Simple Errorless Formulas when Missing Samples

Bernard Lacaze*

14/16 Port Saint-Etienne, 31000 Toulouse, France

Abstract: Even if it is an idealization, the band-limited process is widely taken as model in signal processing and in communications. The classical Shannon formula is exact for the unit rate sampling and a spectral support of 2π-length. It is no longer error-free when one sample is lost and replaced by an estimation, because the set of functions $\{e^{in\omega}, n\epsilon\mathbf{Z}\}$ is free and complete in $\mathbf{L}^2(-\pi,\pi)$. Then, an exact reconstruction can occur only when the process is oversampled. In this context, iterative procedures exist [1], but not analytic formulas, from apart these in [2], which have an uncontrolled convergence. In this paper, we give simple formulas when one or two samples are missing, but which can be generalized to a number of erased samples larger than two. We show that the reintroduction of ignored samples can improve the convergence of the formulas. The links with the Lagrange interpolation formula are highlighted.

Keywords: Stationary processes, irregular sampling, Lagrange interpolation, band-limited processes.

1. INTRODUCTION

Let $\mathbf{Z}=\{Z(t), t\epsilon\mathbf{R}\}$ be a (real or complex) stationary process, owing a regular spectral density $s(\omega)$, such as [3]

$$\mathrm{E}[Z(t)Z^*(t-\tau)] = \int_{-\pi+a}^{\pi-a} e^{i\omega\tau}s(\omega)\,d\omega, 0 < a < \pi.$$
(1)

Such a process can be reconstructed from the series of random samples $\mathbf{Z}'=\{Z(n), n\epsilon\mathbf{Z}\}$, because the "Nyquist condition" is fulfilled [2, 4, 5]. For example, by the so-called Shannon formula

$$Z(t) = \sum_{n\in\mathbb{Z}} \frac{\sin\pi(t-n)}{\pi(t-n)} Z(n).$$
(2)

The formula (2) is exact with respect to the mean-square distance (m.s convergence) but other kinds of convergence can be used [6]. Formula (2) cannot be used when a finite or infinite number of samples is lost, but the necessary information exists for reconstructing $Z(t)$, as long as the mean number of remaining samples is sufficient [7]. General exact formulas can be given [8], completing the sequence of available samples by elements whose influence is negligible. Also, reconstruction by recursive algorithms is currently used [1]. Furthermore, formula (2) belongs to the class of Lagrange interpolation formula (LIF) because of the familiar equality

$$\sin\pi z = \pi z \prod_{n\in\mathbb{Z}^*} \left(1 - \frac{z}{n}\right).$$

Conditions for applying this kind of formulas are well-known, but difficult to be applied, apart from rare cases [9].

*Address correspondence to this author at the 14/16 Port Saint-Etienne, 31000 Toulouse, France; E-mail: bernard.lacaze@tesa.prd.fr

For example, the LIF is divergent when a finite number of samples is missing. But, surprisingly, it is possible to have convergent and available LIF when an infinite number of samples is suppressed. It is the basis of this paper, where we construct simple formulas, which are exact when a finite number N of samples is erased. Analytic expressions are given for $N=1$ and 2, but the method can be iterated for larger values of N. The proofs are given in the Appendix and are based on the complex functions theory, widely used in sampling theory [5].

2. ONE MISSING SAMPLE

We are in the case where the power spectrum is inside $(-\pi+a,\pi-a)$. It is well known that the set of functions $\{exp[it_n\omega], n\epsilon\mathbf{Z}\}$ is complete for sets like $\mathbf{L}^2(-\pi+a,\pi-a)$, for instance when [7].

$$\lim_{|n|\to\infty} \frac{t_n}{n} = \alpha < 1 - \frac{a}{\pi}$$

Therefore, a sufficient amount of information for reconstructing the process is contained in the set of samples $\{Z(n), n\epsilon\mathbf{Z}-\lambda\mathbf{Z}\}$, when l is an integer chosen such as

$$l > \frac{\pi}{a}.$$
(3)

Then, the following sampling formula is available (with respect to the mean-square convergence)

$$Z(t) = \frac{\sin\pi t}{\sin(\pi t/l)} \sum_{k\in\mathbf{J}_l} \frac{(-1)^k \sin(\pi k/l)}{\pi(t-k)} Z(k)$$
(4)

Where $\mathbf{J}_l$ is the set of relative integers k which are not multiple of l, i.e $k=ql+q'$, with $q\epsilon\mathbf{Z}$, $q'\epsilon\{1,2,...,l-1\}$. In this formula, the lost sample $Z(0)$ is not used, but neither are the available samples $Z(ql),|q|=1,2,...$ Obviously, we can take in

