac{1}{\sum_{k\in Z}\hat{\varphi}(\omega+2\pi k)}$	Interpolation $\varphi(k) = \delta_k$
Orthonormal	$\phi = \varphi_{ortho}$	$\frac{1}{\sqrt{\sum_{k\in Z}\lvert\hat{\varphi}(\omega+2\pi k)\rvert^2}}$	Orthonormality $\langle\phi(x),\phi(x-k)\rangle = \delta_k$
Dual	$\overset{\circ}{\varphi} = \varphi_{dual}$	$\frac{1}{\sum_{k\in Z}\lvert\hat{\varphi}(\omega+2\pi k)\rvert^2}$	Biorthonormality $\langle\overset{\circ}{\varphi}(x),\varphi(x-k)\rangle = \delta_k$

not vanish as the sampling step tends to zero (cf., Appendix B).

Another interesting aspect is the time-frequency localization of the basis functions. B-splines are close to optimal because they converge to Gaussians as the order increases; for $n = 3$ (cubic splines), their time-frequency bandwidth product (TFBP) is already within 1% of the limit specified by the uncertainty principle [129]. Furthermore, by using this type of basis function in a wavelet decomposition (see Section V-A), it is possible to trade one type of resolution for the other [31], [129]. This is simply because the TFBP remains a constant irrespective of the scale of the basis function.

F. Sampling and Reproducing Kernel Hilbert Spaces

We now establish the connection between what has been presented so far and Yao and Nashed's general formulation of sampling using reproducing kernel Hilbert spaces (RKHS's) [87], [148].

Definition: A closed vector space V is an RKHS with reproducing kernel $K(x, y)$ if and only if

$$\forall x_0 \in R,\ K(x_0, \cdot) \in V \tag{33}$$

and

$$\forall f \in V,\ \langle K(x, \cdot),\ f(\cdot)\rangle = \int_{y\in R} K(x, y)f(y)\,dy = f(x). \tag{34}$$

The reproducing kernel for a given Hilbert space V is unique; is is also symmetric $K(x, y) = K(y, x)$. For the shift-invariant space $V(\varphi)$ specified by (9), it is given by

$$K(x, y) = \sum_{k\in Z} \varphi(x-k)\overset{\circ}{\varphi}(y-k) \tag{35}$$

where $\overset{\circ}{\varphi}(x)$ is the dual of $\varphi(x)$, as defined in Section III-B.

The concept can be generalized slightly by introducing the notion of a *projection kernel*. We now show how these can be constructed with the help of Theorem 2. We set $\varphi = \varphi_2$ and select a φ_1 such that the invertibility condition in Theorem 2 is satisfied. We can define the equivalent analysis function

$$\tilde{\varphi}(x) = \sum_{k\in Z} q(-k)\varphi_1(x-k) \tag{36}$$

where $q(k)$ is specified as in (25). By construction, this function is biorthogonal to $\varphi = \varphi_2$ (i.e., $\langle\tilde{\varphi}(x-k), \varphi(x-1)\rangle = \delta_{k-l}$). The oblique projection operator (24) may therefore be written in a form similar to (18)

$$\tilde{P}f(x) = \sum_{k\in Z}\langle f(\cdot), \tilde{\varphi}(\cdot-k)\rangle \quad \varphi(x-k).$$

Formally, this is equivalent to

$$\tilde{P}f(x) = \langle f(\cdot), \sum_{k\in Z}\tilde{\varphi}(\cdot-k)\varphi(x-k)\rangle = \langle f(\cdot), \tilde{K}(x, \cdot)\rangle$$

where we have now identified the projection kernel

$$\tilde{K}(x, y) - \sum_{k\in Z}\varphi(x-k)\tilde{\varphi}(y-k). \tag{37}$$

Because we are dealing with a projection operator, we can write the identity

$$\forall f \in V, \langle\tilde{K}(x, \cdot), f(\cdot)\rangle = \int_{y\in R}\tilde{K}(x, y)f(y)dy = f(x) \tag{38}$$

which is the same as (34), except that the projection kernel $\tilde{K}(x, y)$ may be different from $K(x, y)$. Note that it is the inclusion constraint (33) that specifies the reprodcuing kernel in a unique fashion. The requirement $\forall x_0 \in R$, $\tilde{K}(x_0, \cdot) \in V$ is equivalent to the condition $\tilde{\varphi} \in V(\varphi_1)$ [or $V(\varphi - 2)$], which implies that $K(x, y) = \tilde{K}(x, y)$. Thus, the reproducing kernel $K(x, y)$ corresponds to the orthogonal projection.

G. Extension to Higher Dimensions

All results that have been presented carry over directly for the representation of multidimensional signals (or images) $f \in L_2(R^p)$, provided that the sampling is performed on the cartesian grid $\boldsymbol{k} \in Z^p$. This generalization holds because our basic mathematical tool, Poisson's summation formula $\sum_{\boldsymbol{k}\in Z^p} f(\boldsymbol{k}) = \sum_{\boldsymbol{k}\in Z^p}\hat{f}(2\pi\boldsymbol{k})$, remains valid in p dimensions. Thus, we can extend all formulas presented so far by considering the multidimensional time (or space) and frequency variables $\boldsymbol{x} = (x_1, \ldots, x_p) \in R^p$ and $\boldsymbol{w} = (\omega_1, \ldots, \omega_p) \in R^p$, and by replacing simple summations and integrals by multiple ones.

In practice, one often uses separable generating functions of the form

$$\varphi(x_1, \ldots, x_p) = \prod_{i=1}^{p}\varphi(x_i).$$

This greatly simplifies the implementation because all filtering operations are separable. Another advantage of separability is that the one-dimensional Riesz basis condition is equivalent to the multidimensional one. The dual functions remain separable as well.

IV. Controlling the Approximation Error

The projection interpretation of the sampling process that has just been presented has one big advantage: it does not require the band-limited hypothesis and is applicable for any function $f(x)$. Of course, perfect reconstruction is generally not possible when $f(x) \notin V(\varphi)$. It is therefore crucial to

get a good handle on the approximation error. In the classical scheme with ideal anti-aliasing filtering, the error is entirely due to the out-of-band portion of the signal; its magnitude can be estimated simply by integrating the portion of the spectrum above the Nyquist frequency [22], [142]. For more general spline-like spaces $V(\varphi)$, the situation is more complicated. One possibility is to turn to approximation theory and to make use of the general error bounds that have been derived for similar problems, especially in connection with the finite element method [38]–[41], [73], [113]. Specialized error bounds have also been worked out for quasi-interpolation, which is an approximate form of interpolation without any prefilter [40], [74], [109], [110]. Unfortunately, these results are mostly qualitative and not suitable for a precise determination of the approximation error. This has led researchers in signal processing, who wanted a simple way to determine the critical sampling step, to develop an accurate error estimation technique which is entirely Fourier-based [20], [21]. This recent method is easy to apply in practice and yet powerful enough to recover most classical error bounds; it will be described next.

The key parameter for controlling the approximation error is the sampling step T. We therefore consider the rescaled signal space

$$V_T(\varphi) = \left\{ s_T(x) = \sum_{k \in Z} c_T(k) \varphi\left(\frac{x}{T} - k\right) : c_T \in l_2 \right\} \tag{39}$$

where the basis functions are dilated by a factor of T and shifted by the same amount (sampling step). For a given input signal $f(x)$, the interesting question is then to determine the approximation error when the $c_T(k)$'s in (39) are obtained using one of the above algorithms (in Section III-B–D). In this way, we will be able to select a critical sampling step $T_{\max}$ such that the approximation error is below some acceptable threshold. The premise is that the error should decrease and eventually vanish as the sampling step gets smaller; as mentioned before, this is possible only if φ satisfies the partition of unity condition.

A. Calculating the Error in the Frequency Domain

Let Q_T denote a linear approximation operator with sampling step T. The most general mathematical description of a T-shift-invariant[4] approximation operator in $V_T(\varphi)$ is

$$Q_T\{f\} = \sum_{k \in Z} \underbrace{\left(\int f(y) \tilde{\varphi}\left(\frac{y}{T} - k\right) \frac{dy}{T} \right)}_{c_T(k)} \varphi\left(\frac{x}{T} - k\right) \tag{40}$$

where $\tilde{\varphi}$ is a suitable analysis function (or distribution). This corresponds to the general sampling system described in Fig. 2 and includes all the algorithms described so far.

[4]Our definition of T-shift-invariance is that $\forall f \in L_2$, $Q_T\{f(x - kT)\} = Q_T\{f\}(x - kT)$. In other words, Q_T commutes with the shift operator by integer multiples of T.

