

APPROXIMATION PAR DES FRACTIONS RATIONNELLES  
EN NORME  $L^2$   
(APPLICATION AU FILTRAGE NUMÉRIQUE)

par

JAQUES WOLF

(Grenoble)

Soit  $f$  un élément de  $F(\subset L_w^2[-1, +1])$ , espace précisé ultérieurement, que l'on désire approcher par un élément  $g^* \in V_n^m$ -sous-ensemble de l'ensemble des fractions rationnelles dont les degrés des numérateur et dénominateur sont respectivement  $m$  et  $n$ , de telle manière que :

$$\|g^* - f\|_F = \min_{g \in V_n^m} \|g - f\|_F$$

On pose :

$$F = \{f / f(x) \geq 0 \quad \forall x \in [-1, +1], f \in L_w^2[-1, +1], \forall I \neq \emptyset \subset [-1, +1] \\ f(x) \neq 0 \text{ sur } I\}$$

$$V_n^m = \left\{ g/g(x) = \frac{\sum_0^m b_k x^k}{\sum_0^n a_k x^k}, \sum_0^m b_k x^k \geq 0 \quad \forall x \in [-1, +1], \right. \\ \left. \sum_0^n a_k x^k > 0, \quad \forall x \in [-1, +1], \sum_0^n a_i^2 + \sum_0^m b_i^2 = 1 \right\}.$$

1. Existence d'un meilleur approximant

Soit  $c^t = [a^t | b^t] = [a_0, \dots, a_n | b_0, \dots, b_m] \in \mathbb{R}^{n+m+2}$ .

Chercher  $g^* \in V_n^m$  revient à chercher — s'il existe —  $c^* \in \mathfrak{B}_1 \cap \mathfrak{B} \cap \mathcal{Q}$  tel que

$$(1) \quad Q(c) = \int_{-1}^{+1} w(x) \left[ \frac{\sum_{k=0}^m b_k x^k}{\sum_{k=0}^n a_k x^k} - f(x) \right]^2 dx$$

soit minimum pour  $c = c^*$  avec  $\mathfrak{B}_1$  boule unité de  $\mathbf{R}^{n+m+2}$ , et

$$\mathfrak{B} = \{c \in \mathbf{R}^{n+m+2} / \sum b_k x^k \geq 0 \forall x \in [-1, +1]\},$$

$$\mathfrak{A} = \{c \in \mathbf{R}^{n+m+2} / \sum a_k x^k > 0 \forall x \in [-1, +1]\}.$$

On montre [2] que le minimum de  $Q$  — s'il existe — appartient à un compact sur lequel  $Q$  est continue donc :

Il existe une meilleure approximation de  $f \in F$  par un élément de  $V_n^m$ .

### 2. Non unicité (et unicité)

Si  $f \in F$  on ne peut assurer l'unicité du meilleur approximant  $g^* \in V_n^m$  ( $\forall m, n$ ), pour toute fonction poids  $w(x)$ .

Contre-exemple

Soit

$$f(x) = \begin{cases} 1 & x \in \left[-\frac{1}{2}, +\frac{1}{2}\right] \\ 0.1 & x \notin \left[-\frac{1}{2}, +\frac{1}{2}\right] \end{cases}$$

et soit  $w(x) = 1$

On a alors

$$Q(c) = \int_{-1}^{+1} \left[ \frac{b_0 + b_1 u}{a_0 + a_1 u} - f(u) \right]^2 du.$$

Soit  $\varphi^*(u) = \frac{b_0^* + b_1^* u}{a_0^* + a_1^* u}$  un meilleur approximant de  $f$ . Il est immédiat de

voir que  $\varphi^-(u) = \frac{b_0^* - b_1^* u}{a_0^* - a_1^* u}$  est aussi un meilleur approximant de  $f$ .

On a alors : unicité  $\Rightarrow \varphi^*(u) = \varphi^-(u) \forall u$ ; mais  $\varphi^*(u) = \varphi^-(u) \Leftrightarrow \varphi^*(u) = \varphi^*(-u) \Leftrightarrow b_1^* = a_1^* = 0$ ; donc s'il y a unicité on doit avoir  $\varphi^*(u) = b_0$  (c'est-à-dire une constante).