(4) **Z** instead of **J**$_l$ because $\sin(\pi k/l)=0$ outside **J**$_l$. The proof is given in the Appendix.

particularly, the lost sample $Z(0)$ verifies the formula

$$Z(0) = \sum_{k \in \mathbf{J}_l} (-1)^{k+1} \frac{\sin(\pi k/l)}{\pi k/l} Z(k)$$

(5)

This value can be reintroduced in the formula (2). Obviously, the convergence of (4) or (5) depends on the value of l. The larger is l, the slower is the speed of convergence of the series $\mathrm{sinc}(\pi k/l)$. But, at the same time, the rate of forgotten samples (i.e $...Z(-l),Z(0),Z(l),Z(2l)..$) decreases, and then, also the loss of information. Consequently, a compromise has to be found about the value of l.

For example, if $a=0.11\pi$, we have $l \geq 10$. For the minimum value $l=10$, the loss of information is small, but the weights of the $Z(k)$ in (5) vary slowly, as shown in the following table

$Z(1)$	$Z(2)$	$Z(3)$	$Z(4)$	$Z(5)$	$Z(6)$	$Z(7)$	$Z(8)$	$Z(9)$	$Z(11)$
0.99	0.94	0.86	0.82	0.64	0.54	0.37	0.23	0.11	−0.09

$Z(12)$	$Z(13)$	$Z(14)$	$Z(16)$	$Z(18)$	$Z(25)$	$Z(35)$	$Z(45)$
−0.16	−0.20	−0.22	−0.19	−0.10	0.12	−0.09	0.07

This shows a slow convergence of the series (5). Obviously, this fact is explained by a correlation divided between too numerous samples. The correlation coefficient between $Z(0)$ and $Z(1),Z(2),Z(3),Z(4)....$ is equal to $0.12,0.11,0.10,0.09...$ when the spectrum is uniform. Then, distant samples have to intervene in the interpolation formula.

If we take $a=0.34\pi$, then $l \geq 3$. For $l=3$, the loss of the samples $Z(\pm 3),Z(\pm 6)...$ is not negligible, but the weights decrease as follows (the formula is symmetrical)

$Z(1)$	$Z(2)$	$Z(4)$	$Z(5)$	$Z(7)$	$Z(8)$	$Z(10)$	$Z(11)$
0.83	−0.41	0.21	−0.16	0.12	−0.10	0.08	−0.07

In this case, the samples are more linked with their first neighbours. For a uniform spectrum, the first correlations are equal to $0.42,-0.20,-0.01,0.11...$ It seems that the best (the less bad?) convergence appears for the smallest value of l compatible with the power spectrum (i.e verifying (3)). Moreover, it is possible to use several values of l, and to compare the results. A higher confidence will be achieved when the different estimates do not differ significantly.

For $a>0.5\pi$, we can take $l=2$

$$Z(t) = \sum_{n \in \mathbf{Z}} \frac{\sin \frac{\pi}{2}(t-2n-1)}{\frac{\pi}{2}(t-2n-1)} Z(2n+1)$$

(6)

Particularly, $Z(0)$ can be given in function of the $Z(2n+1)$

$$Z(0) = \sum_{n \in \mathbf{Z}} \frac{2(-1)^n}{\pi(2n+1)} Z(2n+1)$$

Actually, (6) is a version of the Shannon formula (2) applied when the sampling period is 2 (with a gap of one unit

on the sampling times). The loss of information is due to the disappearance of the even samples. In this case, to take $l=3$ or 4 seems a better solution. But, when $l=2$, it is possible to reintroduce the samples $Z(-2)$ and $Z(2)$ in the interpolation formula (see the Appendix).

$$Z(0) = -\frac{Z(2)}{2} - \frac{Z(-2)}{2} + \frac{8}{\pi}$$

$$\sum_{n \in \mathbf{Z}} \frac{(-1)^{n+1}}{(2n+1)\left((2n+1)^2 - 4\right)} Z(2n+1)$$

(7)

The coefficients linked to the $Z(n)$ decrease like n^{-3}, which shows a fast convergence. Note that it is an exact formula, which is not the case for interpolations based on splines or wavelets. The first coefficients are

$Z(1)$	$Z(2)$	$Z(3)$	$Z(5)$	$Z(7)$	$Z(9)$
0.85	−0.5	0.17	−0.02	0.008	0.004

It is also possible to reintroduce the samples $Z(\pm 4)$. In this case, the first coefficients become

