

Application des filtres miroirs à un
procédé de codage par découpage en
sous-bandes

A. Roset

80-03

INRS-Télécommunications
a/s Recherches Bell-Northern Ltée
3, Place du Commerce
Ile des Sœurs, Qué.
H3E 1H6

Rapport technique de l'INRS-Télécommunications n° 80-03

Janvier 1980

RESUME

APPLICATION DES FILTRES MIROIRS A UN PROCEDE DE CODAGE PAR DECOUPAGE EN SOUS-BANDES

Ce rapport traite d'une application des filtres miroirs dans un procede de codage par decoupage en 8 sous-bandes d'une largeur de 500 Hz. Ce schema permet de tirer profit des differents niveaux de puissance sur un long terme et des differents seuils de perception de bruit dans chaque bande.

Apres une rapide mise au point par simulation sur ordinateur le codeur a 16 Kb/s obtenu donne des phrases codees tres intelligibles avec conservation de la tonalite mais presentant une legere reverberation.

ABSTRACT

APPLICATION OF QUADRATURE MIRROR FILTERS TO SPLIT BAND VOICE CODING PROCESS

This report discusses an application of quadrature mirror filters for an 8 sub-band coder; this system allows us to take advantage of the differences in the long term power and of the just noticeable noise in each band.

A computer simulation of the coder at 16 Kb/s gives fully intelligible coded sentences which preserve the original tonality but with slight reverberation.

TABLE DES MATIERES

CHAPITRE 1:	<u>Introduction</u>
CHAPITRE 2:	<u>Les moyens disponibles à l'INRS</u>
CHAPITRE 3:	<u>Principes théoriques généraux</u>
	1** remarques théoriques préliminaires
	2** partage en deux bandes avec des filtres miroirs
	3** comportement en cas de bruit additif
	4** partage en plusieurs bandes par des filtres miroirs successifs
CHAPITRE 4:	<u>Réalisation pratique des filtres</u>
	1** réalisation du filtre de base optimal
	2** le système à huit bandes avec ce filtre de base
CHAPITRE 5:	<u>Détermination du rapport signal à bruit subjectif</u>
CHAPITRE 6:	<u>Principe des expérimentations</u>
CHAPITRE 7:	<u>Premières évaluations</u>
	1** premières séries de tests
	2** deuxième série de tests
	3** série de tests théoriques
CHAPITRE 8:	<u>Schéma de codeur à 16 Kb/s</u>
	1** comparaison des configurations à 16 Kb/s
	2** essais avec le codeur
	3** deuxième configuration
	4** application à un codeur à 9,6 Kb/s
CHAPITRE 9:	<u>Conclusion</u>

CHAPITRE 1

INTRODUCTION

Le travail effectué pendant ce stage a permis de soulever et de résoudre en partie les problèmes posés par la transmission de la parole à un débit réduit dont le procédé de codage est basé sur l'utilisation des filtres miroirs.

On trouvera dans les premiers chapitres la description théorique des filtres miroirs et de l'application au système de transmission par codage des sous-bandes. Puis vient ensuite la construction de tels filtres. Enfin les résultats expérimentaux sont regroupés dans les derniers chapitres.

Parallèlement on peut tirer des conclusions sur la perception de bruit dont la répartition spectrale n'est pas uniforme.

CHAPITRE 2

LES MOYENS DISPONIBLES A L'INRS

1**Le matériel:

Pour le traitement de la parole l'INRS dispose d'un PDP 11/45 utilisé en temps partagé, équipé d'un convertisseur analogique digital travaillant sur des mots de 15 bits, d'un "array processor" permettant d'accélérer les gros programmes de calculs et d'une "frame memory" permettant diverses visualisations.

2**Le logiciel:

De très nombreux programmes ont été développés pour utiliser facilement le matériel: un interpréteur de commandes pour le convertisseur (échantillonnage, écoute, stockage, calcul de statistiques), un autre pour visualiser les signaux et les spectrogrammes relatifs à un signal échantillonné avec la possibilité d'écoute simultanée du signal, des sousprogrammes assembleur pour l'utilisation de l'"array processor".

De plus pour le sujet à traiter, plusieurs programmes étaient utilisables: un générateur de bruit blanc multiplicatif, un ensemble de programmes organisant les tests de comparaison de signaux auditifs, un ensemble de programmes permettant le calcul et la visualisation de filtres numériques et de leurs comportements.

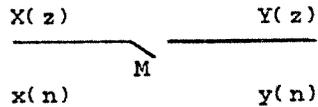
CHAPITRE 3

PRINCIPES THEORIQUES GENERAUX

1**Remarques théoriques préliminaires:

Le sous-échantillonnage:

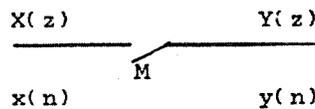
Cette opération consiste à ne garder qu'un échantillon sur M. Cela change la fréquence d'échantillonnage du signal qui passe de F à F/M.



$$\begin{aligned} y(n) &= x(M \cdot n) \\ Y(z) &= 1/M \sum_{m=0, \dots, M-1} X(\exp(2jm\pi/M) \cdot z^{1/M}) \end{aligned}$$

L'interpolation:

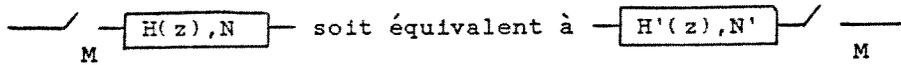
Cette opération consiste à insérer M échantillons entre 2 échantillons initiaux. Cela change la fréquence d'échantillonnage qui passe de F à M.F.



$$\begin{aligned} y(M \cdot p + r) &= x(p) \cdot \delta(r) \\ Y(z) &= X(z^M) \end{aligned}$$

Les problèmes de commutation entre filtrage sous échantillonnage et interpolation:

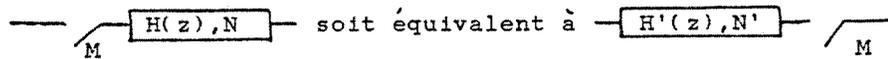
Pour que



il suffit que $H(z) = H'(\exp(2jm\pi/M) \cdot z^{1/N})$ pour $m=0, \dots, N-1$
 soit $h'(M.p+r) = h(p) \cdot \delta(r)$

les coefficients sont obtenus par insertion de M coefficients nuls entre deux coefficients de H. Si H a N coefficients H' en aura M.N-(M-1).

