# Chapter 4
# Function Spaces for Sampling Expansions

**M. Zuhair Nashed and Qiyu Sun**

**Abstract** In this paper, we consider a variety of Hilbert and Banach spaces that admit sampling expansions $f(t) = \sum_{n=1}^{\infty} f(t_n) S_n(t)$, where $\{S_n\}_{n=1}^{\infty}$ is a family of functions that depend on the sampling points $\{t_n\}_{n=1}^{\infty}$ but not on the function $f$. Those function spaces, that arise in connection with sampling expansions, include reproducing kernel spaces, Sobolev spaces, Wiener amalgam space, shift-invariant spaces, translation-invariant spaces, and spaces modeling signals with finite rate of innovation. Representative sampling theorems are presented for signals in each of these spaces. The paper also includes recent results on nonlinear sampling of signals with finite rate of innovation, convolution sampling on Banach spaces, and certain foundational issues in sampling expansions.

## 4.1 Introduction

Series expansions and integral representations of functions and operators play a fundamental role in the analysis of *direct problems* of applied mathematics—witness the role of power series, Fourier series, Karhunen–Loève expansion, eigenfunction expansions of symmetric linear operators, and sampling expansions in signal processing; and the role of Fourier transform, spectral integral representations, and various integral representations in boundary value problems, potential theory, complex analysis, and other areas.

Expansion theorems also play a fundamental role in *inverse problems*. Two important problems discussed below are: (1) the recovery of a function from inner product with a given set of functions (i.e., the moment problem), and (2) the recovery of a function from its values on a subset of its domain (i.e., reconstruction of a function from its samples via sampling expansion).

M.Z. Nashed (✉) • Q. Sun
Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA
e-mail: zuhair.nashed@ucf.edu; qiyu.sun@ucf.edu

One of the important problems in analysis is to expand a given function $f$ in a separable Banach space by a series of the form

$$f(t) = \sum_{n=1}^{\infty} c_n f_n(t), \qquad (4.1)$$

where $\{f_n\}_{n=1}^{\infty}$ is a suitable sequence of functions. This is not always possible. We consider two important cases of such an expansion:

1. Let $H$ be a separable Hilbert space and $\{f_n\}_{n=1}^{\infty}$ be an orthonormal sequence in $H$. If $\{c_n\}_{n=1}^{\infty}$ is a sequence for which the right-hand side of the expansion (4.1) converges to $f$ in $H$, then $c_n = \langle f, f_n \rangle$ for all $n \geq 1$, and hence $c_n, n \geq 1$, are the (generalized) Fourier coefficients. In this case, the series

$$f(t) = \sum_{n=1}^{\infty} \langle f, f_n \rangle f_n(t) \qquad (4.2)$$

is called the *(generalized) Fourier expansion*. This series converges to $f$ if and only if the orthonormal sequence $\{f_n\}$ is complete in the sense that the only function orthogonal to all the $f_n$'s is the zero function. Equivalently, Parseval equality

$$\sum_{n=1}^{\infty} |\langle f, f_n \rangle|^2 = \|f\|^2$$

holds for every $f \in H$.

This expansion theorem is a *direct problem*: Given $f$, find its expansion. The associated inverse problem is the *moment problem*: determine $f$ given moments $\langle f, f_n \rangle, n \in J$ (an index set). Given any sequence of real number $s_n, n \in J$, the existence of the moment problem is whether there exists a function $f$ such that $s_n = \langle f, f_n \rangle, n \in J$, while the uniqueness is whether such a function $f$ is uniquely determined by its moment sequence $s_n, n \in J$.

2. The second type of expansion that is the central theme of this chapter, is what is called a *sampling expansion*:

$$f(t) = \sum_{n=1}^{\infty} f(t_n) S_n(t), \qquad (4.3)$$

where $\{S_n\}_{n=1}^{\infty}$ is called a *sampling sequence* and $\{t_n\}_{n=1}^{\infty}$ is the *sampling set*. The inverse problem for sampling is to determine $f$ from given samples $f(t_n), n \geq 1$.

There have been many advances in sampling theory and its applications to signal and image processing. In the past three decades, many authors developed sampling theorems based on (1) the theory of regular and singular boundary value problems and (2) transforms other than the Fourier transform, including such transforms as the Sturm–Liouville transform, Jacobi transform, and Hankel transform, see [43]. Another main thrust has been in nonuniform sampling for non-bandlimited signals.

In the past 20 years, there have been major advances in sampling theory and its foundational aspects, where methods of functional analysis and harmonic analysis have played pivotal roles. In particular, new directions in sampling theory have been pursued using various function spaces that admit sampling expansions, such as reproducing kernel Hilbert (Banach) spaces, Sobolev spaces, Wiener amalgam space, shift-invariant spaces, translation-invariant spaces, and spaces modeling signals with finite rate of innovation. Another direction of research on sampling in the past decade involves average sampling, convolution sampling, nonlinear sampling, and other fundamental issues in sampling theory. The reader may refer to [1, 6, 7, 9–11, 15–24, 30, 33, 37, 39, 43] for various prospectives on these advances.

The purpose of this chapter is to consider a variety of function spaces mentioned above in which every function admits the sampling expansion (4.3). Representative sampling expansions are presented for signals in each of these spaces. This chapter also includes recent results on nonlinear sampling for signals with finite rate of innovation, convolution sampling on Banach spaces, and certain foundational issues in sampling expansions.

## 4.2 Fourier Series/Fourier Integral Approach to the Whittaker, Shannon, and Kotel'nikov Sampling Theorem

Let $f$ be a bandlimited signal with finite energy; i.e.,

$$f(t) = \int_{-\Omega}^{\Omega} F(\omega) e^{i\omega t} d\omega, \ t \in (-\infty, \infty), \tag{4.4}$$

for some square-integrable function $F$ on $[-\Omega, \Omega], \Omega > 0$. We extend $F(\omega)$ periodically to the real line and expand the extension in complex Fourier series:

$$F(\omega) = \sum_{n=-\infty}^{\infty} c_n \exp(in\pi\omega/\Omega), \ |\omega| < \Omega, \tag{4.5}$$

where

$$c_n = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\omega) \exp(-in\pi\omega/\Omega) d\omega. \tag{4.6}$$

Comparing (4.4) and (4.6) leads to

$$c_n = \frac{1}{2\Omega} f(-n\pi/\Omega). \tag{4.7}$$

Substituting (4.7) in (4.5) gives

$$F(\omega) = \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f(-n\pi/\Omega) \exp(in\pi\omega/\Omega)$$

$$= \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \exp(-in\pi\omega/\Omega), \ |\omega| < \Omega.$$

Substituting this in (4.4) and interchanging the order of integration and summation leads to

$$f(t) = \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \int_{-\Omega}^{\Omega} \exp(-in\pi\omega/\Omega) \exp(i\omega t) d\omega,$$

which yields the celebrated classical expansion of a bandlimited signal $f$:

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}. \tag{4.8}$$

This can be simplified to

$$f(t) = \sin(\Omega t) \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \frac{(-1)^n}{\Omega t - n\pi}.$$

This classical proof is rigorous. The interchange of integration and summation can be easily justified. But the proof is not very revealing: We perform this interchange and a theorem pops up. A theorem is born, but *we did not hear the heartbeat of its proof.*

## 4.3 Properties of the Sinc Function and the Paley–Wiener Space

Let $B_\pi$ consist of all signals that are bandlimited to $[-\pi, \pi]$ (i.e., $\Omega = \pi$ in (4.4)) and have finite energy. The space $B_\pi$ is the same as the Paley–Wiener space of restrictions to the real axis of entire functions of exponential type $\pi$. The Paley–Wiener space $B_\pi$ has many interesting properties that are not exploited or used in the classical proof of the Whittaker–Shannon–Kotel'nikov sampling theorem. Some of these properties are stated in the following theorem.

**Theorem 1.** *(i)* $B_\pi$ *is a reproducing kernel space with kernel*

$$\text{sinc}(s-t) := \frac{\sin \pi(s-t)}{\pi(s-t)}.$$

(ii)   *The sequence $\{S_n\}_{n \in \mathbf{Z}}$, where $S_n(t) := \mathrm{sinc}(t - n)$, is an orthonormal basis for $B_\pi$.*

(iii)  *The sequence $\{S_n\}_{n \in \mathbf{Z}}$ has the discrete orthogonality property:*

$$S_n(m) = \delta_{mn} := \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases}$$

(iv)  $f(\cdot - c) \in B_\pi$ *and* $\|f(\cdot - c)\|_2 = \|f\|_2$ *for all* $f \in B_\pi$ *and* $c \in \mathbf{R}$. *Hence* $B_\pi$ *is unitarily translation-invariant subspace of* $L^2(\mathbf{R})$.