Our goal is now to determine the dependence of the approximation error $\varepsilon_f(T) = \|f - Q_T\{f\}\|_{L_2}$ on the sampling step T. Since the initial starting point of the signal with respect to the origin is arbitrary, we may as well consider an averaged version of the error over all shifts of the input signal $f(x-\tau)$, where τ is the displacement with respect to the sampling grid. A remarkable fact is that this error measure can be computed exactly by simple integration in the frequency domain (see [19, Theorem 2])

$$\begin{aligned} \overline{\varepsilon}_f^2(T) &:= \frac{1}{T} \int_0^T \|f(\cdot - \tau) - Q_T\{f(\cdot - \tau)\}\|^2 \, d\tau \\ &= \int_{-\infty}^{+\infty} E_{\tilde{\varphi}, \varphi}(T\omega) |\hat{f}(\omega)|^2 \frac{d\omega}{2\pi}. \end{aligned} \tag{41}$$

Here, $\hat{f}(\omega)$ is the Fourier transform of the signal to approximate and $E_{\tilde{\varphi}, \varphi}(\omega)$ is the error kernel given by

$$E_{\tilde{\varphi}, \varphi}(\omega) = |1 - \hat{\tilde{\varphi}}^*(\omega)\hat{\varphi}(\omega)|^2 + |\hat{\tilde{\varphi}}(\omega)|^2 \sum_{k \neq 0} |\hat{\varphi}(\omega + 2\pi k)|^2 \tag{42}$$

where $\hat{\varphi}(\omega)$ and $\hat{\tilde{\varphi}}(\omega)$ are the Fourier transform of $\varphi(x)$ and $\tilde{\varphi}(x)$, respectively. This result allows us to predict the general error behavior of an algorithm by simple examination of the function $E_{\tilde{\varphi}, \varphi}(\omega)$.

In the least squares case (see Section III-B), the error kernel reduces to

$$E_{\varphi}(\omega) = 1 - \frac{|\hat{\varphi}(\omega)|^2}{\sum_{k \in Z} |\hat{\varphi}(\omega + 2k\pi)|^2} = 1 - \frac{|\hat{\varphi}(\omega)|^2}{A_{\varphi}(e^{j\omega})} \tag{43}$$

which is consistent with the fact that the orthogonal projection minimizes the approximation error in $V_T(\varphi)$. For the standard Shannon paradigm, which uses ideal analysis and reconstruction filters, we find that $E_{\text{sinc}}(\omega) = 1 - \text{rect}(\omega/(2\pi))$; this confirms the fact that the approximation error is entirely due to the out-of-band portion of the signal.

Even though $\overline{\varepsilon}_f(T)$ is an average measure of the error, it turns out to be an excellent predictor of the true error $\varepsilon_f(T) = \|f - Q_T\{f\}\|_{L_2}$. This is a direct consequence of the following approximation theorem.

Theorem 3 [19]: The L_2-approximation error of the operator Q_T: $L_2 \to V_T(\varphi)$ defined by can be written as

$$\begin{aligned} \varepsilon_f(T) &:= \|f - Q_T\{f\}\| \\ &= \underbrace{\left[\int_{-\infty}^{+\infty} E_{\tilde{\varphi}, \varphi}(T\omega) |\hat{f}(\omega)|^2 \frac{d\omega}{2\pi} \right]^{1/2}}_{\overline{\varepsilon}_f(T)} + \varepsilon_{\text{corr}} \end{aligned} \tag{44}$$

where $\varepsilon_{\text{corr}}$ is a correction term negligible under most circumstances. Specifically, if $f \in W_2^r$ (Sobolev space of order r) with $r > 1/2$, then $|\varepsilon_{\text{corr}}| < \gamma T^r \|f^{(r)}\|$ where γ is some known constant. Moreover, $\varepsilon_{\text{corr}} = 0$, provided that f is bandlimited to $\omega_{\max} = \pi/T$ (Nyquist frequency).

Thus, the estimate $\overline{\varepsilon}_f(T)$ accounts for the dominant part of the true error $\varepsilon_f(T)$, while $\varepsilon_{\text{corr}}$ is merely a perturbation. This latter correction, which may be positive or negative, is guaranteed to vanish provided that f is bandlimited

or at least sufficiently smooth to have $r > 1/2$ derivatives in the L_2-sense (i.e., $f \in W_2^r$). In the latter case, the error can be made arbitrarily small by selecting a sampling step sufficiently small with respect to the smoothess scale of f as measured by $\|f^{(r)}\|$, the norm of its rth derivative.

Thanks to (41) and Theorem 3, the approximation problem has thus been recast in a framework that is reminiscent of filter design and that should be familiar to a signal-processing audience.

B. Approximation Order

In light of Theorem 3, the minimum requirement for the error to vanish as $T \to 0$ is $E_{\tilde{\varphi},\varphi}(0) = 0$, a condition that implies the partition of unity (see Appendix B). More generally, we can predict the rate of decay of the approximation error from the degree of flatness of $E_{\tilde{\varphi},\varphi}(\omega)$ (or from its Taylor series) near the origin. Specifically, if $E_{\tilde{\varphi},\varphi}(\omega) = C^2 \cdot \omega^{2L} + O(\omega^{2L+2})$ as $\omega \to 0$ (because of symmetry all odd powers of ω are zero), then a simple asymptotic argument on (44) yields

$$\|f - Q_T\{f\}\| = \left[\int_{-\infty}^{+\infty} C^2 T^{2L} \omega^{2L} |\hat{f}(\omega)|^2 \frac{d\omega}{2\pi}\right]^{1/2} = C \cdot T^L \cdot \|f^{(L)}\| \quad \text{as} \quad T \to 0 \qquad (45)$$

where we are assuming that $f \in W_2^L$ so that $\|f^{(L)}\|$ is finite. Under the same hypotheses, one can also derive upper bounds of the error having the same form as the right-hand side of (45) but with a larger leading constant, and which are valid for all values of T [21], [38], [135]. This implies that the error decays globally like $O(T^L)$. This rate of decay is called the *order of approximation*; it plays a crucial role in wavelet and approximation theory [73], [111], [112], [115]. Through the Strang–Fix conditions [113], this order property is equivalent to the reproduction of polynomials of degree $n = L - 1$. The order parameter L determines the approximation power of a given function φ. In wavelet theory, it corresponds to the number of vanishing moments of the analysis wavelet [35], [79], [115]; it also implies that the transfer function of the refinement filter in (17) can be factorized as $H(z) = (1+z)^L \cdot Q(z)$. As an example, the B-splines of degree n [see (16)] have an order of approximation $L = n+1$; they are also the shortest and smoothest scaling functions of order L [125].

Relations (41) and (45) provide alternatives to the Nyquist frequency criterion for selecting the appropriate sampling step T. The error will not be zero in general, but it can be made arbitrarily small without any restriction on $f(x)$.

C. Comparison of Approximation Algorithms

The kernel in (42), or the asymptotic relation (45), can be the basis for the comparison (or the design) of approximation/sampling procedures. It is especially interesting to predict the loss of performance when an approximation algorithm such as (30) and (24) is used instead of the optimal least squares procedure (18). As an example, we concentrate on the case of linear splines. Fig. 7 provides a comparison of the error kernel for three standard algorithms: a) sampling without prefiltering, b) sampling with a suboptimal prefilter (simple box function), and c) least squares sampling. The first case corresponds to the standard piecewise linear interpolation $f_{\text{int}}(x) = \sum_k f(k)\beta^1(x-k)$. The second algorithm uses the simplest possible analog prefilter—the box function, or B-spline of degree 0—combined with the digital filtering correction described by Theorem 2; its geometrical interpretation is a projection onto the space of linear splines perpendicular to the space of splines of degree 0. The third optimal algorithm uses the optimal prefilter specified by (20).

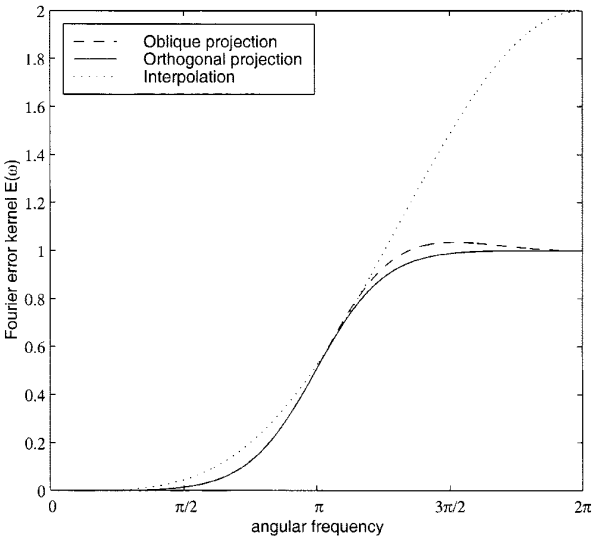

Fig. 7. Error kernels for three linear spline sampling methods: 1) interpolation, 2) oblique projection, and 3) orthogonal projection.

These error kernels all follow a standard pattern: They are close to zero within the Nyquist band and more or less constant outside. As expected, the best algorithm is the third one; in terms of performance, it is the closest to the Shannon paradigm which uses (nonrealizable) ideal filters. The oblique projection is only slightly suboptimal. Sampling without analog prefiltering (interpolation) is by far the least favorable approach. In particular, this method suffers from aliasing; this explains why the signal components above the Nyquist frequency contribute twice to the approximation error: a first time because they cannot be reproduced in $V(\varphi)$, and a second time because of the spectral folding induced by sampling (aliasing).

This comparison clearly emphasizes the importance of prefiltering for the suppression of aliasing. Interestingly, the prefiltering does not need to be optimal—a simple box function, as in algorithm b), may do. In this particular case, the bound constant in (26) is $1/\cos\theta = 1.155$ [136]. This is another indication that this oblique projector is essentially equivalent to the least squares solution; both spline algorithms have been used successfully for image resizing with arbitrary scale factors [72], [133]. The oblique solution is generally simpler and faster to implement.

It is also interesting to look at the asymptotic performance of these algorithms. Their order of approximation is $L = 2$ because they all reproduce linear polynomials [123], [135].

