

**OPTIMISATION DES CONVERTISSEURS ANALOGIQUES-NUMERIQUES  
A MISE EN FORME DU BRUIT EN PRESENCE D'UN  
RETARD DANS LA BOUCLE**

I. LAMBERT<sup>†</sup>, Y. RICHARD\*, G. VEZZOSI<sup>†</sup>, A. DUBREUIL\*

<sup>†</sup> *Laboratoire de Traitement du Signal et de l'Image, Université de Rennes I, 35042 RENNES CEDEX*  
\* *SOREP, Z.I. Bellevue, B.P. 5, 35220 CHATEAUBOURG*

**RÉSUMÉ**

Cet article est consacré à l'étude d'un convertisseur analogique-numérique à suréchantillonnage et mise en forme du bruit lorsque le temps de réponse du CNA présent dans la chaîne de contre-réaction ne peut pas être négligé. On présente une méthode de transfert temps continu-temps discret qui permet, à partir d'un modèle à temps discret optimisé, de spécifier le filtre analogique précédant le CAN ; le filtre analogique optimum dépend du temps de réponse du CNA. La stabilité du convertisseur vis à vis des variations de la réponse du CNA autour de sa valeur nominale est étudiée. L'intervalle de stabilité autour de cette valeur nominale ne dépend que de l'ordre du filtre analogique.

**ABSTRACT**

In this paper an oversampled noise-shaping analog to digital converter is considered when the feedback DAC settling time cannot be neglected. We present a continuous to discrete time transfer method which allows, from an optimized discrete time model, to specify the analog filter preceding the quantizer ; the optimum analog filter depends on the DAC settling time. The stability of the converter towards the DAC settling time variation around its nominal value is studied. The stability interval around this nominal value depends only on the order of the analog filter.

**1. INTRODUCTION**

Un convertisseur analogique-numérique à mise en forme de bruit est un système hybride dont l'entrée est un signal analogique à spectre limité dans  $[-B, B]$ , la sortie un signal quantifié sur un nombre de bits peu élevé à une fréquence d'échantillonnage  $f_e$  très supérieure à  $2B$ , et où se trouvent associés des organes analogiques et numériques (Fig. 1). L'étude d'un tel circuit est souvent réalisée dans la littérature ([1-2]) à partir d'un schéma à temps discret, présenté comme "équivalent", où l'entrée est échantillonnée, et où tous les organes analogiques ont disparu. Cette façon d'opérer peut se justifier lorsque le filtre suivant le sommateur dans la boucle est réalisé avec des capacités commutées. Il n'en va pas de même lorsque le filtre est analogique. Il est alors nécessaire de modéliser le comportement des organes analogiques, en particulier celui du convertisseur numérique-analogique (CNA).

On se propose ici d'étudier et d'optimiser le circuit de mise en forme du bruit en tenant compte du temps de réponse du CNA. Dans la première partie, nous donnons le schéma à temps discret équivalent du convertisseur, obtenu en considérant que la réponse du CNA est celle d'un bloqueur retardé de  $\theta$  et que l'entrée est munie d'un bloqueur synchronisé avec celui du CNA. Le mot équivalent signifie ici que les deux circuits donnent la même réponse pour la même entrée. Nous optimisons le modèle à temps discret, puis nous décrivons une méthode permettant de spécifier le filtre analogique. La deuxième partie est consacrée à l'étude de la stabilité du circuit vis à vis des variations  $\Delta\theta$  de la réponse du CNA autour de sa valeur nominale  $\theta_0$ . Deux structures sont étudiées : l'une comprenant un filtre numérique dans la chaîne de contre-réaction, l'autre n'en comprenant pas.

**2. TRANSFERT TEMPS CONTINU - TEMPS DISCRET**

Le principe général du circuit est donné dans la Fig. 1. Nous supposons que dans ce schéma :

(i) Le CAN équivaut à la donnée d'un échantillonneur, qui multiplie le signal continu par le peigne de Dirac, à la sortie duquel est ajouté le bruit de quantification. Si  $\Delta$  désigne le pas de quantification, ce bruit est supposé représenté par une suite de variables aléatoires ( $v_n$ ) équiparties dans l'intervalle  $[-\Delta/2, \Delta/2]$ , et indépendantes. Le convertisseur apparaît ainsi comme un circuit à deux entrées : la première est le signal à convertir, la seconde un bruit blanc à temps discret, de puissance moyenne  $\sigma_q^2 = \Delta^2/12$ .