On a :

$$Q(b_0, 0, 0) = \min \int_{-1}^{+1} [b_0 - f(u)]^2 du = \|g\|^2 = 0.51005 \quad (b_0^* = 0.5005)$$

et si on calcule

$$Q(0.53, 0.415, 0.5) = \|g\|^2 = 0.5241 < Q(b_0, 0, 0)$$

$\begin{matrix} \uparrow & \uparrow & \uparrow \\ b_0 & b_1 & a_1 \end{matrix}$

(en prenant  $a_0 = 1$  ce qui est équivalent à prendre  $c$  sur la boule unité de  $R^4$ ).

Ceci montre donc que le meilleur approximant de  $f$  (qui est une fonction paire), n'est pas pair, soit tel que  $a_1 = b_1 = 0$ .

Il n'y a donc pas unicité du meilleur approximant.

\* \* \*

Cependant il est assez rapide de voir que si l'on choisit  $m = 0$  et  $n = 1$  il y a unicité du meilleur approximant.

On pose le problème suivant : à quelles conditions le meilleur approximant est-il unique ?

- pour quelle classe de fonctions  $f$  ?
- y-a-t-il unicité si  $m < n$  ?

### 3. Application au filtrage numérique

Construire un filtre numérique consiste à trouver deux jeux de coefficients  $(\beta_0, \beta_1, \dots, \beta_m)$  et  $(\alpha_0, \dots, \alpha_n)$  tels que la réponse en fréquence du filtre [1]

$$H(e^{i\omega\Delta T}) = \frac{\sum_0^m \beta_p e^{-ip\omega\Delta T}}{\sum_0^n \alpha_p e^{-ip\omega\Delta T}}$$

satisfasse un certain nombre de conditions fixées à l'avance.

Le filtre tel que : i)

$$(2) \quad \int_0^\pi w(x) \left[ \left| \frac{\sum_0^m \beta_p e^{-ipx}}{\sum_0^n \alpha_p e^{-ipx}} \right| - f(x) \right]^2 dx \text{ soit minimum.}$$

( $f(x)$  étant un modèle donné à l'avance).

ii) le polynome  $\sum_0^n \alpha_k z^{n-k}$  possède toutes ses racines de module  $< 1$  (condition de stabilité) sera dit filtre optimal au sens de la norme de  $L_w^2[0, \pi]$  ( $w(x) > 0$ );  $n$  est appelé l'ordre du filtre.

L'application des § 1 et § 2 au filtrage numérique est alors immédiate puisque

$$\sum_{p=0}^k \delta_p e^{-ipx} = \sum_{p=0}^k d_p \cos(px) = \sum_0^k h_p \cos^p x$$

et le changement de variable  $\cos x = u$  dans l'intégrale (2) permet de passer à l'intégrale (1), c'est-à-dire à la recherche d'une fraction rationnelle optimale.

*Exemple. Filtrage passe-bas optimaux*

On peut tout d'abord remarquer que le modèle  $h(x)$  du filtre passe-bas idéal n'appartient pas à  $F$ .

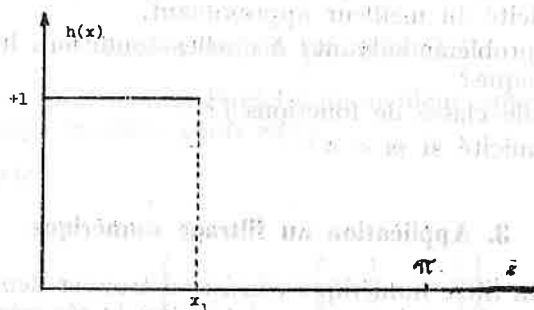


Fig. 3.1

On prend alors pour modèle,  $f(x)$ , d'un filtre passe-bas: ( $\epsilon$  constante  $> 0$  fixée, aussi petite que désirée)

$$f(x) = \begin{cases} 1 & x \in [0, x_1] \\ \frac{(1 - \cos x_1)(1 + \cos(x))}{1 - \cos(x_1) \cos(x)} \epsilon & \text{sinon ce modèle appartient bien à } F \end{cases}$$