Pour que



il suffit que $H(z) = H'(z^M)$
 soit $h(M.p+r) = h'(p) \cdot \delta(r)$

Les coefficients de H' sont obtenus en ne prenant qu'un coefficient tous les M coefficients de H. Si H' a N' coefficients, H doit en avoir M.N'-(M-1).

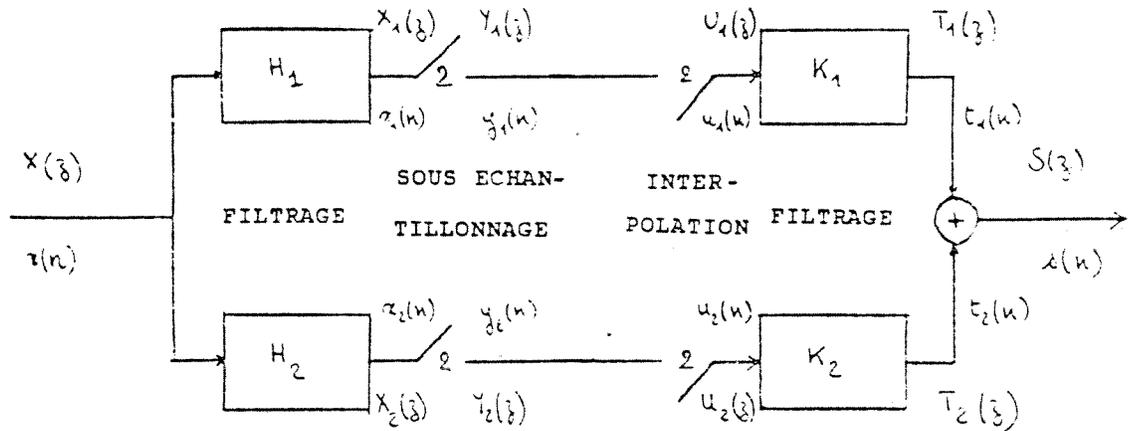
2**Le partage en deux bandes avec des filtres miroirs:

Soient deux filtres H_1 et H_2 miroirs l'un de l'autre:

$$H_2(z) = H_1(-z)$$

et considérons le schéma suivant où K_1 et K_2 seront définis ultérieurement.

SCHEMA DU SYSTEME



$$S(z) = 1/2 \{ H_1(z) \cdot K_1(z) + H_2(z) \cdot K_2(z) \} X(z) + 1/2 \{ H_1(-z) \cdot K_1(z) + H_2(-z) \cdot K_2(z) \} X(-z)$$

Pour éliminer le deuxième terme il suffit de prendre:

$$\begin{aligned} K_1(z) &= H_1(z) & K_1 \text{ est égal à } H_1. \\ K_2(z) &= -H_2(z) & K_2 \text{ est l'opposé de } H_2. \end{aligned}$$

Il reste: $S(z) = 1/2 \{ H_1^2(z) - H_2^2(-z) \} X(z)$

soit $S(\exp(j\omega T)) = 1/2 \{ H_1^2(\exp(j\omega T)) - H_2^2(-\exp(j\omega T)) \} X(\exp(j\omega T))$

en supposant que les filtres sont de réponse impulsionnelle finie d'ordre N:

$$H_i(\exp(j\omega T)) = H_i(\omega) \cdot \exp(-j(N-1)\pi\omega/\omega_s)$$

alors
$$S(\exp(j\omega T)) = 1/2 \{ H_1^2(\omega) - H_1^2(\omega + \omega_s/2) \cdot \exp(j\pi(N-1)) \} \cdot \exp(-j(N-1)\pi\omega/\omega_s) \cdot X(\exp(j\omega T))$$

*****si N est impair la fréquence $\omega = \omega_s/4$ n'est jamais transmise.

*****si N est pair:

$$S(\exp(j\omega T)) = 1/2 \{ H_1^2(\omega) + H_1^2(\omega + \omega_s/2) \} \cdot X(\exp(j\omega T))$$

Le signal est alors reconstruit avec un retard et son amplitude est divisée par deux si l'on a :

$$H_1^2(\omega) + H_1^2(\omega + \omega_s/2) = 1$$

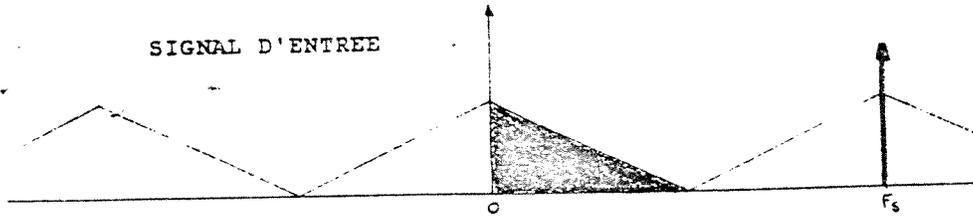
alors $s(n) = 1/2 \cdot x(n-N+1)$

Les diagrammes spectraux suivants illustrent les calculs précédents.

