An Extension of Oblique Projection Sampling Theorem

Akira Hirabayashi† and Michael Unser‡

† Department of Computer Science and Systems Engineering,
Yamaguchi University

‡ Biomedical Imaging Group,
Swiss Federal Institute of Technology Lausanne (EPFL)

Abstract

In this paper, we discuss the sampling problem without a condition that was assumed in conventional sampling theorems. This means that we cannot perfectly reconstruct all functions in the reconstruction space. The perfect reconstruction is possible only for functions in an arbitrary complementary subspace of the intersection of the reconstruction space and the orthogonal complement of the sampling space. We propose a sampling theorem that reconstructs the oblique projection onto the complementary subspace along the orthogonal complement of the sampling space. The sampling theorem guarantees the perfect reconstruction of functions of special interest in the reconstruction space, such as the constant function in image processing applications. In addition, we explain why a conventional sampling theorem is not suitable for the present case.

1 Introduction

Sampling is the process of reconstructing functions of continuous variables from their discrete measurements. The best known result is the sampling theorem for bandlimited functions [1]. For real world signals or images, this theorem is applicable only after a prefiltering since they are never exactly bandlimited. In this case, the bandlimited sampling is an approximation.

There are a number of works that take this approximation theoretic point of view [2, 3, 4, 5, 6]. They can be summarized as follows: Let V_s and V_r be the sampling space and the reconstruction space that are spanned by sampling functions and reconstruction functions, respectively. When we are free to choose either the sampling functions or the reconstruction functions, we can obtain the minimal-error approximation, which is the orthogonal projection of an original function onto V_r, by selecting $V_s = V_r$. When the sampling functions are given a priori, V_s is not equal to V_r in general. In this case, the minimum error approximation can no longer be obtained from the measurements. Then, one possibility is to obtain the oblique projection of the original signal onto V_r along $V_s^{\perp}$ since its measurements are equal to the original ones.

Within these discussions, the condition $V_r \cap V_s^{\perp} = \{0\}$ has been assumed. It ensures that the functions in V_r can be perfectly reconstructed. However, it is easy to display cases where the condition does not hold. For example, let us think of the following situation: The numbers of sampling and reconstruction functions are the same, and these functions are linearly independent. We denote the sampling space by V_{org}, and assume that $V_r \cap V_{org}^{\perp} = \{0\}$. If one of sampling functions is lost in the acquisition process, then the sampling space V_s is now a proper subspace of V_{org}, and the condition $V_r \cap V_s^{\perp} = \{0\}$ does not hold anymore. Another simple example related to the spline sampling theory is shown in Section 2.

In this paper, we discuss the sampling problem for the case where $V_r \cap V_s^{\perp} \neq \{0\}$. In this case, V_r is decomposed into $V_r \cap V_s^{\perp}$ and its arbitrary complementary subspace in V_r, say L. The perfect reconstruction is possible only for functions in L. We first propose a sampling theorem that reconstructs the oblique projection onto L along $V_s^{\perp}$. Then, we propose the main sampling theorem which is the special form of the first sampling theorem with a specific L. It is chosen so that the following two conditions are satisfied. First, if we have some functions of special interest in V_r, and if those functions do not belong to $V_r \cap V_s^{\perp}$, then L contains those functions. The subspace spanned by such functions is denoted by V_i. The second condition is that the rest of $V_i \oplus (V_r \cap V_s^{\perp})$ is its orthogonal complement in V_r, which is denoted by V_c. The main theorem guarantees not only the perfect reconstruction for functions in V_i, but also provides the minimum error approximation for functions in the direct sum of V_c and $V_r \cap V_s^{\perp}$.

In addition, we compare the proposed sampling theorem to the the approach presented in [6]. Although this is the result for the standard case $V_r \cap V_s^{\perp} = \{0\}$, it remains well-defined in the present situation. We investigate the behavior of this type of solution and show that it is less attractive than the present solution in the case where $V_r \cap V_s^{\perp} \neq \{0\}$.