$Z(1)$	$Z(2)$	$Z(3)$	$Z(4)$	$Z(5)$	$Z(7)$
0.91	−0.66	0.54	−0.17	0.04	−0.002

The reintroduction of the samples $Z(\pm 3)$ can also be made when $l=3$, to improve the convergence. Note that the series

$$\sum_{n \in \mathbf{Z}^*} \frac{n \sin \pi(t-n)}{\pi t(t-n)} b_n$$

is a LIF (Lagrange Interpolation Formula) applied to the set $\mathbf{Z}^*$, which converges only if the b_n converge to zero fast enough. It is not the case when $b_n=Z(n)$ because of the stationarity. At the opposite, it is easy to verify that (4) is a (convergent) LIF applied to the set $\mathbf{J}_l$, because, if $f(t)=\sin\pi t/\sin(\pi t/l)$, we have the equalities

$$\begin{cases} f(k) = 0, & f'(k) = \frac{\pi(-1)^k}{\sin(\pi k/l)}, & k \in \mathbf{J}_l \\ Z(t) = \sum_{k \in \mathbf{J}_l} \frac{f(t)}{f'(k)(t-k)} Z(k). \end{cases}$$

Finally, when one sample is lost, it is possible to give exact formulas of reconstruction, with convergence in n^{-q}, with arbitrary q.

3. TWO MISSING SAMPLES

We assume that $Z(0)$ and $Z(b)$ are lost. If the integer l verifies (3) and if $b \in l\mathbf{Z}^*$, then the formulas (4) and (5) are true. Moreover, we have (when $b=ml$)

$$Z(b) = \sum_{k \in \mathbf{J}_l} \frac{(-1)^{k+m(l-1)} l \sin(\pi k/l)}{\pi(ml-k)} Z(k).$$

(8)

When $|b|$ is too small, it is not possible to have simultaneously $b \in l\mathbf{Z}^*$ and $l>\pi/a$. In this situation, we will take l such as

$$l > \frac{2\pi}{a}, \quad b = ml + m', \quad m' \in \{1, 2, .., l-1\}.$$
(9)

Then, the following formula is available (see the Appendix)

$$Z(t) = \frac{\sin \pi t}{\sin \frac{\pi}{l} t \sin \frac{\pi}{l}(t-m')} \sum_{k \in \mathbf{K}_l}$$

$$\frac{(-1)^k \sin\left(\frac{\pi}{l} k\right) \sin \frac{\pi}{l}(k-m')}{\pi(t-k)} Z(k)$$
(10)

where $\mathbf{K}_l$ is the set of integers, which do not cancel the product $\sin(\pi z/l)\sin(\pi(z-m')/l)$, i.e different from kl and of $kl+m'$. Particularly, the lost samples are retrieved from

$$\begin{cases} Z(0) = \frac{l}{\sin \frac{\pi}{l} m'} \sum_{k \in \mathbf{K}_l} \frac{(-1)^k \sin\left(\frac{\pi}{l} k\right) \sin \frac{\pi}{l}(k-m')}{\pi k} Z(k) \\ Z(b) = \frac{l}{\sin \frac{\pi}{l} m'} \sum_{k \in \mathbf{K}_l} \frac{(-1)^{k+b+1} \sin\left(\frac{\pi}{l} k\right) \sin \frac{\pi}{l}(k-m')}{\pi(k-b)} Z(k) \end{cases}$$
(11)

The convergence of (11) is poor. To improve it, we can reintroduce few samples, as explained in the case *N=1*. For example, let us place ourselves in the case *a=0.51* with a flat power spectrum. Let assume that the samples *Z(0)* and *Z(1)* are lost. We can take *l=4* (it is a minimum), with *m=0*, *m′=1*. The coefficients in the (asymmetrical) first formula of (11) are

$Z(-10)$	$Z(-9)$	$Z(-6)$	$Z(-5)$	$Z(-2)$	$Z(-1)$
-0.12	0.14	-0.21	0.25	-0.63	1.27

$Z(2)$	$Z(3)$	$Z(6)$	$Z(7)$	$Z(10)$	$Z(20)$	$Z(30)$
0.63	-0.42	0.21	-0.18	0.13	0.06	0.04

and they decrease like n^{-1} (from the high value *1.27*). If *Z(-3)* and/or *Z(±4)* are reintroduced (the nearest to *Z(0)* suppressed sample(s)), then the coefficients will decrease in n^{-2} (or n^{-3} or n^{-4}), but the weights of the nearest used samples increase surprisingly. For example, if only *Z(-3)* is added, the coefficients become

