A NEW FORMULA FOR LOST SAMPLE RESTORATION

M. Chabert, B. Lacaze

TéSA/Enseeiht/IRIT, 14/16 Port St-Etienne, 31000 Toulouse, France
marie.chabert@enseeiht.fr

ABSTRACT

The reconstruction of oversampled band-limited random signals from incomplete discrete data is addressed. The proposed reconstruction scheme stems from Lagrange interpolation formula such as the Shannon cardinal expansion. However, the proposed formula explicitly takes into account the possible loss of one or more samples. The formula can be fitted to any sample loss or deterioration by a simple time index translation. The reconstruction performance is studied with respect to the number of lost samples, to the number of available samples and to the oversample rate. The proposed scheme, associated to specific interpolation functions, results in a high convergence rate even in the neighborhood of the lost samples providing an accurate sample restoration method.

1. INTRODUCTION

Let $\mathbf{Z} = \{Z(t), t \in \mathbb{R}\}$ denote a zero mean stationary process, with regular band-limited spectrum $s(\omega)$ in $(-\pi, \pi)$ defined by:

$$E[Z(t) Z^*(t-\tau)] = \int_{-\pi}^{\pi} e^{i\omega\tau} s(\omega) d\omega \tag{1}$$

Let consider the periodic oversampling case: the process $\mathbf{Z}$ is assumed sampled above the Nyquist rate i.e. at a rate β larger than 1. The associated sample time sequence is $\{k/\beta, k \in \mathbb{Z}\}$. $\mathbf{Z}$ can be retrieved from a linear combination of the samples $Z(k/\beta), k \in \mathbb{Z}$, for example by the Shannon expansion or cardinal series:

$$Z(t) = \lim_{N\to\infty} \sum_{k=-N}^{N} \frac{\sin\pi(\beta t - k)}{\pi(\beta t - k)} Z\left(\frac{k}{\beta}\right) \tag{2}$$

An estimate of $Z(t)$ from a finite number of samples can be derived by truncation of the previous series:

$$\widehat{Z}_s(t) = \sum_{k=-N}^{N} \frac{\sin\pi(\beta t - k)}{\pi(\beta t - k)} Z\left(\frac{k}{\beta}\right) \tag{3}$$

However, this formula performance is limited by the weak decrease of the cardinal series and drops dramatically when some samples are lost or deteriorated. This paper provides appropriate formulas for the signal reconstruction in the presence of one or more missing samples. The basic idea is a twofold process expansion from available samples on one hand and unavailable periodically spaced samples on the other hand. Under mild conditions on the sampling times, the second term of the expansion can be neglected. Thus, the first term provides an estimate of $Z(t)$ [8]. The main advantage with respect to Shannon estimation is that this formula can be easily fitted to a sample loss or deterioration i.e. to incomplete data. Note that the proposed formula provides a continuous time reconstruction thus it can be used to estimate any unknown value $Z(t)$. In the following, we focus on sample restoration. Consequently, the performance will be studied through point-wise reconstruction error. The method is an alternative to other well-known algorithms generally based on recursive formulas [1]. The paper is organized as follows. Section 2 provides the general reconstruction formula and the conditions on the sampling scheme for convergence. Section 3 develops the particular cases of a single lost sample and of two (adjacent or not) lost samples. Section 4 studies the restoration performance through simulations. Conclusions and perspectives are discussed in section 5.

2. NEW FORMULA FOR SAMPLE RESTORATION

Let $\{t_k, k \in \mathbb{Z}^*\}$ denote the increasing sample time sequence composed of real non-integer values from $-\infty$ to $+\infty$ with $t_{-1} < 0 < t_1$ and:

$$\lim_{|k|\to\infty} \frac{t_k}{k} = \frac{1}{\beta}$$

$$\lim_{N\to\infty} \sum_{k=1}^{N} \frac{1}{t_k} + \frac{1}{t_{-k}} = \gamma$$

where $\beta > 1$ and γ are finite. Note that this sampling sequence generalizes the sequence introduced in the previous section except that the null sampling time has been removed. This modification allows to model sample loss in the following.