(v)   $B_\pi$ *is a shift-invariant subspace of* $L^2(\mathbf{R})$ *generated by the sinc function:*

$$B_\pi = \left\{ \sum_{n \in \mathbf{Z}} c(n) \, \mathrm{sinc}(t - n) : \sum_{n \in \mathbf{Z}} |c(n)|^2 < \infty \right\}.$$

*Proof.*   (i) Take $f \in B_\pi$, and let $F$ be the square-integrable function in (4.4). Then

$$f(t) = \int_{-\pi}^{\pi} e^{it\omega} F(\omega) d\omega, \ t \in (-\infty, \infty). \tag{4.9}$$

This implies that

$$|f(t)| \leq \int_{-\pi}^{\pi} |F(\omega)| d\omega \leq (2\pi)^{1/2} \|F\|_2 = \|f\|_2$$

and

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{it\omega} \left( \int_{-\infty}^{\infty} f(s) e^{-is\omega} ds \right) d\omega \\ &= \int_{-\infty}^{\infty} f(s) \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(t-s)\omega} d\omega \right) ds \\ &= \int_{-\infty}^{\infty} f(s) \frac{\sin \pi(s - t)}{\pi(s - t)} ds, \ t \in (-\infty, \infty). \end{aligned}$$

Hence $B_\pi$ is a reproducing kernel Hilbert space with kernel $\frac{\sin \pi(s-t)}{\pi(s-t)}$.

(ii) By the reproducing property and symmetry of the sinc kernel $k(s,t) := \frac{\sin \pi(s-t)}{\pi(s-t)}$, we have that $\langle S_n, S_m \rangle = k(n,m)$, which takes value one if $m = n$ and zero if $m \neq n$. Hence $S_n, n \in \mathbf{Z}$, is an orthonormal set. The completeness of the orthonormal set $\{S_n\}_{n \in \mathbf{Z}}$ follows from (4.8).

(iii) The discrete orthogonality property is obvious.

(iv) Take $f \in B_\pi$, and let $F$ be the square-integrable function in (4.4). Then it follows from (4.9) that for any $c \in (-\infty, \infty)$,

$$f(t - c) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(t-c)\omega} F(\omega) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{it\omega} F_c(\omega) d\omega,$$

where $F_c(\omega) = e^{-ic\omega}F(\omega)$ is square-integrable. Hence $f(t-c) \in B_\pi$ for all $t \in (-\infty, \infty)$.

(v)  This follows from the conclusion that $\{S_n\}_{n \in \mathbf{Z}}$ is an orthonormal basis for the Paley–Wiener space $B_\pi$.                                                                                                              □

## 4.4  The Engineering Approach to Sampling and Its Mathematical Deficiencies

We now turn to the engineering clever approach to the sampling theorem, see, e.g., [13, 30, 32]. We paraphrase some of the description in [13]. Let us consider what happens when we sample $f(t)$ at uniformly spaced times. If the sampling frequency is $f_s$, then we can model this with a multiplication of $f(t)$ by a train of Dirac impulsed spaces $T_s := 1/f_s$ second apart:

$$f^*(t) := f(t)T_s \sum_{n=-\infty}^{\infty} \delta(t - nT_s) = T_s \sum_{n=-\infty}^{\infty} f(nT_s)\delta(t - nT_s).$$

Mathematicians consider the sequence of sampled values $\{f(nT_s)\}_{n=-\infty}^{\infty}$ as a vector in $\ell_2$ (the space of all square-summable vectors). Electrical engineers like to continue to think of this sequence as a time signal. So to stay in the analog world, they use their beloved "Dirac impulse" as above. Informally, as they assert, "multiplication by an impulse train" in the time domain corresponds to convolution with an impulse train in the frequency domain. If the Fourier transform of the sampled sequence $f^*(t)$ is $F^*(\omega)$, then

$$F^*(\omega) = F(\omega) * \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s) = \sum_{n=-\infty}^{\infty} F(\omega - n\omega_s).$$

Hence, again informally, the Fourier transform of the samples (considered as a time signal in the sense of the above representation) is an infinitely repeated replication of $F(\omega)$ at intervals of $\omega_s := 2\pi f_s$. The portion of the transform between $-\omega_s/2$ and $\omega_s/2$ is called the base band and all the other replication images. If $f(t)$ is bandlimited so that $F(\omega)$ is zero for $|\omega| > 2\pi f_c$, and if $f_s \geq 2f_c$ (Shannon's rate), then there is no overlap between successive replications. We have lost no information in the sampling process; if anything, we have picked up a lot of "superfluous information" at frequencies outside the range of interest.

To recover the original signal, we must remove the replication images. First $F(\omega)$ can be obtained from $F^*(\omega)$ by multiplying it by the characteristic function

$$\chi_{\omega_s/2}(\omega) = \begin{cases} 1 & \text{if } |\omega| \leq \omega_s/2, \\ 0 & \text{if } |\omega| > \omega_s/2. \end{cases}$$

This is done by an analog filter known as an interpolation filter or a low-pass filter. We are now back to $F(\omega)$, and the time function can be recovered via the inverse Fourier transform:

$$f(t) = \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega.$$

In essence engineers view the Shannon sampling theorem in terms of an impulse train and a low-pass filter in the following way [30, 32]:

$$f(t) \xrightarrow{\text{sampler}} f^*(t) \xrightarrow{\text{low pass filter}} f(t).$$

First the signal $f \in B_\pi$ is sampled at the integers to convert $f(t)$ into the impulse train

$$f^*(t) = \sum_{n=-\infty}^{\infty} f(n)\delta(t-n),$$

where $\delta(t-n)$ is the Dirac delta function (impulse) at $t = n$. This is still expressed as an analog signal. This then is transmitted through an ideal low-pass filter which passes all frequencies of absolute value less than $\pi$ and blocks all others. This converts $f^*(t)$ back into $f(t)$.

We denote the sampling map by $S : f \longmapsto f^*$ and the low-pass filter map by $P : f^* \longmapsto f$

$$B_\pi \xrightarrow{S} \boxed{\text{unknown space}} \xrightarrow{P} B_\pi.$$

But the above procedure has some mathematical difficulties which are not resolved in the engineering approach:

- The sampler $S$ takes $f$ into $f^*$, which is out of the space of bandlimited functions; indeed, $f^*$ is not a signal with finite energy.
- In what sense does the impulse train series converge? One may prove that the series converges in the sense of tempered distributions to a generalized function.
- The map $P$ recovers $f$ at least formally since

$$P(\delta(t-n)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\omega t} e^{i\omega n} d\omega = \frac{\sin \pi(t-n)}{\pi(t-n)}.$$

If $P$ is continuous (in some topology), then

$$(Pf^*)(t) = \sum_{n=-\infty}^{\infty} f(n) P(\delta(t-n)) = \sum_{n=-\infty}^{\infty} f(n) \frac{\sin \pi(t-n)}{\pi(t-n)}.$$

However, since $\mathscr{S}'$ (the space of all tempered distributions, see the next section) is not a Hilbert space, we do not know if $P$ is a continuous operator.
- Still another difficulty! $P$ is not even well defined on $\mathscr{S}'$. Indeed,

$$Pg = \mathscr{F}^{-1}(\chi_{[-\pi,\pi]}\hat{g}), \ g \in \mathscr{S}',$$

where $\mathscr{F}^{-1}$ is the inverse Fourier transform of a tempered distribution. So $P$ corresponds under the Fourier transform to the multiplication of the Fourier transform $\hat{g}$ in $\mathscr{S}'$ by the characteristic function of $[-\pi, \pi]$. Unfortunately the characteristic function is not a multiplier in $\mathscr{S}'$. Hence we need to restrict ourselves to a subspace of $\mathscr{S}'$ in which $\chi_{[-\pi,\pi]}$ is a multiplier.

These issues have been resolved by Nashed and Walter [28]. They obtained a rigorous proof of the engineering approach, by considering the sampling map $S$ as an operator from $B_\pi$ to $H^{-1}$ and the filtering map $P$ as an operator (actually an orthogonal projection) from $H^{-1}$ to $B_\pi$:

$$B_\pi \overset{S}{\longmapsto} H^{-1} \overset{P}{\longmapsto} B_\pi,$$

where $H^{-1}$ is a Sobolev space, see the next section for the definition of Sobolev spaces. More importantly, by emulating and extending this proof, they obtained a general unifying approach for sampling theorems in reproducing kernel Hilbert spaces (RKHS) that include many earlier sampling theorems.

