



# Traitement, Synthèse, Technologie et Applications

BIARRITZ — Mai 1984 —

SYNTHESE DE FILTRES NUMERIQUES POUR SIGNAUX VIDEO :  
SPECIFICITE DU PROBLEME ET SOLUTION OPTIMALE  
DIGITAL FILTER DESIGN FOR VIDEO SIGNALS :  
SPECIFICITY OF THE PROBLEM AND OPTIMAL SOLUTION

J.Châtel , P.Siohan

CCETT

Rue du Clos Courtel

Cesson-Séviigné

35510

FRANCE

## RESUME

Bien qu'utilisant des principes connus, le filtrage numérique des signaux de télévision pose des problèmes spécifiques à ce type d'application. Cette spécificité dont les conséquences se font sentir tant au niveau de la construction qu'à celui de la synthèse des filtres, tient d'une part à la nature des signaux traités et d'autre part à des contraintes d'ordre technologique.

Les principaux traits caractéristiques des signaux vidéo sont leur dynamique, la nécessité de traitements à phase linéaire pour un rendu significatif des détails fins sur image, et une étendue spectrale importante. Ces paramètres conduisent, compte tenu des applications habituelles, à synthétiser des filtres numériques non récursifs à coefficients symétriques ou antisymétriques avec des contraintes de gain, et du fait de l'étendue spectrale des technologies rapides s'imposent.

Ces dernières contraintes amènent à rechercher des filtres ayant un nombre de coefficients aussi réduit que possible. De plus, les multiplieurs étant réservés à des applications très particulières, ceux-ci seront réalisés au moyen de mémoires mortes ou mieux encore à l'aide d'additionneurs. Cette dernière solution suppose que l'on soit à même de ramener à un ou deux le nombre de bits égaux à un de la plupart des coefficients.

Les méthodes de synthèse les plus répandues supposent des filtres à coefficients de précision infinie. La quantification avec arrondi de tels coefficients peut conduire à des résultats assez éloignés de l'optimum. Par contre d'autres méthodes moins connues permettent d'obtenir des solutions optimales avec en contre partie des temps de calcul sur ordinateur plus importants. Un algorithme décrit par Kodek effectue la synthèse optimale au sens de Tchebycheff des filtres transversaux à coefficients de longueur finie. Quelques modifications de cet algorithme permettent de satisfaire les contraintes des problèmes posés et d'étendre son champ d'application au cas de la synthèse de transformateurs de Hilbert et de différentiateurs.

Le logiciel mis en oeuvre aboutit à une amélioration très sensible des performances des filtres recherchés du point de vue précision de la réponse et complexité du filtre.

## SUMMARY

Digital filters use principles that have been well established. Their application to television signals however raises some specific problems. This specificity which has consequences on the implementation and design of the filters is due to the nature of the signals and to technological constraints.

The main features of video signals are their dynamic range, the necessity of linear phase processing for a good reproduction of picture details, and a wide spectrum. For most applications, these latter parameters lead to the design of digital transversal filters having symmetrical or antisymmetrical coefficients, together with gain constraints. Additionally due to the wide spectrum, fast technologies are required.

Owing to these constraints filters with the smallest number of coefficients are desirable. Furthermore, assuming that multipliers are devoted to rather specific applications, multiplications will be performed using read only memories or in a simplest way with adders. The latter solution implies that the filter has been designed in such a way that most coefficients have only one or two bits equal to one.

The most widely spread synthesis methods assume an infinite precision for the coefficients. Rounding of such coefficients may sometimes give results significantly different from the optimum. On the other hand there exists some other methods, perhaps less known but better optimized at the expense of an increased computer time. For example the algorithm described by Kodek [1] allows the optimal design in the Tchebycheff sense of transversal filters with finite wordlength coefficients. Through some extensions of this algorithm, it can be adapted to the constraints of different problems and the design of Hilbert transformers and differentiators can be considered.

The final result is a software well adapted to the design problems of video digital filters which gives a better approximation of the desired filtering function together with filters of reduced complexity.