In this case, the general reconstruction formula proposed in [8] expresses as :

$$Z(t) = H_M(t) \left[\sum_{1\le|k|\le M} a_M(t,t_k) Z(t_k) + \sum_{|k|>M} b_M(t,k) Z(k) \right] \tag{4}$$

with $a_M(t,t_k)$, $b_M(t,t_k)$ and H_M defined by:

$$a_M(t,t_k) = \frac{G_M(t_k)}{(t-t_k) F_M'(t_k) \sin\pi t_k}$$

$$b_M(t,k) = \frac{(-1)^k G_M(k)}{\pi(t-k) F_M(k)}$$

$$H_M(t) = \frac{F_M(t)}{G_M(t)} sin\pi t$$

where:

$$F_M(t) = \prod_{1\le|k|\le M} \left(1 - \frac{t}{t_k}\right)$$

$$G_M(t) = \pi t \prod_{1 \leq |k| \leq M-1} \left(1 - \frac{t}{k}\right)$$

F'_M denotes F_M derivative. According to (4), $Z(t)$ can be decomposed into two sums. The first one involves the available samples $Z(t_k)$ and the second one the unavailable regular samples $Z(k)$. The available samples are the measured samples and the so-called unavailable samples are not required in real computations. Indeed, the second sum in (4) is neglected in the following.

The hypotheses on the sampling sequence (2) imply the existence of the following analytic functions:

$$F(t) = \lim_{M \to \infty} F_M(t)$$

$$F'(t) = \lim_{M \to \infty} F'_M(t)$$

$$H(t) = \lim_{M \to \infty} H_M(t)$$

with $H(t) = F(t)$.

Moreover, the second term in equation (4) goes to 0 when M goes to infinity. Consequently, the following general reconstruction formula holds:

$$Z(t) = \lim_{M \to \infty} H_M(t) \left[\sum_{1 \leq |k| \leq M} a_M(t, t_k) Z(t_k) \right] \tag{5}$$

Note that the general formula (4) is a Lagrange interpolation formula. Indeed, the interpolation kernels $H_M(t)a_M(t,t_k)$ and $H_M(t)b_M(t,k)$ are such that:

$$H_M(t_j)a_M(t_j,t_k) = \delta_{jk} \tag{6}$$

$$H_M(j)a_M(j,t_k) = 0 \tag{7}$$

$$H_M(j)b_M(j,k) = \delta_{jk} \tag{8}$$

$$H_M(t_j)b_M(t_j,k) = 0 \tag{9}$$

where δ_{jk} denotes the Kronecker symbol ($\delta_{jk} = 1$ if $j = k$, else $\delta_{jk} = 0$). The reconstruction formula given in Eq. (5) is also a Lagrange interpolation formula when the limit and the sum can be inverted [14]. In this case:

$$Z(t) = \sum_{k \in Z^*} \frac{F(t)}{(t - t_k) F'(t_k)} Z(t_k) \tag{10}$$

Note that the Shannon formula recalled in (2) is a particular case of (5) for a periodic sampling sequence. However the limit and the sum inversion is only possible under very particular conditions on the sampling scheme. Moreover, the proof of the inversion validity is generally untractable. The most detailed theoretical justifications are given in [9]. However, the provided sufficient conditions can hardly be checked in practice, particularly in a random context [5],[8]. For instance, the displacement of only one sample time can call the condition validity in question. Another case of interest is a sample loss. The Lagrange interpolation formula applies in the case of a periodic sampling. However, if a sample is lost, the formula is no longer valid even if the oversampling condition is still fulfilled. Such a scenario can be generalized by considering two sampling sequences: the available and unavailable or lost samples. When one sample goes from one sequence to the other, the required condition validity is also called in question. This paper thus proposes a new reconstruction formula which applies in these cases. As shown in the appendix, the proposed reconstruction formula is based on the development of $e^{i\omega t}$ as a function of ω for each $t \in \mathbb{R}$. Derivations are performed in the analytical function set, under the hypothesis of the complex exponential completeness in the appropriate functional space. [1] provides an overview of the existing iterative and non-iterative reconstruction algorithms. However, to our knowledge, there do not exist other exact reconstruction formulas in the literature.