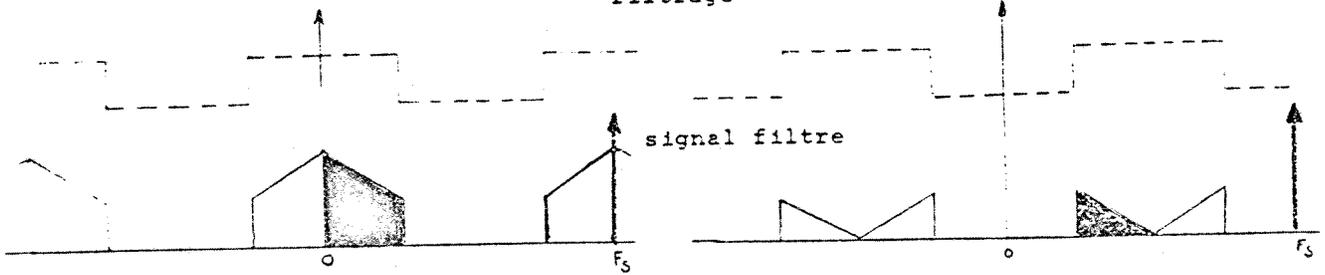
page 3-5: spectrogrammes théoriques du système
page 3-6: spectrogrammes expérimentaux

