

# Une formule de reconstruction exacte pour l'échantillonnage aléatoire

Bernard LACAZE, Marie CHABERT

IRIT/ENSEEIH/TSa  
2, rue Camichel BP 7122 31071 Toulouse Cedex France  
bernard.lacaze@enseeiht.fr, marie.chabert@enseeiht.fr

**Résumé** – Cet article étudie la reconstruction d'un signal à partir d'échantillons prélevés à des instants aléatoires. Une formule de reconstruction exacte est appliquée au cas d'un sur-échantillonnage présentant de la gigue. Les performances de la formule de reconstruction sont étudiées en fonction de la variance de la gigue, du nombre d'échantillons disponibles et du facteur de suréchantillonnage.

**Abstract** – The reconstruction of randomly oversampled band-limited random signals is addressed. An exact reconstruction formula is proposed in the case of periodic oversampling with jitter. The reconstruction formula performance is studied as a function of the jitter variance, the number of available samples and the oversampling factor.

## 1 Introduction

Cet article traite de la reconstruction exacte d'un signal aléatoire à bande limitée, lorsque l'échantillonnage est lui-même aléatoire. Plus précisément, nous considérons le cas d'un échantillonnage périodique perturbé par de la gigue. Une reconstruction exacte n'est alors possible que si le signal est suréchantillonné, à l'exception de cas très particuliers. Le "one-quarter theorem" ou théorème de Kadec [9] correspond à un de ces cas particuliers. L'application de ce théorème suppose l'existence d'une borne supérieure pour l'amplitude de la gigue. Les résultats exposés dans cet article sont plus généraux dans le sens où l'amplitude de la gigue peut être arbitraire, à condition que le signal soit suréchantillonné. La formule utilisée est liée à la formule d'interpolation de Lagrange tout comme la formule de Shannon. Les performances en terme d'erreur de reconstruction normalisée sur une fenêtre d'observation sont étudiées en fonction du nombre d'échantillons disponibles, du facteur de suréchantillonnage moyen et de la variance de la gigue. Les formules proposées montrent un taux de convergence très élevé.

Soit  $Z = \{Z(t), t \in \mathbb{R}\}$  un processus aléatoire stationnaire de moyenne nulle, dont le spectre  $s(\omega)$  est régulier et à bande limitée

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

Considérons la séquence d'échantillonnage  $\{k/\beta, k \in \mathbb{Z}\}$  avec  $\beta > 1$ . La fréquence d'échantillonnage est donc supérieure à celle de Nyquist.  $Z$  peut être reconstruit par combinaison linéaire des échantillons  $Z(k/\beta), k \in \mathbb{Z}$ , en utilisant par exemple la formule de Shannon :

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

Une estimation de  $Z(t)$  à partir d'un nombre fini d'échantillons est obtenue par troncature de la série précédente. Toutefois, les performances de cette formule sont limitées par la décroissance lente de la série cardinale et s'effondrent lorsque l'échantillonnage devient irrégulier, c'est-à-dire lorsque l'on remplace les  $Z(k/\beta)$  par des valeurs du processus en des points différents de  $k/\beta$ .

Cet article met en pratique, dans le cas de l'échantillonnage périodique avec gigue, une formule de reconstruction alternative initialement proposée dans un cadre plus général [8]. L'idée de base est la décomposition d'une formule de reconstruction en deux termes : l'un s'exprimant en fonction des échantillons disponibles, l'autre s'exprimant en fonction d'échantillons pris à des instants périodiques. Le deuxième type d'échantillons, non disponibles en pratique, constitue une série auxiliaire n'intervenant que dans les démonstrations. En effet, moyennant des hypothèses simples sur les instants d'échantillonnage, le deuxième terme de la décomposition peut être négligé. Les échantillons inconnus ne sont donc pas utilisés dans la mise en oeuvre pratique de la formule. Le premier terme fournit alors une estimation de  $Z(t)$  aussi précise qu'on le souhaite [8]. Les performances seront étudiées en fonction de l'erreur d'estimation. L'article s'articule de la façon suivante : la partie 2 donne la formule générale de reconstruction et les conditions nécessaires pour sa convergence. La partie 3 étudie les performances par simulation.