Eq. (5) provides an estimate of $Z(t)$ by truncation of the infinite sum. Indeed, for M large enough, the second sum in (4) can be neglected. First, M is fixed to a sufficiently high value. The corresponding functions $a_M(t,t_k)$ and $H_M(t)$ are then derived (note that simulations provide highly variable results for different values of M). Second, the sum in (5) is truncated to the first $2N$ terms:

$$\widehat{Z}_M(t) = \sum_{1 \leq |k| \leq N} \frac{F(t) G_M\left(\frac{k}{\beta}\right)}{\left(t - \frac{k}{\beta}\right) F'_M\left(\frac{k}{\beta}\right) \sin \pi \frac{k}{\beta}} Z\left(\frac{k}{\beta}\right) \tag{11}$$

with $N \leq M$. The appropriate choice of N as a function of M is studied through simulations in section 4.

The proposed reconstruction scheme considers that the $Z(k)$'s can be neglected for k large enough. As an extension of the proposed reconstruction scheme, the unavailable samples $Z(k)$'s could be replaced in the formulas by the $Z(k/\theta)$'s under the conditions that $\theta > 1$ (oversampling case) and that $\theta < \beta$ (lower sampling rate than the observed sequence). In this case also, the second sum can be neglected for M large enough. However, the reconstruction formula is modified through the definition and values of the function F_M. The auxiliary sampling sequence optimization is a difficult problem which is still to be addressed.

Formula (11) applies in the general frame of irregular and possibly random sampling (jitter, additive random sampling and skip sampling) provided that the sampling conditions (2) can be validated. Furthermore, it provides an alternative estimate of $Z(t)$ in the regular oversampling case defined by $\{t_k = k/\beta, k \in \mathbb{Z}^*\}$ i.e. without the origin time. The origin time suppression models any sample loss by an appropriate time index translation.

β is assumed irrational to prevent the intersection of the two available and unavailable sampling sequences. The formulas for one and two lost samples are given in the following subsections. The formula can be easily generalized when an arbitrary finite number of samples are suppressed. Note that the Papoulis generalized sampling theorem does not hold in the considered scenarios [11].

3. PARTICULAR SAMPLE LOSS CONFIGURATIONS

This section considers the following sampling sequences:

$$\{t_n = n/\beta, n \in \mathbb{Z}^*\} \tag{12}$$

$$\{t_n = n/\beta, n \in \mathbb{Z}^*\} - \left\{\frac{1}{\beta}\right\} \tag{13}$$

$$\{t_n = n/\beta, n \in \mathbb{Z}^*\} - \left\{\frac{k_1}{\beta}\right\}, k_1 \neq \pm 1 \tag{14}$$

which models a single sample loss, two adjacent sample loss and two non-adjacent sample loss respectively.

3.1 One missing sample

Consider the case of a single sample loss or deterioration. In this case, formula (11) applies directly with the following simplifications:

$$F(t) = \lim_{M\to\infty} F_M(t) = \prod_{k=-\infty, k\neq 0}^{\infty} \left(1 - \frac{\beta t}{k}\right) = \frac{sin\pi\beta t}{\pi\beta t} \quad (15)$$

leading to:

$$\widehat{Z}(t) = \frac{\sin\pi\beta t}{\pi\beta t} \sum_{1\leq|k|\leq N} \frac{G_M\left(\frac{k}{\beta}\right) Z\left(\frac{k}{\beta}\right)}{\left(t - \frac{k}{\beta}\right) F'_M\left(\frac{k}{\beta}\right) \sin\left(\pi\frac{k}{\beta}\right)} \quad (16)$$

Particularly, for $t = 0, (16)$ simplifies to:

$$\widehat{Z}(0) = \beta \sum_{1\leq|k|\leq N} \frac{-G_M\left(\frac{k}{\beta}\right)}{kF'_M\left(\frac{k}{\beta}\right) \sin\left(\pi\frac{k}{\beta}\right)} Z\left(\frac{k}{\beta}\right) \quad (17)$$

This formula performs the lost sample restoration.