SPECTROGRAMMES THEORIQUES

SIGNAL D'ENTREE

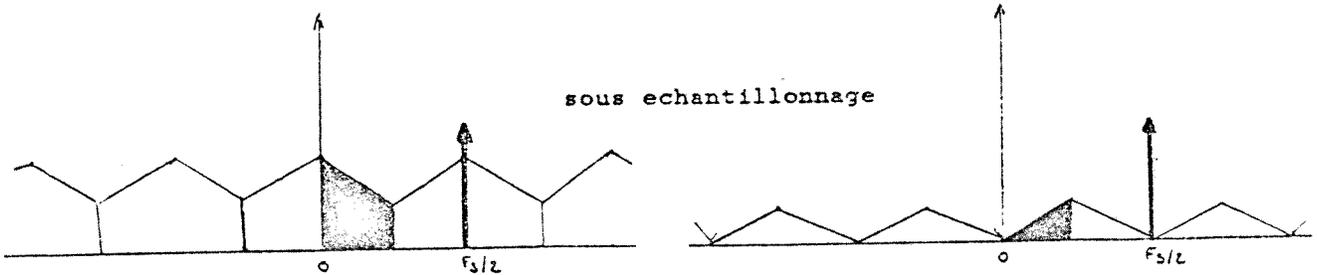


filtrage

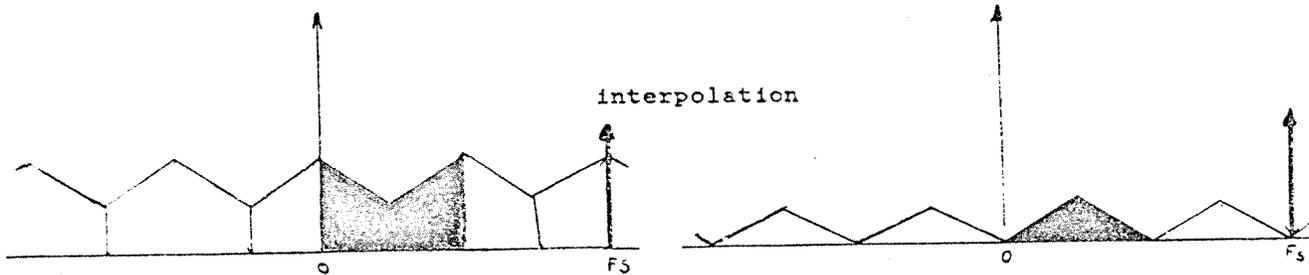


signal filtre

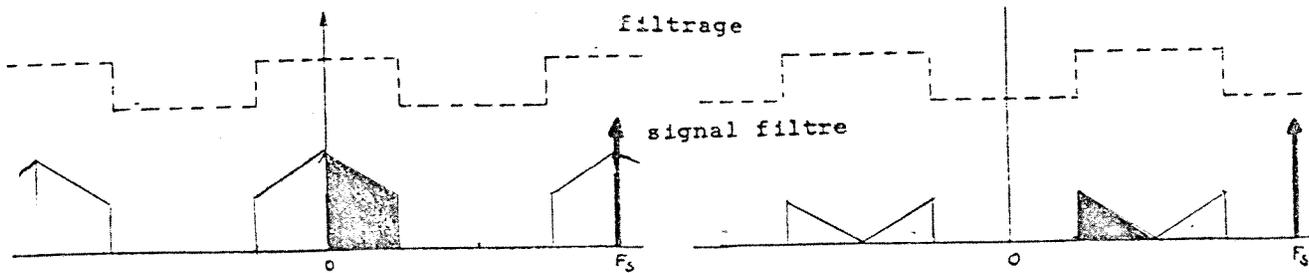
sous echantillonnage



interpolation

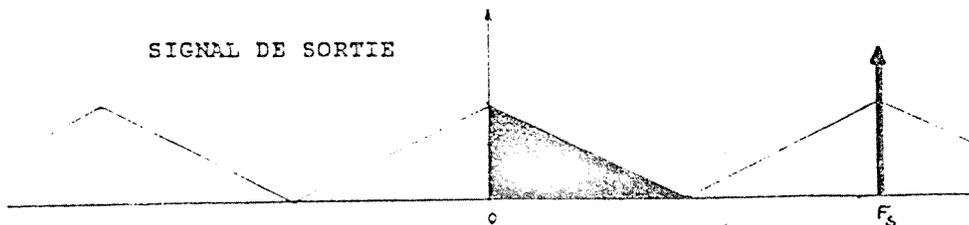


filtrage

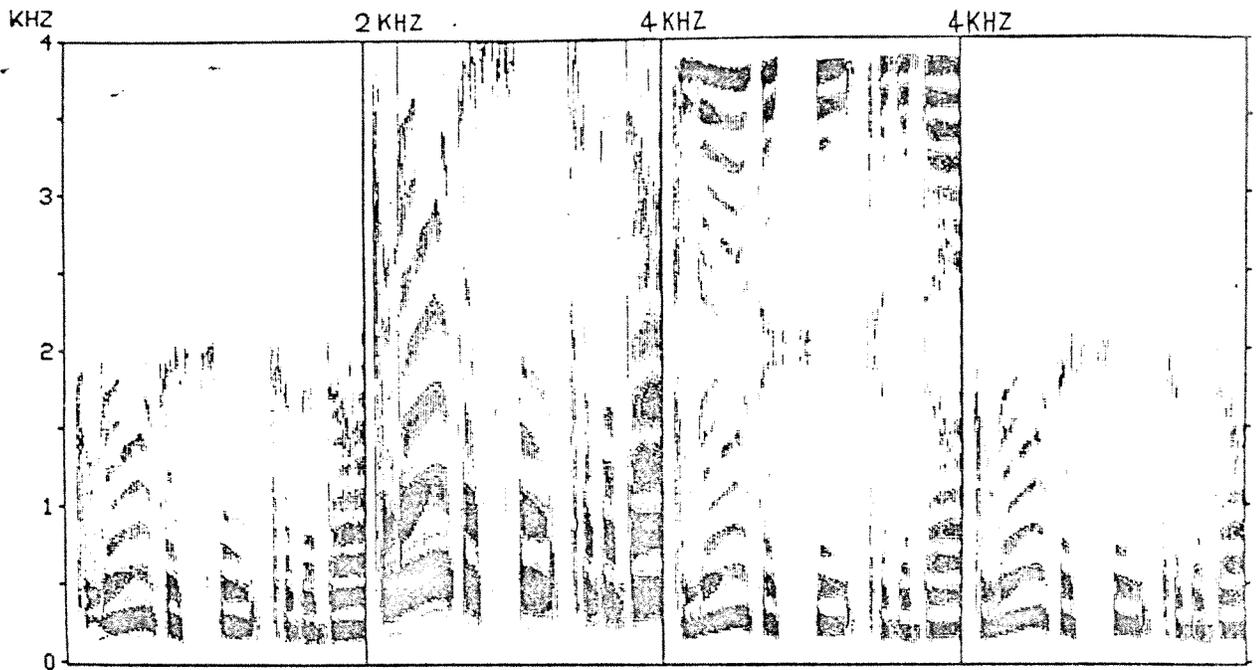


signal filtre

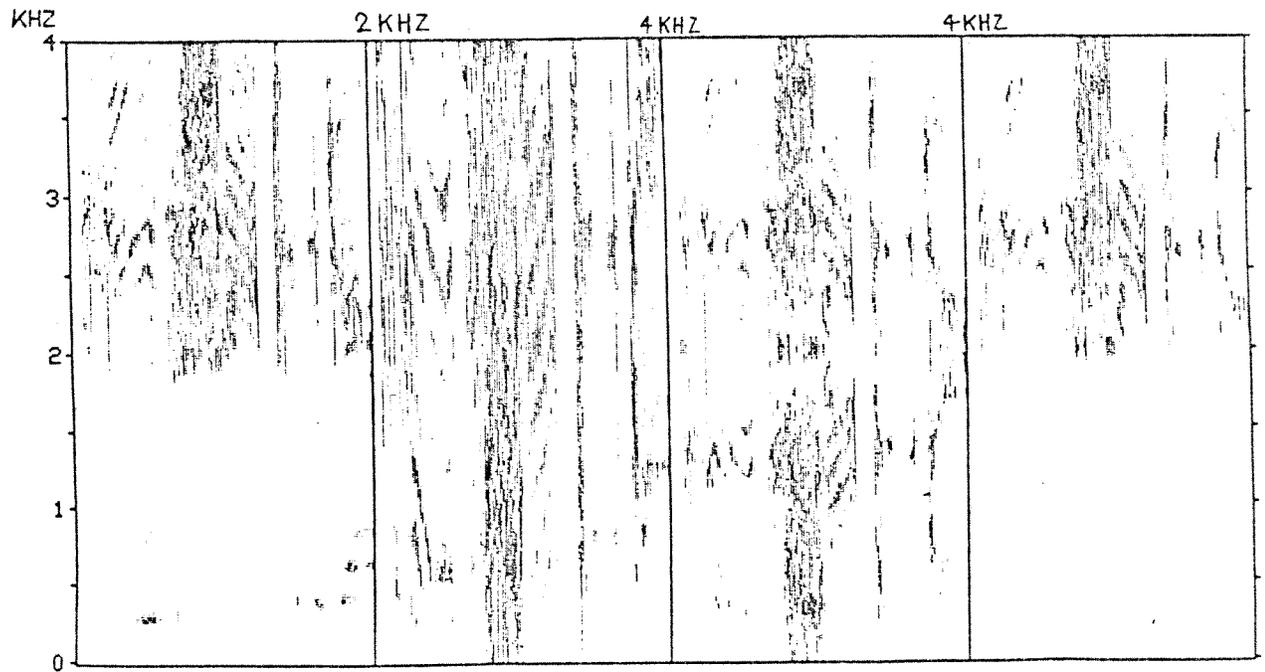
SIGNAL DE SORTIE



SPECTROGRAMMES EXPERIMENTAUX



FILTRE-***-SOUS ECHANTILLONNE-***INTERPOLE-***-FILTRE

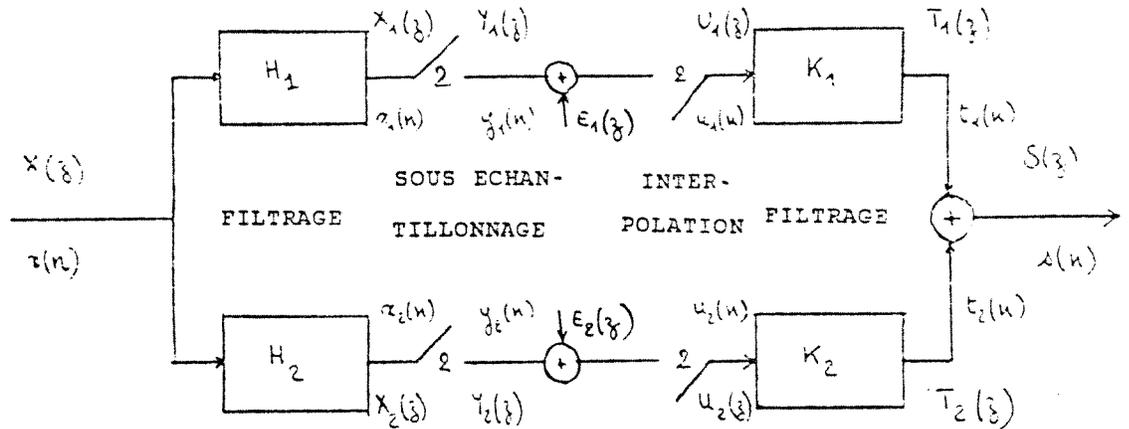


FILTRE-***-SOUS ECHANTILLONNE-***-INTERPOLE-***-FILTRE

3**Comportement en cas de bruit additif:

Reprenons le même schéma mais introduisons deux générateurs indépendants de bruit additif entre le sous échantillonnage et l'interpolation. C'est là en effet que prendront place un codeur, une ligne de transmission, un décodeur.

