

Formules de reconstruction exactes pour processus suréchantillonnés avec échantillons manquants

Bernard Lacaze, Nathalie Thomas

IRIT - ENSEEIHT - TésA, 2 rue Charles Camichel, BP 7122, 31071 Toulouse cedex 7, France

Bernard Lacaze@tesa.prd.fr, Nathalie Thomas@enseeiht.fr

Résumé – L'article décrit une méthode de reconstruction de signaux à bande limitée quand un nombre fini d'échantillons est manquant. La reconstruction est possible à condition que le processus soit suréchantillonné par rapport à la borne de Shannon. Dans ce cas, des formules de reconstruction exactes sont données et leur convergence étudiée. Ces formules sont obtenues en supprimant périodiquement d'autres échantillons. Nous montrons que la réintroduction des échantillons ignorés améliore leur convergence. Des simulations menées sur des signaux de type racine de cosinus surélevé (signaux à bande limitée très utilisés en télécommunications) permettent de valider les formules de reconstruction proposées.

Abstract – The paper describes a reconstruction method for band-limited signals when a finite number of samples is missing. An exact reconstruction can occur if the process is oversampled compared to Shannon bound. In that case exact reconstruction formulas are given and their convergence is studied. To obtain these formulas, some samples are periodically suppressed and it is shown that their reintroduction improves the convergence. Given reconstruction formulas are validated by simulations considering a real raised cosine shaping signal, widely used in telecommunications. This process has a limited and adjustable bandwidth, allowing to reconstruct it even when samples are lost provided it has been oversampled.

1 Introduction

On suppose que $\mathbf{Z} = \{Z(t), t \in \mathbb{R}\}$ est un processus aléatoire stationnaire avec une densité spectrale de puissance régulière sur $[-\pi + a, \pi - a]$ ($a > 0$). $Z(t)$ peut être reconstruit à l'aide de la formule classique

$$Z(t) = \sum_{n \in \mathbb{Z}} \frac{\sin \pi(\frac{t}{T} - n)}{\pi(\frac{t}{T} - n)} Z(nT). \quad (1)$$

à condition que $T < \frac{\pi}{\pi - a}$. Plus généralement, l'ensemble des échantillons $Z(t_n)$, $n \in \mathbb{Z}$ contient une information suffisante pour une parfaite reconstruction linéaire de $Z(t)$ lorsque la suite des t_n est suffisamment régulière et telle que :

$$\lim_{|n| \rightarrow \infty} \frac{t_n}{n} < \frac{\pi}{\pi - a}. \quad (2)$$

Des formules d'interpolation peuvent être construites dans ce cadre très général [2], mais elles sont d'implantation délicate. D'autres stratégies sont possibles dans des cas particuliers, notamment dans le cas de l'échantillonnage périodique avec un nombre fini d'échantillons manquants, situation qui est étudiée dans cet article. Les algorithmes les plus utilisés sont récursifs, et leur propriétés sont bien connues [4]. La stratégie utilisée dans cet article est différente. Elle est basée sur les considérations décrites ci-après.

[7] a construit des formules pour le cas où les échantillons sont pris aux instants distincts $t_{nk} = nT + a_k$, $n \in \mathbb{Z}$, $k = 1, 2, \dots, l$ (dans la nomenclature moderne, il s'agit d'un PNSI pour Periodic Nonuniform Sampling of order 1). Puisqu'un retard est un filtre linéaire invariant, on est aussi dans le cadre de la "généralisation de Papoulis" [1], [8] dans un schéma comprenant l filtres en parallèle. Si l'on prend T

entier et $a_k \in \{1, 2, \dots, T - 1\}$, les t_{nk} sont tous les entiers sauf les multiples de T . Sous réserve que le taux moyen d'échantillonnage soit suffisant (condition de Nyquist), on obtient des formules où l'échantillon $Z(0)$ n'apparaît pas (en admettant que c'est celui-là qui a été effacé). Le prix de cette opération est que l'on néglige les $Z(nT)$, $n \neq 0$, mais l'information qu'ils apportent est redondante. D'autre part, la convergence des formules obtenues est médiocre (le terme de rang n décroît en $1/n$, comme dans les formules d'interpolation de Lagrange).