The leading constant in (45) for linear spline interpolation is $C_{\text{int}} = 9.129 \times 10^{-2}$; this is more than a factor of two above the projection algorithms b) and c), which both achieve the smallest possible constant $C_{\text{opt}} = 3.727 \times 10^{-2}$. More generally, it has been shown that the performances of Lth order orthogonal and oblique projectors are asymptotically equivalent, provided that the analysis and synthesis functions are biorthogonal and that $\tilde{\varphi}$ satisfies the partition of unity [123]. Under those hypotheses, the leading constant in (45) is given by

$$C_L = \frac{1}{L!}\left(2\sum_{k=1}^{+\infty}|\hat{\varphi}^{(L)}(2\pi k)|^2\right)^{1/2} \tag{46}$$

where $\hat{\varphi}^{(L)}(\omega)$ denotes the Lth derivative of the Fourier transform of φ. A simpler and more direct formula in terms of the refinement filter h is available for wavelets [20].

D. Comparison of Approximation Spaces

The other interesting issue that we may address using the above approximation results is the comparison of approximation subspaces. Indeed, it is desirable to have some quantitative criteria for selecting a generating function φ. In particular, we would like to identify functions that have good approximation properties and that are reasonably short to allow rapid computation. To factor out the effect of the algorithm, we will base the comparison on the least squares procedure characterized by the kernel (43).

As a first example, we have plotted the error kernels for splines of degrees $n = 0$, to 3 in Fig. 8. This graph clearly shows that, for signals that are predominantly low-pass (i.e., with a frequency content within the Nyquist band), the error tends to be smaller for higher order splines. Of course, the price to pay for better performance is the larger support of the basis functions, which may also induce Gibbs oscillations [63]. As n increases, the spline approximation converges to Shannon's band-limited solution [130]. Since the convergence happens quite rapidly, there is usually not much benefit beyond quintic splines.

Based on the knowledge of these kernels, it is easy, using (41), to predict the behavior of the error as a function of T for some given input function $f(x)$. The corresponding log–log plot for the approximation of the Gaussian test function is given in Fig. 9. We observe that a polynomial spline approximation of degree n provides an asymptotic decay of $(n+1)\cdot$ 20 dB per decade, which is consistent with (45).

The next graph in Fig. 10 provides a performance comparison of four piecewise cubic polynomial kernels of same support $W = 4$: a) Keys' interpolation function [66], which is frequently used for image interpolation, b) a cubic Lagrange-like interpolation [97], c) the cubic B-spline [59], [128], and d) the optimal O-Moms function recently derived in [18]. The last three generating functions are of order $L = 4$ and are members of the so-called Moms (maximum order minimum support) family; the cubic spline is the smoothest member while the O-Moms is the one with the smallest asymptotic constant (46), which explains its better performance. The least favorable case is Keys' interpolating

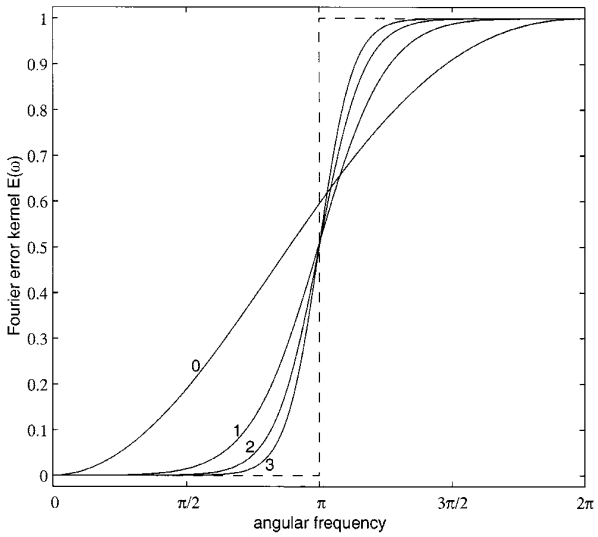

Fig. 8. Frequency plot of the error kernels for the least-squares spline approximations of degree $n = 0, 1, 2, 3$. Below the Nyquist frequency $\omega = \pi$, the kernels (and therefore the errors) tend to be smaller for splines of higher degree. The dotted line corresponds to the Shannon solution with ideal low-pass prefilter.

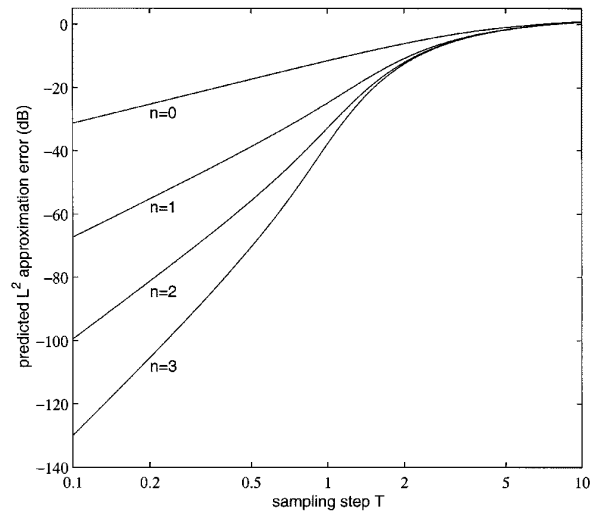

Fig. 9. Approximation error as a function of the sampling step T for the least squares approximation of the function $f(x) = e^{-x^2/2}$ with splines of degree $n = 0$ to 3.

function [with optimal parameter $a = -(1/2)$], which happens to have one less order of approximation. This example demonstrates that enforcing the interpolation constraint [cases a) and b)] is detrimental to overall performance. Comparisons of this nature are quite relevant because all four kernels have the same computational complexity. It thus appears to be more advantageous to use noninterpolating functions such as the B-splines (if derivatives are required), or the O-Moms; this is especially true for medical imaging applications where quality is a key concern [89], [120].

It is only recently that researchers have realized there may be a lot to gain from relaxing the usual interpolation constraint. Keys' short cubic convolution is still considered

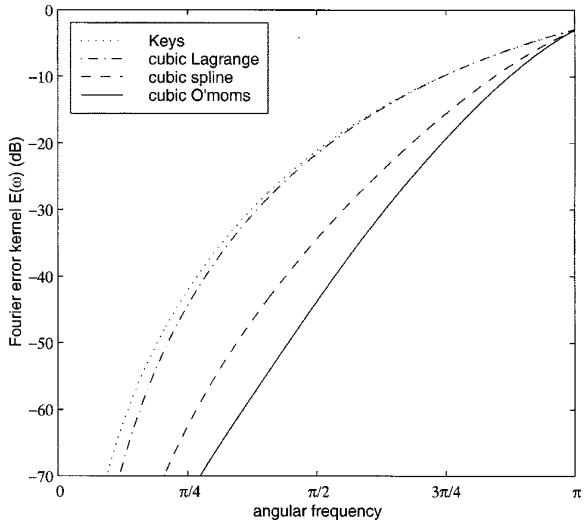

Fig. 10. Frequency plot of the error kernels for four piecewise cubic generating functions of equal support: 1) Keys', 2) Langrange-like interpolator, 3) cubic B-spline, and 4) cubic O-Moms.

the state-of-the art interpolation method in image processing [91], [94], but the situation is likely to change. There is still room for optimization and design work in order to find the basis functions that give the best quality for a given computational budget (or support).

V. Variations and Extensions of Sampling Theory

In this section, we briefly mention some related topics (such as wavelets), which can be thought of as variations and extensions of sampling theory. Our intention here is not to be exhaustive but rather to bring to the attention of the reader some interesting variations of Shannon's theory, while providing pointers to the appropriate literature. We have also tried to give an up-to-date account of the current research in the field, which is greatly influenced by wavelets. What is essential to a subject like sampling is the communication taking place between engineering and mathematical communities. One of the places where this has been happening recently is the International Workshop on Sampling Theory and Applications, held biannually since 1995 [1], [2].

For a more classical perspective on sampling, we refer to Jerri's excellent tutorial article, which gives a very comprehensive view of the subject up to the mid 1970's [62]. Another useful source of information are the survey articles that appeared in the mathematical literature [26], [27], [57].

A. Wavelets

In Section II, we have already encountered the scaling function φ, which plays a crucial role in wavelet theory. There, instead of a single space $V(\varphi) = V_0$, one considers a whole ladder of rescaled subspaces $V_i = V_{T=2^i}(\varphi) = \mathrm{span}\{\varphi_{i,k}\}_{k\in Z}$ using the standard notation $\varphi_{k,i} = 2^{-i/2}\varphi(x/2^i - k)$. If φ satisfies the two-scale relation (17), then these spaces are nested and form a multiresolution analysis of L_2 [80], [81], [85]. The wavelet function $\psi(x)$, which may be represented as $\psi(x/2) = \sum_{k\in Z} g(k)\varphi(x-k)$, is designed to generate a Riesz basis of the difference spaces $W_i = V_{i-1} - V_i$; these are also rescaled versions of each other but their pairwise intersections are all $\{0\}$—in contrast with the V_i's, which are included hierarchically. Since the multiresolution is dense in L_2 (i.e., we can approximate any L_2-function as closely as we wish by letting the scale $a = 2^i$ go to zero), it is thus possible to represent any function $f(x) \in L_2$ in terms of the wavelet expansion

$$f(x) = \sum_{k\in Z}\sum_{i\in Z} \langle f, \tilde{\psi}_{k,i}\rangle \psi_{k,i} \tag{47}$$

where k and i are the position and scale indexes, respectively. This is quite similar to the sampling expansion (18), except that (47) includes an additional summation over the scale index i. The wavelet expansion works because the analysis and synthesis wavelets $\tilde{\psi}$ and ψ generate a biorthogonal basis of L_2 such that $\langle \psi_{k,i}, \psi_{l,j}\rangle = \delta_{k-l,i-j}$ [32].