Sans nuire à la généralité, nous considérons que la sortie du CAN échantillonne la sortie  $u(t)$  du filtre analogique aux instants entiers ( $f_e = 1$ ). La sortie du CAN peut ainsi s'écrire :

$$y_d(t) = \sum_n [u(n) + q(n)] \delta(t-n) , \quad (1)$$

où  $q(n)$  est le bruit de quantification.

(ii) Le CNA est modélisé par un bloqueur retardé de  $\theta$  ( $0 \leq \theta < 1$ ) par rapport aux instants d'échantillonnage du CAN (hypothèse réaliste). Le retard  $\theta$  tient compte du temps de calcul du filtre numérique et du temps de réponse du CNA. Si  $p(t)$  désigne l'indicatrice de l'intervalle  $]0, 1]$ , la sortie  $w(t)$  du CNA vaut ainsi :

$$w(t) = \sum_n v(n) p(t-\theta-n) . \quad (2)$$

(iii) Le filtre analogique est rationnel, et le degré du numérateur est inférieur ou égal au degré  $N$  du dénominateur. Bien que cette hypothèse ne soit pas nécessaire (voir par exemple les relations (7) ci dessous), nous supposons pour simplifier qu'il n'a que des pôles simples. Sa fonction de transfert ( $f_t$ ) admet alors la décomposition :



$$F(p) = C + \sum_{n=1}^N \frac{\alpha_n}{p - p_n}, \quad (3)$$

à laquelle correspond la réponse percussionnelle (rp)

$$f(t) = C \delta(t) + \sum_{n=1}^N \alpha_n e^{p_n t}. \quad (4)$$

La définition d'un modèle à temps discret équivalent au circuit de la Fig. 1 demande de préciser le mode de conversion du signal continu d'entrée en la suite de ses échantillons. Comme le filtre  $F(p)$  est analogique, il est naturel de considérer que la conversion est réalisée avec un bloqueur. Nous supposons dans la suite ce bloqueur identique à celui représentant la réponse du CNA, ce qui va permettre de simplifier la représentation. Puis, nous vérifierons que les performances ne sont pas modifiées en l'absence du bloqueur d'entrée.

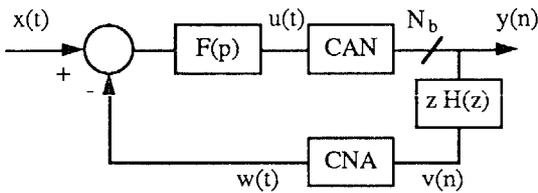


Fig. 1 : Principe de la conversion A/N à mise en forme du bruit ; (la boucle est suivie d'un filtre passe-bas et d'un décimateur qui ne sont pas représentés).

## 2.1 Cas du convertisseur avec un bloqueur à l'entrée

Puisque le bloqueur d'entrée est identique à celui représentant le CNA, on peut les remplacer par un bloqueur unique placé derrière le sommateur (Fig. 2a). La mise en série du bloqueur retardé de  $\theta$ , du filtre analogique et de l'échantillonneur du CAN équivaut alors à un filtre à temps discret, de fonction de transfert  $z^{-1}F_d(z)$  (Fig. 2b), avec

$$F_d(z) = f_0 + f_1 z^{-1} + \dots$$

Les coefficients  $f_n$  sont obtenus en convoluant la porte retardée  $p(t-\theta)$  avec la rp  $f(t)$ , et en échantillonnant le résultat aux instants entiers. D'après la relation (4) donnant la rp de  $f(t)$ , on obtient :

$$F_d(z) = A + C + \sum_{n=1}^N \frac{\beta_n}{1 - e^{p_n} z^{-1}}, \quad (5-a)$$