## 4.5   Function Spaces: Tempered Distributions, Sobolev Spaces and Reproducing Kernel Hilbert Spaces

### 4.5.1   The Space of Tempered Distributions

Let $\mathscr{S}$ be the space of all rapidly decreasing $C^\infty$ functions on the real line **R**, i.e., functions that satisfy

$$|g^{(k)}(t)| \leq C_{p,k}(1+|t|)^{-p}, \ t \in \mathbf{R},$$

for all $p, k = 0, 1, 2, \ldots$. Convergence on $\mathscr{S}$ may be defined by endowing $\mathscr{S}$ with the seminorms:

$$\mu_{p,k} := \sup_{t \in \mathbf{R}} (1+|t|)^p |g^{(k)}(t)|.$$

Then $g_n \to g$ in $\mathscr{S}$ whenever

$$(1+|t|)^p \big(g_n^{(k)}(t) - g^{(k)}(t)\big) \to 0$$

uniformly in $t \in \mathbf{R}$ for each $p$ and $k \geq 0$ as $n \to \infty$. The set $\mathscr{S}$ is dense in $L^2 := L^2(\mathbf{R})$ (the space of all square-integrable functions on the real line). We observe that compactly supported $C^\infty$ functions are contained in $\mathscr{S}$, and that the space $\mathscr{S}$ is complete with respect to the convergence of semi-norms $\mu_{p,k}, p, k \geq 0$.

A tempered distribution is an element in the dual space $\mathscr{S}'$ of $\mathscr{S}$, i.e., $\mathscr{S}'$ consists of all continuous linear functionals on $\mathscr{S}$. The definition of Fourier transform

$$\mathscr{F}f(\omega) := \int_{-\infty}^{\infty} f(t)\mathrm{e}^{-it\omega}\mathrm{d}t, \quad f \in \mathscr{S}$$

may be extended from $\mathscr{S}$ to $\mathscr{S}'$. The following examples of Fourier transforms on $\mathscr{S}'$ are needed in the derivation of a rigorous setting for the proof of the engineering approach:

$$\mathscr{F}\big(\delta(t-\alpha)\big) = e^{-i\omega\alpha},$$

$$\mathscr{F}\left(\sum_{n\in\mathbf{Z}} \delta(t-2\pi n)\right) = \sum_{n\in\mathbf{Z}} \delta(\omega-n),$$

$$\mathscr{F}\left(\sum_{n\in\mathbf{Z}} a_n e^{int}\right) = 2\pi \sum_{n\in\mathbf{Z}} a_n\delta(\omega-n),$$

and

$$\mathscr{F}\left(\sum_{n\in\mathbf{Z}} a_n e^{int} \chi_{[-\pi,\pi]}(t)\right) = \sum_{n\in\mathbf{Z}} a_n \frac{\sin\pi(\omega-n)}{\pi(\omega-n)},$$

where $\{a_n\}_{n\in\mathbf{Z}} \in \ell^2$.

### 4.5.2  Sobolev Spaces

An important Hilbert space structure on certain subsets of $\mathscr{S}'$ is provided by a class of Sobolev spaces. For $r \in \mathbf{R}$, the Sobolev space $H^r$ consists of all tempered distributions $f \in \mathscr{S}'$ such that

$$\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2(\omega^2+1)^r d\omega < \infty.$$

The inner product of $f$ and $g$ in $H^r$ is defined by

$$\langle f,g\rangle_r := \int_{-\infty}^{\infty} \hat{f}(\omega)\overline{\hat{g}(\omega)}(\omega^2+1)^r d\omega,$$

where $\hat{f} := \mathscr{F}f$ is the Fourier transform of $f$. The space $H^r$ is complete with respect to this inner product. For $r = 0$, $H^0$ is just $L^2(\mathbf{R})$ by Parseval identity. For $r = 1, 2, \ldots$, $H^r$ is the usual Sobolev space of functions that are $(r-1)$-times differentiable and whose $r$th derivative is in $L^2(\mathbf{R})$. For $r = -1, -2, \ldots$, $H^r$ contains all tempered distributions with point support of order $r$. Thus the Dirac delta $\delta \in H^{-1}$, and $\delta'$, the distributional derivative of $\delta$, belongs to $H^{-2}$.

### 4.5.3  Reproducing Kernel Hilbert Spaces

A Hilbert space $H$ of complex-valued functions on a set $\Omega$ is called a RKHS if all the evaluation functionals $H \ni f \longmapsto f(t) \in \mathbf{C}$ are continuous (bounded) for each fixed $t \in \Omega$; i.e., there exist a positive constant $C_t$ for each $t \in \Omega$ such that $|f(t)| \leq C_t\|f\|$ for all $f \in H$.

By Riesz representation theorem, for each $t \in \Omega$ there exists a unique element $k_t$ such that $f(t) = \langle f, k_t \rangle$ for all $f \in H$. The *reproducing kernel* $k(\cdot, \cdot) : \Omega \times \Omega \longmapsto \mathbf{C}$ of a RKHS $H$ is defined by $k(s,t) = \langle k_s, k_t \rangle, s, t \in \Omega$.

We summarize some basic properties of RKHS that are particularly relevant to signal processing, wavelet analysis, and approximation theory:

- $k(s,t) = \overline{k(t,s)}$ for all $t, s \in \Omega$.
- $k(s,s) \geq 0$ for all $s \in \Omega$. Furthermore, if $k(t_0, t_0) = 0$ for some $t_0 \in \Omega$, then $f(t_0) = 0$ for all $f \in H$.
- $|k(s,t)| \leq \sqrt{k(s,s)} \sqrt{k(t,t)}$ for all $s,t \in \Omega$.
- The reproducing kernel $k(s,t)$ on $\Omega \times \Omega$ is a nonnegative definite symmetric kernel. Conversely by Aronszajn–Moore theorem, every nonnegative definite symmetric function $k(\cdot, \cdot)$ on $\Omega \times \Omega$ determines a unique Hilbert space $H_k$ for which $k(\cdot, \cdot)$ is a reproducing kernel [5]. Here a complex-valued function $F$ on $\Omega \times \Omega$ is said to be *positive definite* if for any $n$ points $t_1, \ldots, t_n \in \Omega$, the matrix $A := (F(t_i, t_j))_{1 \leq i,j \leq n}$ is nonnegative definite, i.e., $u^* A u = \sum_{i,j=1}^{n} \overline{u_i} F(t_i, t_j) u_j \geq 0$ for all $u = (u_1, \ldots, u_n) \in \mathbf{C}^n$.
- A closed subspace $\tilde{H}$ of a RKHS $H$ is also a RKHS. Moreover, the orthogonal projector $P$ of $H$ onto $\tilde{H}$ and the reproducing kernel $\tilde{k}(s,t)$ of the RKHS $\tilde{H}$ are related by $Pf(s) = \langle f, \tilde{k}_s \rangle, s \in \Omega$ for all $f \in H$ where $\tilde{k}_s = Pk_s$.
- If a RKHS space $H$ with kernel $k(\cdot, \cdot)$ has direct orthogonal decomposition $H = H_1 \oplus H_2$ for some complementary orthogonal closed subspaces $H_1$ and $H_2$, then $k = k_1 + k_2$, where $k_1, k_2$ are reproducing kernels of the reproducing kernel Hilbert spaces $H_1$ and $H_2$, respectively.
- In a RKHS, the element representing a given bounded linear functional $\phi$ can be expressed by means of the reproducing kernel: $\phi(f) = \langle f, h \rangle$, where $h = \phi(k)$. Similarly for a bounded linear operator $L$ on $H$ to $H$, we have that $Lf(t) = \langle Lf, h \rangle = \langle f, L^*h \rangle$.
- Every finite-dimensional function space is a RKHS $H$ with reproducing kernel $k(s,t) = \sum_{i=1}^{n} u_i(s) \overline{u_i(t)}$, where $\{u_i\}_{i=1}^{n}$ is an orthonormal basis for $H$. (Notice that the sum in the above definition of the kernel $k$ is invariant under the choice of orthonormal basis).
- The space $W^1$, that contains all functions $f \in L^2[0,1]$ such that $f$ is absolutely continuous and $f'$, which exists almost everywhere, is in $L^2[0,1]$, and $f(0) = 0$, is a RKHS with kernel $k(s,t) = \min(s,t)$ under the inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)\mathrm{d}t$.
- Sobolev space $H^s, s > 1/2$, is a reproducing kernel Hilbert space.
- Let $H$ be a separable RKHS, then its reproducing kernel $k(\cdot, \cdot)$ has the expansion:

$$k(s,t) = \sum_{n=1}^{\infty} \varphi_n(t) \overline{\varphi_n(s)},$$

where $\{\phi_n\}_{n=1}^{\infty}$ is an orthonormal basis for $H$. We remark that for a general separable Hilbert space $H$, $\sum_{n=1}^{\infty} \varphi_n(t) \overline{\varphi_n(s)}$ is not a reproducing kernel and also

that $\phi_n$'s do not generally corresponds to sampling expansions. If they do, i.e.,if $\varphi_n(t) = k(t_n, t)$ for some sequence $\{t_n\}$, then we have that $f(t) = \sum_{n=1}^{\infty} f(t_n)\varphi_n(t)$, this constitutes a sampling theorem.