$Z(-20)$	$Z(-8)$	$Z(-7)$	$Z(-6)$	$Z(-5)$	$Z(-3)$
0.01	-0.10	0.14	-0.21	0.38	1

$Z(-2)$	$Z(-1)$	$Z(2)$	$Z(3)$	$Z(6)$	$Z(7)$	$Z(20)$
1.91	$-1.91.$	-0.38	0.21	-0.07	0.05	0.008

As in the case *N=1*, the application of these formulas to the unused samples can allow to verify them. Furthermore, the method can be iterated as long as the set of erased samples remains finite. Of course, the formulas become more and more complicated. As in the case *N=1*, (10) is a convergent LIF.

When available samples are not reintroduced to improve the convergence of interpolations, we are in a sampling scheme, which was used for example in [10-12], where the used samples are members of some sets in the form $\{nl+l',\ n \in Z\}$.

4. CONCLUSION

The recovery of missing samples is a problem, which can be treated separately [1] or inside the frame of the irregular or nonuniform sampling [8]. In this paper, we study bandlimited stationary processes, with sampling rate above the Nyquist bound. The problem is to reconstruct exactly the process when a finite number of samples are erased. In a similar way, we have to find the exact value of the missing samples in a (linear) function of the available samples. We notice that it can be worthwhile to construct interpolation formulas which do not take into account periodic sequences of samples containing erased samples. This method has at least three benefits: (theoretically) exact reconstruction, possible verification with unused samples, or by using different formulas. This last aim can be reached by changing the period of unused samples, and /or adding samples. Actually, the obtained formulas are Lagrange interpolations, which are not true when all available samples are used.

5. APPENDIX

1) We look at the complex function *f(z)* defined by

$$f(z) = \frac{e^{iz\omega} \sin(\pi z/l)}{(z-t) \sin \pi z}$$

f(z) is integrated on the square C_n, *n* large enough, with horizontal sides at the ordinates $\pm(nl+\frac{1}{2})$ and centred at the origin. Using the theorem of residues [13], we obtain (with order one poles on $\mathbf{J}_l$)

$$\frac{1}{2i\pi} \int_{C_n} f(z)\, dz = \frac{e^{i\omega t} \sin(\pi t/l)}{\sin \pi t} + \sum_{k \in \mathbf{J}_l} \frac{(-1)^k e^{ik\omega} \sin(\pi k/l)}{\pi(k-t)}$$
(12)

If we take into account the behaviour of *f(z)* on C_n:

$$\begin{cases} \left| f\left(x \pm i\left(nl+\frac{1}{2}\right)\right) \right| \approx \frac{1}{n} \exp\left[\pm nl\omega + (1-l)\pi n\right] \\ \left| f\left(\pm\left(nl+\frac{1}{2}\right)+iy\right) \right| \approx \frac{1}{n} \exp\left[-\omega y + \left(\frac{1}{l}-1\right)|y|\right] \end{cases}$$

and if we suppose that $\omega \in [-\pi(l-1)/l,\ \pi(l-1)/l]$ then

$$\lim_{n \to \infty} \frac{1}{2i\pi} \int_{C_n} f(z)\, dz = 0$$
(13)

From (12) and (13) we deduce the equality

$$e^{i\omega t} = \frac{\sin \pi t}{\sin(\pi t/l)} \sum_{k \in \mathbf{J}_l} \frac{(-1)^k \sin(\pi k/l)}{\pi(t-k)} e^{i\omega k}$$

The interval $[-\pi(l-1)/l,\ \pi(l-1)/l]$ contains $[-\pi+a,\ \pi-a]$, because of (3). The familiar isometry between functional spaces yields (4), replacing in this formula $e^{i\omega t}$ by $Z(t)$ [3].

2) Let assume that *N=1*, *a>π/2*, *l=2*. With the same kind of integration contour, we consider the complex function

$$f(z) = e^{i\omega z} \frac{\sin(\pi z/2)}{(z-t)(z^2-4) \sin \pi z}$$

Where the poles ± 2 reappear. With such $f(z)$, we obtain, for $\omega \in [-\pi/2, \pi/2]$

$$\frac{e^{i\omega t}}{\cos \pi t/2} = -\frac{t+2}{4}e^{2i\omega} + \frac{t-2}{4}e^{-2i\omega} +$$

$$\frac{2}{\pi}\sum_{n\in\mathbf{Z}}\frac{(-1)^n(t^2-4)e^{(2n+1)i\omega}}{(2n+1-t)\left((2n+1)^2-4\right)}$$

then (7) for $t=0$.