L'apport de cet article consiste à construire des formules exactes qui convergent avec un taux réglable (n'importe quelle puissance de n). Pour ce faire, on réintroduit un certain nombre d'échantillons négligés, par exemple $Z(T)$ et $Z(-T)$ pour une convergence en n^{-3} . Le procédé peut être généralisé avec le même avantage dans le cas où le nombre d'échantillons effacés est quelconque.

2 Un échantillon manquant

2.1 Formule de reconstruction

Supposons que $Z(0)$ soit l'échantillon manquant (l'hypothèse de stationnarité justifie ce choix). Pour

$$l > \frac{\pi}{a}, l \in \mathbb{N}, \quad (3)$$

(cette hypothèse assure que le nombre d'échantillons retenus est suffisant), on a :

$$Z(t) = \frac{\sin \pi t}{\sin(\pi t/l)} \sum_{k \in J_l} \frac{(-1)^k \sin(\pi k/l)}{\pi(t - k)} Z(k), \quad (4)$$

où $J_l = \{k \in \mathbb{Z}, k \notin l\mathbb{Z}\}$. Cette formule est basée sur $l - 1$ suites périodiques de périodes l , et donc s'inscrit à la fois

dans le cadre étudié par Yen [7] et dans celui étudié par Papoulis [1], [8]. C'est aussi la formule d'interpolation de Lagrange adaptée à la suite J_l , mais l'on sait combien il est difficile d'en vérifier les conditions de validité (données dans [3]). C'est un cas particulier d'échantillonnage multipériodique traité dans [8]. Dans toutes ces références, les processus (ou les fonctions) vérifient la condition de Nyquist en bande de base. Notons qu'il existe d'autres formules utilisant tous les échantillons disponibles, très générales dans [2], [11], ou basées sur des partitions particulières de \mathbb{Z} en sous-ensembles adaptés dans [10].

La formule (4) ne contient pas $Z(0)$ (l'échantillon manquant), ainsi que les $Z(nl)$, $n \in \mathbb{Z}$ (qui sont disponibles). Il est intéressant de mesurer l'erreur en fonction de l lorsqu'on utilise seulement un nombre fini de termes dans la somme (ce qui est le cas dans la réalité).

Soit $\mathbf{Z} = \{Z(t), t \in \mathbb{R}\}$ un signal de type cosinus surélevé très utilisé dans le domaine des télécommunications : ce signal est à bande limitée et sa bande peut être ajustée grâce à un paramètre appelé roll off [6]. Ce paramètre a été choisi de manière à obtenir une densité spectrale de puissance de support $[-\pi + a, \pi - a]$, avec $\pi - a = 0.75\pi$. La figure 1 trace une partie d'une réalisation de \mathbf{Z} ainsi que sa densité spectrale de puissance.

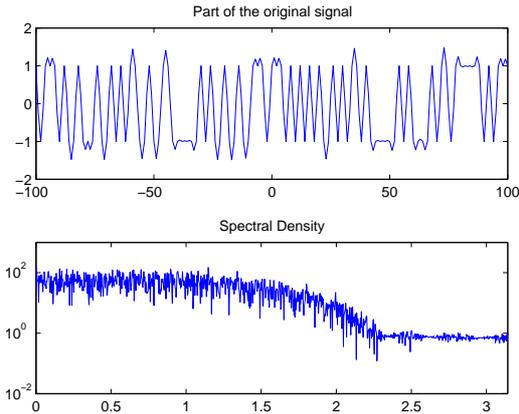


FIG. 1 – $Z(t)$ et densité spectrale de puissance associée.