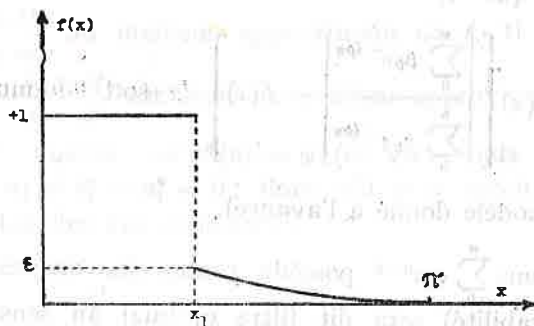


Fig. 3.2

Soit  $F_{x_1}$  le filtre optimal correspondance à modèle. On a le résultat suivant :

**THÉORÈME.** On pose  $m = n$ , et  $w(x) = 1 + \text{tg}^2(x/2)$ . Soient  $(\beta^1)^t = [\beta_0^1, \dots, \beta_m^1]$  et  $(\alpha^1)^t = [\alpha_0^1, \dots, \alpha_m^1]$  les coefficients du filtre optimal passe-bas  $F_{x_1}$ . Soient  $(\beta^2)^t$  et  $(\alpha^2)^t$  les coefficients du filtre optimal passe-bas  $F_x$ . Il existe une matrice  $M$  à  $(m + 1)$  lignes et à  $(m + 1)$  colonnes telle que

$$\beta^2 = M\beta^1 \quad \text{et} \quad \alpha^2 = M\alpha^1$$

*Remarque.* A partir d'un filtre  $F_{x_1}$  fixé, pour un ordre donné, on peut construire avec des transformation simples tous les autres filtres passe-bas ou bien passe-haut de même ordre.

Pour construire des filtres passe-bande (ou multi-bande) à partir de ce filtre  $F_{x_1}$ , dit filtre de base le problème n'est pas encore entièrement résolu car on ne sait pas s'il y a unicité dans le cas  $m < n$ .

Ces filtres passe-bande, quoique excellents comme le montre la figure 5.3, ne peuvent être appelés optimaux.

**4. Calcul numérique d'un meilleur approximant**

Le calcul de la meilleure approximation — c'est-à-dire du filtre optimal — est difficile pour plusieurs raisons :

- a) la fonction à minimiser ne possède pas de "bonnes propriétés", qu'on espère trouver dans les problèmes de minimisation c'est-à-dire forme quadratique ou fonction convexe.
- b) on doit minimiser la fonction  $Q(c)$  sous un ensemble de contraintes en nombre infini.

Dans le cas de minimisation avec contraintes on a choisi une méthode de pénalisation intérieure. Les résultats ont été assez décevants : lenteur de la convergence (le filtre obtenu ne présentant pas certaines propriétés d'alternance du filtre optimal).

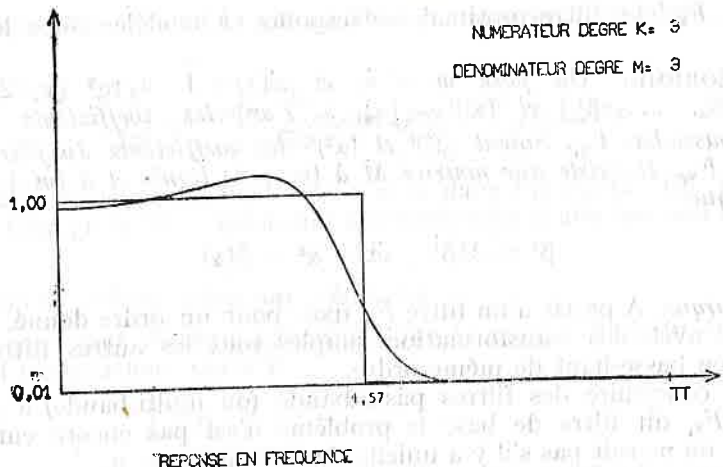
Cependant dans le cas où  $n = 0$  (pas de dénominateur), une méthode particulièrement appropriée a été mise au point [3].