While a detailed discussion of wavelets is beyond the scope of this paper, we want to point out that the analysis tools and mathematics are essentially the same as those used in the modern formulations of sampling theory. In this sense, wavelets have had a very positive feedback on sampling and have contributed to a revival of the subject. The reader who wishes to learn more about wavelets is referred to the standard texts [35], [79], [115], [139].

B. Generalized (or Multichannel) Sampling

In 1977, Papoulis introduced a powerful extension of Shannon's sampling theory, showing that a band-limited signal could be reconstructed exactly from the samples of the response of m linear shift-invariant systems sampled at $1/m$ the reconstruction rate [90]. The main point is that there are many different ways of extracting information from a signal—a reconstruction is generally possible provided there are as many measurements as there are degrees of freedom in the signal representation. If the measurements are performed in a structured manner, then the reconstruction process is simplified: for example, in the Papoulis framework, it is achieved by multivariate filtering [23], [83]. Typical instances of generalized sampling are interlaced and derivative sampling [75], [149], both of which are special cases of Papoulis' formulation. While the generalized sampling concept is relatively straightforward, the reconstruction is not always feasible because of potential instabilities [24], [30], [82].

More recently, Papoulis' theory has been generalized in several directions. While still remaining with band-limited functions, it was extended for multidimensional, as well as multicomponent signals [25], [101]. Djokovic and Vaidyanathan applied similar ideas for the reconstruction of functions in certain wavelet spaces [44]; they concentrated on the special cases of interlaced sampling, sampling of a function and of its derivative, and reconstruction from local averages. A further step was taken by Unser and Zerubia who generalized the approach for a reconstruction in the space $V(\varphi)$ without any constraining hypothesis on the

input signal [137]. Instead of an exact reconstruction, which is obviously not possible as soon as $f \notin V(\varphi)$, they looked for a solution that is consistent in the sense that it yields the exact same measurements if it is reinjected into the system. Their key result is a multivariate extension of the sampling theorem described in Section III-C. The computational solution takes the form of a multivariate filterbank and is compatible with Papoulis' theory in the special case $\varphi = \mathrm{sinc}$. These authors also looked at performance issues and derived general formulas for the condition number of the system, as well as error bounds for the comparison with the least squares solution [136]. Janssen and Kalker carried out an explicit stability analysis for the particular case of interlaced sampling with a reconstruction in the space of piecewise linear splines [61].

Recent applications of generalized sampling include motion-compensated deinterlacing of televison images [11], [121], and super-resolution [107], [138]. The latter is an attempt to reconstruct a higher resolution image from a series of lower resolution ones, which are shifted slightly with respect to each other.

C. Finite Elements and Multiwavelets

An interesting generalization of (9) is to consider m generating functions $\varphi_1(x), \ldots, \varphi_m(x)$ instead of a single one; this corresponds to the finite element—or multiwavelet—framework. To obtain a signal representation that has the same sampling density as before, the multifunctions are translated by integer multiples of m

$$f(x) = \sum_{k \in Z} \sum_{i=1}^{m} c_i(k) \varphi_i(x - mk). \tag{48}$$

In the finite-elements method, the basis functions are typically chosen as short as possible and with minimal overlap, to facilitate the inversion of the matrices involved in the numerical solution of differential equations [114], [118]. In this area of application, the approximation order of the representation is the key parameter and the φ_i's do not need to be particularly smooth [113].

In the more recent multiwavelet constructions, the multiscaling functions satisfy a vector two-scale relation—similar to (17)—that involves a matrix refinement filter instead of a scalar one [4], [45], [51], [56], [116]. One of the primary motivation for this kind of extension is to enable the construction of scaling functions and wavelets that are symmetric (or antisymmetric), orthonormal, and compactly supported. These are properties that cannot be enforced simultaneously in the conventional wavelet framework, with the notable exception of the Haar basis [33].

For multiwavelet applications, one of the key issues is the appropriate determination of the $c_i(k)$'s in (48) for the initial representation of the signal at the finest scale available—the so-called *initialization* problem [140], [146]. The situation is very much the same as in the scalar case, when the generating function is noninterpolating (see Section III-D). Given the equidistant samples (or measurements) of a signal $f(x)$, the expansion coefficients are usually obtained through an appropriate digital prefiltering procedure (analysis filterbank) [54], [140], [146]. The initialization step—or prefiltering—can be avoided for the class of so-called balanced multiwavelets [71], [103]. Recently, Selesnick has designed multiscaling functions that are interpolating, orthonormal, and compactly supported [104]; these are the vector counterparts of the interpolating function φ_{int} in Section III-D. A simultaneous fulfillment of all these properties is not possible in the scalar case, except for the Haar scaling function, as proven by Xia [147].

Because of the importance of the finite elements in engineering, the quality of this type of approximation has been studied thoroughly by approximation theorists [64], [73], [111]. In addition, most of the results presented in Section IV are also available for the multifunction case [19].

D. Frames

The notion of frame, which generalizes that of a basis, was introduced by Duffin and Schaffer [47]; it plays a crucial role in nonuniform sampling [12]. A frame is essentially a set of functions that span the whole space, but are not necessarily linearly independent [3]. To be more precise, a sequence $\{\varphi_k(x)\}_{k \in Z}$ with $\varphi_k \in V$ constitutes a frame of the function space V if there exist two strictly positive constants A and B (frame bounds) such that

$$\forall f \in H,\ A \cdot \|f\|^2 \leq \sum_k |\langle f, \varphi_k \rangle|^2 \leq B \cdot \|f\|^2. \tag{49}$$

This implies that there is no function $f \in V$ (except zero) that is orthogonal to all frame components simultaneously, which ensures a complete coverage of the space V. The main difference with the Riesz basis condition (10) is that the frame definition allows for redundancy: there may be more template functions than are strictly necessary. In terms of sampling, this translates into situations where one has more measurements (or samples) than the minimum requirement. This is especially advantageous for reducing noise [119]. Practically, the signal reconstruction is obtained from the solution of an overdetermined system of linear equations [92]; see also [34], for an iterative algorithm when A is close to B.

When dealing with frames, the important fact to realize is that the signal representation in terms of the φ_k's is generally not unique. However, there is one representation (the minimum norm solution) that is especially attractive because it yields the same type of expansion as a Riesz basis: $f = \sum_k \langle f, \tilde{\varphi}_k \rangle \varphi_k$ where $\tilde{\varphi}_k = S^{-1}\varphi_k$ is the so-called *dual frame*—S^{-1} is the inverse of the *frame operator* S, defined as $Sf = \sum_k \langle f, \varphi_k \rangle \varphi_k$ [3], [47]. In particular, when the frame is tight—i.e., $A = B$—the operator S^{-1} is a simple renormalization, and one gets

$$\forall f \in H,\ f = \frac{1}{A} \sum_{k \in Z} \langle f, \phi_k \rangle \phi_k = \frac{1}{A} Sf \tag{50}$$

which is almost the same formula as (9), except for the normalization by A. When all vectors have a unit norm, then the factor A also gives the redundancy of the frame [34].

E. Irregular Sampling

Irregular or nonuniform sampling constitutes another whole area of research that we mention here only briefly to make the connection with what has been presented. There are essentially two strategies for finding a solution: 1) by considering the same kind of "shift-invariant" spaces as in the uniform case and by fitting the model to the measurements or 2) by defining new basis functions (or spaces) that are better suited to the nonuniform structure of the problem. The two approaches are not incompatible; for instance, one may very well construct nonuniform bases (or even frames) for any of the shift-invariant spaces $V(\varphi)$.

1) Irregular Sampling in Shift-Invariant Spaces: The problem that has been studied most extensively is the recovery of a band-limited function from its nonuniform samples $\{f(x_k)\colon x_k \in X\}$ [12], [48], [52], [70], [96], [102]. A set X for which a stable reconstruction is possible for all $f \in V(\varphi)$ is called a *set of sampling* for $V(\varphi)$. The stability requirement is important because there exist sets of samples that uniquely determine a band-limited function but for which the reconstruction is unstable—this happens, for example, when the samples are all bunched together. One of the deepest and most difficult results in this area is the Beurling–Landau density theorem [17], [53], [70]. In its simplifed version, the theorem states that all sets of sampling for the space of band-limited functions $V(\mathrm{sinc})$ must have a (lower) sampling density $D(X) \geq 1$—roughly speaking, $D(X)$ represents the (minimum) average number of samples per unit length in X. Conversely, if the Beurling density of the set X is $D(X) > 1$, then X is a set of sampling for $V(\mathrm{sinc})$, which means that that a perfect reconstruction of a band-limited function from its nonuniform samples in X is possible [17]. Efficient numerical methods for performing such reconstructions are described in [49] and [50]. More recently, researchers have extended these techniques to the more general wavelet and spline-like spaces $V(\varphi)$ [5], [29], [76], [77]. Aldroubi and Gröchenig derived generalized versions of the Beurling–Landau theorems based on an appropriate definition of the sampling density D_φ [6]. Specifically, they showed that the condition $D_\varphi(X) \geq 1$ is necessary to have a set of sampling for the general spaces $V(\varphi)$. Conversely, they also proved that the condition $D_\varphi(X) > 1$ is sufficient to have a set of sampling for the polynomial splines. The reconstruction can be achieved using the iterative algorithm described in [5].