SYNTHESE DE FILTRES NUMERIQUES POUR SIGNAUX VIDEO :  
SPECIFICITE DU PROBLEME ET SOLUTION OPTIMALE

DIGITAL FILTER DESIGN FOR VIDEO SIGNALS :  
SPECIFICITY OF THE PROBLEM AND OPTIMAL SOLUTION

J.Châtel , P.Siohan

## I - INTRODUCTION

La synthèse des filtres numériques a fait au cours des dix dernières années l'objet d'une activité intense, le nombre d'articles publiés sur le sujet en témoigne. Pour les filtres à réponse impulsionnelle finie (RIF) qui sont ici notre sujet de préoccupation, ce n'est qu'assez progressivement que sont apparues des méthodes de synthèse structurées, une étape importante ayant été franchie avec la synthèse de filtres à ondulation constante obtenue par la minimisation de la norme infinie de l'erreur sur la réponse amplitude fréquence.

Ces méthodes qui pour un nombre de coefficients donné conduisent à un filtre de réponse optimale déterminent en fait un algorithme de filtrage optimal. Pour mettre en oeuvre ce filtre, on est cependant amené à en quantifier les coefficients et donc à s'écarter de cet optimum. Pour un optimum plus proche des applications, ce n'est que beaucoup plus récemment que sont apparues des méthodes de synthèse en nombres entiers conduisant généralement à des solutions sensiblement plus performantes que la simple quantification des coefficients obtenus par les méthodes précédentes. Toutefois, bien que se rapprochant davantage des applications certaines méthodes de calcul de coefficients en nombres entiers gardent un caractère assez général. Pour que les résultats d'une synthèse de filtre soient immédiatement applicables il faut en effet que la méthode utilisée autorise toutes les contraintes du problème et permette le calcul de coefficients n'ayant qu'un nombre limité d'éléments binaires. Cet aspect prendra un caractère d'autant plus fort que l'on sera conduit à une réalisation en logique câblée.

On présente ici une méthode de synthèse de filtres qui peut s'adapter aux contraintes d'une application particulière, en l'occurrence celles des signaux vidéo. On en montre la souplesse et les performances après avoir précisément caractérisé et justifié les contraintes survenant dans ce type d'application.

## II - CARACTERES SINGULIERS DU FILTRAGE NUMERIQUE DE SIGNAUX VIDEO

Les particularismes du filtrage numérique des signaux vidéo tiennent en premier lieu à la nature de ces signaux et aux contraintes qu'ils induisent, puis par voie de conséquence, à des contraintes d'ordre technologiques.

Les signaux vidéo tels qu'ils sont habituellement traités dans les systèmes de télévision numérique se présentent sous la forme de trois composantes: luminance et différences de couleur, numérisées chacune à raison de 8 bits par échantillon. L'encombrement spectral de ces composantes conduit à des fréquences d'échantillonnage qui dépassent une dizaine de mégahertz pour la composante de luminance et avoisinent la moitié de cette valeur pour les différences de couleur. Lorsque le signal est traité sous forme composite, il arrive fréquemment que les fréquences d'échantillonnage utilisées atteignent 15 à 20 MHz.

A ces paramètres fondamentaux essentiellement significatifs d'un point de vue technologique doivent être associées des considérations sur la nature de l'information traitée. Une image est en effet avant tout autre chose une information de nature spatiale caractérisée globalement par des zones quasi uniformes et par des contours. Du traitement et de la resti-

tution des uns et des autres dépendra la qualité de l'image obtenue.

Pour les zones quasi uniformes c'est l'aspect quantification et précision de l'amplitude du signal qui se trouve mis en avant : dans ces zones l'observateur sera sensible à des petites variations de niveau. Ceci se traduit pour les filtres numériques par des contraintes sur le gain, tant pour assurer la transmission parfaite de la composante continue que pour éviter des bruits résiduels tels que ceux pouvant intervenir lors de traitements de changement de fréquence d'échantillonnage.

La qualité de l'image dans les zones comportant des contours ou des détails fins sera quant à elle dépendante de l'impression de netteté de l'image restituée. Dans ces zones la phase des transitions joue un rôle particulièrement important. Il est non seulement nécessaire de maintenir la coïncidence de phase des transitions des trois composantes de l'image, mais également dans une certaine mesure de maintenir la forme des transitions. Un filtrage passe-bas par exemple devra avoir un certain caractère homothétique, en influant essentiellement sur le temps de montée d'une transition plutôt que sur sa forme. Ces considérations conduisent à utiliser des filtres ayant une réponse phase-fréquence linéaire. Cette propriété sera par principe garantie avec des filtres RIF moyennant la symétrie ou l'antisymétrie de leurs coefficients.

D'une façon générale les filtres RIF répondent à l'équation aux différences suivante :

$$y(k) = \sum_{i=-n}^n h_i x(k-i) \quad (1)$$

dans laquelle  $x(k)$  et  $y(k)$  représentent respectivement les séquences d'entrée et de sortie du filtre et les termes  $h_i$  sa réponse impulsionnelle (la fréquence d'échantillonnage étant normalisée à 1).