- If the reproducing kernel $k(s,t)$ of a RKHS $H$ is continuous on $\Omega \times \Omega$, then $H$ is a space of continuous functions (uniformly continuous on a bounded $\Omega$). This follows from

$$|f(t) - f(s)| = |\langle f, k_t - k_s \rangle| \leq \|f\|\|k_t - k_s\|$$

and $\|k_t - k_s\|^2 = k(t,t) - 2k(t,s) + k(s,s)$ for all $s,t \in \Omega$.

- Strong convergence in a RKHS $H$ implies pointwise convergence and uniform convergence on compact sets, because

$$|f(t) - f_n(t)| = |\langle f - f_n, k_t \rangle| \leq \|f - f_n\|\sqrt{k(t,t)}.$$

- $L_2[a,b]$, the space of all square-integrable functions on the interval $[a,b]$, is not a RKHS. Indeed, point evaluation is not well defined. Each function $f \in L^2[a,b]$ is actually an equivalence class of functions equal to each other almost everywhere. Thus the "value" at a point has no meaning since any point has measure zero.

## 4.6   Rigorous Justification of the Engineering Approach to Sampling

This section involves a search for function spaces in which the mathematical difficulties described in Sect. 4.3 are resolved. Clearly we want to work with a *subspace H* of the space $\mathscr{S}'$ of tempered distributions. We require that $\delta \in H$ and the convolution signal of the impulse train must converge in $H$ under a mild condition on $\{f(t_n)\}_{n=-\infty}^{\infty}$. As remarked in Sect. 4.4, the characteristic function is not a multiplier in the space $\mathscr{S}'$, and hence the space $H$ must also have the property that the characteristic function is a multiplier in $H$. Finally, the sampling map $S$ and the low-pass filter map $P$ must be well defined on the appropriate spaces, and their composition $PS$ is the identity map:

- The characteristic functions are multipliers in the space of Fourier transforms of the elements in $H^r(\mathbf{R})$, which can be identified with the image under the Fourier transform of $L_r^2(\mathbf{R})$, the space of square-integrable functions with respect to the measure $d\mu(x) = (1 + |x|^2)^r dx$. So we may consider $H^r$ with $r > -1/2$; specifically we take $H^{-1}$, since $\delta \in H^{-1}$.
- We consider the sampling map $Sf = f^*$ as a map onto $H^{-1}$. Then the partial sums of the sampled impulse train $\sum_{n=-N}^{N} f(n)\delta(t-n)$ belong to $H^{-1}$ and converges in the norm of $H^{-1}$ to $f^*$. This proposition does not require the signal to be bandlimited, but the next result does.

- When we consider projector onto the space $H^{-1}$, $B_\pi \subset L_2 \subset H^{-1}$ and $B_\pi$ is closed in the topology of $H^{-1}$. Hence we can define the orthogonal projection of $H^{-1}$ onto $B_\pi$. The reproducing kernel of $B_\pi$ enables us to compute this projection easily. In fact,

$$P\delta_a(t) = k(t,a), \ t \in \mathbf{R}$$

and

$$Pf^* = \lim_{N \to \infty} \sum_{n=-N}^{N} f(n)k(\cdot, n)$$

in the norm of $H^{-1}$ [28]. The above series converges to $Pf^*(t)$ by the continuity of the orthogonal projector and to $f(t)$ by the sampling theorem. Thus $Pf^* = f$.

The above ideas provide a mathematical proof of the engineering arguments, but they also suggest important extensions to other RKHS, as discussed in [28].

## 4.7 Sampling in Reproducing Kernel Hilbert Spaces

At the outset, the expansions (4.2) and (4.3) mentioned in introduction section appear markedly different, or at least not seem to be related. The expansion (4.2) holds for any complete orthonormal sequence in a separable Hilbert space, and sampling points $\{t_n\}$ could have no meaning in this context. On the other hand, the sampling expansion does not require the sampling sequence $\{S_n(t)\}_{n=1}^{\infty}$ to be orthonormal.

The expansions (4.2) and (4.3) can indeed be related. In finite-dimensional spaces the expansions are based on different choices of orthonormal basis. One with respect to inner product of two continuous functions, and the other is based on discrete orthogonality or biorthogonality of the sequences.

The expansions (4.2) and (4.3) may also be related in some special cases of orthonormal sequences in certain infinite-dimensional spaces. Let $f$ be a signal defined on the real line. Suppose that there exists a reproducing kernel $k(t,s)$ such that for some real numbers $\{t_n\}_{n=1}^{\infty}$, the sequence $f_n(s) = k(t_n,s), n \geq 1$, is a complete orthonormal sequence, then $c_n = \langle f, f_n \rangle = f(t_n)$. The series expansion $\sum_{n=1}^{\infty} c_n f_n$ then becomes a sampling expansion, i.e., it states how to recover $f(t)$ from the sample values $\{f(t_n)\}_{n=1}^{\infty}$. For example, for $\pi$-bandlimited signals, $k(t,s) = \frac{\sin \pi(t-s)}{\pi(t-s)}$ and $S_n(t) = k(t,n)$ is an orthonormal basis for $B_\pi$, and the expansion result mentioned above reduces to the Whittaker–Shannon–Kotel'nikov sampling expansion [42].

In [28], the authors introduced an approach that provided general sampling theorem for functions in RKHS. The sampling theorems are for functions in a general RKHS with reproducing kernel $k(t,s)$, which is closed in the Sobolev space $H^{-p}, p > 1/2$. The sampling functions $S_n, n \geq 1$, need not be an orthogonal system, and the theory allows nonorthogonal sampling sequences. Then the system

$\{S_n := k(t_n, \cdot)\}_{n=1}^{\infty}$ has to satisfy a biorthogonality condition, i.e., $S_n(t_m) = \delta_{mn}$ for all $m, n \geq 1$, and the sampling points must satisfy a density-type condition. This general setup includes sampling theorems related to other transforms than the Fourier transform as in the classical theory. For example, Sturm–Liouville, Jacobi, and Laguerre transforms are among the examples discussed in [28], as well as sampling using frames. Also for the orthogonal case, several error analyses, such as truncation, aliasing, jittering, and amplitude error, are discussed in details.

Now we state a representative sampling theorem for signals in a reproducing kernel Hilbert space [28].

**Theorem 2.** *Let $H \subset L^2(\mathbf{R})$ be a reproducing kernel Hilbert space that is closed in the Sobolev space $H^{-1}$ and under differentiation. Assume that its reproducing kernel $k(s,t)$ is continuous and has the zero sequence $\{t_n\}$ which is a set of uniqueness for $H$, and assume that $\{t_n\}$ tends to infinity as $n$ tends to infinity. If $f \in H$ satisfies $f(t)/k(t,t) = O(t^{-2})$, then the sampled sequence*

$$f^*(t) = \sum_n \frac{f(t_n)}{k(t_n, t_n)} \delta(t - t_n)$$

*converges in the sense of $H^{-1}$ and its orthogonal projection onto $H$ equals to $f(t)$, and the series*

$$f(t) = \sum_n \frac{f(t_n)}{k(t_n, t_n)} k(t_n, t)$$

*converges uniformly on sets for which $k(t,t)$ is bounded.*

## 4.8   Sampling in Shift-Invariant Spaces

A *shift-invariant space* generated by a square-integrable function $\phi$ is given by

$$V_2(\phi) := \left\{ \sum_{n \in \mathbf{Z}} c_n \phi(\cdot - n) : \sum_{n \in \mathbf{Z}} |c_n|^2 < \infty \right\}.$$

Shift-invariant spaces have been shown to be realistic for modeling signals with smoother spectrum, and also suitable for taking into account real acquisition and reconstruction devices. The notion of shift-invariant spaces arises in approximation theory, wavelet theory, and sampling theory.

For the generator $\phi$ of a shift-invariant space, we usually assume that $\{\phi(\cdot - n) : n \in \mathbf{Z}\}$ consisting of all integer shifts of the generator $\phi$ is a Riesz basis for $V_2(\phi)$; i.e., there exist positive constants $A$ and $B$ such that

$$A \sum_{n \in \mathbf{Z}} |c_n|^2 \leq \left\| \sum_{n \in \mathbf{Z}} c_n \phi(\cdot - n) \right\|_2^2 \leq B \sum_{n \in \mathbf{Z}} |c_n|^2.$$

The Paley–Wiener space $B_\pi$ is a shift-invariant space generated by the sinc function $\mathrm{sinc}(t)$, and the integer shifts of the sinc function form an orthonormal basis for the Paley–Wiener space $B_\pi$. The following is a representative sampling theorem for signals in a shift-invariant space $V_2(\phi)$ established in [40].