$$\beta_n = \frac{\alpha_n}{p_n} e^{-p_n \theta} [e^{p_n} - 1], \quad (5-b)$$

$$A = \sum_{n=1}^N \frac{\alpha_n}{p_n} [e^{-p_n \theta} - 1]. \quad (5-c)$$

Les relations dépendent explicitement du retard  $\theta$  du bloqueur, qui intervient au numérateur de  $F_d(z)$ , mais pas au dénominateur (celui-ci ne dépend que des pôles de  $F(p)$ ). De plus, ces relations peuvent être inversées. Supposons donnés un retard  $\theta$  et une  $f_t F_d(z) = P(z)/Q(z)$  telle que  $\deg P < \deg Q = N$  et n'ayant que des pôles simples. Il suffit, pour obtenir  $F(p)$ , de décomposer  $F_d(z)$  en éléments simples

$$F_d(z) = \sum_{n=1}^N \frac{\beta_n}{1 - r_n z^{-1}}, \quad (6)$$

puis de calculer les pôles de  $F(p)$  par  $p_n = \log r_n$ , les coefficients  $\alpha_n$  par (5-b), la constante  $A$  par (5-c), puis  $C = -A$ . Le passage de  $F_d(z)$  à  $F(p)$  peut également s'effectuer par un calcul des résidus.

On obtient :

$$C = \sum_n \text{Res} \left\{ F_d(z) \frac{1}{z} \frac{z^\theta - 1}{z - 1}, r_n \right\}, \quad (7-a)$$

$$F(p) - C = \sum_n \text{Res} \left\{ F_d(z) \frac{1}{z} \frac{z^\theta}{z - 1} \frac{\log z}{p - \log z}, r_n \right\}. \quad (7-b)$$

Ces dernières expressions s'appliquent même lorsque  $F_d(z)$  a des pôles multiples. La sommation est alors étendue à tous les pôles distincts de  $F_d(z)$ . Par exemple, si  $B$  tend vers zéro, on sait [1] que  $F_d(z)$  tend vers  $(1-z^{-1})^{-N}$ . Les relation (7) nous donnent alors le filtre analogique correspondant :

$$C = \theta (\theta+1) \dots (\theta+N-1) / N!, \quad (8-a)$$

$$F(p) - C = \text{Res} \left\{ \frac{z^{N-1+\theta}}{(z-1)^{N+1}} \frac{\log z}{p - \log z}, 1 \right\}. \quad (8-b)$$

Pour  $N=3$ , on obtient :

$$F(p) - C = \frac{1}{p^3} + \frac{1}{p^2} (\theta+1) + \frac{1}{p} \left( \frac{3\theta^2 + 6\theta + 2}{6} \right). \quad (9)$$

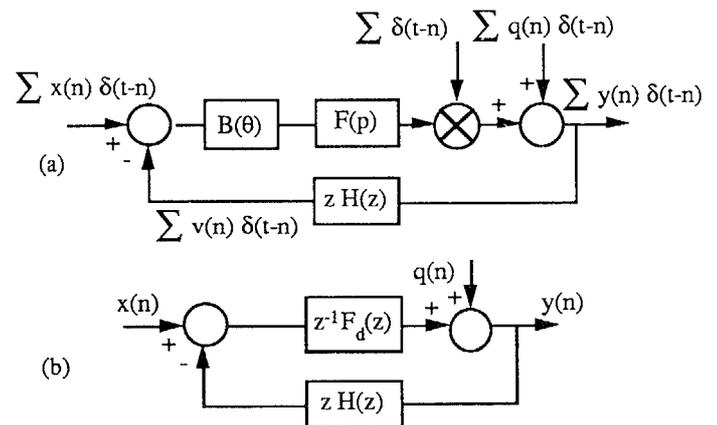


Fig. 2 (a) Schéma équivalent du convertisseur avec un bloqueur à l'entrée ; (b) modèle à temps discret équivalent.

## 2.2 Optimisation et rapport S/B

Dans le schéma de la Fig. 2b, le transfert pour le bruit de quantification et l'entrée valent respectivement :

$$T_q(z) = \frac{1}{1 + F_d(z) H(z)}, \quad T_x(z) = z^{-1} F_d(z) T_q(z). \quad (10)$$