Considérons $l \geq 5$ pour respecter la condition (3). La figure 2 trace l'erreur quadratique moyenne normalisée (EQM) $J(0)$ entre le point reconstruit à l'instant $t = 0$, nommé $\hat{Z}(0)$, et la vraie valeur $Z(0)$, en fonction du nombre d'échantillons N considérés dans la somme de l'équation (4) :

$$J(0) = \frac{E[(Z(0) - \hat{Z}(0))^2]}{E[Z(0)^2]} \quad (5)$$

La meilleure convergence apparaît pour la plus petite valeur de l compatible avec la densité spectrale de puissance du signal, c'est-à-dire vérifiant la condition (3). En effet, la vitesse de convergence associée à l'équation (4) dépend de la valeur de l : plus l est grand plus lente est la vitesse de convergence vers 0 de la fonction sinus cardinal $\frac{\sin(\pi k/l)}{\pi k/l}$. Ainsi, les valeurs élevées de l nécessitent de considérer des échantillons plus éloignés dans la formule d'interpolation pour obtenir une estimation précise de $Z(0)$.

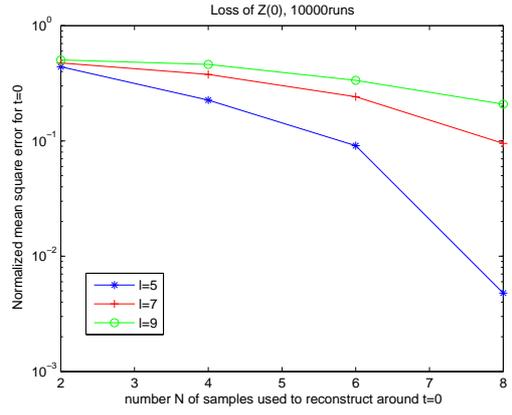


FIG. 2 – Erreur quadratique moyenne normalisée à l'instant $t = 0$.

La figure 3 présente quelques reconstructions du signal initial en utilisant $l = 5$ et en prenant en compte N échantillons dans le terme de droite de l'équation (4). Pour $N > 20$, on peut constater que le signal original est bien reconstruit.

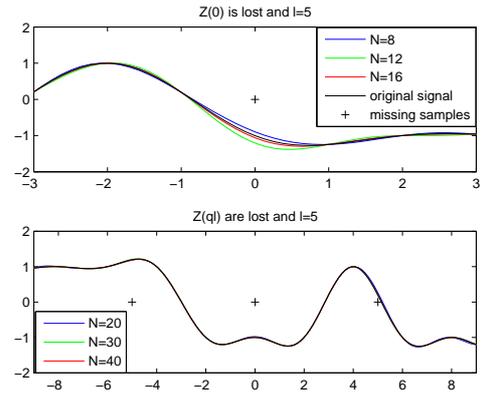


FIG. 3 – Parties du signal initial et du signal reconstruit en utilisant différents nombres d'échantillons N pour la reconstruction.

3 Amélioration de la convergence

Une fois la valeur de l déterminée, il est possible d'améliorer la convergence de la formule donnée par l'équation (4) en réintroduisant certains échantillons ignorés. Soit un processus suréchantillonné avec $a > 0.5\pi$, conduisant au choix de $l = 2$. Si l'on considère que l'échantillon à $t = 0$ est manquant, l'équation (4) peut s'écrire :

$$Z(t) = \sum_{n \in \mathbb{Z}} \frac{\sin \frac{\pi}{2}(t - 2n - 1)}{\frac{\pi}{2}(t - 2n - 1)} Z(2n + 1), \quad (6)$$

c'est-à-dire pour $t = 0$:

$$Z(0) = \sum_{n \in \mathbb{Z}} \frac{2(-1)^n}{\pi(2n + 1)} Z(2n + 1). \quad (7)$$

Considérons, par exemple, les échantillons $Z(2)$ et $Z(-2)$ qui ont été supprimés pour obtenir l'équation (7) mais

qui ne sont pas manquants. Si on les réintroduit dans la formule d'interpolation, $Z(0)$ peut être reconstruit par :

$$Z(0) = -\frac{Z(2)}{2} - \frac{Z(-2)}{2} + \frac{8}{\pi} \sum_{n \in \mathbb{Z}} \frac{(-1)^{n+1} Z(2n+1)}{(2n+1)((2n+1)^2 - 4)}. \quad (8)$$