SCHEMA DU SYSTEME:



$$S(z) = 1/2 \{ H_1(z) \cdot K_1(z) + H_2(z) \cdot K_2(z) \} X(z) + 1/2 \{ H_1(-z) \cdot K_1(z) + H_2(-z) \cdot K_2(z) \} X(-z) + K_1(z) \cdot E_1(z^2) + K_2(z) \cdot E_2(z^2)$$

Une différence entre l'entrée X(z) et la sortie S(z) peut provenir de chacun des 3 termes de cette équation.

Le premier terme:

Par construction $H_2(z) = H_1(-z)$ exactement (construction par simple changement de signe de coefficients) d'où

$$H_1(z) \cdot K_1(z) + H_2(z) \cdot K_2(z) = H_1^2(z) + H_1^2(-z)$$

Ce terme donne un bruit multiplicatif $\xi(w)$ donné par

$$H_1^2(w) + H_1^2(w + w_s/2) = 1 + \xi(w)$$

Cette déformation du signal d'entrée apparait sur la courbe de réponse fréquentielle du système.

Le deuxième terme:

Par construction des filtres ce terme s'annule exactement.

Le troisième terme:

On suppose que les bruits sont blancs stationnaires, de covariance χ_i et de variance σ_i^2 . Après interpolation le signal n'est plus stationnaire et l'on doit utiliser la fonction d'autocorrelation $\phi(m+q, m)$.

après interpolation:

$$\phi_{u_i}(2.p+r+q, 2.p+r) = \sigma_i^2 \cdot \delta(q) \cdot \delta(r) \quad \text{où } m=2.p+r$$

après filtrage:

$$\phi_{T_i}(2.p+r+q, 2.p+r) = \sigma_i^2 \cdot \sum_{b=0}^{N/2-1} h_i(2.b+r+q) \cdot h_i(2.b+r)$$

En se servant de la symétrie des coefficients: $h_i(j) = h_i(N-1-j)$ on montre aisément que:

$$\sum_{b=0}^{N/2-1} h_i(2.b+r+q) \cdot h_i(2.b+r)$$

est indépendant de r ($r=0, 1$) pour les points où q est pair mais pour les q impairs $\phi_{T_i}(2.p+r+q, 2.p+r)$ est dépendant de r et ainsi le signal n'est plus stationnaire.

A la sortie $\phi_s(m+q, m) = \phi_{T_1}(m+q, m) + \phi_{T_2}(m+q, m)$ et $h_2(j) = (-1)^j \cdot h_1(j)$.

$$\phi_s(2.p+r+q, 2.p+r) = \sum_{b=0}^{N/2-1} h_i(2.b+r+q) \cdot h_i(2.b+r) \cdot (\sigma_1^2 + (-1)^j \sigma_2^2)$$

Dans le cas particulier où $\sigma_1 = \sigma_2$ on retrouve un signal stationnaire $\phi_s(2.p+r+q, 2.p+r)$ ne dépendant plus de r (les points où q est impair donnent une valeur nulle).