Pour optimiser le circuit, nous commençons par imposer au transfert entrée-sortie  $T_x(z)$  de valoir  $z^{-1}$ . Ceci entraîne que  $F_d(z)$  et  $H(z)$  sont liés par la condition  $H(z) + 1/F_d(z) = 1$ , donc

$$T_q(z) = 1 - H(z). \quad (11)$$

Puis nous supposons que  $H(z)$  est un filtre FIR

$$H(z) = a_1 z^{-1} + \dots + a_N z^{-N}, \quad (12)$$

que nous calculons de façon à minimiser la puissance moyenne du bruit en sortie dans la bande du signal. Comme cette puissance vaut  $\sigma_q^2 P_N(B)$ , avec

$$P_N(B) = \int_{-B}^{+B} |1 - H(e^{2i\pi f})|^2 df, \quad (13)$$

le filtre  $H(z)$  optimum n'est autre que le meilleur prédicteur à un pas pour le bruit blanc dans la bande  $[-B, B]$ . Le filtre  $F_d(z)$  correspondant a ses pôles simples et vaut

$$F_d(z) = \frac{1}{1 - H(z)} = \sum_{n=1}^N \frac{\beta_n}{1 - r_n z^{-1}} \quad (14)$$

Pour un retard  $\theta$  donné, il reste à appliquer les formules (3-5-14) à l'envers pour en déduire le filtre analogique  $F(p)=F(p, \theta)$ . Le rapport Signal/Bruit correspondant vaut

$$\left[ \frac{S}{B} \right] = \frac{\sigma_x^2}{\sigma_q^2} \frac{1}{P_N(B)}, \quad (15)$$

où  $\sigma_x^2$  est la puissance moyenne du signal.

### 2.3 Cas du convertisseur sans bloqueur d'entrée

Dans le convertisseur de la Fig. 1, l'absence du bloqueur ne modifie pas le transfert du bruit de quantification, qui continue d'être donné par la relation (11). Le transfert entrée-sortie peut s'obtenir en sortant de la chaîne d'action de la Fig. 1 le bloc filtre analogique + échantillonneur pour le mettre avant le sommateur. On retrouve alors ce bloc dans la chaîne de contre-réaction (Fig. 3), et la sortie du circuit aux instants entiers prend la forme:

$$y(n) = T_q(z) \left( q(n) + \int_0^n f(n-s) x(s) ds \right) \quad (16)$$

A l'échantillonnage près, la fonction de transfert de l'entrée vaut donc :

$$T_x(p) = T_q(e^p) F(p, \theta) \quad (17)$$

Dans le produit du second membre, les pôles de  $F(p)$  sont compensés par les zéros de  $T_q(e^p)$ , donc la fonction  $T_x(p)$  est une fonction entière. On a de plus d'après (5) :  $F(0)=F_d(1)$ , donc  $T_x(0)=1$ . Supposant l'entrée strictement limitée à la bande  $[-B, B]$ , l'étude de la distorsion se ramène alors au tracé des courbes de gain et de phase de la fonction  $T_x(2i\pi f)$  dans la bande du signal. La Fig. 4 donne un exemple de telles courbes, pour  $N=3$  et  $B=0,05$ . On y voit que la distorsion est négligeable. Ce type de résultat peut être prévu. Pratiquement  $B$  est toujours inférieur à  $0,05$ , et on peut avoir une expression approchée de  $T_x(p)$  en remplaçant dans la relation (17) les fonctions  $F(p)$  et  $T_q(e^p)$  par leur limite quand  $B$  tend vers zéro (cf (8)). Pour  $N=3$  on obtient par exemple

$$T_x(p) = 1 + \left(\theta - \frac{1}{2}\right)p + \frac{1}{2} \left(\theta^2 - \theta + \frac{1}{6}\right)p^2, \quad (18)$$

ce qui est proche du retard pur  $e^{-p(1/2-\theta)}$ .

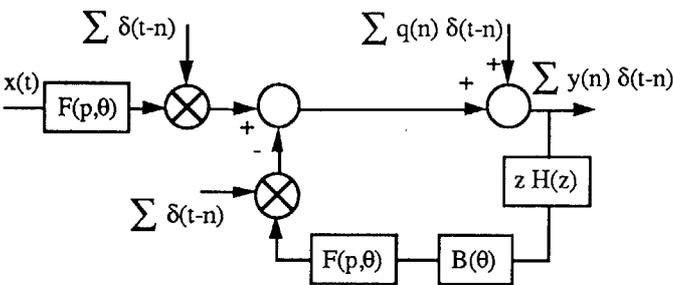


Fig. 3 : Schéma équivalent du convertisseur sans bloqueur d'entrée.