1.1 Mathematical Preliminaries

The following notations are used in this paper. $\mathbf{C}^N$ and $\mathbf{C}^K$ stand for the N-dimensional and K-dimensional Hermitian space, respectively. Let $\{\boldsymbol{e}_n^{(N)}\}_{n=1}^N$ and $\{\boldsymbol{e}_k^{(K)}\}_{k=1}^K$ be the standard bases in $\mathbf{C}^N$ and $\mathbf{C}^K$, respectively. That is, $\boldsymbol{e}_n^{(N)}$ and $\boldsymbol{e}_k^{(K)}$ are the N-dimensional and K-dimensional vectors consisting of zero elements except for the n-th and k-th elements equal to 1, respectively. Let I be the identity operator.

The orthogonal complement of a closed subspace S is denoted by $S^{\perp}$. $\mathcal{R}(A)$ and $\mathcal{N}(A)$ stand for the range and the null space of an operator A, respectively. Let A^* be the adjoint operator.

Let $A^{\dagger}$ and P_S denote the Moore-Penrose generalized inverse of an operator A and the orthogonal projection operator onto S, respectively. It holds that

$$AA^{\dagger} = P_{\mathcal{R}(A)}, \quad A^{\dagger}A = P_{\mathcal{R}(A^*)}, \tag{1}$$

$$A^{\dagger} = A^*(AA^*)^{\dagger} = (A^*A)^{\dagger}A^*, \tag{2}$$

$$P_S(AP_S)^{\dagger} = (AP_S)^{\dagger} \text{ and } (P_SA)^{\dagger}P_S = (P_SA)^{\dagger}. \tag{3}$$

2 Problem Formulation

Let f be an original signal defined on the domain $\mathcal{D}$. We assume that f lies in a Hilbert space H. The measurements of f, denoted by $c_1[n]$ $(n = 1, 2, \ldots, N)$, are given by the inner product of f with the sampling functions $\{\psi_n\}_{n=1}^N$:

$$c_1[n] = \langle f, \psi_n \rangle. \tag{4}$$

The N-dimensional vector whose n-th element is $c_1[n]$ is denoted by $\boldsymbol{c}_1$. Let A_s be the operator that maps f into $\boldsymbol{c}_1$:

$$\boldsymbol{c}_1 = A_s f. \tag{5}$$

The reconstructed signal $\tilde{f}$ is given by a linear combination of reconstruction functions, $\{\varphi_k\}_{k=1}^K$:

$$\tilde{f} = \sum_{k=1}^{K} c_2[k]\varphi_k. \tag{6}$$

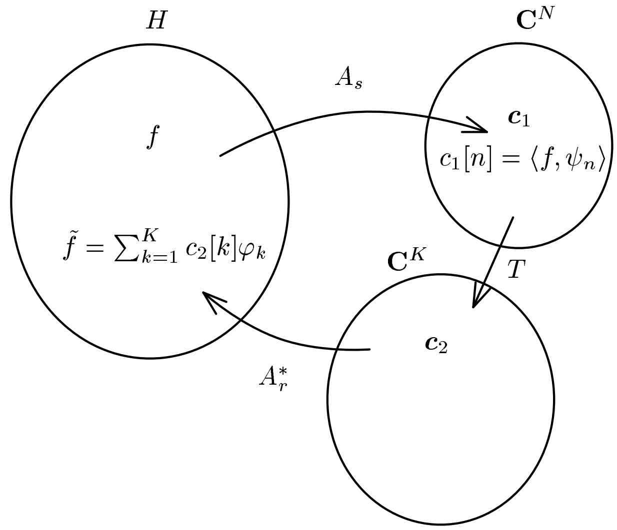

Figure 1: Sampling and Reconstruction Formulation.