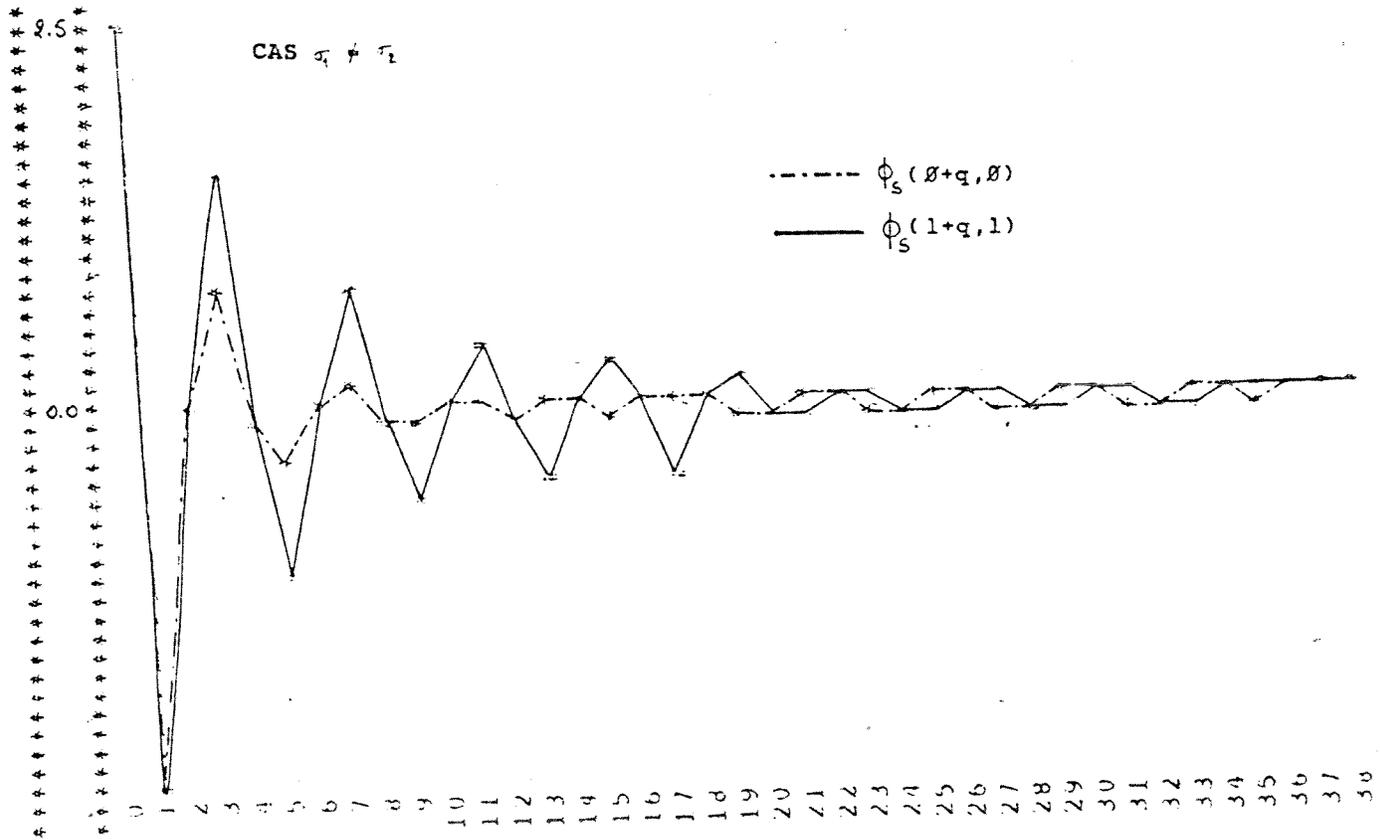
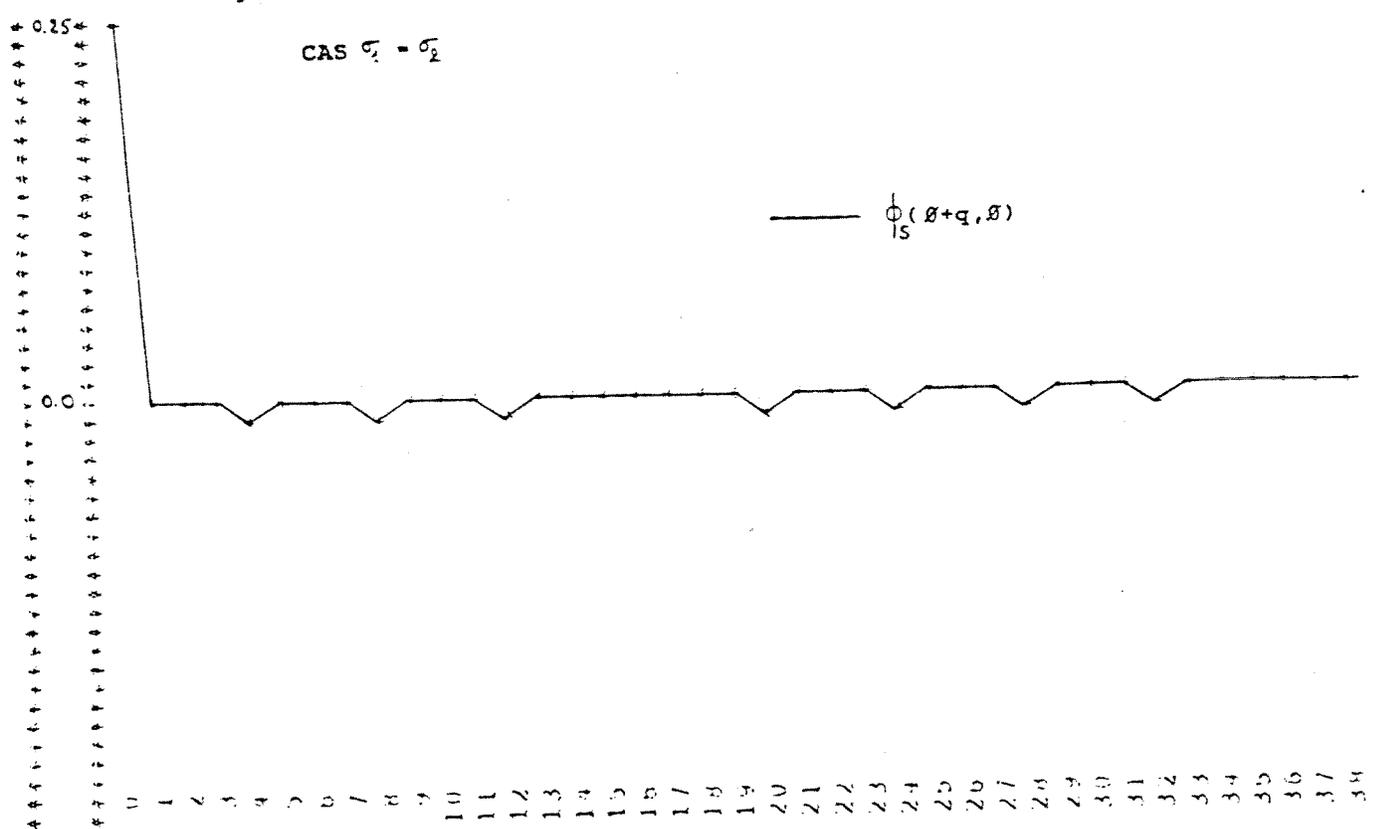
Mais dans le cas où $\sigma_1 \neq \sigma_2$ le signal n'est pas stationnaire. Il est donc difficile de parler du spectre du bruit à la sortie. On peut seulement considérer $\phi_s(m+q, m)$ pour $r=0$ et $r=1$ dans deux cas: (courbes pages suivantes)

dirac: ** $\sigma_1 = \sigma_2$ la fonction $\chi(q)$ s'approche de la fonction de dirac dans ce cas particulier le bruit est blanc.

** $\sigma_1 \neq \sigma_2$ comme prévu $\phi_s(1+q, 1)$ est différent de $\phi_s(0+q, 0)$.

Le travail dans le domaine des z doit être mené avec une extrême prudence.

FONCTIONS D'AUTOCORRELATION



4**Le partage en plusieurs bandes par des filtres miroirs successifs:

Le partage en deux bandes ne permet pas de tirer pleinement profit du système. Une redécomposition permet d'obtenir 4 bandes puis 8 bandes le système est alors représenté par la figure page suivante. On obtient ainsi après le partage, 8 bandes d'une largeur de $f/16$ échantillonnées à $f/8$. Pour permettre une adaptation aisée aux programmes de traitement existants on transforme la structure d'arbre en une structure de branches parallèles représentée par la deuxième figure de la page suivante grâce aux formules d'équivalences rappelées.