### 3. SENSIBILITE DU CONVERTISSEUR AU RETARD DU CNA

Supposons le convertisseur optimisé pour un retard du CNA égal à une valeur nominale  $\theta_0$  ( $F(p) = F(p, \theta_0)$ ). Si le CNA effectivement utilisé a un retard  $\theta = \theta_0 + \Delta\theta$  différent de  $\theta_0$ , les transferts du bruit de quantification et du signal sont modifiés. Le schéma équivalent est en effet celui de la Fig. 3 dans lequel  $F(p, \theta)$  est remplacé par  $F(p, \theta_0)$ . On voit apparaître en contre-réaction la chaîne bloqueur retardé de  $\theta + F(p, \theta_0) +$  échantillonneur. Le filtre à temps discret  $z^{-1}F_d(z)$  équivalent à cette chaîne est obtenu à partir des relations (5) dans lesquelles, les coefficients  $\alpha_n, p_n$  et  $C$  étant connus, on remplace  $\theta$  par  $\theta_0 + \Delta\theta$ . Un calcul rapide permet de montrer que

$$F_d(z) = F_d(z) + \Delta F_d(z), \quad (19)$$

dans lequel  $\Delta F_d(z)$  dépend de  $\Delta\theta$  mais pas de  $\theta_0$  :

$$\Delta F_d(z) = (1 - z^{-1}) \sum_{n=1}^N \frac{e^{-p_n \Delta\theta} - 1}{r_n - 1} \frac{\beta_n r_n}{1 - r_n z^{-1}}, \quad (20)$$

où les coefficients  $\beta_n$  et  $r_n$  sont définis en (14).

La fonction de transfert du bruit de quantification devient

$$T_q(z) = \frac{1}{1 + (F_d(z) + \Delta F_d(z))H(z)} = \frac{1 - H(z)}{1 + \Delta F_d(z)H(z)(1 - H(z))} \quad (21)$$

Celle de l'entrée continue d'être donnée (sans bloqueur d'entrée, et à l'échantillonnage près) par :

$$T_x(p) = T_q(e^p) F(p, \theta_0)$$

Comme précédemment les pôles de  $F(p, \theta_0)$  sont compensés par les zéros de  $T_q(e^p)$ . Mais la fonction  $T_q(z)$  ayant maintenant un dénominateur fonction de  $\Delta\theta$ , le convertisseur peut devenir instable. Nous avons représenté à la Fig. 5 les bornes de stabilité de  $\Delta\theta$  en fonction de la fréquence maximale  $B$  du signal d'entrée. Les bornes ont été calculées en prenant un pas sur  $\Delta\theta$  de  $\pm 10^{-4}$  à partir de  $\Delta\theta=0$ , et en appliquant le test de Schur au dénominateur de  $T_q(z)$ . On note que augmenter le degré du filtre analogique accroît simultanément la précision nécessaire à apporter sur le retard du CNA. Par exemple, si  $B = 0,05$  et  $\theta_0 = 0,66$  l'intervalle de stabilité sera  $[0,579, 0,893]$  pour un filtre d'ordre  $N = 2$ , et  $[0,653, 0,674]$  pour  $N = 4$ .

La sortie du circuit aux instants entiers continue d'être donnée par la relation (16) dans laquelle  $T_q(z)$  est calculée avec (21). Dans la bande du signal,  $H(e^{2i\pi f})$  est proche de 1,  $1 - H(e^{2i\pi f})$  est proche de zéro, le terme correctif  $\Delta F_d(z) H(z) (1 - H(z))$  qui apparaît dans  $T_q(z)$  peut être négligé, donc le transfert est le même que si  $\Delta\theta = 0$ . Il s'ensuit que tant que  $\Delta\theta$  reste dans l'intervalle de stabilité, le rapport Signal/Bruit est proche de l'optimum (15).