On note que les poids de $Z(2n+1)$ décroissent en n^{-3} au lieu de n dans l'équation (7), ce qui conduit à une convergence plus rapide. De la même manière, il est également possible de réintroduire $Z(4)$ et $Z(-4)$:

$$Z(0) = -\frac{2}{3}(Z(2) + Z(-2)) - \frac{1}{6}(Z(4) + Z(-4)) + \frac{128}{\pi} \sum_{n \in \mathbb{Z}} \frac{(-1)^n Z(2n+1)}{(2n+1)((2n+1)^2 - 4)((2n+1)^2 - 16)}, \quad (9)$$

ce qui conduit à une convergence en n^{-5} au lieu de n^{-3} ... Bien entendu, la réintroduction d'échantillons peut être réalisée pour n'importe quelle valeur de l afin d'améliorer la vitesse de convergence. Notons que l'équation (8) est une formule exacte, ce qui n'est pas le cas des interpolations basées sur les splines ou les ondelettes.

4 Deux échantillons manquants

Pour $l > \frac{2\pi}{a}$ (cette hypothèse assure que le nombre d'échantillons retenus est suffisant), on a :

$$Z(t) = \frac{\sin(\pi t)}{\sin(\frac{\pi}{l} t) \sin(\frac{\pi}{l} (t - m'))} \sum_{k \in J_l} \frac{(-1)^k \sin(\frac{\pi}{l} k) \sin(\frac{\pi}{l} (k - m'))}{\pi(t - k)} Z(k), \quad (10)$$

où $J_l = \{k \in \mathbb{Z} / k \neq ml, k \neq ml + m'\}$. En considérant que $Z(0)$ and $Z(m')$ sont les deux échantillons manquants, ils peuvent être retrouvés en utilisant :

$$Z(0) = \frac{l}{\sin \frac{\pi m'}{l}} \sum_{k \in J_l} \frac{(-1)^k \sin \frac{\pi k}{l} \sin \frac{\pi(k-m')}{l}}{\pi k} Z(k),$$

$$Z(m') = \frac{l}{\sin \frac{\pi m'}{l}} \sum_{k \in J_l} \frac{(-1)^{k+m'+1} \sin \frac{\pi k}{l} \sin \frac{\pi(k-m')}{l}}{\pi(k - m')} Z(k).$$

La figure 4 montre un exemple de reconstruction du signal de départ dans le cas où deux échantillons sont manquants. Le signal de départ est un signal de type cosinus surélevé dont la densité spectrale de puissance a un support de $(-\pi + a, \pi - a)$ avec $\pi - a = 0.5\pi$. $Z(0)$ et $Z(3)$ sont manquants et l a été fixé à 5 de manière à respecter la condition de reconstruction ($l > \frac{2\pi}{a}$). Les échantillons manquants ont été mis à 0 dans la figure du milieu, le signal original est en haut, le signal reconstruit en bas. $N = 40$ échantillons ont été utilisés pour approcher (10). Comme dans le cas d'un seul échantillon manquant, la convergence de (10) peut être améliorée en réintroduisant des échantillons. Notons qu'en l'absence d'échantillons réintroduits le schéma d'échantillonnage proposé se ramène à celui étudié dans [7], [5], [9].

Dans le cas de M échantillons manquants, la formule permettant de reconstruire le signal d'origine est donnée en annexe (équation (14)). La figure 5 donne un exemple de reconstruction pour $M = 6$. l a été fixé à 13 de manière à respecter la condition de reconstruction : $l > \frac{M\pi}{a}$.