The first paper on irregular sampling in terms of frames was [15]. In particular, Benedetto *et al.* analyzed the role of the coefficients as sampled values [12], [15]. They also gave frame sampling theorems for non-band-limited functions, allowing for a quantitative means of measuring aliasing.

2) Nonuniform Splines and Radial Basis Functions: Another fruitful approach to irregular sampling is to use specially tailored basis functions such as the nonuniform splines [37], [100]. The B-spline formalizm, in particular, is well suited for practical applications: the underlying B-spline basis functions are compactly supported and can be constructed systematically using de Boor's recursion [36], [37]. The expansion coefficients of the model are then determined from the solution of a sparse (band-diagonal) system of equations [37], [43]. One remarkable theoretical result, which makes the connection with Shannon's theory, is that the method can be used to recover a band-limited function from its nonuniform samples; the key theorem is that the nonuniform polynomial spline interpolant converges to the band-limited function as the order of the spline goes to infinity [78].

In higher dimensions, the B-spline formalism is no longer applicable unless the grid is separable. A more general approach is to use radial basis functions [93], which are closely related to splines as well [46]. The radial basis function model has the form

$$f_{\mathrm{int}}(\boldsymbol{x}) = \sum_k c(k)\rho(\|\boldsymbol{x} - \boldsymbol{x}_k\|) \tag{51}$$

where the basis functions $\rho(\|\boldsymbol{x} - \boldsymbol{x}_k\|)$ are centered on the sampling points $\boldsymbol{x}_k \in R^p$; they are radially symmetric because they depend on the distance $r_k = \|\boldsymbol{x} - \boldsymbol{x}_k\|$ only. Typical examples of radial functions are $\rho(r) = |r|$ (membrane spline) and $\rho(r) = r^p \log|r|$ (p even) or $\rho(r) = r^{4-p}$ (p odd). This latter choice yields the interpolant whose Laplacian energy is minimum (thin plate splines) [46]; it is the natural variational extension of the cubic splines in multiples dimensions. The expansion coefficients in (51) are determined by solving a linear system of equations which expresses the interpolation constraint $f_{\mathrm{int}}(\boldsymbol{x}_k) = f(\boldsymbol{x}_k)$ [28], [84]. Often, the solution includes an additional low-order polynomial term that is constrained to be orthogonal to the rest of the expansion. Micchelli has proven that the radial basis function interpolation problem with an arbitrary set of sampling points has a unique solution for a relatively wide class of positive, increasing radial functions $\rho(r)$ [86].

At first sight, the representation (51) looks rather similar to (9) for it also involves the shifts of a single template. However, the nature of the basis functions is fundamentally different. In the first case, $\varphi(x)$ is an L_2-function that is well localized (typically, compactly supported). Practically, this gets translated into a sparse and well-conditioned interpolation problem. In the second case, $\rho(r)$ is unbounded at infinity and is certainly not in L_2. Thus, the structure of the interpolation equations is not so favorable, which makes working with radial basis functions more delicate and much less efficient computationally. This appears to be the price to pay for their generality.

VI. Conclusion

Fifty years later, Shannon's sampling theory is still alive and well. It treats such a fundamental problem, with so many practical repercussions, that it is simply unavoidable. The sampling theorem is part of the basic knowledge of every engineer involved with digital signals or images. The subject is far from being closed and its importance is most likely to grow in the future with the ever-increasing trend of replacing analog systems by digital ones; typical application areas are

communications including the Web, sound, television, photography, multimedia, medical imaging, etc.

Recently, sampling has experienced a strong research revival, especially after it was realized that some of mathematics developed for wavelets (Riesz bases, projection operators) were the right tools for this problem as well. This has motivated many researchers with different background in engineering (e.g., signal and image processing) and mathematics (harmonic analysis, mathematical physics, and approximation theory) to pick up the subject, and has resulted in a substantial advancement of the field—especially its theoretical aspects. Many of the results reviewed in this paper have a potential for being useful in practice because they allow for a realistic modeling of the acquisition process and offer much more flexibility than the traditional band-limited framework. The newer formulations of sampling tend to give better practical results because the solutions are designed to be optimal in a well-defined sense (e.g., least squares).

Last, we believe that the general unifying view of sampling that has emerged during the past decade is beneficial because it offers a common framework for understanding—and hopefully improving—many techniques that have been traditionally studied by separate communities. Areas that may benefit from these developments are analog-to-digital conversion, signal and image processing, interpolation, computer graphics, imaging, finite elements, wavelets, and approximation theory.

APPENDIX A
SAMPLING FORMULAS

We briefly review two alternative ways of understanding the basic sampling formulas which are at the heart of Shannon's theory. To simplify the argument, we use a normalized time scale with a sampling step $T = 1$.

A. Sampling and Dirac Distributions

It is a common engineering practice to model the sampling process by a multiplication with a sampling sequence of Dirac impulses

$$\sum_{k\in Z} \delta(x-k) \overset{\text{Fourier}}{\longleftrightarrow} 2\pi \sum_{m\in Z} \delta(\omega+2\pi m). \tag{A.1}$$

The corresponding sampled signal representation is

$$fs(x) = \sum_{k\in Z} f(k)\delta(x-k). \tag{A.2}$$

In the Fourier domain, multiplication corresponds to a convolution, which yields

$$\hat{f}_\delta(\omega) = \hat{f}(\omega) * \sum_{m\in Z} \delta(\omega+2\pi m) = \sum_{m\in Z} \hat{f}(\omega+2\pi m) \tag{A.3}$$

where the underlying Fourier transforms are to be taken in the sense of distributions. Thus, the sampling process results in a periodization of the Fourier transform of f, as illustrated in Fig. 1(b). The reconstruction is achieved by convolving the sampled signal $f_\delta(x)$ with the reconstruction function φ

$$f_{\text{rec}}(x) = (f_\delta * \varphi)(x) = \sum_{k\in Z} f(k)\varphi(x-k). \tag{A.4}$$

In the Fourier transform domain, this gives

$$\hat{f}_{\text{rec}}(\omega) = \hat{\varphi}(\omega) \cdot \sum_{m\in Z} \hat{f}(\omega+2\pi m). \tag{A.5}$$

Thus, as illustrated in Fig. 1(c), we see that a perfect reconstruction is possible if $\hat{\varphi}(\omega)$ is an ideal low-pass filter [e.g., $\varphi(x) = \text{sinc}(x)$] and $\hat{f}(\omega) = 0$ for $|\omega| > \pi$ (Nyquist criterion).

If, on the other hand, f is not bandlimited, then the periodization of its Fourier transform in (A.5) results in spectral overlap that remains after postfiltering with $\hat{\varphi}(\omega)$. This distortion, which is generally nonrecoverable, is called *aliasing*.

B. Sampling and Poisson's Summation Formula

The standard form of Poisson's summation formula is (see [117])

$$\sum_{k\in Z} g(k) = \sum_{m\in Z} \hat{g}(2\pi m) \tag{A.6}$$

where $\hat{g}(\omega)$ denotes the Fourier transform of the continuous time function $g(x) \in L_1$. The reader is referred to [13], [16], or [65] for a rigorous mathematical treatment.

Considering the function $g(x) = f(x)e^{-\omega_0 x}$, the Fourier transform of which is $\hat{g}(\omega) = \hat{f}(\omega+\omega_0)$ (modulation property), we get

$$\sum_{k\in Z} f(k)\, e^{-j\omega_0 k} = \sum_{m\in Z} \hat{f}(\omega_0+2\pi m) = F(e^{j\omega_0}). \tag{A.7}$$

This is precisely the discrete-time Fourier transform of the sequence $\{f(k)\}$ with ω_0 as the frequency variable. The central term of (A.7) is identical to (A.3), which means that the 2π-periodic functions $F(e^{j\omega})$ and $\hat{f}_\delta(\omega)$ are in fact equivalent, even though they have very different interpretations—the former is a discrete-time Fourier transform, while the latter is a continuous-time one.

The last step in this formulation is to derive the Fourier transform of the reconstructed signal

$$\hat{f}_{\text{rec}}(\omega) = \int_{-\infty}^{+\infty} \left(\sum_{k\in Z} f(k)\varphi(x-k) \right) e^{-j\omega x}\, dx.$$

Exchanging the order of integration and making the change of variable $y = x - k$, we get

$$\begin{aligned} \hat{f}_{\text{rec}}(\omega) &= \sum_{k\in Z} f(k)\, e^{-j\omega k} \int_{-\infty}^{+\infty} \varphi(y)\, e^{-j\omega y}\, dy \\ &= F(e^{j\omega}) \cdot \hat{\varphi}(\omega). \end{aligned} \tag{A.8}$$

Together with (A.7), (A.8) is equivalent to (A.5).