3.2 Two missing samples

Now suppose that two samples have been lost. The first one is at the index $t_{k_0} = 0$ without loss of generality and the second one is such that $t_{k_1} = \frac{k_1}{\beta}$. The general formulas apply with slight modifications:

$$\widehat{Z}(t) = \sum_{1\leq|k|\leq N, k\neq k_1} \frac{F^1(t)\, G_M(t_k)}{(t - t_k)\, F_M^{1\prime}(t_k) \sin\pi t_k} Z(t_k) \quad (18)$$

with:

$$F_M^1(t) = \prod_{1\leq|k|\leq M, k\neq k_1} \left(1 - \frac{\beta t}{k}\right) = \frac{F_M^1(t)}{1 - \frac{\beta t}{k_1}} \quad (19)$$

hence

$$F^1(t) = \lim_{M\to\infty} F_M^1(t) = \frac{1}{1 - \frac{\beta t}{k_1}} \frac{sin\pi\beta t}{\pi\beta t} \quad (20)$$

$$\widehat{Z}(0) = \beta \sum_{1\leq|k|\leq N, k\neq k_1} \frac{-G_M\left(\frac{k}{\beta}\right)}{kF_M^{1\prime}\left(\frac{k}{\beta}\right) \sin\left(\pi\frac{k}{\beta}\right)} Z\left(\frac{k}{\beta}\right) \quad (21)$$

Now consider the particular case of two adjacent samples with indices $t_{k_0} = 0$ and $t_{k_1} = \frac{k_1}{\beta} = \pm\frac{1}{\beta}$. This configuration is the worst scenario for sample $Z(0)$ restoration. For simplicity and without loss of generality let $t_{k_1} = \frac{1}{\beta}$. The previous formulas apply when replacing k_1 by 1.

4. SIMULATIONS

The restoration performance is studied through the normalized mean squared error between the original lost sample and the reconstructed sample value defined by :

$$J(0) = \frac{E\left[\left(Z(0) - \hat{Z}(0)\right)^2\right]}{E\left[Z(0)^2\right]} \quad (22)$$

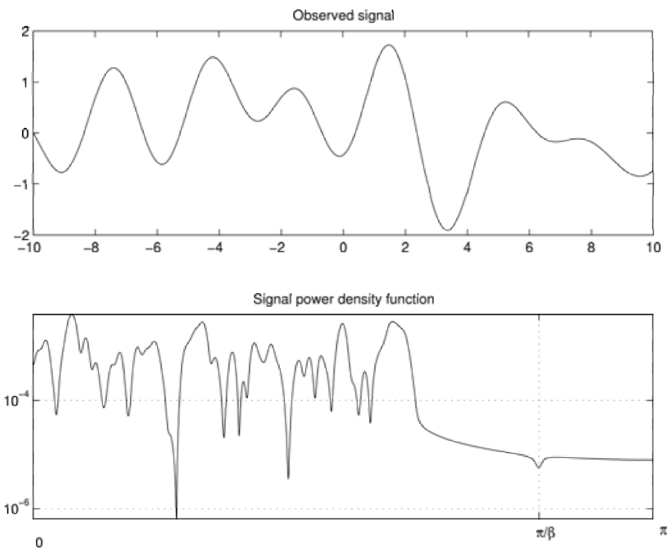

Figure 1: Original signal and associated power spectral density

$J(0)$ is estimated through n_R runs. The observed signal is a filtered white Gaussian noise with initial spectrum strictly contained in $(-\pi + a, +\pi - a)$ with $a > 0$ oversampled at a rate $\beta > 1$. Figure 1 shows a signal run and the associated power spectrum density for $a = \frac{\pi}{4}$ and $\beta = 1.225$. First, the restoration performance is studied as a function of the number of available samples for a given sampling rate β. M samples are assumed available and the reconstruction formula performance is studied using $N \leq M$ samples and an order M approximation for the interpolation functions. The proposed formulas obtained for different values of M: $M = N$, $M = N+1$, $M = N+2$, $M = 2N$ and $M = 3N$ are compared. Figures 2 and 3 show that the most powerful restoration is obtained for $M = N$ (use of the whole available samples) in (4). Moreover, the proposed formula leads to a decreasing mean square error with a very high convergence rate. The same behavior is observed in the case of one or two lost samples, with a larger error in the second case.

Simulations reveal the numerical instability of the interpolation functions for small M particularly through $F_M(t)$. This results in a high normalized error for small N values. Hence a minimum number of available samples is required. Fig.4 displays the original and reconstructed signal over $[-3T, 3T]$ for increasing values of N and for $M = N$. The performance has also been studied as a function of the sampling rate β for a given number of the available samples such that $N = M = 8$. As β increases, the convergence rate of the proposed formula increases.