c) en réalité on peut se ramener à un problème sans contraintes en les inconnues  $[a_0, \dots, a_n, \beta_0, \dots, \beta_m]$  (les inconnues du dénominateur n'étant pas actives). Pour cette minimisation sans contraintes la méthode de Fletcher et Powell s'est révélée très efficace.

Les résultats obtenus par cette méthode sont ceux des exemples du paragraphe § 5.

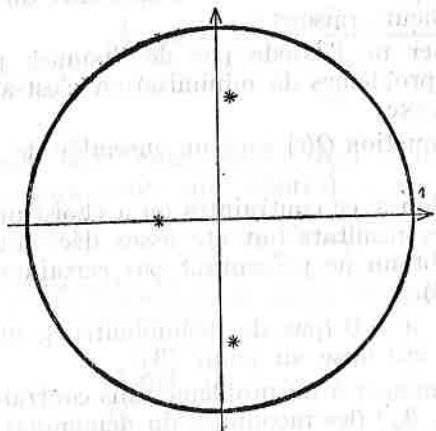
**5. Exemples**

La figure 5.1 montre le filtre passe-bas, de base, d'ordre 3 qui permet de calculer tous les autres filtres passe-bas optimaux d'ordre 3 et des filtres passe-haut ou passe-bande. Sur chaque figure est dessiné :



NUMERATEUR DEGRE K= 3  
DENOMINATEUR DEGRE M= 3

REPOSE EN FREQUENCE



EMPLACEMENT DES POLES

COEFFICIENTS DU FILTRE :

ALPHA :

- 0 1,000000
- 1 0,186489
- 2 0,371160
- 3 0,129237

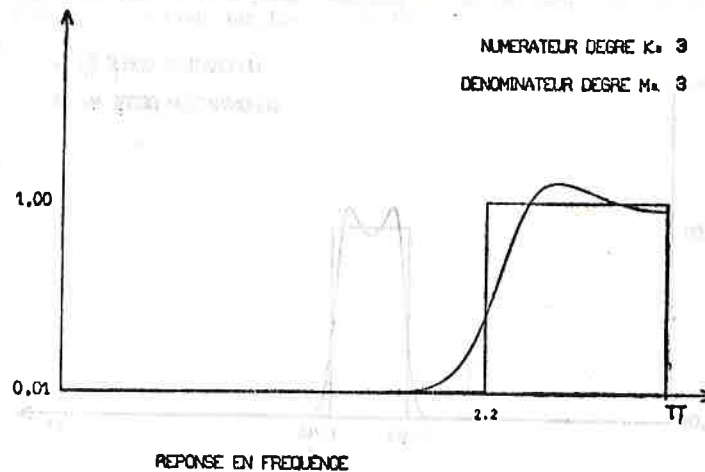
BETA :

- 0 0,270017
- 1 0,554612
- 2 0,557008
- 3 0,272414

Fig. 5.1

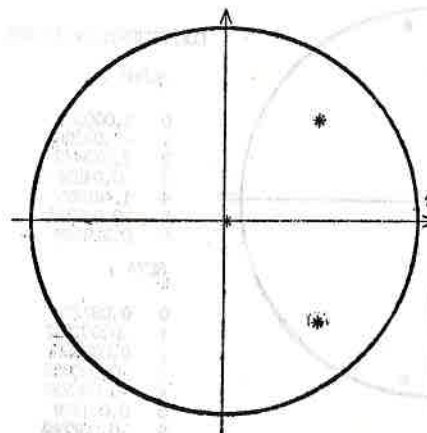
- le modèle (carré du module de la réponse en fréquence)
- le carré du module de la réponse du filtre optimal
- la répartition des pôles à l'intérieur du cercle unité (ce qui permet d'avoir une estimation de la durée de la phase transitoire due au filtre).
- les coefficients cherchés du filtre étudié.