APPENDIX B
NECESSITY AND SUFFICIENCY OF THE PARTITION OF UNITY

Our goal is to find conditions on φ such that the approximation error vanishes as $T \to 0$ for all $f \in W_2^1$ (the set

of L_2-functions that are differentiable once). We make use of Theorem 3 to get the asymptotic form of the error as the sampling step gets sufficiently small

$$\lim_{T\to 0} \|f - Q_T\{f\}\|^2 = \lim_{T\to 0} \int_{-\infty}^{+\infty} E_{\tilde{\varphi},\varphi}(T\omega)|\hat{f}(\omega)|^2 \frac{d\omega}{2\pi}.$$

Note that $E_{\tilde{\varphi},\varphi}(\omega)$ is bounded if $\hat{\tilde{\varphi}}(\omega)$ is bounded and φ satisfies the Riesz condition (11). Thus, we can apply Lebesgue's dominated convergence theorem and move the limit inside the integral

$$\begin{aligned}\lim_{T\to 0} \|f - Q_T\{f\}\|^2 &= \int_{-\infty}^{+\infty} \lim_{T\to 0} E_{\tilde{\varphi},\varphi}(T\omega)|\hat{f}(\omega)|^2 \frac{d\omega}{2\pi} \\ &= E_{\tilde{\varphi},\varphi}(0)\cdot\|f\|^2.\end{aligned}$$

This final manipulation requires that $E_{\tilde{\varphi},\varphi}(\omega)$ be continuous at $\omega = 0$. Consequently, under suitable technical conditions, the requirement $\forall\, f \in W_2^1$, $\lim_{T\to 0}\|f - Q_T\{f\}\|^2 = 0$ is equivalent to $E_{\tilde{\varphi},\varphi}(0) = 0$. Using (42), we obtain

$$E_{\tilde{\varphi},\varphi}(0) = |1 - \hat{\tilde{\varphi}}(0)\hat{\varphi}(0)|^2 + |\hat{\tilde{\varphi}}(0)|^2 \sum_{n\neq 0} |\hat{\varphi}(2\pi n)|^2 = 0.$$

This expression is a sum of positive terms; it can obviously vanish only if all are zero. In particular, we need to have $\hat{\tilde{\varphi}}(0)\hat{\varphi}(0) = 1$, which is possible only if both factors are nonzero. We can therefore renormalize φ such that $\hat{\varphi}(0) = \int \varphi(x)\,dx = 1$. Thus, the conditions that need to be satisfied are

$$\hat{\tilde{\varphi}}(0) = 1 \quad \text{and} \quad \hat{\varphi}(2\pi n) = \delta_n. \tag{B.1}$$

The second part is the same as the right-hand side of (14); it is equivalent to the partition of unity condition $\sum_k \varphi(x+k) = 1$ (in the sense of distributions).

Note that a more restrictive form of the above equivalence for $\tilde{\varphi}(x) = \delta(x)$ can be found in [95]. This result is also closely related to the Strang–Fix conditions of order one [113]—which give a strict equivalence between the partition of unity (reproduction of the constant) and a first order of approximation; i.e., $\|f - Q_T\{f\}\| = O(T^{-1})$, as $T \to 0$. However, the converse part of the Strang–Fix equivalence (see Section IV-B) requires some additional decay condition on φ: compact support [113], or at least some inverse polynomial decay [19]. Here, we have considered the weakest possible assumption—the continuity of $E_{\tilde{\varphi},\varphi}(\omega)$ around $\omega = 0$—which allows for slower rates of decay. Examples of admissible generating functions, which are not covered by the Strang–Fix theory, are the fractional B-splines of degree $-(1/2) < \alpha < 0$; these satisfy the partition of unity but have a fractional order of approximation $\alpha + 1 < 1$ [134].

ACKNOWLEDGMENT

The author wishes to thank A. Aldroubi and T. Blu for mathematical advice. He is also grateful to T. Blu, P. Thévenaz, and four anonymous reviewers for their constructive comments on the manuscript.

REFERENCES

[1] *Proc. SAMPTA'97*, Avero, Portugal, 1997.
[2] *Proc. SAMPTA'99*, Loen, Norway, 1999.
[3] A. Aldroubi, "Portraits of frames," in *Proc. Amer. Math. Soc.*, vol. 123, 1995, pp. 1661–1668.
[4] ——, "Oblique and hierarchical multiwavelet bases," *Appl. Computat. Harmonic Anal.*, vol. 4, no. 3, pp. 231–263, 1997.
[5] A. Aldroubi and H. Feichtinger, "Exact iterative reconstruction algorithm for multivariate irregularly sampled functions in spline-like spaces: The L_p-theory," in *Proc. Amer. Math. Soc.*, vol. 126, 1998, pp. 2677–2686.
[6] A. Aldroubi and K. Gröchenig, "Beurling–Landau-type theorems for non-uniform sampling in shift invariant spline spaces," *J. Fourier Anal. Appl.*, vol. 6, no. 1, pp. 91–101, 2000.
[7] A. Aldroubi and M. Unser, "Families of wavelet transforms in connection with Shannon sampling theory and the Gabor transform," in *Wavelets—A Tutorial in Theory and Applications*, C. K. Chui, Ed. San Diego, CA: Academic, 1992, pp. 509–528.
[8] ——, "Families of multiresolution and wavelet spaces with optimal properties," *Numer. Funct. Anal. Optimizat.*, vol. 14, no. 5-6, pp. 417–446, 1993.
[9] ——, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Funct. Anal. Optimizat.*, vol. 15, no. 1&2, pp. 1–21, 1994.
[10] A. Aldroubi, M. Unser, and M. Eden, "Cardinal spline filters: Stability and convergence to the ideal sinc interpolator," *Signal Process.*, vol. 28, no. 2, pp. 127–138, 1992.
[11] E. Bellers and G. de Haan, "New algorithm for motion estimation on interlaced video," in *Proc. SPIE Visual Communication and Image Processing*, vol. 3309, 1998, pp. 111– 121.
[12] J. J. Benedetto, "Irregular sampling and frames," in *Wavelets—A Tutorial in Theory and Applications*, C. K. Chui, Ed. Boca Raton, FL: CRC Press, 1992, pp. 445–507.
[13] ——, *Harmonic Analysis and Applications*. Boca Raton, FL: CRC Press, 1997.
[14] ——, "Frames, sampling, and seizure prediction," in *Advances in Wavelets*, K.-S. Lau, Ed. New York: Springer-Verlag, 1998, pp. 1–15.
[15] J. J. Benedetto and W. Heller, "Irregular sampling and the theory of frames," *Math. Note*, vol. 10, pp. 103–125, 1990.
[16] J. J. Benedetto and G. Zimmermann, "Sampling operators and the Poisson summation formula," *J. Fourier Anal. Applicat.*, vol. 3, pp. 505–523, 1997.
[17] A. Beurling, *The Collected Works of Arne Beurling*. Boston, MA: Birkhauser, 1989.
[18] T. Blu and M. Unser, "Minimum support interpolators with optimum approximation properties," in *Proc. Int. Conf. Image Processing*, vol. III, Santa Barbara, CA, 1998, pp. 242–245.
[19] ——, "Approximation error for quasi-interpolators and (multi-) wavelet expansions," *Appl. Computat. Harmonic Anal.*, vol. 6, no. 2, pp. 219–251, 1999.
[20] ——, "Quantitative Fourier analysis of approximation techniques—Part II: Wavelets," *IEEE Trans. Signal Processing*, vol. 47, pp. 2796–2806, Oct. 1999.
[21] ——, "Quantitative Fourier analysis of approximation techniques—Part I: Interpolators and projectors," *IEEE Trans. Signal Processing*, vol. 47, pp. 2783–2795, Oct. 1999.
[22] J. L. Brown, "On the error in reconstructing a nonbandlimited function by means of the band pass sampling theorem," *J. Math. Anal. Appl.*, vol. 18, pp. 75–84, 1967.
[23] ——, "Multi-channel sampling of lowpass-pass signals," *IEEE Trans. Circuits Syst.*, vol. 28, no. 2, pp. 101–106, 1981.
[24] J. L. Brown Jr. and S. D. Cabrera, "On well-posedness of the Papoulis generalized sampling expansion," *IEEE Trans. Circuits Syst.*, vol. 38, no. 5, pp. 554–556, 1991.
[25] J. L. Brown Jr. and K. Sa-ngsari, "Sampling reconstruction of N-dimensional band-limited images after multilinear filtering," *IEEE Trans. Circuits Syst.*, vol. 36, no. 7, pp. 1035–1038, 1989.
[26] P. L. Butzer, "A survey of the Whittaker–Shannon sampling theorem and some of its extensions," *J. Math. Res. Exposit.*, vol. 3, pp. 185–212, 1983.
[27] P. L. Butzer and R. L. Stens, "Sampling theory for not-necessarily band-limited functions: A historical overview," *SIAM Rev.*, vol. 34, pp. 40–53, 1992.

[28] J. C. Carr, W. R. Fright, and R. K. Beatson, "Surface interpolation with radial basis functions for medical imaging," *IEEE Trans. Med. Imag.*, vol. 16, no. 1, pp. 96–107, 1997.

[29] W. Chen, S. Itoh, and J. Shiki, "Irregular sampling theorems for wavelet subspaces," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1131–1142, 1998.

[30] K. F. Cheung and R. J. Marks II, "Ill-posed sampling theorems," *IEEE Trans. Circuits Syst.*, vol. CAS-32, no. 5, pp. 481–484, 1985.

[31] C. K. Chui, *Wavelets: A Mathematical Tool for Signal Analysis*. Philadelphia, PA: SIAM, 1997.