On montre facilement que moyennant une propriété de symétrie ( $h_i = h_{-i}$ ) ou d'antisymétrie ( $h_i = -h_{-i}$ ) des coefficients de tels filtres, leurs réponses en fréquence peuvent s'exprimer sous la forme d'un terme réel éventuellement affecté d'un facteur exponentiel complexe significatif de la phase. On aura par exemple :

$$H(e^{j2\pi f}) = h_0 + 2 \sum_{i=1}^n h_i \cos 2\pi f i \quad (2)$$

Du fait de la symétrie des coefficients, l'expression du signal filtré pourra être réécrite sous une forme du type suivant :

$$y(k) = h_0 x(k) + \sum_{i=1}^n h_i (x(k+i) + x(k-i)) \quad (3)$$

De cette expression on déduit immédiatement la structure de filtres numériques presque universellement utilisée pour le filtrage de signaux vidéo. L'exemple qui en est donné à la figure 1 met en outre à profit la nécessité d'introduire dans le circuit des registres destinés à mémoriser les résultats intermédiaires du calcul.

A ces considérations générales sur les signaux vidéo et les propriétés qu'ils imposent aux filtres numériques qui leur sont destinés, s'ajoutent des contraintes d'ordre technologique liées à la valeur élevée des fréquences d'échantillonnage utilisées. Du fait de celles-ci, les opérateurs programmables sont exclus pour un fonctionnement en temps réel. Le seul niveau

SYNTHESE DE FILTRES NUMERIQUES POUR SIGNAUX VIDEO :  
 SPECIFICITE DU PROBLEME ET SOLUTION OPTIMALE  
 DIGITAL FILTER DESIGN FOR VIDEO SIGNALS :  
 SPECIFICITY OF THE PROBLEM AND OPTIMAL SOLUTION  
 J.Châtel , P.Siohan

de programmabilité permis se limitera tout au plus au chargement des coefficients du filtre dans un opérateur câblé, mais cette solution est si lourde qu'on la réserve à des applications très particulières.

Ces aspects technologiques et notamment la nécessité de mise en oeuvre de logiques câblées amènent, plus que pour tout autre application, à réduire le plus possible la complexité des filtres construits. On se donnera donc pour contrainte sévère la limitation du nombre de coefficients du filtre.

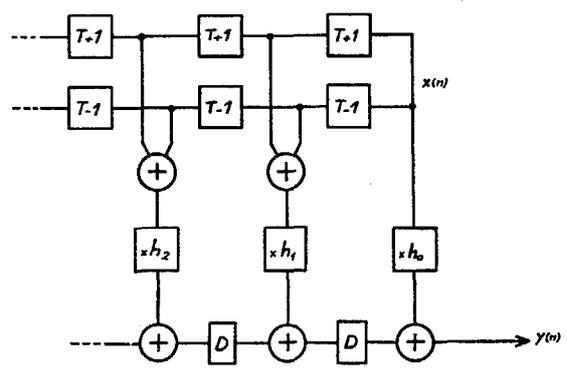


Figure 1 : Structure d'un filtre RIF à phase linéaire  
 D : Registre à décalage  
 T-1 : Élément de retard

Pour mettre matériellement en place ces coefficients, trois solutions sont envisageables :

- On peut utiliser des multiplieurs. Leur intérêt est de permettre une modification des coefficients. Ce sont cependant des circuits encombrants et coûteux, de sorte qu'on ne les attribuera qu'aux applications nécessitant une modification dynamique des coefficients.
- On peut également utiliser des mémoires mortes programmables dans lesquelles on charge le résultat de la multiplication du mot d'entrée par le coefficient à lui appliquer. Bien que n'ayant pas la souplesse de la précédente, cette solution n'impose que peu de contraintes au niveau des coefficients. Ce type de circuit, d'un coût très inférieur aux multiplieurs reste assez lourd de mise en oeuvre du fait de la programmation qu'il nécessite.
- Enfin la solution la plus avantageuse d'un point de vue matériel consiste à réaliser les coefficients au moyen d'additionneurs. C'est au moyen de ces circuits qu'on atteint les vitesses les plus élevées et que l'on obtient les mises en oeuvre les moins onéreuses. C'est par conséquent ce type de solution que l'on retiendra pour la grande majorité des applications aux signaux vidéo. En contrepartie on est amené pour utiliser cette méthode, à réduire à un ou deux le nombre d'éléments binaires de chaque coefficient lors de la synthèse du filtre.