La figure 5.2 montre un filtre passe-bas de pulsation de coupure (normalisée) : 2,2. Ce filtre optimal est obtenu à partir du filtre précédent 5.1 par utilisation du théorème du § 3 (c'est-à-dire sans aucun calcul de minimisation).



NUMERATEUR DEGRE K= 3  
DENOMINATEUR DEGRE M= 3

REPOSE EN FREQUENCE



EMPLACEMENT DES POLES

COEFFICIENTS DU FILTRE :

ALPHA :

- 0 1,000000
- 1 0,987370
- 2 0,526511
- 3 0,006211

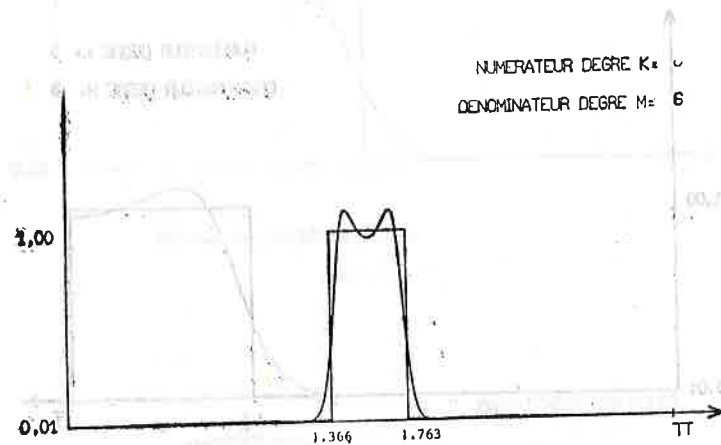
BETA :

- 0 0,142329
- 1 -0,116619
- 2 0,118103
- 3 -0,144415

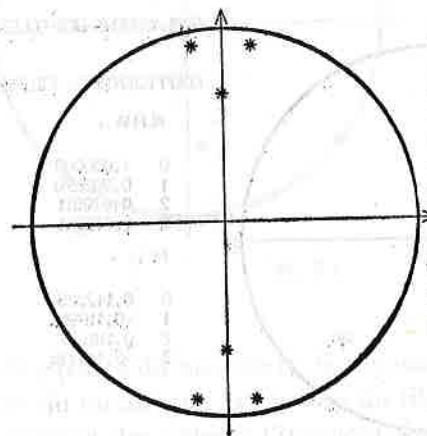
Fig. 5.2

La figure 5.3 montre un filtre passe-bande obtenu comme le précédent à partir du filtre de base, mais cette fois-ci l'ordre est multiplié par 2 l'ordre est donc 6.

Ces différents résultats montrent que ces filtres optimaux, au sens de la norme  $L^2$ , sont au moins aussi bons que les filtres plus classiques (Butterworth, Tchebycheff, ...) obtenus avec d'autres types d'approximation.



REPOSE EN FREQUENCE



EMPLACEMENT DES POLES

COEFFICIENTS DU FILTRE

## ALPHA :

0	1,000000
1	-0,032567
2	2,070443
3	-0,045374
4	1,461995
5	-0,015715
6	0,328289

## BETA :

0	0,067297
1	-0,001532
2	0,036685
3	-0,000006
4	-0,036239
5	0,001538
6	-0,067746

Fig. 5.3

## BIBLIOGRAPHIE

- [1] Rader-Gold, *Digital filter design techniques in the frequency domain*. Proceedings of the IEEE Vol. 55, n° 2 (1967).
- [2] Wolf, J., *Meilleure approximation en filtrage numérique*. C.R.A.S. Paris t. 276 série A (1973).
- [3] Wolf, J., *Minimisation d'une forme quadratique avec une infinité de contraintes*. Colloque C.N.R.S. : La Colle sur Loup (juin 1973).

Requ le 7. I. 1978

Université de Grenoble