**Theorem 3.** *Let $\phi$ be a real continuous function such that $\sup_{t\in\mathbf{R}}|\phi(t)|(1+|t|)^{1+\varepsilon} < \infty$ for some $\varepsilon > 0$, $\hat{\phi}^*(\omega) := \sum_{n\in\mathbf{Z}}\phi(n)e^{-in\omega} \neq 0$ for all $\omega \in \mathbf{R}$, and $\{\phi(t-n): n\in\mathbf{Z}\}$ is an orthonormal basis for $V_2(\phi)$. Then any signal $f \in V_2(\phi)$ can be stably reconstructed from its samples $\{f(n)\}_{n\in\mathbf{Z}}$ on the integer lattice. Moreover,*

$$f(t) = \sum_{n\in\mathbf{Z}} f(n)\tilde{\phi}(t-n) \quad \text{for all } f \in V_2(\phi),$$

*where $\tilde{\phi} \in V_2(\phi)$ is defined by $\hat{\tilde{\phi}}(\omega) = \hat{\phi}(\omega)/\hat{\phi}^*(\omega), \omega \in \mathbf{R}$.*

The reader may refer to [1, 4, 35, 37, 41] and references therein for some fundamental issues to sampling theory in shift-invariant spaces.

## 4.9 Sampling in Unitarily Translation-Invariant Hilbert Spaces

In this section, we consider sampling theorems on a unitarily translation-invariant RKHS generated from a single function. To be more specific, the RKHS $H_\phi$ has the reproducing kernel

$$k_\phi(t,s) = \int_{-\infty}^{\infty} \phi(u-t)\phi(u-s)\mathrm{d}u \qquad (4.10)$$

generated by a function $\phi \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$, whose Fourier transform does not have real zeros. Examples of such a generating function $\phi$ includes $(\sigma^2+t^2)^{-1}$, $e^{-\sigma^2 t^2}$, and $e^{-\sigma|t|}$ where $\sigma > 0$. The RKHS $H_\phi$ with reproducing kernel $k_\phi$ in (4.10) is given by

$$H_\phi = \{f : \|f\|_{H_\phi} < \infty\}, \qquad (4.11)$$

where

$$\|f\|_{H_\phi} := \left(\frac{1}{2\pi}\int_{\mathbf{R}}|\hat{f}(\omega)|^2/|\hat{\phi}(\omega)|^2\mathrm{d}\omega\right)^{1/2}.$$

**Theorem 4 ([38]).** *Let $\phi$ be an integrable function on the real line such that its Fourier transform $\hat{\phi}$ does not have real zeros and $\int_{\mathbf{R}}|\phi(t)|^2(1+t^2)^\alpha \mathrm{d}t < \infty$ for some $\alpha > 1$, $k_\phi$ be the reproducing kernel in (4.10), and $\cdots < \lambda_{-2} < \lambda_{-1} < \lambda_0 = 0 < \lambda_1 < \lambda_2 < \cdots$ be sampling points with $\lambda_{j+1} - \lambda_j \geq \varepsilon > 0$ for all $j \in \mathbf{Z}$. Denote by $\mathcal{X}$ the closed subspace of the reproducing kernel Hilbert space $H_\phi$ in (4.11) spanned by $k_\phi(\cdot,t_j), j \in \mathbf{Z}$. Then the sampling operator*

$$\mathcal{X} \ni f \longmapsto (f(t_j))_{j\in\mathbf{Z}}$$

*is stable in the sense that there exist positive constants A and B such that*

$$A\|f\|_{H_\phi} \le \|(f(t_j))_{j\in\mathbf{Z}}\|_2 \le B\|f\|_{H_\phi} \quad \text{for all } f \in \mathcal{X}.$$

*Moreover, the sampling expansion*

$$f(t) = \sum_{j\in\mathbf{Z}} f(t_j)\frac{k_\phi(t,t_j)}{k_\phi(0,0)}$$

*is valid for all $f \in \mathcal{X}$.*

## 4.10   Sampling Signals with Finite Rate of Innovation

Signals with *finite rate of innovation* are those signals that can be determined by finitely many samples per unit of time [39]. The concept of signals with finite rate of innovation was introduced and studied by Martin Vetterli and his school. Prototype examples of signals with finite rate of innovation include delta pulses, narrow pulses in ultrawide band communication, mass spectrometry data in medical diagnosis, and splines with (non-)uniform knots. They also include bandlimited signals and time signals in shift-invariant spaces, which are discussed in the previous sections.

A common feature of signals with finite rate of innovation is that they have a parametric representation with a finite number of parameters per unit time. So we may model a signal $f$ with finite rate of innovation as a superposition of impulse response of varying positions, amplitudes, and widths [34], i.e.,

$$f(t) = \sum_{\lambda\in\Lambda} c_\lambda \phi_\lambda(t-\lambda), \tag{4.12}$$

where each $\lambda \in \Lambda$ represents the innovative position of the signal, $\phi_\lambda$ is the impulse response of the signal-generating device at the innovative position $\lambda$, and $c_\lambda$ is the amplitude of the signal at the innovation position $\lambda$. Thus the function space

$$V_p(\Phi) := \left\{ \sum_{\lambda\in\Lambda} c(\lambda)\phi_\lambda(\cdot-\lambda) : (c(\lambda))_{\lambda\in\Lambda} \in \ell^p(\Lambda) \right\}, 1 \le p \le \infty, \tag{4.13}$$

could be suitable for modeling signals with finite rate of innovation.

Sampling theory for signals with finite rate of innovation has been demonstrated to be important for accurate time estimation of ultraband communication, registration of multiview images, pattern recognition, quantification of spectra, etc. The following is a sampling theorem for signals with finite rate of innovation when the innovative position of the signal and the impulse response of the signal-generating device at the innovative position are given [8,33]:

**Theorem 5.** *Let $\Lambda, \Gamma$ be relatively separated subsets of $\mathbf{R}$, $\Phi = \{\phi_\lambda : \lambda \in \Lambda\}$ be a family of continuous functions on $\mathbf{R}$ such that $\sup_{t \in \mathbf{R}, \lambda \in \Lambda} |\phi_\lambda(t)|(1+|t|)^\alpha < \infty$ for some $\alpha > 1$, and the space $V_2(\Phi)$ be as in (4.13). Assume that $\Phi$ is a Riesz basis of $V_2(\Phi)$, and that $\Gamma$ is a stable ideal sampling set for $V_2(\Phi)$, i.e., there exist positive constants $A$ and $B$ such that*

$$A\|f\|_2 \leq \|(f(\gamma))_{\gamma \in \Gamma}\|_2 \leq B\|f\|_2 \quad \text{for all } f \in V_2(\Phi).$$

*Then there exists a displayer $\tilde{\Psi} = \{\tilde{\psi}_\gamma : \gamma \in \Gamma\}$ such that*

$$\sup_{t \in \mathbf{R}, \gamma \in \Gamma} |\psi_\gamma(t)|(1+|t|)^\alpha < \infty$$

*and*

$$f(t) = \sum_{\gamma \in \Gamma} f(\gamma)\tilde{\psi}_\gamma(t-\gamma) \quad \text{for all } f \in V_2(\Phi).$$

Now we consider nonlinear and highly challenging problem of how to identify innovative positions and amplitudes of a signal with finite rate of innovation. For the stability for identification, the innovation positions should be separated from each other and the amplitudes at innovation positions should be above a certain level. So we may model those signals with finite rate of innovation as superposition of impulse response of active and nonactive generating devices located at unknown neighbors of a uniform grid. In [36] we assume, after appropriate scaling, that signals live in a perturbed shift-invariant space

$$V_{\infty,\oslash} := \left\{ \sum_{n \in \mathbf{Z}} c_n \varphi(\cdot - n - \sigma_n) \ : \ (c_k)_{k \in \mathbf{Z}} \in \ell^\infty_\oslash(\mathbf{Z}) \right\}$$

with unknown perturbation $\sigma := (\sigma_k)_{k \in \mathbf{Z}}$, where

$$\ell^\infty_\oslash(\mathbf{Z}) = \left\{ c := (c_k)_{k \in \mathbf{Z}} : \ \|c\|_{\ell^\infty_\oslash} := \sup_{c_k \neq 0} |c_k| + |c_k|^{-1} < \infty \right\}.$$

A negative result for sampling in the perturbed shift-invariant space $V_{\infty,\oslash}$ is that not all signals in such a space can be recovered from their samples if $\varphi$ satisfies the popular Strang-Fix condition. The reason is that one cannot determine the jitter $\sigma_0$ of the signal $\sum_{k \in \mathbf{Z}} \varphi(\cdot - k - \sigma_0), \sigma_0 \in \mathbf{R}$, as it has constant amplitudes and is identical for all $\sigma_0 \in \mathbf{R}$. On the positive side, it is shown in [36] that any signal $h$ in a perturbed shift-invariant space with unknown (but small) jitters can be recovered exactly from its average samples $\langle h, \psi_m(\cdot - k)\rangle, 1 \leq m \leq M, k \in \mathbf{Z}$, provided that the generator $\varphi$ of the perturbed shift-invariant space and the average samplers $\psi_m, 1 \leq m \leq M$, satisfy the following condition:

$$\text{rank}\left( \begin{matrix} [\widehat{\nabla\varphi}, \widehat{\psi_1}](\xi) & \cdots & [\widehat{\nabla\varphi}, \widehat{\psi_M}](\xi) \\ [\hat{\varphi}, \widehat{\psi_1}](\xi) & \cdots & [\hat{\varphi}, \widehat{\psi_M}](\xi) \end{matrix} \right) = 2 \quad \text{for all } \xi \in [-\pi, \pi]. \tag{4.14}$$

Here the *bracket product* $[f,g]$ of two square-integrable functions $f$ and $g$ is given by $[f,g](\xi) = \sum_{l \in \mathbf{Z}} f(\xi + 2l\pi)\overline{g(\xi + 2l\pi)}$.