Le problème de la synthèse de filtres numériques pour signaux vidéo se ramène donc en résumé à la recherche de filtres à réponse impulsionnelle finie, à coefficients symétriques ou antisymétriques, tenant compte non seulement de la quantification des coefficients mais réduisant ceux-ci dans les limites du possible, à un ou deux éléments binaires égaux à plus ou moins un. Cette synthèse doit en outre être effectuée avec au moins une contrainte de gain.

III.- METHODE DE SYNTHESE

Comme on le souhaite pour satisfaire aux contraintes des applications, la méthode de synthèse utilisée prend en compte la limitation du nombre de bits des coefficients du filtre. Pour la synthèse de filtres optimaux à coefficients quantifiés des méthodes spécifiques ont été proposées, qui s'avèrent plus performantes que les logiciels classiques de programmation linéaire mixte [3], elles associent des techniques de séparation et évaluation progressive (S.E.P ou en anglais Branch and Bound) à des algorithmes de programmation linéaire [4], ou, comme c'est le cas ici à l'algorithme de Remez [1]. Ce dernier a en effet l'avantage d'être plus rapide et, utilisé conjointement à une technique de S.E.P prenant en compte un certain nombre de contraintes [2], conduit à une méthode qui répond aux besoins spécifiques des problèmes posés.

III.1 - Formulation du problème

La recherche des coefficients quantifiés optimaux pour un filtre RIF à phase linéaire se ramène à un problème d'approximation où l'erreur minimale pondérée  $\delta_e$  est définie par :

$$\delta_e = \min_{\{\alpha_i^{(q)}\}} \max_{f \in F} |\hat{W}(f) | \hat{D}(f) - P_n(f) | \quad (4)$$

avec 
$$P_n(f) = \sum_{i=0}^{n-1} \alpha_i \cos 2\pi i f \quad (5)$$

F est le domaine de fréquences où l'approximation est calculée.

$\{\alpha_i^{(q)}\}$  est un ensemble de valeurs discrètes qui résulte de la quantification.

$\hat{D}(f)$  et  $\hat{W}(f)$  sont des fonctions réelles et continues sur F qui dépendent respectivement de la fonction désirée D(f) et de la fonction de pondération positive W(f).

Il existe quatre types de filtre RIF à phase linéaire, pour chacun d'entre eux le tableau n°1 résume les différentes relations qui permettent d'aboutir au problème d'approximation précédent.

	Cas 1	Cas 2	Cas 3	Cas 4
	N impair L=0	N pair L=0	N impair L=1	N pair L=1
n	$\frac{N+1}{2}$	$\frac{N}{2}$	$\frac{N-1}{2}$	$\frac{N}{2}$
$\hat{D}(f)$	D(f)	$\frac{D(f)}{\cos \pi f}$	$\frac{D(f)}{\sin 2 \pi f}$	$\frac{D(f)}{\sin \pi f}$
$\hat{W}(f)$	W(f)	W(f) cos $\pi f$	W(f) sin 2 $\pi f$	W(f) sin $\pi f$
$\{\alpha_i^{(q)}\}$	$\{i q\}$ si $i=0$ $\{2i q\}$ si $i \neq 0$ $-i_n \leq i \leq i_n$	$\{2^g \alpha_i\}$ $i q$ ; $-i_n \leq i \leq i_n$ $i_n = 2^{b-1} - 1$		

Tableau n°1 : Expressions de n,  $\hat{D}(f)$ ,  $\hat{W}(f)$ ,  $\{\alpha_i^{(q)}\}$  pour les 4 types de filtres numériques RIF à phase linéaire.



On peut noter dans le tableau n°1 que :

- N désigne la longueur (ou l'ordre) du filtre.
- L est un paramètre égal à 0 pour une réponse impulsionnelle symétrique et à 1 pour une réponse anti-symétrique.
- q est le pas de quantification, et est égal à  $1/2^{b-1}$ , b étant le nombre de bits (signe inclus).
- Dans le cas 1, l'ensemble  $\{\alpha_i^{(q)}\}$  est obtenu en supposant que les coefficients quantifiés du filtre  $\{h_i^{(q)}\}$  se situent dans l'intervalle  $[-1, +1]$ . Dans les autres cas cette hypothèse ne suffit pas pour déterminer l'ensemble  $\{\alpha_i^{(q)}\}$  qui va alors dépendre des coefficients  $\alpha_i$  solutions du problème d'approximation continu, d'où l'introduction du facteur d'échelle  $2^g(\alpha_i)$  ( $g(\alpha_i)$  étant une valeur entière positive ou nulle).