Let A_r be the operator that maps f into $\sum_{k=1}^K \langle f, \varphi_k \rangle \boldsymbol{e}_k^{(K)}$. The K-dimensional vector consisting of $c_2[k]$ is denoted by $\boldsymbol{c}_2$. Then, it holds that

$$\tilde{f} = A_r^* \boldsymbol{c}_2. \tag{7}$$

Let T be a $K \times N$ matrix that maps $\boldsymbol{c}_1$ to $\boldsymbol{c}_2$:

$$\boldsymbol{c}_2 = T\boldsymbol{c}_1. \tag{8}$$

Then, Eqs.(7), (8), and (5) yield

$$\tilde{f} = A_r^* T A_s f. \tag{9}$$

With this formulations, the sampling problem becomes a problem of finding T so that $\tilde{f}$ satisfies some criterion. The sampling and reconstruction formulation is illustrated in Fig.1.

Let V_s and V_r be subspaces spanned by $\{\psi_n\}_{n=1}^N$ and $\{\varphi_k\}_{k=1}^K$, respectively. They are called the *sampling space* and the *reconstruction space*, respectively. It holds that

$$V_s = \mathcal{R}(A_s^*), \tag{10}$$

$$V_r = \mathcal{R}(A_r^*). \tag{11}$$

So far, the sampling problem was discussed under the assumption that

$$V_r \cap V_s^{\perp} = \{0\} \tag{12}$$

[4, 5, 6]. This ensures that any f in V_r can be perfectly reconstructed from the measurements $\boldsymbol{c}_1$ by choosing T in Eq.(8) appropriately. However, it is easy to find a case where Eq.(12) does not hold, as mentioned in the introduction. Let us present another example. The B-splines of degree 0 and 1, denoted by $\beta^0(x)$ and $\beta^1(x)$, are defined by

$$\beta^0(x) = \begin{cases} 1 \ (0 \leq x < 1), \\ 0 \ (x < 0, x \geq 1), \end{cases} \tag{13}$$

and

$$\beta^1(x) = (\beta^0 * \beta^0)(x), \tag{14}$$

respectively, where $*$ is the convolution operator.

Example 1 *Assume that $N = K$ and the domain $\mathcal{D} = [0, N]$. Let ψ_n and φ_k be*

$$\psi_n(x) = \beta^0(x - n + 1), \tag{15}$$

$$\varphi_k(x) = \beta^1_K(x - k + 1), \tag{16}$$

respectively, where

$$\beta^1_K(x) = \beta^1(x) + \beta^1(x - K). \tag{17}$$

If N is even, then $\tilde{f}$ in Eq.(7) with

$$\boldsymbol{c}_2 = (1, -1, \ldots, 1, -1) \tag{18}$$

belongs to $V_s^\perp$.

Even in this simple example, Eq.(12) does not hold. This further motivates our investigation of the case

$$V_r \cap V_s^\perp \neq \{0\}. \tag{19}$$

3 New Sampling Theorem

Eq.(19) implies that all f in V_r cannot be perfectly reconstructed. Indeed, f in $V_r \cap V_s^\perp$ yields $\tilde{f} = 0$ because $A_s f = 0$. The perfect reconstruction is possible only for functions in an arbitrary complementary subspace of $V_r \cap V_s^\perp$ in V_r. It is denoted by L, and satisfies

$$L \oplus (V_r \cap V_s^\perp) = V_r, \tag{20}$$

where $\oplus$ denotes the direct sum.

We assume one more condition in this paper, which is

$$V_r + V_s^\perp = H. \tag{21}$$

For example, if the dimension of V_s is less than or equal to that of V_r, then Eq.(21) holds.

Eqs.(20) and (21) imply that H can be decomposed into the direct sum of

$$H = L \oplus V_s^\perp. \tag{22}$$

Then, we can define the oblique projection operator onto L along $V_s^\perp$, which is denoted by $P_{L,V_s^\perp}$. We can achieve

$$\tilde{f} = P_{L,V_s^\perp} f \tag{23}$$

for any f in H. In fact, the following theorem holds.