5. CONCLUSION

The reconstruction of a random process from its sample has been addressed in the oversampling case. The particular application of lost sample restoration has been studied. The proposed formula stems from Lagrange interpolation formula but explicitly takes into account the possible loss of one or more samples. The basic principle is the decomposition of the signal expansion into two series. The first one uses the

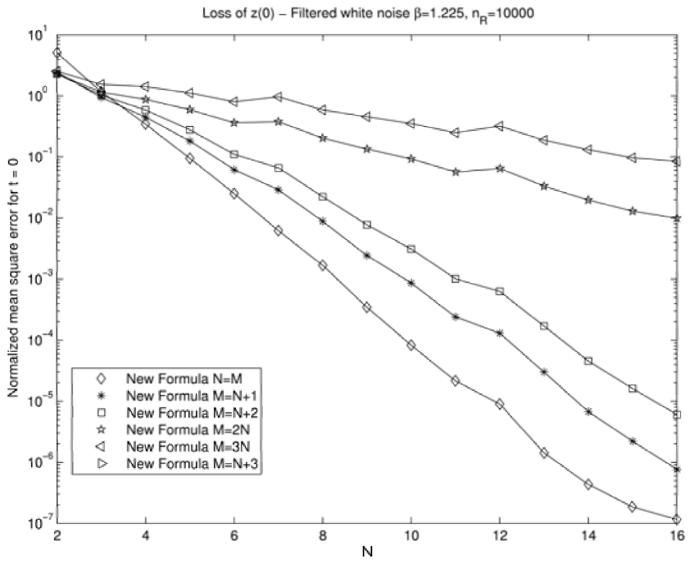

Figure 2: Case of one missing sample

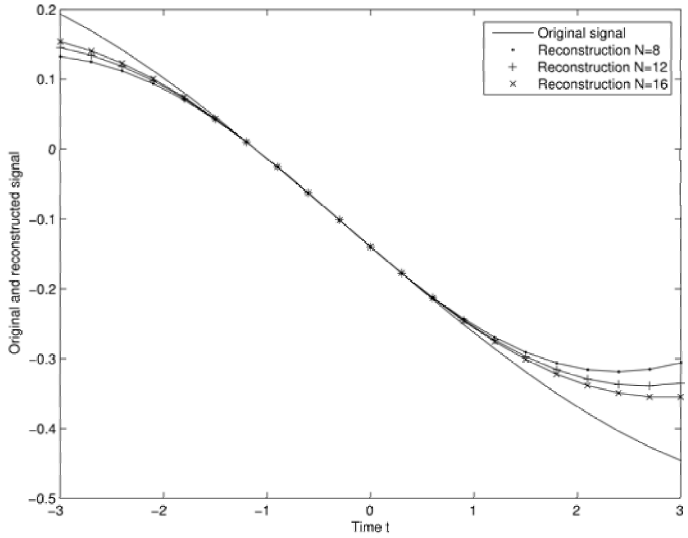

Figure 4: Original and reconstructed signal for different values of N

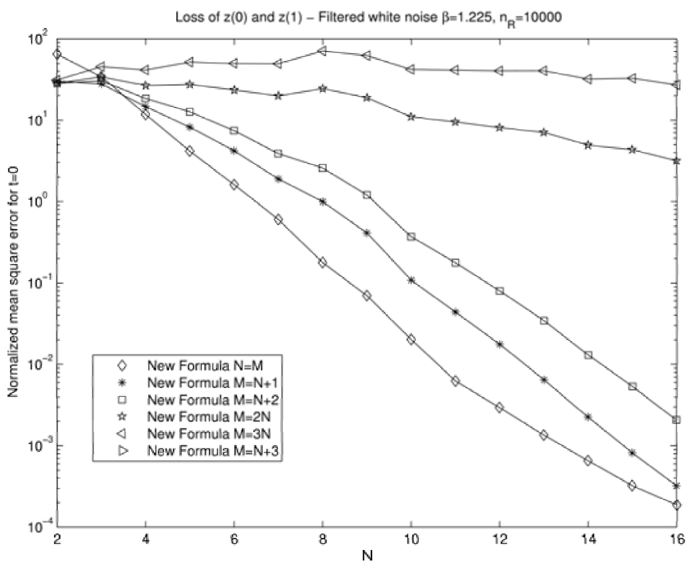

Figure 3: Case of two adjacent missing samples

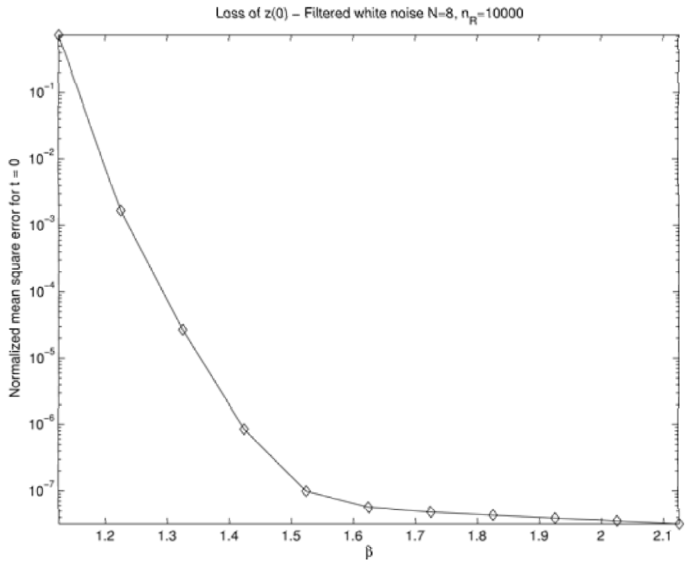

Figure 5: Influence of the over-sampling rate

available samples and the other one an auxiliary sample sequence. The performance has been studied through simulations. The proposed scheme shows a very high convergence rate even in the case of adjacent sample loss. This paper proposes $\{Z(k), k \in \mathbb{Z}\}$ as the auxiliary sample sequence. However, other auxiliary sample sequences could be proposed in future works. Indeed, the optimization of the auxiliary sample sequence is a difficult problem which is beyond the scope of this paper.

6. APPENDIX

Assume that the Fourier transform of $K_Z(\tau) = E[Z(t)Z^*(t-\tau)]$ verifies (1). If $e^{i\omega t}$ can be developed in some sense as follows:

$$e^{i\omega t} = \sum_{n=-\infty}^{n=+\infty} \alpha_n(t)\, e^{i\omega t_n} \tag{23}$$