Du fait de la nature discrète des coefficients  $\{\alpha_i^{(q)}\}$ , l'algorithme de Remez\* ne peut pas s'appliquer directement pour résoudre le problème décrit par les relations (4) et (5), mais il peut cependant intervenir de manière importante.

### III.2 - Description générale de l'algorithme

L'algorithme peut être considéré comme un arbre où à chaque niveau on fixe un coefficient  $\alpha_i$  de  $P_n(f)$  à une valeur discrète  $\alpha_i^{(q)}$ . Il en découle un nouveau problème qui peut être résolu par l'algorithme de Remez si la condition de Haar\* est respectée, ceci implique de déterminer les coefficients séquentiellement en commençant par  $\alpha_{n-1}^{(q)}$ .

Ainsi le problème à résoudre à un niveau  $\ell$  quelconque ( $1 \leq \ell \leq n-1$ ) sera :

$$\delta_\ell = \min_{\{\alpha_i\}; 0 \leq i \leq n-\ell-1} \max_{f \in F} \hat{W}(f) |\hat{D}_\ell(f) - P_{n-\ell}(f)| \quad (6)$$

$$\text{avec } \hat{D}_\ell(f) = \hat{D}(f) - \sum_{i=n-\ell}^{n-1} \alpha_i^{(q)} \cos 2\pi i f \quad (7)$$

$$\text{et } P_{n-\ell}(f) = \sum_{i=0}^{n-\ell-1} \alpha_i \cos 2\pi i f \quad (8)$$

Les relations précédentes montrent que  $\min \delta_\ell$  est une fonction croissante de  $\ell$ . Si on considère que l'erreur d'arrondi  $\delta_a$  est le point de départ de l'algorithme, lorsque à un niveau  $\ell$ ,  $\delta_\ell > \delta_a$ , il est inutile d'examiner les branches qui en découlent.

Une première solution est obtenue quand  $n$  est inférieur à  $\delta_a$ . Cette nouvelle valeur appelée erreur optimale courante, et notée  $\delta_{0,c}$  se substitue à  $\delta_a$  dans la recherche de l'optimum.

\*Remez : l'algorithme d'échange de Remez est un algorithme itératif qui permet de réaliser la meilleure approximation au sens de Tchebycheff d'une fonction  $\hat{D}(f)$  par un polynôme  $P_n(f)$  (5). Il utilise la caractérisation d'alternance de l'erreur pour obtenir les coefficients continus optimaux du polynôme.

\*Haar : la condition de Haar implique qu'un polynôme  $P_n(f)$  appartenant à un sous-espace vectoriel de dimension  $n$  dans l'espace des fonctions continues admette au plus  $n-1$  zéros.

Lorsque tous les noeuds de l'arbre ont été explorés  $\delta_e = \delta_{0,c}$ , et au chemin optimal qui conduit à cette erreur correspond un ensemble de coefficients optimaux de  $\{\alpha_i^{(q)}\}$ . Le mode d'exploration de l'arbre dépend de la stratégie adoptée dans l'algorithme S.E.P la méthode utilisée ici est dite de "profondeur d'abord".

### III.3 - Optimisation avec contraintes

Lorsque des contraintes interviennent sur les coefficients quantifiés du filtre, l'algorithme précédent ne peut s'appliquer directement. Dans ce cas il faut inclure dans la procédure de S.E.P, les contraintes correspondantes sur les coefficients du polynôme. L'ensemble des contraintes qui interviennent dans les applications décrites ici, concernent des filtres de type 1 (voir tableau n°1) et peuvent être classées comme suit :

#### a) Contraintes d'égalité

- Dans le cas de filtres demi-bande, un coefficient sur deux est nul, et, le coefficient central vaut 0.5. Lors de leur synthèse on imposera donc :

$$\begin{aligned} \alpha_i^{(q)} &= 0 & i = 2p \quad (p \neq 0) \\ 2\alpha_0^{(q)} &= 1 \end{aligned} \quad (9)$$

- De manière plus générale, l'objectif d'interpolation (ou de réduction) par un facteur 2 peut être atteint en appliquant les 2 contraintes qui suivent :

$$\begin{aligned} \sum_{i=0}^p \alpha_{2i}^{(q)} &= 1 & ; \quad p \leq \frac{n-1}{2} \\ \sum_{i=0}^{p'} \alpha_{2i+1}^{(q)} &= 1 & ; \quad p' \leq \frac{n-2}{2} \end{aligned} \quad (10)$$

On peut noter d'une part que la réalisation de ces 2 dernières égalités permet de satisfaire la contrainte de gain unité, d'autre part qu'une généralisation des relations précédentes (9), (10) est applicable à des filtres d'interpolation (ou de réduction) par des rapports entiers différents de 2 [2].