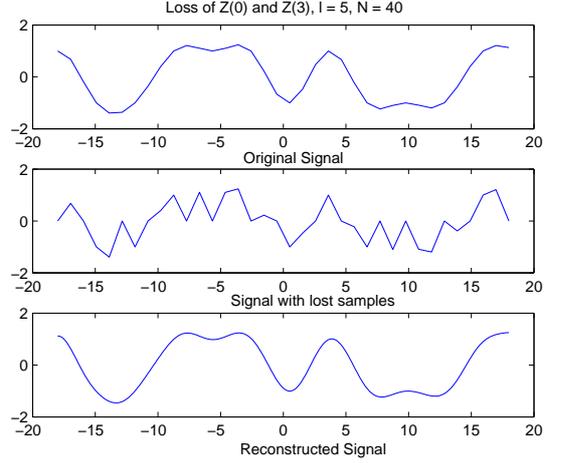


FIG. 4 – Exemple de reconstruction quand deux échantillons sont manquants

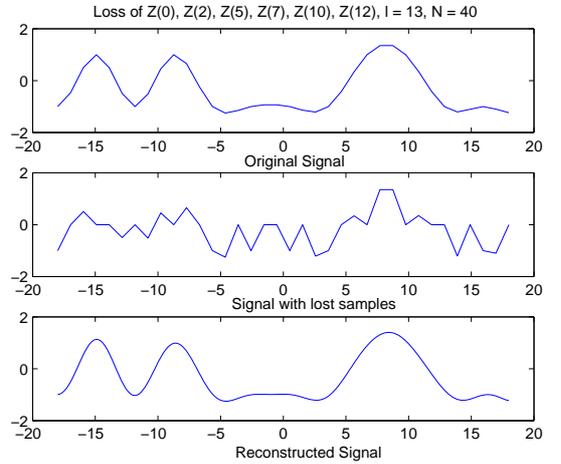


FIG. 5 – Exemple de reconstruction quand six échantillons sont manquants

5 Conclusions

Cet article traite du problème de reconstruction d'un processus à bande limitée quand un nombre fini d'échantillons est manquant. La reconstruction est possible à condition que le processus soit suréchantillonné par rapport à la borne de Shannon. Dans ce cas, des formules de reconstruction exactes sont données. Ces formules sont obtenues en supprimant périodiquement d'autres échantillons. Nous montrons que leur réintroduction améliore la convergence des formules de reconstruction. Ils permettent également de les vérifier. Les formules de reconstruction obtenues sont validées par l'intermédiaire de simulations. Ces simulations considèrent un signal en cosinus surélevé très utilisé en télécommunications et qui a la particularité de posséder un spectre à support borné et de largeur ajustable.

6 Annexes

6.1 Un échantillon manquant

Soit la fonction complexe $f(z)$ définie par :

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

Soit C_n un carré centré sur l'origine et de taille $\pm(nl + \frac{1}{2})$ où n est suffisamment grand. Le théorème des résidus [12] conduit à :

$$\frac{1}{2i\pi} \int_{C_n} f(z) dz = \frac{e^{i\omega t} \sin(\pi t/l)}{\sin \pi t} + \sum_{k \neq pl, |k| < nl, p \in \mathbb{Z}} \frac{(-1)^k e^{ik\omega} \sin(\pi k/l)}{\pi(k-t)} \quad (11)$$

En prenant en compte le comportement de $f(z)$ sur C_n :

$$|f(x \pm i(nl + \frac{1}{2}))| \sim \frac{1}{n} \exp[\pm nl\omega + (1-l)\pi n]$$

$$|f(\pm(nl + \frac{1}{2}) + iy)| \sim \frac{1}{n} \exp[-\omega y + (\frac{1}{l} - 1)|y|]$$

et en supposant que $\omega \in]-\pi(1 - \frac{1}{l}), \pi(1 - \frac{1}{l})[$, alors

$$\lim_{n \rightarrow \infty} \frac{1}{2i\pi} \int_{C_n} f(z) dz = 0. \quad (12)$$

En utilisant (11) et (12) on obtient :

$$e^{i\omega t} = \frac{\sin \pi t}{\sin(\pi t/l)} \sum_{k \in J_l} \frac{(-1)^k \sin(\pi k/l)}{\pi(t-k)} e^{i\omega k}. \quad (13)$$

Si la condition (3) est respectée, l'intervalle $]-\pi(1 - \frac{1}{l}), \pi(1 - \frac{1}{l})[$ contient $[-\pi + a, \pi - a]$. En utilisant l'isométrie fondamentale, c'est-à-dire en remplaçant $e^{i\omega t}$ dans (13) par $Z(t)$, on obtient la formule (4) [13].