Theorem 1 *The oblique projection in Eq.(23) is reconstructed by Eq.(6) if and only if T in Eq.(8) is given by*

$$T = (A_r^*)^\dagger P_{L,V_s^\perp} A_s^\dagger + Y - A_r A_r^\dagger Y A_s A_s^\dagger, \tag{24}$$

where Y is an arbitrary operator from $\mathbf{C}^N$ to $\mathbf{C}^K$.

(Proof) Assume that $\tilde{f}$ is the oblique projection in Eq.(23). Then, it follows from Eq.(9) that

$$A_r^* T A_s f = P_{L,V_s^\perp} f \quad \text{for any } f \in H, \tag{25}$$

which is equivalent to

$$A_r^* T A_s = P_{L,V_s^\perp}. \tag{26}$$

Since L is a subspace in V_r, it follows from Eq.(11) that

$$\mathcal{R}(A_r^*) = V_r \supset L = \mathcal{R}(P_{L,V_s^\perp}). \tag{27}$$

Further, it holds that

$$\mathcal{N}(A_s) = \mathcal{N}(P_{L,V_s^\perp}). \tag{28}$$

These two equations imply that Eq.(26) always has a solution, and its general form is given by Eq.(24) [7].

Conversely, assume that T is given by Eq.(24). It follows from Eq.(27) that

$$A_r^*(A_r^*)^\dagger P_{L,V_s^\perp} = P_{L,V_s^\perp}.$$

Further, it follows from Eq.(28) that

$$P_{L,V_s^\perp} A_s^\dagger A_s = P_{L,V_s^\perp}.$$

Hence, from Eqs.(9) and (24), we have

$$\tilde{f} = A_r^* T A_s f = A_r^*(A_r^*)^\dagger P_{L,V_s^\perp} A_s^\dagger A_s f = P_{L,V_s^\perp} f,$$

which implies Eq.(23). ■

Based on Theorem 1, we propose a special sampling theorem with a specific subspace L. It is chosen so that the following two conditions are satisfied. First, if we have some functions of special interest in V_r, and if those functions do not belong to $V_r \cap V_s^\perp$, then L includes those functions. For example, the constant function is very important in image processing applications. Let $\{\phi_i\}_{i=1}^I$ be such functions in V_r. Let V_i be the subspace spanned by $\{\phi_i\}_{i=1}^I$ and A_i the operator defined by

$$A_i f = \sum_{i=1}^{I} \langle f, \phi_i \rangle \boldsymbol{e}_i^{(I)}. \tag{29}$$

The subscript i means 'interest'.

If the direct sum of $V_r \cap V_s^\perp$ and V_i is equal to V_r, that is, if

$$V_i \oplus (V_r \cap V_s^\perp) = V_r, \tag{30}$$

then L is uniquely determined so that $L = V_i$. Otherwise, we need a second condition. Here, we adopt the condition that the remainder of $V_i \oplus (V_r \cap V_s^{\perp})$ is its orthogonal complement in V_r, which is denoted by V_c. The subscript c means 'complement'.

Consequently, the subspace L is determined so that

$$L = V_i \oplus V_c, \tag{31}$$

and the application of Theorem 1 yields to our next result.

Theorem 2 *The oblique projection of any f in H onto L in Eq.(31) along $V_s^{\perp}$ is reconstructed by Eq.(6) if and only if T in Eq.(8) is given by*

$$T = (A_r^*)^{\dagger}(A_s W)^{\dagger} + Y - A_r A_r^{\dagger} Y A_s A_s^{\dagger}, \tag{32}$$

where W is an operator defined by

$$W = A_i^{\dagger} A_i + P_{\mathcal{R}(P_{V_r} V_s)} - (P_{\mathcal{R}(P_{V_r} V_s)} A_i^*)(P_{\mathcal{R}(P_{V_r} V_s)} A_i^*)^{\dagger}. \tag{33}$$

In order to prove Theorem 2, we use the following lemma. Proofs of all lemmas are deferred the appendix.