#### b) Contraintes d'inégalité

Dans les bandes de transition des filtres, il est parfois indispensable pour respecter un gabarit de contraindre de manière ponctuelle les variations de la réponse en fréquence. Ce qui se traduit dans l'algorithme S.E.P par l'inclusion d'une ou plusieurs contraintes du type suivant :

$$m_j \leq \sum_{i=0}^{n-1} \alpha_i^{(q)} \cos 2\pi i f_j \leq M_j ; f_j \in F \quad (11)$$

#### c) Contraintes de codage

Certains codages particuliers des coefficients des filtres permettent d'éviter l'emploi de multiplieurs, les multiplications étant alors décomposées en un nombre réduit d'additions et de décalages. La manière la plus simple est de se limiter à 1 bit à +1 ou -1 par coefficient, mais il est également possible d'accepter 2 bits à +1 ou -1, l'expression codée des coefficients étant :

$$h_i = \sum_{k=1}^2 S_k(i) 2^{r_k(i)} \quad (12)$$

avec  $S_k(i) = +1, 0, -1$  et  $r_k(i)$  un entier dans l'intervalle  $[-(b-1), 0]$ .

L'arbre de l'algorithme est alors généré de manière à limiter le domaine de recherche à des solutions du type précédent.

## IV - INTERPOLATION

Les changements de fréquence d'échantillonnage sont des opérations assez fréquentes dans le traitement numérique des signaux vidéo. Dans ce domaine les filtres demi-bande (ou plus généralement  $N^{\text{ième}}$ -Bande) présentent de plus des caractéristiques qui facilitent leur réalisation.



SYNTHESE DE FILTRES NUMERIQUES POUR SIGNAUX VIDEO :  
SPECIFICITE DU PROBLEME ET SOLUTION OPTIMALE  
DIGITAL FILTER DESIGN FOR VIDEO SIGNALS :  
SPECIFICITY OF THE PROBLEM AND OPTIMAL SOLUTION  
J.Châtel , P.Siohan

**Exemple 1 :** Il s'agit danc ce 1er exemple de réaliser la synthèse d'un filtre demi-bande de longueur 19 dont les coefficients sont codés sur 9 bits. Les contraintes à vérifier sont d'une part la (9) car il s'agit d'un filtre demi-bande et d'autre part la (12) pour aboutir à une réalisation simplifiée. La fréquence d'échantillonnage est de 18 MHz, et la bande de transition a une largeur de 3 MHz. Les différents résultats sont reportés à la figure n°2, on peut noter le gain important près de 5,5 dB dans la bande atténuée (de 34.44 dB à 39.80 dB) dû à l'optimisation.

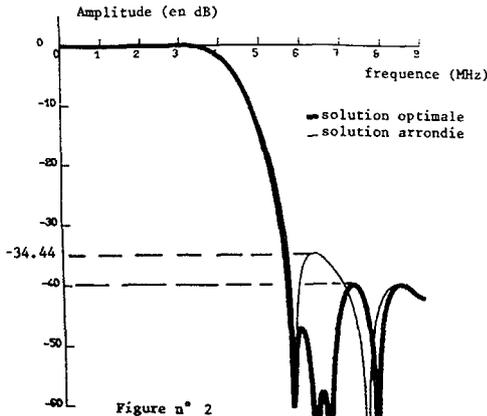


Figure n° 2  
Filtres demi-bande avec N=19, b=9

Coefficients (non nuls) des filtres multipliés par  $2^8=256$  et codés avec 2 puissances de 2

Solution arrondie	Solution optimale
$h_9 = +2^0 = h_{-9}$	$h_9 = +2^1 = h_{-9}$
$h_7 = -2^2+2^0 = h_{-7}$	$h_7 = -2^2-2^0 = h_{-7}$
$h_5 = +2^3+2^0 = h_{-5}$	$h_5 = +2^3+2^1 = h_{-5}$
$h_3 = -2^4-2^3 = h_{-3}$	$h_3 = -2^4-2^3 = h_{-3}$
$h_1 = +2^6+2^4 = h_{-1}$	$h_1 = +2^6+2^4 = h_{-1}$
$h_0 = +2^7$	$h_0 = +2^7$