**Theorem 6.** *Let $\varphi$ and $\psi_1, \ldots, \psi_M$ satisfy (4.14) and have the following regularity and decay properties:*

$$\sup_{t \in \mathbf{R}} \left( |\varphi(t)| + |\varphi'(t)| + |\varphi''(t)| + \sum_{m=1}^{M} |\psi_m(t)| \right) (1 + |t|)^{\alpha} < \infty \qquad (4.15)$$

*for some $\alpha > 1$. Then for any $L \geq 1$, there exists a positive number $\delta_1 \in (0, 1/2)$ such that any signal $h(t) = \sum_{k \in \mathbf{Z}} c_k \varphi(t - k - \sigma_k)$ in the space $V_{\infty, \oslash}$ with $\|(c_k)_{k \in \mathbf{Z}}\|_{\ell^{\infty}_{\oslash}} \leq L$ and $\|(\sigma_k)_{k \in \mathbf{Z}}\|_{\infty} \leq \delta_1$ could be reconstructed from its average sample data $\langle h, \psi_m(\cdot - k) \rangle, 1 \leq m \leq M, k \in \mathbf{Z}$, in a stable way.*

## 4.11 Sampling in Reproducing Kernel Banach Subspaces of $L^p$

Let $1 \leq p \leq \infty$. A bounded linear operator $T$ on $L^p(\mathbf{R})$ is said to be an *idempotent operator* if $T^2 = T$. Denote the range space of the idempotent operator $T$ on $L^p(\mathbf{R})$ by $V_p$; i.e.,

$$V_p := \{Tf : f \in L^p(\mathbf{R})\}. \qquad (4.16)$$

The Paley–Wiener space, finitely generated shift-invariant spaces, $p$-integrable spline spaces, spaces modeling signals with finite rate of innovation, and $L^p$ itself are the range space of some idempotent operators.

Denote the *Wiener amalgam space* by

$$W^1 := \left\{ f \in L^1(\mathbf{R}) : \|f\|_{W^1} := \left\| \sup_{-1/2 \leq z < 1/2} |f(\cdot + z)| \right\|_1 < \infty \right\}$$

and the *modulus of continuity* of a kernel function $K$ on $\mathbf{R} \times \mathbf{R}$ by

$$\omega_{\delta}(K)(s,t) := \sup_{-\delta \leq z_1, z_2 \leq \delta} |K(s + z_1, t + z_2) - K(x,y)|.$$

A *sampling set* $\Gamma$ in this chapter means a relatively separated discrete subset of $\mathbf{R}$; i.e.,

$$B_{\Gamma}(\delta) := \sup_{t \in \mathbf{R}} \sum_{\gamma \in \Gamma} \chi_{[-\delta, \delta]}(t - \gamma) < \infty \qquad (4.17)$$

for some $\delta > 0$, where $\chi_E$ is the characteristic function on a set $E$. A sampling set $\Gamma$ is said to have *gap* $\delta > 0$ if

$$A_{\Gamma}(\delta) := \inf_{t \in \mathbf{R}} \sum_{\gamma \in \Gamma} \chi_{[-\delta, \delta]}(t - \gamma) \geq 1 \qquad (4.18)$$

[1, 3, 4]. If we assume that the idempotent operator $T$ is an integral operator

$$Tf(s) = \int_{\mathbf{R}} K(s,t)f(t)\mathrm{d}t, \ \ f \in L^p(\mathbf{R}), \tag{4.19}$$

whose measurable kernel $K$ has certain off-diagonal decay and regularity,

$$\left\| \sup_{z \in \mathbf{R}} |K(\cdot + z, z)| \right\|_{W^1} < \infty \tag{4.20}$$

and

$$\lim_{\delta \to 0} \left\| \sup_{z \in \mathbf{R}} |\omega_\delta(K)(\cdot + z, z)| \right\|_{W^1} = 0, \tag{4.21}$$

then,

- $V_p$ is a reproducing kernel subspace of $L^p(\mathbf{R})$; i.e., for any $t \in \mathbf{R}$ there exists a positive constant $C_t$ such that

$$|f(t)| \leq C_t \|f\|_{L^p(\mathbf{R})} \quad \text{for all } f \in V_p. \tag{4.22}$$

- The kernel $K$ satisfies the "reproducing kernel property":

$$\int_{\mathbf{R}} K(s,z)K(z,t)\mathrm{d}z = K(s,t) \quad \text{for all } s,t \in \mathbf{R}. \tag{4.23}$$

- $K(\cdot,t) \in V$ for any $t \in \mathbf{R}$.
- $V_p := \{\sum_{\lambda \in \Lambda} c(\lambda)\phi_\lambda(t - \lambda) : (c(\lambda))_{\lambda \in \Lambda} \in \ell^p(\Lambda)\}$, where $\Lambda$ is a relative separated discrete subset of $\mathbf{R}$ and $\Phi = \{\phi_\lambda\}_{\lambda \in \Lambda} \subset V_p$ is localized in the sense that there exists a function $h$ in the Wiener amalgam space $W^1$ such that $\phi_\lambda$ is dominated by $h$ for every $\lambda \in \Lambda$, i.e.,

$$|\phi_\lambda(t)| \leq h(t) \quad \text{for all } \lambda \in \Lambda \text{ and } t \in \mathbf{R}. \tag{4.24}$$

- Signals in $V_p$ have finite rate of innovation.
- For $p = 2$, an idempotent operator $T$ with kernel $K$ satisfying symmetric condition $K(x,y) = \overline{K(y,x)}$ is a projection operator onto a closed subspace of $L^2$. In this case, the idempotent operator $T$ and its kernel $K$ is uniquely determined by its range space $V_2$ onto $L^2$.

The following sampling problem in the reproducing kernel space $V_p$ is established in [25].

**Theorem 7.** *Let $1 \leq p \leq \infty$, $T$ be an idempotent integral operator whose kernel $K$ satisfies* (4.20) *and* (4.21)*, $V$ be the reproducing kernel subspace of $L^p(\mathbf{R})$ associated with the operator $T$, and $\delta_0 > 0$ be so chosen that*

$$r_0 := \left\| \sup_{z \in \mathbf{R}} |\omega_{\delta_0/2}(K)(\cdot + z, z)| \right\|_{L^1(\mathbf{R})} < 1. \tag{4.25}$$

*Then any signal f in V can be reconstructed in a stable way from its samples* $f(\gamma), \gamma \in \Gamma$, *taken on a relatively separated subset* $\Gamma$ *of* **R** *with gap* $\delta_0$.

Similar conclusion to the one in the above sampling theorem has been established in [12, Sect. 7.5] when the kernel $K$ of the idempotent operator $T$ satisfies the symmetric condition $K(x,y) = \overline{K(y,x)}$.

## 4.12   Convolution Sampling in Reproducing Kernel Banach Subspaces of $L^p$

In this section, we consider convolution sampling for signals in certain reproducing kernel subspaces of $L^p, 1 \le p \le \infty$. Here *convolution sampling* of a signal is ideal sampling of the convoluted signal taken on a sampling set. Precisely, given an integrable convolutor $\psi$ and a sampling set $\Gamma$, the convolution sampling of a signal $f$ includes two steps: Convoluting $\psi$ with the signal $f$,

$$\psi * f(t) := \int_{-\infty}^{\infty} f(s)\psi(t-s)\mathrm{d}s,$$

and then sampling the convoluted signal $\psi * f$ at the sampling set $\Gamma$,

$$f \overset{\text{convoluting}}{\longmapsto} f * \psi \overset{\text{sampling}}{\longmapsto} \{(f * \psi)(\gamma)\}_{\gamma \in \Gamma}.$$

The data obtained by the above convolution sampling procedure is given by $\{(f * \psi)(\gamma)\}_{\gamma \in \Gamma}$. In [27], it is shown that any signal in the reproducing kernel subspace $V_p$ associated with an idempotent operator can be stably reconstructed from its convolution samples taken on a sampling set with small gap if and only if the convolution procedure is stable on that space.