Lemma 1 *The subspace V_c is given by*

$$V_c = P_{V_r} V_s \cap V_i^{\perp}. \tag{34}$$

(Proof of Theorem 2) It is obvious that the first term in the right-hand side of Eq.(33) is the orthogonal projection operator onto V_i. The rest of the right-hand side has the following meaning. Theorem 4 in [8] implies that $P_{V_r} V_s$ can be decomposed into the orthogonal direct sum of

$$P_{V_r} V_s = P_{P_{V_r} V_s} V_i \oplus (P_{V_r} V_s \cap V_i^{\perp}). \tag{35}$$

Hence, the orthogonal projection operator onto $P_{V_r} V_s \cap V_i^{\perp}$ is given by the difference of those onto $P_{V_r} V_s$ and $P_{P_{V_r} V_s} V_i$, which is the rest of the right-hand side of Eq.(33). Since V_i and $P_{V_r} V_s \cap V_i^{\perp}$ are perpendicular to each other, the sum of these two orthogonal projection operators becomes that onto L in Eq.(31).

Proposition 1 in [6] implies that $W(A_s W)^{\dagger} A_s$ is the oblique projection operator onto L along $V_s^{\perp}$. Since W is the orthogonal projection operator, it follows from Eq.(3) that

$$W(A_s W)^{\dagger} = (A_s W)^{\dagger}. \tag{36}$$

Then, the first term of Eq.(24) yields

$$(A_r^*)^{\dagger} P_{L, V_s^{\perp}} A_s^{\dagger} = (A_r^*)^{\dagger} (A_s W)^{\dagger} A_s A_s^{\dagger}. \tag{37}$$

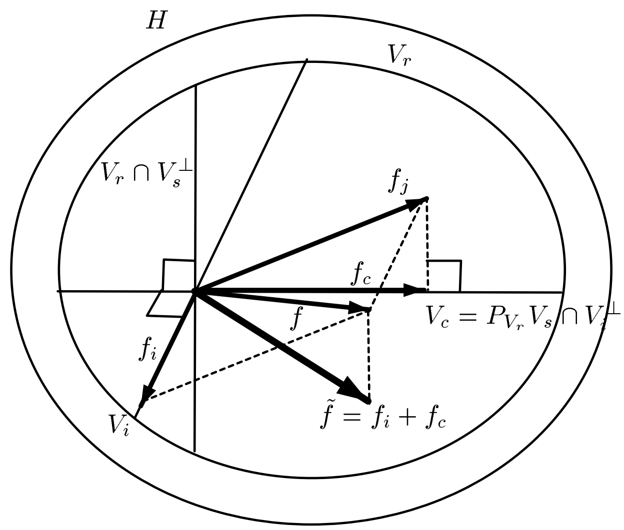

Figure 2: Reconstruction $\tilde{f}$ by Theorem 2. Note that V_i and $V_r \cap V_s^{\perp}$ are not perpendicular in general.

Further, it follows from Eq.(2) that

$$\begin{aligned}(A_s W)^{\dagger} A_s A_s^{\dagger} &= [\{(A_s W)^*(A_s W)\}^{\dagger}(A_s W)^*] A_s A_s^{\dagger} \\ &= \{(A_s W)^*(A_s W)\}^{\dagger} W A_s^* A_s A_s^{\dagger} \\ &= \{(A_s W)^*(A_s W)\}^{\dagger} W A_s^* \\ &= (A_s W)^{\dagger},\end{aligned}$$

so that

$$(A_s W)^{\dagger} A_s A_s^{\dagger} = (A_s W)^{\dagger}. \tag{38}$$

The use of Theorem 1 together with Eqs.(37) and (38) finally yields Eq.(32). ■

In order to verify our second condition of the orthogonal complement, let us decompose f in L into two components of