[32] A. Cohen, I. Daubechies, and J. C. Feauveau, "Bi-orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

[33] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.

[34] ——, "The wavelet transform—Time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, Sept. 1990.

[35] ——, *Ten Lectures on Wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992.

[36] C. de Boor, "On calculating with B-splines," *J. Approximat. Theory*, vol. 6, pp. 50–62, 1972.

[37] ——, *A Practical Guide to Splines*. New York: Springer-Verlag, 1978.

[38] ——, "Quasi-interpolation and approximation power of multivariate splines," in *Computation of Curves and Surfaces*, W. Dahmen *et al.*, Eds. Dordrecht, The Netherlands: Kluwer, 1990, pp. 313–345.

[39] C. de Boor, R. A. DeVore, and A. Ron, "Approximation from shift-invariant subspaces of $L_2(Rd)$," *Trans. Amer. Math. Soc.*, vol. 341, no. 2, pp. 787–806, 1994.

[40] C. de Boor and G. Fix, "Spline approximation by quasi-interpolants," *J. Approximat. Theory*, vol. 8, pp. 19–45, 1973.

[41] C. de Boor and R.-Q. Jia, "Controlled approximation and characterization of the local approximation order," in *Proc. Amer. Soc.*, vol. 95, 1985, pp. 547–553.

[42] G. Deslauriers and S. Dubuc, "Symmetric iterative interpolation processes," *Constr. Approx.*, vol. 5, pp. 49–68, 1989.

[43] P. Diercks, *Curve and Surface Fitting With Splines*. Oxford, U.K.: Oxford Univ. Press, 1995.

[44] I. Djokovic and P. P. Vaidyanathan, "Generalized sampling theorems in multiresolution subspaces," *IEEE Trans. Signal Processing*, vol. 45, pp. 583–599, Mar. 1997.

[45] G. C. Donovan, J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Construction of orthogonal wavelets using fractal interpolation functions," *SIAM J. Math. Anal.*, vol. 27, no. 4, pp. 1158–1192, 1996.

[46] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller, Eds. Berlin, Germany: Springer-Verlag, 1977, pp. 85–100.

[47] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series," *Trans. Amer. Math. Soc.*, vol. 72, pp. 314–366, 1952.

[48] H.-.G. Feichtinger and K. Gröchenig, "Irregular sampling theorems and series expansions of band-limited functions," *J. Math. Anal. Appl.*, vol. 167, no. 2, pp. 530–556, 1992.

[49] ——, "Iterative reconstruction of multivariate band-limited functions from irregular sampling values," *SIAM J. Math. Anal.*, vol. 23, no. 1, pp. 244–261, 1992.

[50] H. G. Feichtinger, K. Gröchenig, and T. Ströhmer, "Efficient numerical methods in non-uniform sampling theory," *Numerische Mathematik*, vol. 69, no. 4, pp. 423–440, 1995.

[51] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *J. Approx. Theory*, vol. 78, no. 3, pp. 373–401, 1994.

[52] K. Gröchenig, "Reconstruction algorithms in irregular sampling," *Math. Comp.*, vol. 59, no. 199, pp. 181–194, 1992.

[53] K. Gröchenig and H. Razafinjatovo, "On Landau's necessary density conditions for sampling and interpolation of band-limited functions," *J. London Math. Soc. (2)*, vol. 54, no. 3, pp. 557–565, 1996.

[54] D. P. Hardin and D. W. Roach, "Multiwavelet prefilters—1: Orthogonal prefilters preserving approximation order $p \leq 2$," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 8, pp. 1106–1112, 1998.

[55] G. H. Hardy, "Notes of special systems of orthogonal functions—IV: The orthogonal functions of Whittaker's series," in *Proc. Camb. Phil. Soc.*, vol. 37, 1941, pp. 331–348.

[56] C. Heil, G. Strang, and V. Strela, "Approximation by translates of refinable functions," *Numer. Math.*, vol. 73, no. 1, pp. 75–94, 1996.

[57] J. R. Higgins, "Five short stories about the cardinal series," *Bull. Amer. Math. Soc.*, vol. 12, pp. 45–89, 1985.

[58] ——, *Sampling Theory in Fourier and Signal Analysis*. Oxford, U.K.: Clarendon, 1996.

[59] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, no. 6, pp. 508–517, 1978.

[60] A. J. E. M. Janssen, "The Zak transform and sampling theorems for wavelet subspaces," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3360–3364, 1993.

[61] A. J. E. M. Janssen and T. Kalker, "A note on Unser–Zerubia generalized sampling theory applied to the linear interpolator," *IEEE Trans. Signal Processing*, vol. 47, pp. 2332–2335, Aug. 1999.

[62] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, pp. 1565–1596, Nov. 1977.

[63] ——, *The Gibbs Phenomenon in Fourier Analysis, Splines, and Wavelet Approximations*. Boston, MA: Kluwer Academic, 1998.

[64] R.-Q. Jia and J. Lei, "Approximation by multi-integer translates of functions having global support," *J. Approximat. Theory*, vol. 72, no. 2-23, 1993.

[65] Y. Katznelson, *An Introduction to Harmonic Analysis*. New York: Dover, 1976.

[66] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 6, pp. 1153–1160, 1981.

[67] V. A. Kotel'nikov, "On the transmission capacity of "ether" and wire in electrocommunications," *Izd. Red. Upr. Svyazzi RKKA (Moscow)*, 1933.

[68] ——, "Reprint: On the transmission capacity of "ether" and wire in electrocommunications," in *Modern Sampling Theory: Mathematics and Applications*, J. J. Benedetto and P. J. S. G. Ferreira, Eds. Boston, MA: Birkhauser, 2000, to be published.

[69] E. Kreyszig, *Introduction to Functional Analysis with Applications*. New York: Wiley, 1978.

[70] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, vol. 117, pp. 37–52, 1967.

[71] J. Lebrun and M. Vetterli, "Balanced multiwavelets theory and design," *IEEE Trans. Signal Processing*, vol. 46, pp. 1119–1125, Apr. 1998.

[72] C. Lee, M. Eden, and M. Unser, "High quality image re-sizing using oblique projection operators," *IEEE Trans. Image Processing*, vol. 7, pp. 679–692, May 1998.

[73] W. A. Light, "Recent developments in the Strang–Fix theory for approximation orders," in *Curves and Surfaces*, P. J. Laurent, A. L. Méhauté, and L. L. Schumaker, Eds. Boston, MA: Academic, 1991, pp. 285–292.

[74] W. A. Light and E. W. Cheney, "Quasi-interpolation with translates of a function having non-compact support," *Constructive Approximat.*, vol. 8, pp. 35–48, 1992.

[75] D. A. Linden, "A discussion of sampling theorems," *Proc. IRE*, vol. 47, pp. 1219–1226, 1959.

[76] Y. Liu, "Irregular sampling for spline wavelet subspaces," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 623–627, 1996.

[77] Y. M. Liu and G. G. Walter, "Irregular sampling in wavelet subspaces," *J. Fourier Anal. Appl.*, vol. 2, no. 2, pp. 181–189, 1995.

[78] Y. Lyubarskii and W. R. Madych, "The recovery of irregularly sampled band limited functions via tempered splines," *J. Functional Analysis*, vol. 125, no. 1, pp. 201–222, 1994.

[79] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1998.

[80] ——, "Multiresolution approximations and wavelet orthogonal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, no. 1, pp. 69–87, 1989.

[81] ——, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.

[82] R. J. Marks II, "Noise sensitivity of band-limited signal derivative interpolation," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 31, pp. 1028–1032, 1983.

[83] ——, *Introduction to Shannon Sampling Theory*. New York: Springer-Verlag, 1991.

[84] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," *Zeitschrift für Angewandte Mathematik und Physik*, vol. 30, pp. 292–304, 1979.