H_{ijk} est le produit de trois filtres: $\text{---} \boxed{H_i} \text{---} \boxed{H'_j} \text{---} \boxed{H''_k} \text{---}$

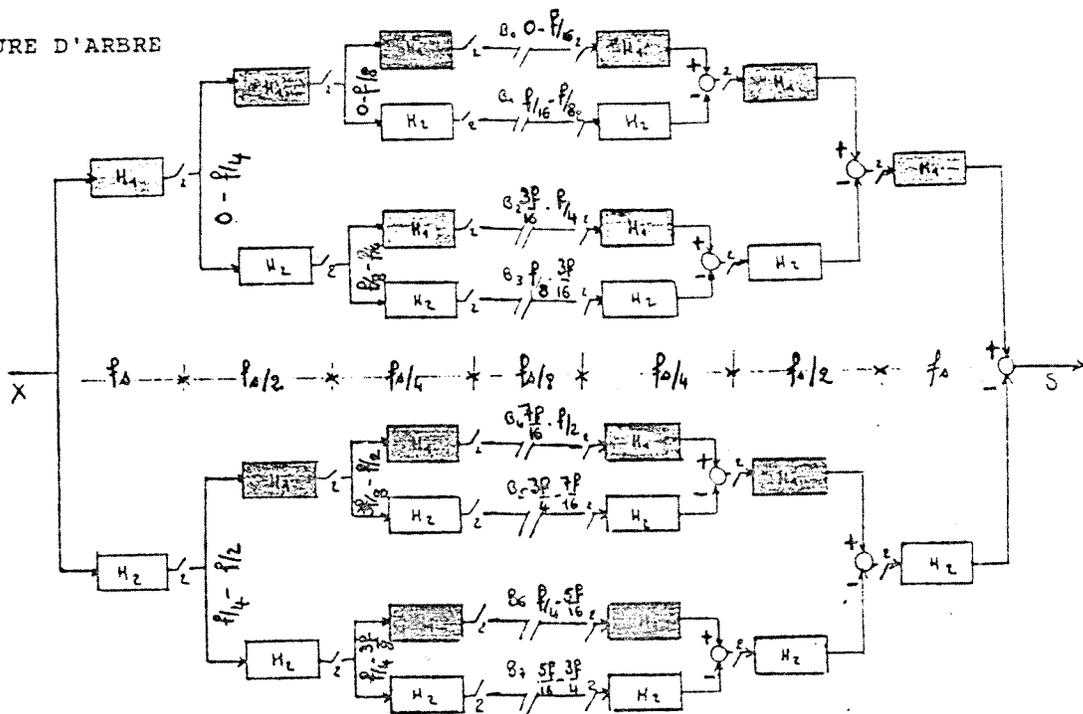
H'_j est obtenu par insertion de 1 zero entre les coefficients de H_j

H''_k est obtenu par insertion de 3 zeros entre les coefficients de H_k

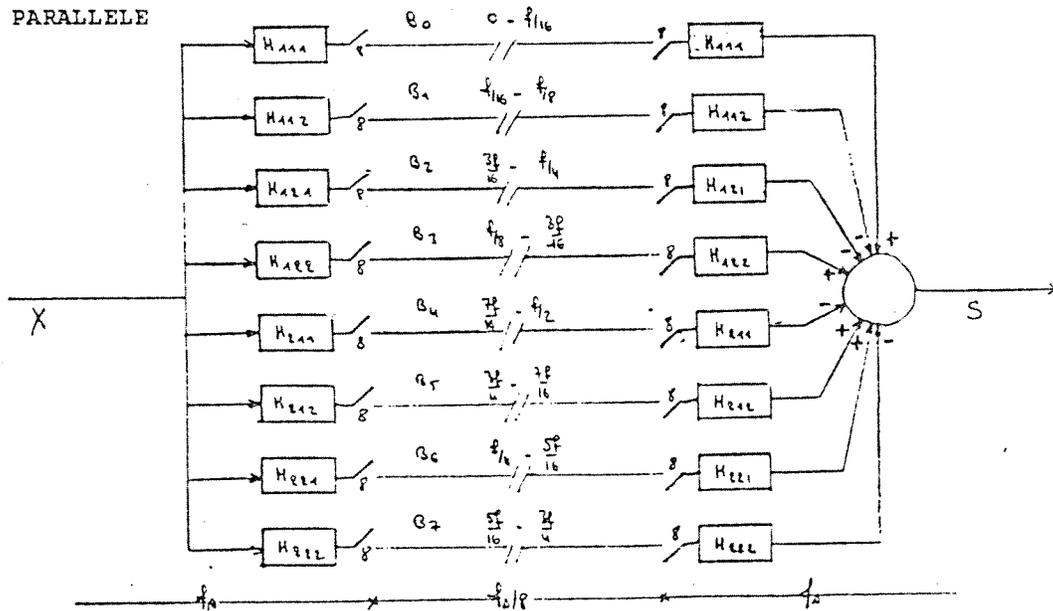
Il est intéressant de noter que par l'effet des sous échantillonnages successifs la bande $(7F/16-F/2)$, la plus aigüe par exemple est transmise a travers les filtres H_3 puis H_1 puis H_1 .

SCHEMAS GENERAUX DU SYSTEME ETUDIE

STRUCTURE D'ARBRE



STRUCTURE PARALLELE



CHAPITRE 4

REALISATION PRATIQUE DES FILTRES

1**Réalisation du filtre de base optimal:

Les conditions à satisfaire:

Lors du sous-échantillonnage à 1 Khz il faut que la largeur de la bande soit réduite à 500 hz. Pour cela il faut prendre un filtre à coupure à 0,25 (en fréquence normalisée) la plus raide possible. Il faut donc trouver un filtre numérique passe-bas à réponse impulsionnelle finie, quadratiquement symétrique, d'affaiblissement entre 0,25 et 0,5 (en fréquence normalisée) le plus grand possible (au minimum 30 DB) et de pente de coupure la plus forte possible.