$$f = f_i + f_j, \tag{39}$$

where f_i and f_j lie in V_i and $\{V_c \oplus (V_r \cap V_s^{\perp})\}$, respectively. The first component f_i is perfectly reconstructed because it is in $V_i \subset L$. On the other hand, the second component f_j is not perfectly reconstructed, but projected onto V_c. The projection is denoted by f_c. Then, the reconstruction for f is given by

$$\tilde{f} = f_i + f_c. \tag{40}$$

Note that, since V_c is perpendicular to $V_r \cap V_s^{\perp}$, f_c is the minimum error approximation of f_j in V_c. This guarantees that our sampling theorem provides a good approximation. These interpretations are illustrated in Fig.2.

In the extreme case of $V_i = \{0\}$, the subspace L in Eq.(31) becomes

$$L = V_c, \tag{41}$$

which is the orthogonal complement of $V_r \cap V_s^{\perp}$ in V_r. Then, Theorem 2 reduces to the following special form.

Corollary 1 *The oblique projection of any f in H onto L in Eq.(41) along $V_s^{\perp}$ is reconstructed by Eq.(6) if and only if T in Eq.(8) is given by*

$$T = (A_r^*)^{\dagger}(A_s P_{V_r})^{\dagger} + Y - A_r A_r^{\dagger} Y A_s A_s^{\dagger}. \quad (42)$$

(Proof) In the case of $V_i = \{0\}$, the operator W in Eq.(33) reduces to

$$W = P_{\mathcal{R}(P_{V_r} V_s)}. \quad (43)$$

Hence, it follows from Eqs.(1) and (10) that

$$\begin{aligned} A_s W &= A_s(P_{V_r} A_s^*)(P_{V_r} A_s^*)^{\dagger} \\ &= (P_{V_r} A_s^*)^*(P_{V_r} A_s^*)(P_{V_r} A_s^*)^{\dagger} \\ &= A_s P_{V_r}, \end{aligned}$$

which implies Eq.(32) yields Eq.(42). ■

The predominant term of the operator T in either Theorem 2 or Corollary 1, is the first one. It involves the factors $A_s W$ and A_r which are operators from the function space H to the vector spaces $\mathbf{C}^N$ and $\mathbf{C}^K$, respectively, and which are therefore not suitable for computer calculations. Hence, by using Eq.(2), we convert the first term into

$$(A_r^*)^{\dagger}(A_s W)^{\dagger} = (A_r A_r^*)^{\dagger} A_r W A_s^* (A_s W A_s^*)^{\dagger}. \quad (44)$$

Note that the operator $A_r A_r^*$, $A_r W A_s^*$, and $A_s W A_s^*$ are the $K \times K$, $K \times N$, and $N \times N$ matrices, respectively. Hence, they can be calculated numerically in this form. A similar conversion is possible for Eq.(42). In this case, the second term becomes $A_r A_s^*$.

4 Discussion

The result proposed in [6] can be applied to the present case. We compare it to our result. First, let us review the result.

Proposition 1 *[6] Any $f \in H$ can be consistently reconstructed from the measurements $\mathbf{c}_1$ by Eq.(6) if T in Eq.(8) is given by*

$$T = (A_s A_r^*)^{\dagger}, \quad (45)$$

and the consistent reconstruction is unique.

This is the result for the case of Eq.(12). However, the operator in Eq.(45) is well-defined even if Eq.(12) does not hold. The behavior of the operator for the case of Eq.(19) is clarified in the following theorem.

Theorem 3 *The oblique projection of any f in H onto $\mathcal{R}(A_r^* A_r A_s^*)$ along $V_s^{\perp}$ is reconstructed by Eq.(6) if and only if T in Eq.(8) is given by*

$$T = (A_s A_r^*)^{\dagger} + Y - A_r A_r^{\dagger} Y A_s A_s^{\dagger}. \quad (46)$$

In order to prove Theorem 3, we use the following two lemmas.