**V - SELECTION DE BANDE**

Les filtres sont les plus souvent utilisés pour de la sélection de bande, et leurs réponses en fréquence tenues de respecter des gabarits très stricts, comme c'est le cas dans l'exemple qui suit :

**Exemple 2 :** On cherche à réaliser la synthèse d'un filtre passe-bas effectuant en outre une interpolation par un facteur 2, sa longueur est de 11 avec un codage des coefficients sur 8 bits. Les spécifications en décibels pour l'amplitude et en mégahertz pour la fréquence sont :

(La fréquence d'échantillonnage est de 13.5 MHz)  
Bande passante  $0 \leq f \leq 0.5$  ;  $-0.5 \leq 20 \log D(f) \leq -0.2$   
Bande atténuée  $3.4 \leq f \leq 6.75$  ;  $20 \log D(f) \leq -30$

Dans la bande de transition on impose :

pour $f = 1.3$	20 log D(f) $\geq -3$	
$f = 1.5$	20 log D(f) $\leq -3$	(13)
$f = 3$	20 log D(f) $\leq -20$	

Les contraintes prises en compte par l'algorithme S.E.P sont donc : les contraintes d'égalité (10) pour réaliser l'interpolation et les 3 contraintes d'inégalité du type (11) avec des valeurs  $f_j, m_j, M_j$  calculées pour respecter les relations (13). L'objectif de la synthèse est d'obtenir le filtre ayant une atténuation maximale et qui bien sûr vérifie les contraintes du problème.

Les résultats obtenus par l'algorithme optimal sont comparés (tableau n°2) à ceux dus à une autre quantification qui permet également d'obtenir le gabarit souhaité.

On peut noter que l'optimisation apporte un gain de plus de 3 dB et que les valeurs entières des coefficients permettent d'envisager des réalisations simplifiées avec un codage sur 7 bits.

Type de quantification	Coefficients du filtre multipliés par $2^7=128$	Atténuation (en dB)
Sous-optimale	-1 -2 3 16 30 36	32.22
Optimale	-1 -1 4 16 29 34	35.31

Tableau n° 2 : Comparaison de 2 quantifications, optimale et sous-optimale, pour le filtre passe-bas de l'exemple 2.

A la figure n°3 sont représentées les courbes en fréquence pour chacune des solutions.

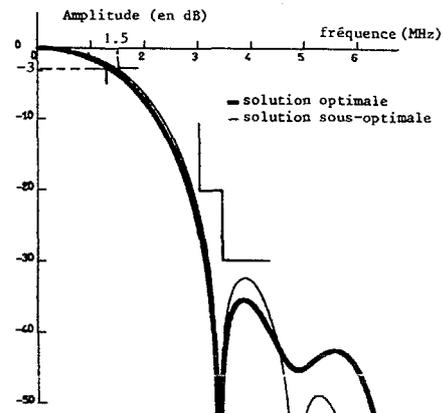


Figure n° 3 : Réponses en fréquence pour un passe-bas (N=11, b=8)

**VI - AUTRES APPLICATIONS**

La méthode présentée s'applique également à d'autres types de filtres, ainsi à des différentiateurs ou à des transformateurs de Hilbert [6] .

**Exemple 3 :** On effectue la synthèse d'un différentiateur large bande, de pente égale à 1, de longueur 16, avec 8 bits par coefficients. Les résultats obtenus sont reportés au tableau n°3, les coefficients optimaux étant dans ce cas particulier ceux qui minimisent l'erreur relative.

Type de solution	Coefficients du filtre multipliés par $2^7=128$	Erreur relative
Arrondie	0 0 0 0 -1 1 -3 26	0.2635
Optimale	0 0 0 0 0 1 -3 26	0.0798

Tableau n° 3 : Comparaison des solutions quantifiées, arrondie et optimale, pour un différentiateur large bande.

Il apparaît que pour chaque type de solution une réalisation câblée est envisageable, même à des fréquences élevées, et que encore une fois les performances en filtre optimal améliorent considérablement la solution arrondie. Les courbes d'erreur relative sont comparées à la figure n°4, ci-après.

Les transformateurs de Hilbert permettent la mise en quadrature de 2 signaux, et leur réalisation sous forme de filtres numériques RIF à phase linéaire présente un intérêt particulier car l'erreur sur le déphasage est alors nulle.