**Theorem 8.** *Let* $1 \le p \le \infty, \psi_1, \ldots, \psi_L$ *be integrable functions on the real line,* $V_p$ *be the reproducing kernel subspace of* $L^p$ *in* (4.16)*, and set* $\Psi = (\psi_1, \ldots, \psi_L)^T$. *Assume that the kernel* $K$ *of the idempotent operator* $T$ *associated with the reproducing kernel space* $V_p$ *satisfies* (4.20) *and* (4.21)*. Then the following two statements are equivalent:*

*(i)* $\Psi$ *is a stable convolutor on* $V_p$*; i.e.,*

$$0 < \inf_{g \in V_p, \|g\|_p = 1} \sum_{l=1}^{L} \|\psi_l * g\|_p \le \sup_{g \in V_p, \|g\|_p = 1} \sum_{l=1}^{L} \|\psi_l * g\|_p < \infty.$$

*(ii)* $\Psi$ *is a stable convolution sampler on* $V_p$ *for all sampling sets having sufficiently small gap; i.e., there exists* $\delta_0 > 0$ *such that*

$$0 < \inf_{0 \neq f \in V_p} \frac{\sum_{l=1}^{L} \left\| \left( \psi_l * f(\gamma) \right)_{\gamma \in \Gamma} \right\|_p}{\|f\|_p} \leq \sup_{0 \neq f \in V_p} \frac{\sum_{l=1}^{L} \left\| \left( \psi_l * f(\gamma) \right)_{\gamma \in \Gamma} \right\|_p}{\|f\|_p} < \infty$$

*holds for any sampling set* $\Gamma$ *satisfying* $1 \leq A_\Gamma(\delta) \leq B_\Gamma(\delta) < \infty$ *for some* $\delta \in (0, \delta_0)$.

The equivalence in Theorem 8 was considered in [4] under the assumption that the reproducing kernel space $V_p$ is a finitely-generated shift-invariant space.

## 4.13 Reproducing Kernel Hilbert Space Induced by Sampling Expansions

As indicated earlier, both the sampling map

$$f \longmapsto (f(t_n))_{n=1}^{\infty}$$

and the inverse map

$$(f(t_n))_{n=1}^{\infty} \longmapsto f$$

need to be continuous in a setting where sampling expansions are to be used. Thus the evaluation functional $E_t f := f(t)$ needs to be continuous for all $t$. Equivalently, the signal resides in a RKHS, even though this RKHS may not explicitly be identified. In [28], the authors have shown that, under very mild conditions, many versions of sampling theorems hold for RKHS. In [29], the authors asked whether an RKHS exists for each sampling theorem and showed that the answer is affirmative when a sampling sequence satisfies minimal properties. The starting point is an abstract notion of a sampling expansion.

**Definition:** Let $f$ be a function belonging to a class $\mathscr{F}$ of continuous functions on $\Omega \subset \mathbf{R}$. A sampling theorem is associated with $\mathscr{F}$ if there is a sequence of sampling pairs $\{(S_n, t_n)\}$ of functions $S_n \in \mathscr{F}$ and points $t_n \in \Omega$ such that

- $S_n(t_k) = \delta_{nk}$, where $\delta_{nk}$ is the Kronecker delta.
- For each $f \in \mathscr{F}$, the sequence $\{f(t_n)\} \in \ell^2$, i.e., $\sum_n |f(t_n)|^2 < \infty$.
- The set $\{t_n\}$ is a set of uniqueness for $\mathscr{F}$.
- For each $\{b_n\} \in \ell^2$ the series $\sum_n b_n S_n(t)$ converges pointwise in $\Omega$.

Then the authors construct a RKHS associated with sampling expansion as follows [29]:

**Theorem 9.** *Let $H_0$ be the Hilbert space consisting of $\mathscr{F}$ with the inner product $\langle f,g \rangle_0 = \sum_n f(t_n)\overline{g(t_n)}$. Then $H_0$ satisfies the following:*

- *$H_0$ is a reproducing kernel Hilbert space with $k(t,s) = \sum_n \overline{S_n(t)}S_n(s)$ as its reproducing kernel.*
- *$\{S_n\}$ is an orthogonal basis for $H_0$.*
- *The sampling expansion $g(t) = \sum_n g(t_n)S_n(t)$ holds for any $g \in H_0$.*

## 4.14   Sampling in Reproducing Kernel Banach Spaces

A *reproducing kernel Banach space* is a Banach space $B$ of functions on a set $\Omega$ such that the evaluation functions $f \to f(t)$ is continuous for each $t \in \Omega$ [5]. The range space $V_p$ of an idempotent integral operator is a reproducing kernel Banach space when the kernel of the idempotent operator satisfies certain regularity conditions. In this section, we investigate sampling in a reproducing kernel Banach space.

Let $1 \leq p \leq \infty$ and $B$ be a Banach space with norm denoted by $\|\cdot\|_B$. We say that a countable subset $\Lambda$ of $\Omega$ is a *p-sampling set* for the Banach space $B$ if

$$0 < \inf_{f \in B, \|f\|_B = 1} \|(f(\lambda))_{\lambda \in \Lambda}\|_p \leq \sup_{f \in B, \|f\|_B = 1} \|(f(\lambda))_{\lambda \in \Lambda}\|_p < \infty, \qquad (4.26)$$

and a countable collection of elements $g_\lambda, \lambda \in \Lambda$, in the dual space of $B$ to be a *p-frame* if

$$0 < \inf_{\|f\|_B = 1} \|(g_\lambda(f))_{\lambda \in \Lambda}\|_p \leq \sup_{\|f\|_B = 1} \|(g_\lambda(f))_{\lambda \in \Lambda}\|_p < \infty, \qquad (4.27)$$

i.e., the analysis operator $T : B \ni f \longmapsto (g_\lambda(f))_{\lambda \in \Lambda} \in \ell^p$ is bounded from both above and below [2]. Similarly to sampling in a RKHS, for a reproducing kernel Banach space $B$ of functions on a set $\Omega$, a countable subset $\Lambda$ of $\Omega$ is a *p*-sampling set for the space $B$ if and only if the corresponding evaluation functionals $h_\lambda, \lambda \in \Lambda$, form a *p*-frame for the space $B$. Moreover, in [14] it is shown that a reconstruction formula always exists.

**Theorem 10.** *Let $1 \leq p, q \leq \infty$ satisfy $1/p + 1/q = 1$, $B$ be a reproducing kernel Banach space of functions on a set $\Omega$, and $\Lambda \subset \Omega$ be a p-sampling set. Then there exists a collection of functions $S_\lambda(t), \lambda \in \Lambda$, such that:*

- *$(S_\lambda(t))_{\lambda \in \Lambda}$ is q-summable for every $t \in \Omega$.*
- *$(\eta_\lambda)_{\lambda \in \Lambda}$ is a p-frame for the range space of the sampling operator $S : B \ni f \longmapsto (f(\lambda))_{\lambda \in \Lambda} \in \ell^p$, where $\eta_\lambda = (S_{\lambda'}(\lambda))_{\lambda' \in \Lambda}$.*

- *Every signal f in the reproducing kernel Banach space B has the following sampling expansion:*

$$f(t) = \sum_{\lambda \in \Lambda} f(\lambda) S_\lambda(t), \ t \in \Omega,$$

  *with the pointwise convergence.*

## 4.15   Average Sampling in $L^2$

In this section, we consider very general sampling procedure where the samples are obtained by inner products between time signal and sampling functionals. More precisely, given a time signal $f$ living in a Hilbert space $H$, its average sample $y_\gamma$ at the location $\gamma \in \Gamma$ is obtained by taking the inner product between the signal $f$ and the sampling functional $\psi_\gamma(\cdot - \gamma)$ at the location $\gamma$; i.e., the sampling procedure on $H$ via the average sampler $\Psi = (\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$ is a linear operator from $H$ to $\ell^2(\Gamma)$:

$$S : H \ni f \longmapsto \{y_\gamma := \langle f, \psi_\gamma(\cdot - \gamma) \rangle\}_{\gamma \in \Gamma} \in \ell^2(\Gamma). \tag{4.28}$$

We restrict ourselves to consider *well-localized samplers* $\Psi = (\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$, which means that $\Gamma$ is a relatively separated subset of **R** and the sampling functionals $\psi_\gamma$ are dominated by a function $h$ in the Wiener amalgam space $W^1$; i.e., $|\psi_\gamma(t)| \leq h(t)$ for all $t \in \mathbf{R}$ and $\gamma \in \Gamma$. The reasons for considering well-localized samplers are twofold:

- At each position $\gamma \in \Gamma$, we locate an acquisition device, and hence it is reasonable to assume that there are finitely many such acquisition devices in any unit intervals, which in turn implies that $\Gamma$ is relatively separated.
- We use the sampling functional $\psi_\gamma$ to reflect the characteristic of the acquisition device at the location $\gamma$, and hence the sampling functional $\psi_\gamma$ should essentially be supported in a neighborhood of the sampling location $\gamma$, which can be described by the dominance by a function $h$ with fast decay at infinity.