then the random process $Z(t)$ can be retrieved from the following expansion:

$$Z(t) = \sum_{n=-\infty}^{n=+\infty} \alpha_n(t)\, Z(t_n) \tag{24}$$

Let consider the function $\gamma_M(z)$ defined for $t \notin \mathbb{Z} \bigcup \mathbf{t}$ by:

$$\gamma_M(z) = \frac{e^{i\omega z}}{(z-t)sin\pi z} \frac{G_M(z)}{F_M(z)} \tag{25}$$

The proof of the reconstruction formula stems from the residue theorem [12]. This theorem allows to integrate $\gamma_M(z)$ on a closed curve Γ_d. Let Γ_d denote a square centered at the origin which perpendicularly cuts the axes at a distance $d+\frac{1}{2}$ from the origin.

$$\frac{1}{2i\pi}\int_{\Gamma_d}\gamma_M(z) = Res[\gamma_M,t] + \sum_{M\leq|k|\leq d} Res[\gamma_M,k] + \sum_{k=-M}^{k=M} Res[\gamma_M,t_k] \quad (26)$$

Let consider d_M such that, for $d \geq d_M$, the subsequence $\{t_k, k \leq M+1\}$ is strictly included in Γ_d for $d > d_M$. Note that $\frac{G_M(z)}{F_M(z)}$ is uniformly bounded on Γ_d for $d > d_M$ since $G_M(z)$ and $F_M(z)$ are polynomial with respective degrees $2(M-1)$ and $2M$. Under the additional condition that $|\omega| < \pi$, for any M:

$$\lim_{d\to\infty}\frac{1}{2i\pi}\int_{\Gamma_d}\gamma_M(z) = 0 \quad (27)$$

Residues are easy to derive since the singular points of γ_M are poles with order 1. Finally, $e^{i\omega t}$ can be expanded as follows:

$$e^{i\omega t} = sin\pi t\frac{G_M(t)}{F_M(t)}\left[\sum_{1\leq|k|\leq M}\frac{G_M(t_k)}{(t-t_k)F'_M(t_k)\sin\pi t_k}e^{i\omega t_k} + \sum_{|k|>M}\frac{(-1)^k G_M(k)}{\pi(t-k)F_M(k)}e^{i\omega k}\right] \quad (28)$$