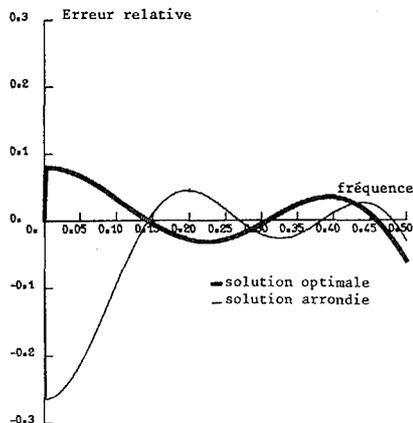


Figure n°4 : Courbes d'erreur relative d'un différentiateur d'ordre 25 avec des coefficients codés sur 8 bits

## VII - CONCLUSION

La méthode de synthèse proposée montre sa souplesse d'adaptation à toutes les contraintes d'un problème donné, en particulier à celles s'associant à une recherche des moyens de mise en oeuvre les plus simples. Elle s'avère d'une grande utilité pratique puisque la réponse du filtre synthétisé est celle du filtre qui sera effectivement câblé et que le résultat obtenu est donné sous une forme propre à une réalisation immédiate.

D'un point de vue performances, malgré le nombre réduit de coefficients, on ne peut que remarquer l'amélioration très significative apportée aux caractéristiques d'un filtre de longueur donnée par l'utilisation de la méthode proposée plutôt que par la simple quantification avec arrondi de coefficients obtenus à précision infinie.

## REFERENCES BIBLIOGRAPHIQUES

- 1 D.Kodek "An algorithm for the design of optimal finite wordlength FIR digital filters", Proceedings IEEE ICASSP-1980.
- 2 P.Siohan, A.Benslimane "Synthèse des filtres numériques non récursifs à phase linéaire et coefficients de longueur finie : différentes approches et applications" A paraître dans les Annales des Télécommunications.
- 3 D.Kodek "Design of optimal finite wordlength FIR digital filters using integer programming techniques", IEEE transactions on ASSP, n°3 - Juin 1980
- 4 JP.Marques de Sà "A new design method of optimal finite wordlength linear phase FIR digital filters" IEEE transactions on ASSP, n°4 - Août 1983.
- 5 RK.Patney, SC.Dutta Roy "Design of linear phase FIR filters using pseudo-boolean methods" IEEE transactions on circuits and systems, n°4 - Avril 1979.
- 6 LR.Rabiner, B.Gold "Theory and application of digital signal processing" Prentice Hall 1975.

Exemple 4 : On veut réaliser la synthèse d'un transformateur de Hilbert ayant 10 coefficients ( $N=20$ ) codés chacun sur 8 bits et dont la bande de transition en fréquence normalisée va de 0.0 à 0.05. Les différents résultats sont comparés au niveau des coefficients et de l'erreur absolue (tableau n°4)

Type de solution	Coefficients du filtre multipliés par $2^7=128$	Erreur absolue
Arrondie	2 2 3 4 5 7 10 15 26 81	0.0424
Optimale	1 2 3 4 5 7 10 15 27 81	0.0325

Tableau n° 4 : Comparaison des solutions arrondie et optimale de l'exemple 4

Les courbes d'erreur sont représentées à la figure n°5.

Contrairement à l'exemple précédent le choix d'un nombre pair de coefficients n'est en général pas le meilleur pour un transformateur de Hilbert. L'intérêt de cet exemple est davantage de montrer l'efficacité de la méthode par rapport à d'autres, en particulier celle de la référence [5] qui dans ce cas précis donne une erreur de 0.03315.

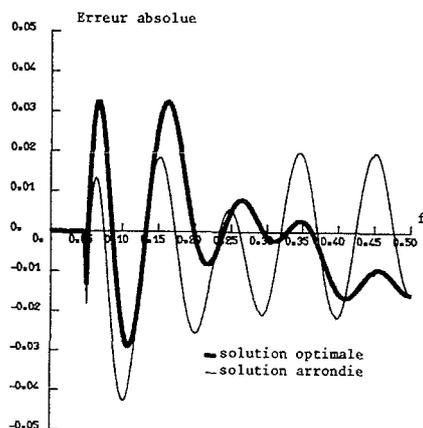


Figure n°5 : Courbes d'erreur absolue d'un transformateur de Hilbert ayant des coefficients codés sur 8 bits