It is well known that signals with finite energy do not have finite rate of innovation. In [26], we show that any signal $f$ with finite energy could be determined by its samples $\langle f, \psi_\gamma(\cdot - \gamma) \rangle, \gamma \in \Gamma$ for some well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$, but could not be recovered in a stable way from the samples $\langle f, \psi_\gamma(\cdot - \gamma) \rangle, \gamma \in \Gamma$ for any well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$.

**Theorem 11.** *(i)   There is a well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$ such that any function $f \in L^2$ is uniquely determined by its samples $\langle f, \psi_\gamma(\cdot - \gamma) \rangle, \gamma \in \Gamma$.*
*(ii)   There does not exist a well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$ such that the sampling operator $S$ in (4.28) is stable for $H = L^2$ in the sense that there exist positive constants $A$ and $B$ such that $A\|f\|_2^2 \leq \sum_{\gamma \in \Gamma} |\langle f, \psi_\gamma(\cdot - \gamma) \rangle|^2 \leq B\|f\|_2^2$ for all $f \in L^2$.*

We remark that functions $\psi_\gamma, \gamma \in \Gamma$, in the well-localized sampler in the first conclusion of Theorem 11 cannot be selected to be supported in a fixed compact set, but it is possible to let elements $\psi_\gamma, \gamma \in \Gamma$, in the well-localized sampler to be independent on $\gamma \in \Gamma$. This is closely related to the spectral problem: the density of the collection of exponentials $\{\exp(i\gamma t)\}_{\gamma \in \Gamma}$ in a weighted $L^2$ space [31]. In [26], we conjecture that there is not a determining sampler $\{\psi_\gamma(\cdot - \gamma) \mid \gamma \in \Gamma\}$ such that $\|\psi_\gamma\|_2 = 1$ and $|\psi_\gamma(x)| \leq C\exp(-\varepsilon|x|)$ for some positive constants $C, \varepsilon$ and a relatively separated subset $\Gamma$ of $\mathbf{R}$.

**Dedication.** This chapter is dedicated to Professor Gilbert Walter on the occasion of his 80th birthday:

- In appreciation of his friendship and important contributions to Mathematical Analysis and Applications.
- With admiration of the novel and clever ways in which he has brought together ideas from classical and modern analysis to advance our understanding of generalized functions, wavelets, and signal processing.

# References

1. Aldroubi A, Gröchenig K (2001) Nonuniform sampling and reconstruction in shift-invariant space, SIAM Rev 43:585–620
2. Aldroubi A, Sun Q, Tang W-S (2001) $p$-frames and shift-invariant subspaces of $L^p$. J Fourier Anal Appl 7:1–21
3. Aldroubi A, Sun Q, Tang W-S (2004) Nonuniform average sampling and reconstruction in multiply generated shift-invariant spaces. Constr Approx 20:173–189
4. Aldroubi A, Sun Q, Tang W-S (2005) Convolution, average sampling, and a Calderon resolution of the identity for shift-invariant spaces. J Fourier Anal Appl 22:215–244
5. Aronszajn N (1950) Theory of reproducing kernels. Trans Am Math Soc 68:337–404
6. Benedetto JJ, Ferreira PJSG (eds.) (2001) Modern sampling theory: mathematics and applications. Birkhuse, Boston
7. Benedetto JJ, Zayed AI (eds.) (2003) Sampling, wavelets, and tomography. Birkhauser, Boston
8. Bi N, Nashed MZ, Sun Q (2009) Reconstructing signals with finite rate of innovation from noisy samples. Acta Appl Math 107:339–372.
9. Butzer PL (1983) A survey of the Whittaker-Shannon sampling theorem and some of its extensions. J Math Res Exposition 3:185–212
10. Butzer PL, Splettstö$\beta$er W, Stens RL (1988) The sampling theorem and linear prediction in signal analysis. Jahresber Deutsch Math-Verein 90:1–60
11. Butzer PL, Stens RL (1992) Sampling theory for not necessarily band-limited functions: a historical overview. SIAM Rev 34:40–53
12. Fornasier M, Rauhut H (2005) Continuous frames, function spaces, and the discretization problem. J Fourier Anal Appl 11:245–287
13. Haddad RA, Parsons TW (1991) Digital signal processing: theory, applications and hardware. Computer Science Press
14. Han D, Nashed MZ, Sun Q (2009) Sampling expansions in reproducing kernel Hilbert and Banach spaces. Numer Funct Anal Optim 30:971–987

15. Higgins JR (1985) Five short stories about the cardinal series. Bull Am Math Soc 12:45–89
16. Higgins JR (1996) Sampling theory in fourier and signal analysis volume 1: foundations. Oxford University Press, Oxford
17. Higgins JR, Stens RL (2000) Sampling theory in fourier and signal analysis: volume 2: advanced topics. Oxford Science Publications, Oxford
18. Ismail M, Nashed Z, Zayed A, Ghaleb A (eds.) (1995) Mathematical analysis, wavelets and signal processing (contemporary mathematics), vol 190. American Mathematical Society, Providence, RI
19. Jerri JA (1977) The Shannon sampling theorem—its various extensions and applications: a tutorial review. Proc IEEE 65:1565–1596
20. Larson D, Massopust P, Nashed Z, Nguyen MC, Papadakis M, Zayed A (eds.) (2008) Frames and operator theory in analysis and signal processing (Contemporary Mathematics), vol 451. American Mathematical Society, Providence, RI
21. Marks II RJ (1991) Introduction to shannon sampling and interpolation theory. Springer, Berlin
22. Marks RJ (1993) Advanced topics in shannon sampling and interpolation theory. Springer, Berlin, Heidelberg
23. Marvasti FA (ed) (2001) Nonuniform Sampling: theory and practice (information technology: transmission, processing, and storage). Plenum Pub Corp, New York
24. Nashed MZ, Scherzer O (eds.) (2002) Inverse problems, image analysis and medical imaging (Contemporary Mathematics), vol 313. American Mathematical Society, Providence, RI
25. Nashed MZ, Sun Q (2010) Sampling and reconstruction of signals in a reproducing kernel subspace of $L^p(R^d)$. J Funct Anal 258:2422–2452
26. Nashed MZ, Sun Q, Tang W-S (2009) Average sampling in $L^2$. Can Acad Sci Paris, Ser I 347:1007–1010
27. Nashed MZ, Sun Q, Xian J Convolution sampling and reconstruction of signals in a reproducing kernel subspace. Proc Am Math Soc (to appear)
28. Nashed MZ, Walter GG (1991) General sampling theorems for functions in reproducing kernel Hilbert spaces. Math Contr Signals Syst 4:363–390
29. Nashed MZ, Walter GG (1995) Reproducing kernel Hilbert spaces from sampling expansions. Contemporary Math 190:221–226
30. Oppenheim AV, Schafer RW (1989) Digital signal processing. Prentice-Hall Inc., New Jersey
31. Olevskii A, Ulanovskii A (2004) Almost integer translates. Do nice generators exist? J Fourier Anal Appl 10:93–104
32. Papoulis A (1977) Generalized sampling expansion. IEEE Trans Circ Syst 24:652–654
33. Sun Q (2006) Non-uniform sampling and reconstruction for signals with finite rate of innovations. SIAM J Math Anal 38:1389–1422
34. Sun Q (2008) Frames in spaces with finite rate of innovation. Adv Comput Math 28:301–329
35. Sun Q (2010) Local reconstruction for sampling in shift-invariant spaces. Adv Comput Math 32:335–352
36. Sun Q (2011) Localized nonlinear functional equations and two sampling problems in signal processing. submitted
37. Unser M (2000) Sampling—50 years after Shannon. Proc IEEE 88:569–587
38. van der Mee C, Nashed MZ, Seatzu S (2003) Sampling expansions and interpolation in unitarily translation invariant reproducing kernel Hilbert spaces. Adv Comput Math 19:355–372
39. Vetterli M, Marziliano P, Blu T (2002) Sampling signals with finite rate of innovation. IEEE Trans Signal Process 50:1417–1428
40. Walter GG (1992) A sampling theorem for wavelet subspaces. IEEE Trans Inform Theor 38:881–884
41. Walter GG, Shen X (2000) Wavelets and other orthogonal systems, 2nd edn. CRC Press, USA
42. Yao K (1967) Applications of reproducing kernel Hilbert spaces-bandlimited signal models. Inform Contr 11:429–444
43. Zayed A (1993) Advances in Shannon's sampling theory. CRC Press, USA