## 2 Nouvelle formule de reconstruction pour l'échantillonnage aléatoire

Soit  $t = \{t_k, k \in \mathbb{Z}^*\}$  la séquence croissante de  $-\infty$  à  $+\infty$  des instants d'échantillonnage aléatoires vérifiant les

conditions :

$$\lim_{|k| \rightarrow \infty} \frac{t_k}{k} = \frac{1}{\beta}, \quad \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{1}{t_k} + \frac{1}{t_{-k}} < \infty \quad (3)$$

Dans le cas de la gigue, les instants d'échantillonnage peuvent être modélisés par :

$$t_k = \frac{k}{\beta} + g_k \quad (4)$$

où les  $g_k$  sont connus. On peut, pour des raisons de clarté de la présentation, supposer les  $g_k$  uniformément bornés. En effet, la condition  $|g_k| < 1/2\beta$  permet de ne pas avoir d'interversion d'indices. Toutefois, la formule reste valable quelle que soit l'amplitude de la gigue. La formule générale de reconstruction proposée dans [4] pour un échantillonnage vérifiant (3) a pour expression :

$$Z(t) = H_M(t) \left[ \sum_{|k| \leq M} a_M(t, t_k) Z(t_k) + \sum_{|k| > M} b_M(t, k) Z(k) \right] \quad (5)$$

avec  $a_M(t, t_k)$ ,  $b_M(t, t_k)$  et  $H_M$  définis par :

$$a_M(t, t_k) = \frac{G_M(t_k)}{(t-t_k)F'_M(t_k) \sin \pi t_k}, \quad b_M(t, k) = \frac{(-1)^k G_M(k)}{\pi(t-k)F_M(k)},$$

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

où :

$$F_M(t) = \prod_{0 \leq |k| \leq M} \left(1 - \frac{t}{t_k}\right),$$

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

$F'_M$  est la dérivée de  $F_M$ . D'après (5),  $Z(t)$  peut être décomposé selon deux séries. La première utilise les échantillons mesurés  $Z(t_k)$ , la deuxième des échantillons réguliers non disponibles  $Z(k)$  (on suppose que les  $t_k$  ne peuvent pas être des entiers). Les hypothèses sur la séquence d'échantillonnage (3) impliquent l'existence des fonctions analytiques suivantes :

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

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

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

De plus, le second terme de l'équation (5) tend vers 0 quand  $M$  tend vers l'infini (la démonstration de ce résultat, basée sur le théorème des résidus, est donnée en annexe). Par conséquent, la formule de reconstruction générale conduit à :

$$Z(t) = \lim_{M \rightarrow \infty} H_M(t) \left[ \sum_{0 \leq |k| \leq M} a_M(t, t_k) Z(t_k) \right] \quad (6)$$

formule dans laquelle les  $Z(k)$  n'apparaissent plus. Notons que la formule (5) est une formule d'interpolation de Lagrange. En effet, les noyaux d'interpolation  $H_M(t)a_M(t, t_k)$  et  $H_M(t)b_M(t, k)$  sont tels que :

$$H_M(t_j)a_M(t_j, t_k) = \delta_{jk}, \quad H_M(j)a_M(j, t_k) = 0,$$

$$H_M(j)b_M(j, k) = \delta_{jk}, \quad H_M(t_j)b_M(t_j, k) = 0 \quad (7)$$

où  $\delta_{jk}$  représente le symbole de Kronecker ( $\delta_{jk} = 1$  si  $j = k$ ,  $\delta_{jk} = 0$  sinon). La formule (6) devient aussi une formule d'interpolation de Lagrange quand il est possible d'invertir la limite et la somme [14] :

$$Z(t) = \sum_{k \in Z^*} \frac{F(t)}{(t-t_k)F'(t_k)} Z(t_k) \quad (8)$$

Or inverser la limite et la somme n'est possible que sous des conditions très particulières portant sur la séquence des instants d'échantillonnage. Les arguments théoriques les plus complets sont développés dans [9]. Toutefois, les conditions requises s'avèrent la plupart du temps impossibles à vérifier en pratique, en particulier dans un contexte aléatoire [4], [5]. Le déplacement d'un seul échantillon peut les remettre en question. En revanche, la formule étudiée dans cet article s'applique à une grande variété de cas incluant l'échantillonnage périodique avec perte d'échantillons et diverses modèles d'échantillonnage aléatoire.