6.2 Un échantillon manquant, amélioration de la convergence

En supposant que $a > 0.5\pi$ et $l = 2$, avec le même contour d'intégration que dans 6.1, on peut considérer la fonction complexe suivante en ajoutant les pôles ± 2 :

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

et on obtient pour $\omega \in [-\frac{\pi}{2}, \frac{\pi}{2}]$:

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

En appliquant l'isométrie fondamentale :

$$Z(t) = \cos\left(\frac{\pi t}{2}\right) \left\{ -\frac{t+2}{4} Z(2) + \frac{t-2}{4} Z(-2) + \frac{2}{\pi} \sum_{n \in \mathbb{Z}} \frac{(-1)^n (t^2-4) Z(2n+1)}{(2n+2-t)((2n+1)^2-4)} \right\}$$

Pour $t = 0$, on obtient l'équation (8).

6.3 M échantillons manquants

Supposons maintenant que M échantillons sont manquants, aux positions $ml + m'_k$, avec $l \in \mathbb{N}$, $m \in \mathbb{Z}$ et $m'_k \in \{1, 2, \dots, l-1\}$ pour $k = 1, \dots, M-1$. La condition de reconstruction est $l > M\pi/a$ et la fonction complexe à considérer est donnée par :

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

En supposant que les m'_k sont différents et en appliquant le théorème des résidus à $f(z)$, cela conduit à la généralisation suivante des équations (4) ($M = 1$) et (10) ($M = 2$) :

$$Z(t) = \frac{\sin \pi t}{\sin \frac{\pi t}{l} \prod_{n=1}^{M-1} \sin \frac{\pi}{l}(t - m'_n)} \sum_{k \in J_l} \frac{(-1)^k \sin \frac{\pi k}{l} \prod_{n=1}^{M-1} \sin \frac{\pi}{l}(k - m'_n)}{\pi(t-k)} Z(k), \quad (14)$$

où $J_l = \mathbb{Z} - \cup_{k=1}^{M-1} (l\mathbb{Z} + m'_k)$ est l'ensemble des pôles (d'ordre 1) de $f(z)$. Dans ce cas également, il est possible d'améliorer la convergence en réintroduisant des échantillons.

Références

- [1] A. Papoulis, "Probability, Random Variables and Stochastic Processes," *Mc Graw-Hill*, 1991.
- [2] B. Lacaze, "Reconstruction formula for irregular sampling," *Sampling Theory Sign. Image Proc.*, vol. 4, no. 1, pp. 33-43, Feb. 2005.
- [3] B. Levin, "Distribution of Zeros of Entire Functions," *American Mathematical Society Press*, 1964.
- [4] P. Ferreira, "Iterative and Noniterative Recovery of Missing Samples for 1-D Band-Limited Signals," *Sampling Theory and Practice*, Plenum Publishing Corporation, pp. 235-282, 2001.
- [5] J. R. Higgins, "Some gap samplings for multiband signals," *Signal Processing*, vol. 12, pp. 313-319, Dec. 1987.
- [6] J. G. Proakis, "Digital Communications," *Mc Graw Hill*, pp. 556-561, 2001.
- [7] J. L. Yen, "On Nonuniform Sampling of Bandwidth-Limited Signals," *IRE Trans. on Circ. Th.*, vol. 3, no. 12, pp. 251-257, 1956.
- [8] B. Lacaze, "A Theoretical Exposition of Stationary Processes Sampling," *Sampling Theory in Signal and Image Processing*, vol. 3, pp. 201-230, 2005.
- [9] B. Lacaze, "About a Multiperiodic Sampling Scheme," *IEEE Sig. Proc. Lett.*, vol. 6, no. 12, pp. 307-308, 1999.
- [10] B. Lacaze and M. Chabert, "Improved reconstruction formulas for oversampled processes and functions," *Signal Processing*, vol. 86, pp. 2827-2835, 2006.
- [11] B. Lacaze, "The Ghost Sampling Sequence method," *Sampling Theory in Signal and Image Processing*, vol. 1, pp. 385-398, 2009.
- [12] V. I. Smirnov, "A Course of Higher Mathematics," *Pergamon Press*, 1964.
- [13] H. Cramer and M. R. Leadbetter, "Stationary and related stochastic processes," *Wiley*, 1966.