for $-\pi < \omega < \pi$ and $t \notin \mathbb{Z}\bigcup\mathbf{t}$. Since the convergence of the infinite sum is uniform in ω on any $\Delta \subset [-\pi,\pi]$, Eq. (4) holds for each M. The second sum can be neglected for M large enough. Indeed, for a given t, for instance near 0, the condition $\frac{t_k}{k}$ goes to $\frac{1}{\beta}$ assures that the time t will be surrounded by the set of t_k and that the integers k will be far away. Then the corresponding terms can be neglected. For example, if $\beta = 1.7$, the interval $(-10,10)$ contains 34 instants t_k, and the nearest terms in the second sum is for $Z(36)$. These terms and the following ones have no influence on $Z(0)$. From a mathematical point of view, the proof can be deduced from the inequality

$$\sup_{|x|>M}\left|\prod_{1\leq|m|\leq M}\frac{1-\frac{x}{m}}{1-\frac{x}{t_m}}\right| < \eta \quad (29)$$

Consequently, b_M can be expressed as:

$$b_M(t) = \sum_{|m|>M}\frac{(-1)^m u(m)}{t-m}\frac{m}{1-\frac{m}{t_0}}\frac{1}{(1-\frac{m}{t_{-1}})(1-\frac{m}{t_1})}\prod_{1\leq|l|\leq M}\frac{1-\frac{m}{l}}{1-\frac{m}{t_{l+sgnl}}} \quad (30)$$

where $sgnl = 1$ for $l > 0$ and $sgnl = -1$ for $l > 0$. After slight changes in the order of $\mathbf{t}$, Eq. (29) leads to:

$$|b_M(t)| = \sum_{|m|>M}\frac{\alpha}{m^3} \quad (31)$$

for bounded $u(t)$ and $M > M_0$. Consequently, $|b_M(t)|$ can be viewed as the remainder of a convergent series, which leads to:

$$\lim_{M\to\infty}|b_M(t)| = 0 \quad (32)$$

REFERENCES

[1] P. J. S. G. Ferreira, *Iterative and Noniterative Recovery of Missing Samples for 1-D Band-Limited Signals*, in Nonuniform Sampling, ed. F. Marvasti, Kluwer Ac. (2001).

[2] J. R. Higgins, *Some Gap Sampling Series for Multiband Signals,* Signal Proc. 12 (1987) 313-319.

[3] A. J. Jerri, *The Shannon sampling theorem. Its various extensions and applications.* A tutorial review, Proc. IEEE 65 (11) (1977) 1565-1596.

[4] B. Lacaze, *Reconstruction of Stationary Processes sampled at Random Times*, in Nonuniform Sampling-Theory and Practice, Marvasti ed. (2001) 361-390.

[5] B. Lacaze, *A Theoretical Exposition of Stationary Processes Sampling,* Samp. Th. in Sign. Im. Proc. 4 (3) (2005) 201-230.

[6] B. Lacaze, *Stationary clock changes on stationary processes,* Signal Proc. 55 (2) (1996) 191-206.

[7] B. Lacaze, M. Chabert, *Improved reconstruction formulas for oversampled processes and functions,* Signal Proc. (2006).

[8] B. Lacaze, *Reconstruction formula for irregular sampling,* Samp. Th. in Sign. Im. Proc. 4 (1) (2005) 33-43.

[9] N. Levinson, *Gap and density theorems,* Am. Math. Soc. Coll. Pub. Vol. 26 (1940).

[10] B. Ja. Levin, *Zeros of Entire Functions,* Am. Math. Soc. (1964).

[11] A. Papoulis, *Signal Analysis,* McGraw-Hill, (1977).

[12] W. Rudin, *Real and complex analysis,* McGraw-Hill, (1966).

[13] J. L. Yen, *On Nonuniform Sampling of Bandwith-Limited Signals,* IRE Trans. on Circ. Th., 1956, CT-3 (12) 251-257.

[14] A. I. Zayed, P. L. Butzer, *Lagrange Interpolation and Sampling Theorems*, in Nonuniform Sampling, ed. F. Marvasti, Kluwer Ac. (2001).