Lemma 2 *The function space H can be decomposed into the direct sum of*

$$H = \mathcal{R}(A_r^* A_r A_s^*) \oplus V_s^{\perp}. \quad (47)$$

Lemma 3 *The operator $A_r^*(A_s A_r^*)^{\dagger} A_s$ is the oblique projection operator onto $\mathcal{R}(A_r^* A_r A_s^*)$ along $V_s^{\perp}$.*

(Proof of Theorem 3) If we substitute $P_{L,V_s^{\perp}}$ by $A_r^*(A_s A_r^*)^{\dagger} A_s$ in Eq.(24), it follows from Eqs.(1) and (3) that

$$\begin{aligned} &(A_r^*)^{\dagger} P_{L,V_s^{\perp}} A_s^{\dagger} = (A_r^*)^{\dagger} A_r^* (A_s A_r^*)^{\dagger} A_s A_s^{\dagger} \\ &= P_{\mathcal{R}(A_r)}(P_{\mathcal{R}(A_s)} A_s A_r^* P_{\mathcal{R}(A_r)})^{\dagger} P_{\mathcal{R}(A_s)} \\ &= (P_{\mathcal{R}(A_s)} A_s A_r^* P_{\mathcal{R}(A_r)})^{\dagger} = (A_s A_r^*)^{\dagger}, \end{aligned}$$

i.e.,

$$(A_r^*)^{\dagger} P_{L,V_s^{\perp}} A_s^{\dagger} = (A_s A_r^*)^{\dagger}. \quad (48)$$

Hence, Lemma 3 implies that Theorem 1 reduces to the above theorem. ■

Theorem 3 means that the reconstruction $\tilde{f}$ by $(A_s A_r^*)^{\dagger}$ lies in $\mathcal{R}(A_r^* A_r A_s^*)$. Note that this subspace changes depending on the sampling operator A_s^*. There is not a clear reason to use this subspace. Further, the perfect reconstruction for functions in V_i is not guaranteed, while it is guaranteed in Theorem 2. Hence, to use $(A_s A_r^*)^{\dagger}$ is not suitable for the case of Eq.(19).

One might say that we should use some regularization technique for the case of Eq.(19). This is a reasonable assertion. However, the technique does not guarantee the perfect reconstruction for functions in V_i, either.

5 Conclusion

In this paper, we discussed the sampling problem for the case where the intersection of the reconstruction space V_r and the orthogonal complement of the sampling space V_s is not empty. This is a situation that may happen in practice, as examplified in Sections 1 and 2. In this case, we cannot reconstruct all functions in V_r perfectly. The perfect reconstruction is possible only for functions in an arbitrary complementary subspace of $V_r \cap V_s^{\perp}$ in V_r, which is denoted by

L. We first proposed a sampling theorem that reconstructs the oblique projection onto L along $V_s^{\perp}$. Then, we proposed the main sampling theorem which was the special form of the first sampling thoerem with a special L. It was chosen so that (1) L contains a subspace V_i spanned by functions of special interest in V_r, and (2) the remainder of $V_i \oplus (V_r \cap V_s^{\perp})$ is its orthogonal complement in V_r, which is denoted by V_c. The main theorem guarantees not only the perfect reconstruction for functions in V_i, but also yields the best approximation for functions in the direct sum of V_c and $V_r \cap V_s^{\perp}$. We also clarified the fact that the solutions given by conventional projection-based sampling theorems are not suitable for situations where $V_r \cap V_s^{\perp} \neq \{0\}$, even when the problem is not ill-posed.

Acknowledgment

This research was partially supported by the Venture Business Laboratory, Yamaguchi University, Japan.