### 4. OPTIMISATION D'UN CONVERTISSEUR A/N SIMPLIFIE

Nous étudions maintenant les performances du convertisseur A/N à mise en forme du bruit dans le cas où la boucle ne comprend pas de filtre numérique dans la chaîne de contre-réaction ; le CAN et le CNA ont ainsi la même résolution. Comme précédemment, nous effectuons l'optimisation en supposant la présence à l'entrée d'un bloqueur (fictif) identique à celui du CNA, puis nous vérifions que ce bloqueur n'a en fait pas d'importance. Si nous ôtons le filtre numérique dans la Fig. 1, le modèle à temps discret équivalent est donné par la Fig. 2b avec  $zH(z)=1$ . Les transferts du



bruit de quantification et de l'entrée valent alors respectivement :

$$T_q(z) = \frac{1}{1 + z^{-1} F_d(z)}, \quad T_x(z) = z^{-1} F_d(z) T_q(z). \quad (22)$$

Contrairement au cas précédent, on ne peut plus demander à la fois à  $T_x(z)$  de valoir  $z^{-1}$  et à  $T_q(z)$  d'être un filtre FIR. Nous demandons à  $T_q(z)$  d'être le filtre FIR  $1-H(z)$  où  $H(z)$  est le prédicteur à un pas du bruit blanc dans la bande. Le filtre  $F_d(z)$  correspondant vaut alors :

$$F_d(z) = \frac{z H(z)}{1 - H(z)} = \sum_{n=1}^N \frac{\beta_n r_n}{1 - r_n z^{-1}}, \quad (23)$$

où les coefficients  $\beta_n$  et  $r_n$  déterminent la décomposition en éléments simples de  $(1-H(z))^{-1}$ . La démarche est ensuite la même que précédemment :

. Définition du filtre analogique : appliquer (3-5-23) à l'envers. Si  $B$  tend vers 0, la limite du filtre analogique vaut  $F(p) = C + (F(p)-C)$ , avec

$$C = (\theta+1)(\theta+2)\dots(\theta+N) / N! - 1, \quad (24-a)$$

$$F(p) - C = \text{Res} \left\{ \frac{z^N - (z-1)^N}{(z-1)^{N+1}} z^\theta \frac{\log z}{p - \log z}, 1 \right\}. \quad (24-b)$$

Par exemple, pour  $N=3$ , on trouve :

$$F(p) - C = \frac{1}{p^3} + \frac{\theta+2}{p^2} + \frac{1}{p} \left( \frac{3\theta^2 + 12\theta + 11}{6} \right). \quad (25)$$

. Fonction de transfert entrée-sortie sans bloqueur d'entrée et distorsion :

$$T_x(p) = T_q(e^p) F(p, \theta).$$

Si  $B$  tend vers 0, on obtient pour  $N=3$

$$T_x(p) = 1 + \left(\theta + \frac{1}{2}\right)p + \frac{1}{2} \left(\theta^2 + \theta + \frac{1}{6}\right)p^2, \quad (26)$$

qui montre que la distorsion est acceptable (voir aussi la Fig.4).

. Stabilité : si le filtre analogique est calculé pour le retard  $\theta_0$  et que le retard effectif du CNA vaut  $\theta = \theta_0 + \Delta\theta$ , la ft du bruit de quantification a un dénominateur. On trouve

$$T_q(z) = \frac{1 - H(z)}{1 + \Delta F_d(z) z^{-1} (1 - H(z))}, \quad (27)$$

$$\Delta F_d(z) = (1 - z^{-1}) \sum_{n=1}^N \frac{e^{-p_n \Delta\theta} - 1}{r_n - 1} \frac{\beta_n r_n^2}{1 - r_n z^{-1}}. \quad (28)$$

L'intervalle de stabilité autour de  $\theta_0$  est légèrement supérieur à celui du convertisseur précédent.

. Rapport S/B dans l'intervalle de stabilité : proche de l'optimum (15).

### 5. CONCLUSION

Le temps de réponse du CNA est un paramètre important à considérer dans la conception d'un convertisseur analogique-numérique à mise en forme du bruit utilisant un filtre analogique, car il peut entraîner l'instabilité du circuit. Nous avons montré comment, pour un temps de réponse nominal  $\theta_0$ , calculer le filtre analogique, et nous avons étudié la stabilité autour de sa valeur nominale. Il apparaît que l'intervalle de stabilité décroît quand l'ordre du filtre analogique augmente. Tant qu'on reste dans l'intervalle de stabilité, le rapport S/B est proche de l'optimum. Les conclusions sont les mêmes si le convertisseur ne contient pas de filtre numérique.