3) When two samples are lost, we take $f(z)$ as

$$f(z) = e^{i\omega z}\frac{\sin\frac{\pi}{l}z\sin\frac{\pi}{l}(z-m')}{(z-t)\sin\pi z}$$

Where $m' \in \{1,2,...,l-1\}$. The singularities of $f(z)$ are real poles of order one at the integers different of kl and $kl+m'$, $k \in Z$. The theorem of residues leads to

$$\begin{cases}\lim_{n\to\infty}\frac{1}{2i\pi}\int_{C_n}f(z)\,dz = \frac{e^{i\omega t}\sin\left(\frac{\pi}{l}t\right)\sin\left(\frac{\pi}{l}(t-m')\right)}{\sin\pi t} + \\ \cdots\sum_{k\in\mathbf{K}_l}\frac{(-1)^k e^{ik\omega}\sin\left(\frac{\pi}{l}k\right)\sin\left(\frac{\pi}{l}(k-m')\right)}{\pi(k-t)} = 0\end{cases}$$

Where the sequence C_n goes through the abscissa axis at appropriate points.

4) If $N>2$ samples are lost, we will take $l>N\pi/a$, and the corresponding sampling times t_k will be noted

$$t_k = m_k l + m'_k, \quad k = 1, 2, .., N, \quad m'_k \in \{1, 2, .., l-1\},$$

assuming that the m'_k are different (it is not difficult to withdraw this hypothesis).

We apply the residue's theorem to the complex function

$$f(z) = \frac{e^{i\omega z}}{(z-t)\sin\pi z}\prod_{k=1}^{N}\sin\frac{\pi}{l}(z-m'_k).$$

Therefore, we generalize (10) to

$$Z(t) = \frac{\sin\pi t}{\prod_{n=1}^{N}\sin\frac{\pi}{l}(t-m'_n)}\sum_{k\in\mathbf{K}_l}\frac{(-1)^k\prod_{n=1}^{N}\sin\frac{\pi}{l}(k-m'_n)}{\pi(t-k)}Z(k)$$

Where $\boldsymbol{K}_l = \mathbf{Z} - U_{k=1...N}(\mathbf{Z}+m'_k)$ is the set of order one poles of $f(z)$. Like soon explained, the convergence rate is improved, reintroducing samples.

REFERENCES

[1]   P. J. S. G. Ferreira, "Iterative and noniterative recovery of missing samples for 1-D band-limited signals," In *Nonuniform Sampling*, F. Marvasti, Ed., NewYork: Kluwer Academic/Plenum Publishers, 2001.

[2]   B. Lacaze, "Reconstruction formula for irregular sampling, Sampling," *Theory Signal Image Process*, vol. 4, No. 1, pp. 33-43, 2005.

[3]   H. Cramer, and M. R. Leadbetter, *Stationary and related stochastic processes*, New York: Wiley, 1966.

[4]   A. J. Jerri, "The Shannon sampling theorem. Its various extensions and applications. A tutorial review," *Proceedings of IEEE*, vol. 65, No. 11, 1977, pp.1565-1596.

[5]   J. R. Higgins, "Sampling Theory in Fourier and Signal Analysis," USA: Oxford University Press, 1996.

[6]   T. Pogany, "Almost sure sampling restoration of bandlimited stochastic signals," ln: J. R. Higgins, R. L. Stens, Eds., *Sampling Theory in Fourier and Signals Analysis: Advanced Topics*, USA: Oxford University Press, 1999, pp. 203-232.

[7]   B. Ja. Levin, "Zeros of entire functions," *Transactions on Mathematical Monographs 5*, Providence, Rhode Island: American Mathematical Society, 1964, p. 493.

[8]   B. Lacaze, "A Theoretical exposition of stationary processes sampling," *Sampling Theory Signal Image Process*, vol. 4, No. 3, pp. 201-230, 2005.

[9]   K. Seip, "An irregular sampling theorem for functions bandlimited in a generalized sense," *SIAM Journal Applied of Mathematics*, vol. 47, No. 5, pp. 1112-1116, 1987.

[10]   J. L. Yen, "On Nonuniform Sampling of Bandwidth-Limited Signals," *IRE Transactions on Circuit Theory*, vol. 12, *CT-3*, pp. 251-257, 1956.

[11]   J. R. Higgins, "Some gap samplings for multiband signals," *Signal Processing*, vol. 12, No. 3, pp. 313-319, 1987.

[12]   V. I. Smirnov, *A Course of Higher Mathematics*, USA: Pergamon Press, 1964.