A.1 Proof of Lemma 1

Theorem 4 in [8] implies that V_r can be decomposed into the orthogonal direct sum of

$$V_r = P_{V_r} V_s \oplus (V_r \cap V_s^{\perp}).$$

That is, $P_{V_r} V_s$ is the orthogonal complement of $V_r \cap V_s^{\perp}$ in V_r. V_c is the intersection of the orthogonal complements of V_i and $V_r \cap V_s^{\perp}$. Hence, V_c is given by Eq.(34). ∎

A.2 Proof of Lemma 2

Let u be an element in $\mathcal{R}(A_r^* A_r A_s^*)^{\perp} \cap V_s$. Then, it holds that

$$\|A_r u\|^2 = \langle A_r^* A_r u, u \rangle = 0,$$

which implies

$$\mathcal{N}(A_r) \supset \mathcal{R}(A_r^* A_r A_s^*)^{\perp} \cap V_s. \tag{49}$$

Taking orthogonal complement of the equation yields

$$V_r \subset \mathcal{R}(A_r^* A_r A_s^*) + V_s^{\perp}. \tag{50}$$

Hence, it follows from Eq.(21) that

$$H = V_r + V_s^{\perp} \subset \mathcal{R}(A_r^* A_r A_s^*) + V_s^{\perp},$$

which implies

$$H = \mathcal{R}(A_r^* A_r A_s^*) + V_s^{\perp}. \tag{51}$$

Let u be an element in $\mathcal{R}(A_r^* A_r A_s^*) \cap V_s^{\perp}$. Then, u satisfies $A_s u = 0$, and there exist some v such that $u = A_r^* A_r A_s^* v$. It holds that

$$0 = \langle A_s u, v \rangle = \langle A_s A_r^* A_r A_s^* v, v \rangle = \|A_r A_s^* v\|^2.$$

Hence, $A_r A_s^* v = 0$ which implies $u = 0$. Therefore, the right-hand side of Eq.(51) is a direct sum, and Eq.(47) holds. ∎

A.3 Proof of Lemma 3

It is obvious from [6] that the operator is projection along $V_s^{\perp}$. Hence, we show that

$$\mathcal{R}(A_r^* (A_s A_r^*)^{\dagger} A_s) = \mathcal{R}(A_r^* A_r A_s^*). \tag{52}$$

It holds that

$$\begin{aligned} &\mathcal{R}(A_r^* (A_s A_r^*)^{\dagger} A_s) \subset \mathcal{R}(A_r^* (A_s A_r^*)^{\dagger}) \\ &= \mathcal{R}(A_r^* (A_s A_r^*)^*) = \mathcal{R}(A_r^* A_r A_s^*) \\ &= \mathcal{R}(A_r^* (A_s A_r^*)^{\dagger} A_s A_r^* (A_s A_r^*)^*) \\ &\subset \mathcal{R}(A_r^* (A_s A_r^*)^{\dagger} A_s), \end{aligned}$$

which implies Eq.(52). ∎

References

[1] C. Shannon, "Communications in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, 1949.

[2] H. Ogawa, "A unified approach to generalized sampling theorems," in *Proc. of ICASSP'86, 1986 IEEE-IECEJ-ASJ Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 3, Tokyo, 1986, pp. 1657–1660.

[3] ——, "A generalized sampling theorem," *Electronics and Communications in Japan, Part.3*, vol. 72, no. 3, pp. 97–105, 1989.

[4] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 2915–2925, November 1994.

[5] M. Unser, "Sampling—50 Years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.

[6] Y. C. Eldar, "Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors," *The Journal of Fourier Analysis and Applications*, vol. 1, no. 9, pp. 77–96, 2003.

[7] A. Albert, *Regression and the Moore-Penrose Pseudoinverse.* New York: Academic Press, 1972.

[8] H. Ogawa and S. Hara, "Properties of partial projection filter," *The Transactions of The Institute of Electronics, Information and Communication Engineers A*, vol. J71-A, no. 2, pp. 527–534, 1988, (in Japanese).