### Références

- [1] S.K. TEWKSBURRY, R.W. HALLOCK, "Oversampled, Linear Predictive and Noise-Shaping Coders of Order  $N>1$ ", IEEE Trans. Circuits Syst., vol CAS 25, NO 7, July 1978.
- [2] J.C. CANDY, "A Use of Double Integration in Sigma-Delta Modulation", IEEE Trans. Commun., vol COM-33, NO3, March 1985.

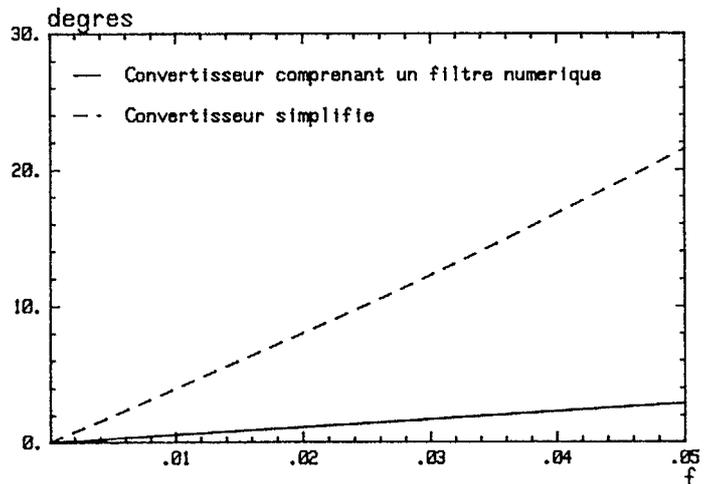
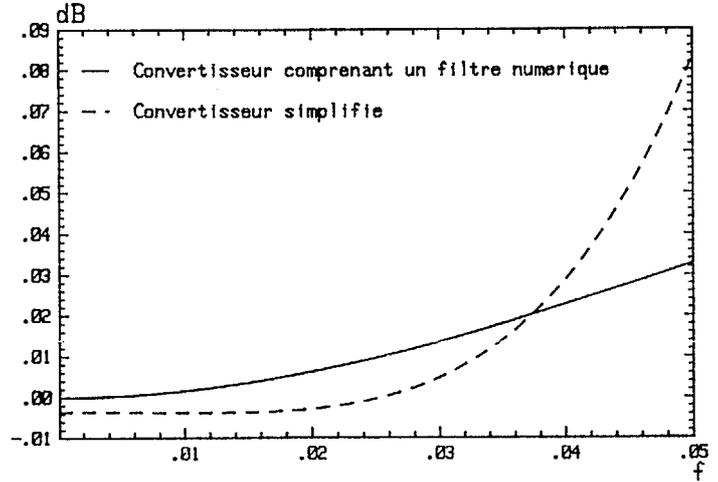


Fig. 4 : (a) gain (b) phase de la fonction  $T_x(p) = (1-H(e^p))F(p, \theta_0)$  pour  $N=3$ ,  $B=0,05$ ,  $\theta_0=0,66$ .

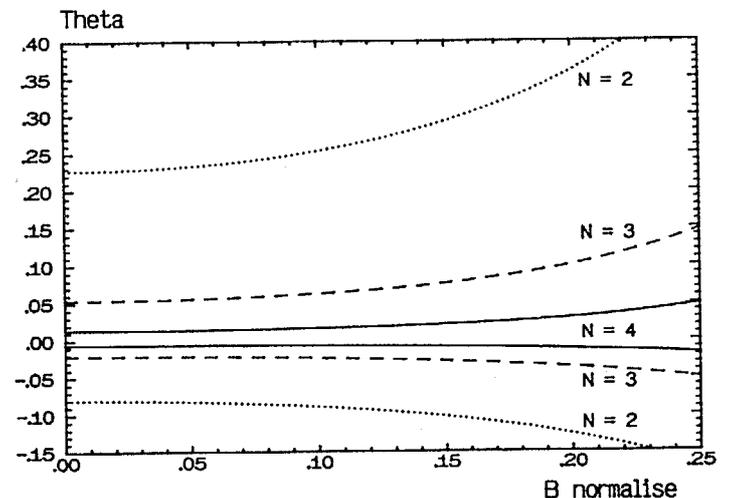


Fig. 5 : Bornes supérieure et inférieure de l'intervalle de stabilité de  $\Delta\theta$  en fonction de B et N dans le cas du convertisseur comprenant le filtre numérique.