[85] Y. Meyer, *Ondelettes et Opérateurs—I : Ondelettes*. Paris, France: Hermann, 1990.
[86] C. A. Micchelli, "Interpolation of scattered data: Distance matrices and conditionally positive definite functions," *Constructive Approximat.*, vol. 2, pp. 11–22, 1986.
[87] M. Z. Nashed and G. G. Walter, "General sampling theorems for functions in reproducing kernel Hilbert spaces," *Math. Control, Signals, Syst.*, vol. 4, pp. 373–412, 1991.
[88] H. Nyquist, "Certain topics in telegraph transmission theory," *Trans. Amer. Inst. Elect. Eng.*, vol. 47, pp. 617–644, 1928.
[89] J. L. Ostuni, A. K. S. Santha, V. S. Mattay, D. R. Weinberger, R. L. Levin, and J. A. Frank, "Analysis of interpolation effects in the reslicing of functional MR-images," *J. Comput. Assist. Tomography*, vol. 21, no. 5, pp. 803–810, 1997.
[90] A. Papoulis, "Generalized sampling expansion," *IEEE Trans. Circuits Syst.*, vol. 24, pp. 652–654, Nov. 1977.
[91] J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imaging*, vol. MI-2, no. 1, pp. 31–39, 1983.
[92] S. C. Pei and M. H. Yeh, "An introduction to discrete finite frames," *IEEE Signal Processing Mag.*, vol. 14, no. 6, pp. 84–96, 1997.
[93] M. J. D. Powell, "The theory of radial basis function approximation in 1990," in *Advances in Numerical Analysis—II: Wavelets, Subdivision Algorithms and Radial Functions*, W. A. Light, Ed. Oxford, U.K.: Oxford Univ. Press, 1992, pp. 105–210.
[94] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1991.
[95] S. Ries and R. L. Stens, "Approximation by generalized sampling series," in *Constructive Theory of Functions*. Sofia, Bulgaria: Bulgarian Academy of Sciences, 1984, pp. 17–37.
[96] G. Say Song and I. G. H. Ong, "Reconstruction of bandlimited signals from irregular samples," *Signal Process.*, vol. 46, no. 3, pp. 315–29, 1995.
[97] A. Schaum, "Theory and design of local interpolators," *CVGIP: Graph. Models Image Process.*, vol. 55, pp. 464–481, 1993.
[98] I. J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, vol. 4, pp. 45–99, 112–141, 1946.
[99] ——, *Cardinal Spline Interpolation*. Philadelphia, PA: Society of Industrial and Applied Mathematics, 1973.
[100] L. L. Schumaker, *Spline Functions: Basic Theory*. New York: Wiley, 1981.
[101] D. Seidner, M. Feder, D. Cubanski, and S. Blackstock, "Introduction to vector sampling expansion," *IEEE Signal Processing Lett.*, no. 5, pp. 115–117, May 1998.
[102] K. Seip, "An irregular sampling theorem for functions bandlimited in a generalized sense," *SIAM J. Appl. Math.*, vol. 47, no. 5, pp. 1112–1116, 1987.
[103] I. W. Selesnick, "Multiwavelet bases with extra approximation properties," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 2898–2908, 1998.
[104] ——, "Interpolating multiwavelets bases and the sampling theorem," *IEEE Trans. Signal Processing*, vol. 47, no. 6, pp. 1615–1621, 1999.
[105] C. E. Shannon, "Communication in the presence of noise," in *Proc. IRE*, vol. 37, 1949, pp. 10–21.
[106] ——, "Classic paper: Communication in the presence of noise," *Proc. IEEE*, vol. 86, no. 2, pp. 447–457, 1998.
[107] H. Shekarforoush, M. Berthod, and J. Zerubia, "3D super-resolution using generalized sampling expansion," in *Proc. Int. Conf. Image Processing*, vol. II, Washington, DC, October 23–26, 1995, pp. 300–303.
[108] D. Slepian, "On bandwidth," *Proc. IEEE*, vol. 64, no. 3, pp. 292–300, 1976.
[109] R. L. Stens, "Error estimates for sampling sums based on convolution integrals," *Information Control*, vol. 45, pp. 37–47, 1980.
[110] ——, "Approximation of functions by Whittaker's cardinal series," in *General Inequalities*, W. Walter, Ed. Basel: Birkhäuser, 1984, pp. 137–149.
[111] G. Strang, "The finite element method and approximation theory," in *Numerical Solution of Partial Differential Equations—II*, B. Hubbard, Ed. New York: Academic, 1971, pp. 547–583.
[112] ——, "Wavelets and dilation equations: A brief introduction," *SIAM Rev.*, vol. 31, pp. 614–627, 1989.
[113] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," *Constructive Aspect of Functional Analysis*, pp. 796–830, 1971.
[114] ——, *An Analysis of the Finite Element Method*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[115] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge, 1996.
[116] G. Strang and V. Strela, "Orthogonal multiwavelets with vanishing moments," *Opt. Eng.*, vol. 33, no. 7, pp. 2104–2107, 1994.
[117] R. Strichartz, *A Guide To Distribution Theory and Fourier Transforms*. Boca Raton, FL: CRC Press, 1994.
[118] B. Szabo and I. Babuska, *Finite Element Analysis*. New York: Wiley, 1991.
[119] A. Teolis and J. J. Benedetto, "Local frames and noise reduction," *Signal Processing*, vol. 45, no. 3, pp. 369–387, 1995.
[120] P. Thévenaz, T. Blu, and M. Unser, "Image interpolation and resampling," in *Handbook of Medical Image Processing*, I. Bankman, Ed. New York: Academic, 2000, ch. 25.
[121] G. Thomas, "A comparison of motion-compensated interlace-to-progressive conversion methods," *Signal Process. Image Commun.*, vol. 12, no. 3, pp. 209–29, 1998.
[122] M. Unser, "A general Hilbert space framework for the discretization of continuous signal processing operators," in *Proc. SPIE Conf. Wavelet Applications in Signal and Image Processing III*, San Diego, CA, July 9–14, 1995, pp. 51–61.
[123] ——, "Approximation power of biorthogonal wavelet expansions," *IEEE Trans. Signal Processing*, vol. 44, no. 3, pp. 519–527, 1996.
[124] ——, "Quasi-orthogonality and quasi-projections," *Appl. Computat. Harmonic Anal.*, vol. 3, no. 3, pp. 201–214, 1996.
[125] ——, "Ten good reasons for using spline wavelets," in *Proc. SPIE Conf Wavelet Applications in Signal and Image Processing V*, San Diego, CA, Aug. 6–9, 1997, pp. 422–431.
[126] ——, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, vol. 16, no. 6, pp. 22–38, 1999.
[127] M. Unser and A. Aldroubi, "A general sampling theory for non-ideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, pp. 2915–2925, Nov. 1994.
[128] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 277–285, Mar. 1991.
[129] ——, "On the asymptotic convergence of B-spline wavelets to Gabor functions," *IEEE Trans. Inform. Theory*, vol. 38, pp. 864–872, Mar. 1992.
[130] ——, "Polynomial spline signal approximations: Filter design and asymptotic equivalence with Shannon's sampling theorem," *IEEE Trans. Inform. Theory*, vol. 38, pp. 95–103, Jan 1992.
[131] ——, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Trans. Signal Processing*, vol. 41, pp. 834–848, Feb. 1993.
[132] ——, "B-spline signal processing: Part I—Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821–833, Feb. 1993.
[133] ——, "Enlargement or reduction of digital images with minimum loss of information," *IEEE Trans. Image Process.*, vol. 4, pp. 247–258, Mar 1995.
[134] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Rev.*, vol. 42, no. 1, pp. 43–67, Mar. 2000.
[135] M. Unser and I. Daubechies, "On the approximation power of convolution-based least squares versus interpolation," *IEEE Trans. Signal Processing*, vol. 45, pp. 1697–1711, July 1997.
[136] M. Unser and J. Zerubia, "Generalized sampling: Stability and performance analysis," *IEEE Trans. Signal Processing*, vol. 45, pp. 2941–2950, Dec. 1997.
[137] ——, "A generalized sampling without bandlimiting constraints," *IEEE Trans. Circuits Syst.*, vol. 45, pp. 959–969, Aug. 1998.
[138] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *Comput. Vision, Graph., Image Process.*, vol. 54, no. 2, pp. 181–186, 1991.
[139] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[140] M. J. Vrhel and A. Aldroubi, "Projection based prefiltering for multiwavelet transforms," *IEEE Trans. Signal Processing*, vol. 46, pp. 3088–3092, Nov. 1998.
[141] G. G. Walter, "A sampling theorem for wavelet subspaces," *IEEE Trans. Inform. Theory*, vol. 38, pp. 881–884, Mar. 1992.
[142] P. Weiss, "An estimate of the error arising from misapplication of the sampling theorem," *Notices Amer. Math. Soc.*, vol. 10, p. 351, 1963.
[143] E. T. Whittaker, "On the functions which are represented by the expansion of interpolating theory," in *Proc. R. Soc. Edinburgh*, vol. 35, 1915, pp. 181–194.

[144] J. M. Whittaker, "The Fourier theory of the cardinal functions," in *Proc. Math. Soc. Edinburgh*, vol. 1, 1929, pp. 169–176.
[145] A. D. Wyner and S. Shamai, "Introduction to 'Communication in the presence of noise' by C.E. Shannon," *Proc. IEEE*, vol. 86, pp. 442–446, Feb. 1998.
[146] X. G. Xia, "A new prefilter design for discrete multiwavelet transforms," *IEEE Trans. Signal Processing*, vol. 46, pp. 1558–1570, June 1998.
[147] X. G. Xia and Z. Zhang, "On sampling theorem, wavelets, and wavelet transforms," *IEEE Trans. Signal Processing*, vol. 41, pp. 3524–3535, Dec. 1993.
[148] K. Yao, "Applications of reproducing kernel Hilbert spaces—Bandlimited signal models," *Inform. Control*, vol. 11, pp. 429–444, 1967.
[149] J. L. Yen, "On the nonuniform sampling of bandwidth limited signals," *IRE Trans. Circuit Theory*, vol. CT-3, pp. 251–257, 1956.

Michael Unser (M'89–SM'94–F'99) received the M.S. (*summa cum laude*) and Ph.D. degrees in electrical engineering from the Swiss Federal Institute of Technology in Lausanne (EPFL), Switzerland, in 1981 and 1984, respectively.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD. He is now Professor and Head of the Biomedical Imaging Group at EPFL. His main research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing. He is the author of more than 80 published journal papers in these areas. He serves as regular Chair for the SPIE conference on Wavelet Applications in Signal and Image Processing, held annually since 1993.

Dr. Unser is an Associate Editor for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He is on the editorial boards of several other journals, including *Signal Processing,* IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1995), and IEEE SIGNAL PROCESSING LETTERS (1994–1998). He received the IEEE Signal Processing Society's 1995 Best Paper Award.