Dans un premier temps,  $M$  est fixé à une valeur suffisamment grande pour que la deuxième somme dans (5) puisse être négligée. Les fonctions correspondantes  $a_M(t, t_k)$  et  $H_M(t)$  sont calculées pour cette valeur de  $M$ . Enfin, l'Eq. (6) fournit une estimation de  $Z(t)$  par troncature de la somme infinie aux  $2N+1$  premiers termes avec  $N \leq M$  :

$$\hat{Z}_M(t) = \sum_{0 \leq |k| \leq N} \frac{F(t)G_M(t_k)}{(t-t_k)F'_M(t_k) \sin \pi t_k} Z(t_k) \quad (9)$$

### 3 Etude des performances

Les performances sont étudiées en fonction de l'erreur de reconstruction normalisée sur la fenêtre  $[-\frac{1}{2}, \frac{1}{2}]$ . L'erreur est estimée à partir de  $n_R$  réalisations. Le signal observé est un bruit blanc gaussien filtré de spectre initialement contenu dans  $(-\pi + a, \pi - a)$  avec  $a > 0$  sur-échantillonné avec un facteur  $\beta > 1$  et une gigue uniformément répartie sur  $(-\sigma_g \sqrt{3}, \sigma_g \sqrt{3})$  de variance  $\sigma_g^2$ . Les performances sont tout d'abord étudiées en fonction du nombre d'échantillons disponibles pour un facteur de sur-échantillonnage et une variance de la gigue donnés. Le problème du choix de  $N$  en fonction de  $M$  est également étudié par comparaison des résultats obtenus pour  $M = N$ ,  $M = N + 1$ ,  $M = N + 2$ ,  $M = 2N$  et  $M = 3N$ . Les figures 2 et 3 montrent que la meilleure estimation est obtenue pour  $M = N$  dans (5). De plus, la formule proposée montre une très grande vitesse de convergence.

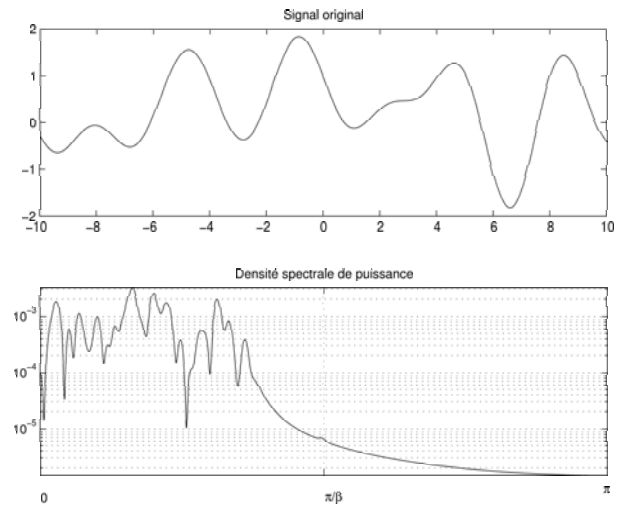


FIG. 1 – Signal original et densité spectrale associée.

La figure 4 représente le signal original et le signal reconstruit sur la fenêtre  $[-3T, 3T]$  pour des valeurs crois-

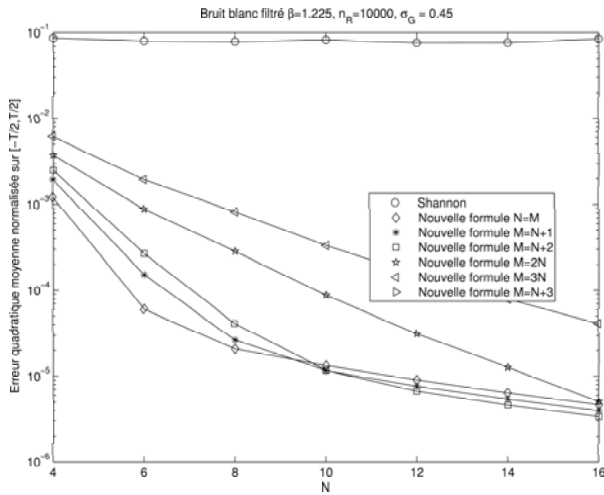


FIG. 2 – Etude des performances.

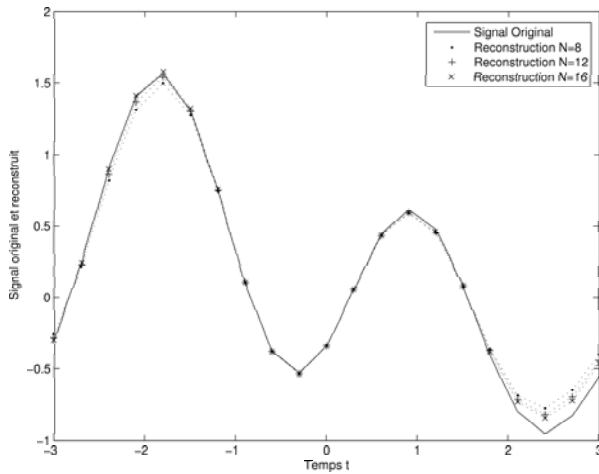


FIG. 3 – Signal original et reconstruit.

santes de  $N$  et pour  $M = N$ . Les performances ont aussi été étudiées en fonction du facteur de sur-échantillonnage  $\beta$  pour un nombre d'échantillons disponibles donné  $N = M = 8$ . Lorsque  $\beta$  augmente, la convergence de la formule s'améliore.

## 4 Conclusion

Cet article étudie la reconstruction d'un processus aléatoire sur-échantillonné en présence de gigue. Le principe est de développer le signal en deux séries. La première utilise les échantillons disponibles, la deuxième une suite auxiliaire d'échantillons non disponibles. Les performances ont été étudiées par simulations en fonction des différents paramètres du problème (variance de la gigue, facteur de sur-échantillonnage, nombre d'échantillons disponibles). Dans tous les cas, la vitesse de convergence de l'algorithme est remarquable.

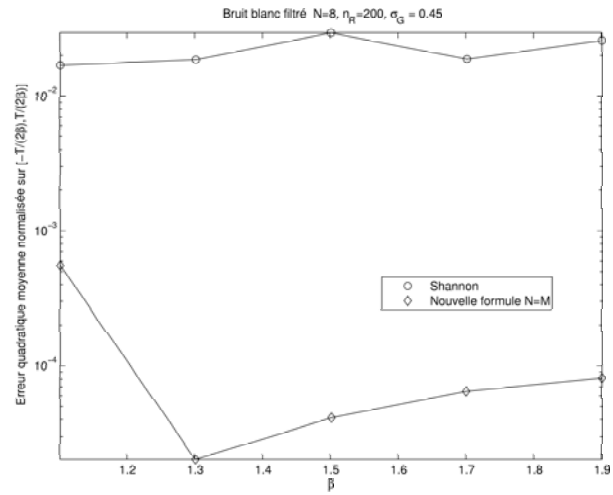


FIG. 4 – Erreur en fonction du facteur de suréchantillonnage.

## 5 Annexe

Supposons que la transformée de Fourier de  $K_Z(\tau) = E[Z(t)Z^*(t-\tau)]$  vérifie (1). Si  $e^{i\omega t}$  peut être développée en un certain sens (par exemple uniformément sur  $[-\pi + \varepsilon, \pi - \varepsilon]$ ) sous la forme

$$e^{i\omega t} = \sum_{n=-\infty}^{n=+\infty} \alpha_n(t) e^{i\omega t_n} \quad (10)$$

alors le processus  $Z(t)$  peut être reconstruit par :

$$Z(t) = \sum_{n=-\infty}^{n=+\infty} \alpha_n(t) Z(t_n) \quad (11)$$

Considérons la fonction  $\gamma_M(z)$  définie pour  $t \notin \mathbb{Z} \cup \mathbf{t}$  par :

$$\gamma_M(z) = \frac{e^{i\omega z}}{(z-t)\sin\pi z} \frac{G_M(z)}{F_M(z)} \quad (12)$$

La démonstration de la formule de reconstruction est basée sur le théorème des résidus [12]. Ce théorème permet d'intégrer  $\gamma_M(z)$  sur le contour fermé  $\Gamma_d$ , qui est le carré centré à l'origine et qui coupe les axes perpendiculairement à la distance  $d + \frac{1}{2}$  de l'origine.

$$\frac{1}{2i\pi} \int_{\Gamma_d} \gamma_M(z) dz = \text{Res}[\gamma_M, t] + \sum_{M \leq |k| \leq d} \text{Res}[\gamma_M, k] + \sum_{k=-M}^{k=M} \text{Res}[\gamma_M, t_k] \quad (13)$$

Considérons  $d_M$  tel que, pour  $d \geq d_M$ , l'ensemble  $\{t_k, k \leq M+1\}$  est strictement inclus dans  $\Gamma_d$  pour  $d > d_M$ . Notons que  $\frac{G_M(z)}{F_M(z)}$  est uniformément bornée sur  $\Gamma_d$  pour  $d > d_M$  puisque  $G_M(z)$  et  $F_M(z)$  sont des polynômes de degrés respectifs  $2(M-1)$  et  $2M$ . Sous la condition supplémentaire  $|\omega| < \pi$  :

$$\lim_{d \rightarrow \infty} \frac{1}{2i\pi} \int_{\Gamma_d} \gamma_M(z) dz = 0 \quad (14)$$

Les résidus sont faciles à calculer puisque les points singuliers de  $\gamma_M$  sont des pôles d'ordre 1. Finalement,  $e^{i\omega t}$  se développe de la manière suivante :

$$e^{i\omega t} = \sin \pi t \frac{G_M(t)}{F_M(t)} \times \left[ \sum_{1 \leq |k| \leq M} \frac{G_M(t_k)}{(t - t_k) F'_M(t_k) \sin \pi t_k} e^{i\omega t_k} + \sum_{|k| > M} \frac{(-1)^k G_M(k)}{\pi (t - k) F_M(k)} e^{i\omega k} \right] \quad (15)$$

pour  $-\pi < \omega < \pi$  et  $t \notin \mathbb{Z} \cup \mathbf{t}$ . Puisque la convergence est uniforme en  $\omega$  sur n'importe quel intervalle  $\Delta \subset [-\pi, \pi]$ , l'équation (5) est vraie pour tout  $M$ . La seconde somme peut être négligée pour  $M$  suffisamment grand. Evidemment, pour  $t$  donné, par exemple proche de 0, la condition  $\frac{t_k}{k} \rightarrow_{k \rightarrow \infty} \frac{1}{\beta}$  assure que  $t$  sera entouré par un ensemble de  $t_k$  ( $1 \leq |k| \leq M$ ) et que les entiers  $k$  ( $|k| > M$ ) en seront très éloignés. Alors, les termes correspondants à ces entiers pourront être négligés. Par exemple, pour  $\beta = 1, 7$ , l'intervalle  $(-10, 10)$  contient 34 instants  $t_k$ , et le terme le plus proche dans la seconde somme est  $Z(\pm 36)$ . Ces termes et les suivants n'ont évidemment aucune influence sur  $Z(0)$ . D'un point de vue mathématique, la preuve peut être déduite de l'inégalité

$$\sup_{|x| > M} \left| \prod_{1 \leq |m| \leq M} \frac{1 - \frac{x}{m}}{1 - \frac{x}{t_m}} \right| < \eta \quad (16)$$

En conséquence,  $b_M$  s'écrit :

$$b_M(t) = \sum_{|m| > M} \frac{(-1)^m u(m)}{t - m} \frac{m}{1 - \frac{m}{t_0}} \frac{1}{(1 - \frac{m}{t_{-1}})(1 - \frac{m}{t_1})} \prod_{1 \leq |l| \leq M} \frac{1 - \frac{m}{l}}{1 - \frac{m}{t_{l+sgnl}}} \quad (17)$$

où  $sgnl = 1$  pour  $l > 0$  et  $sgnl = -1$  pour  $l < 0$ . Après des changements légers dans l'ordre des  $t_k$  l'équation (16) mène à l'inégalité :

$$|b_M(t)| = \sum_{|m| > M} \frac{\alpha}{m^3} \quad (18)$$

lorsque  $u(t)$  est bornée et  $M > M_0$ . En conséquence,  $|b_M(t)|$  peut être vue comme le reste d'une série convergente et donc

$$\lim_{M \rightarrow \infty} |b_M(t)| = 0 \quad (19)$$

## Références

- [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 Multi-band 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 Processing, 86 (2006), 2827-2835.
- [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 Bandwidth-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).