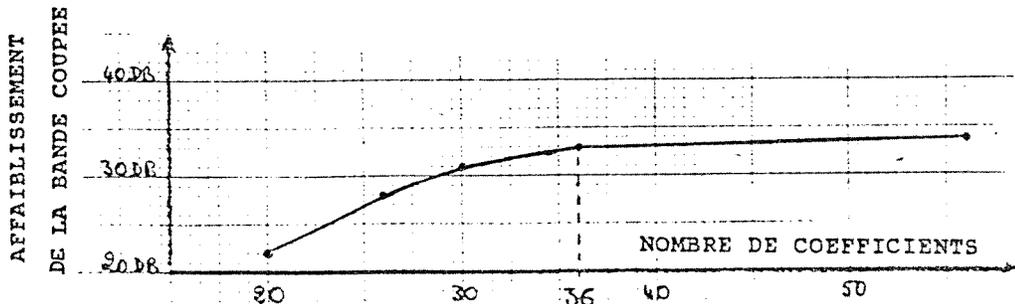
La réalisation:

Le programme existant de calcul des coefficients des filtres numériques permet d'atteindre le gabarit demandé avec une déviation qu'il garantit la plus faible possible étant donné les conditions imposées (gabarit, nombre de coefficients). De nombreux essais n'ont pas permis de trouver de cette manière un filtre suffisamment symétrique donnant un système ayant une courbe de réponse convenable (les ondulations de la courbe de réponse atteignant 6 DB au moins).

Il a fallu procéder à une modification du programme et introduire une deuxième optimisation sur le gabarit d'entrée en lui imposant seulement d'avoir un affaiblissement de 0 DB sur une première bande de 0 à A (A valant de 0.22 à 0.24 en fréquence normalisée) puis ensuite d'être symétrique.

Le nombre de coefficients a été fixé à 36 à la suite de nombreux essais qui ont permis de comparer différents filtres. Le filtre ayant 36 coefficients se trouve être optimal en considérant l'affaiblissement dans la bande coupée.

AFFAIBLISSEMENT EN FONCTION DU NOMBRE DE COEFFICIENT DU FILTRE

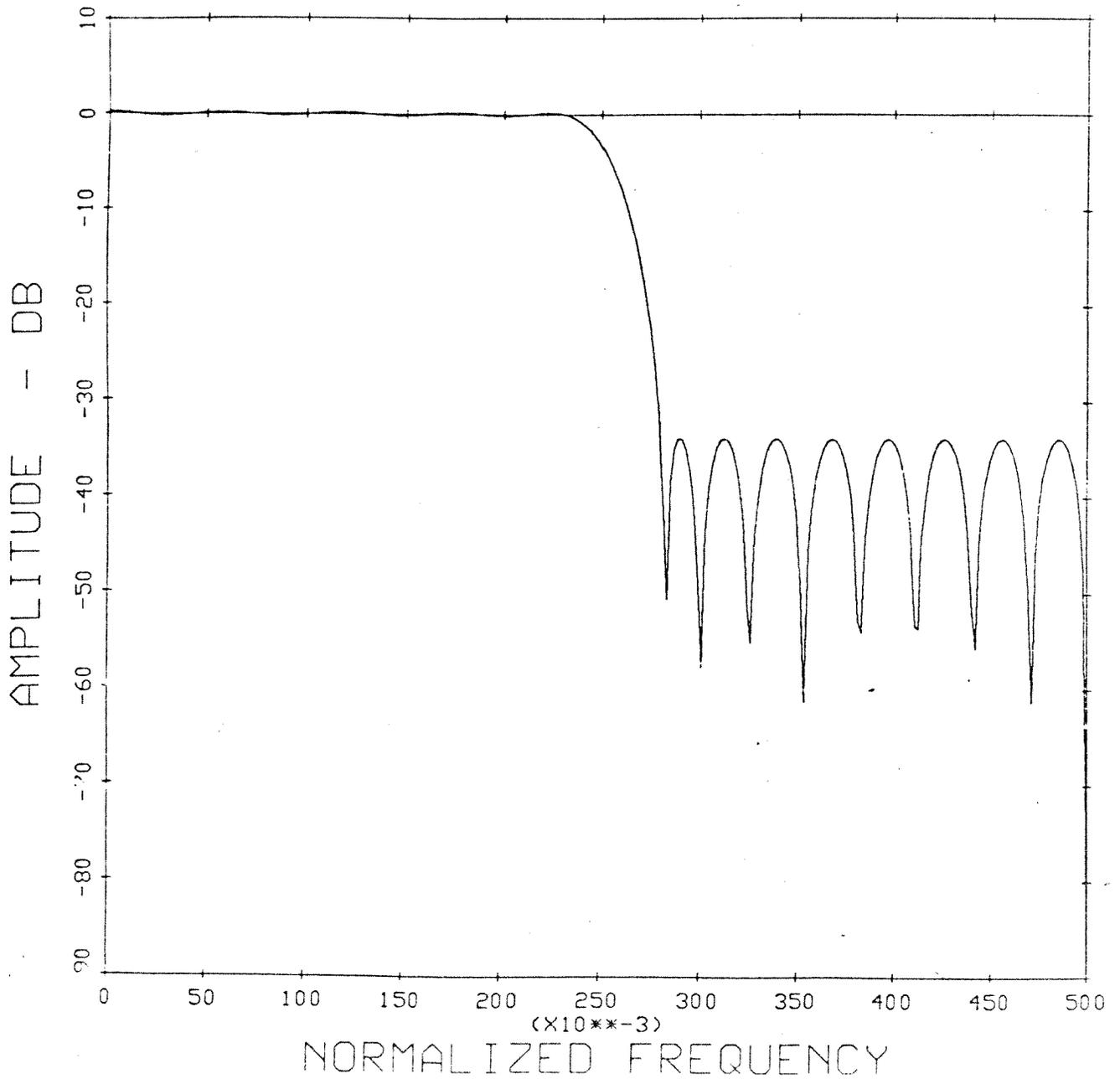


L'affaiblissement dans la bande coupée ainsi que la raideur de la pente de coupure du filtre peuvent être mis en évidence rapidement par l'observation de la fonction de transfert du produit $H(z).H(-z)$. Cette fonction passe par le point -6 DB pour $F=0.25$ et descend de part et d'autre au niveau de l'affaiblissement du filtre. Plus la coupure est raide plus la largeur du pic de cette fonction de transfert est étroite.

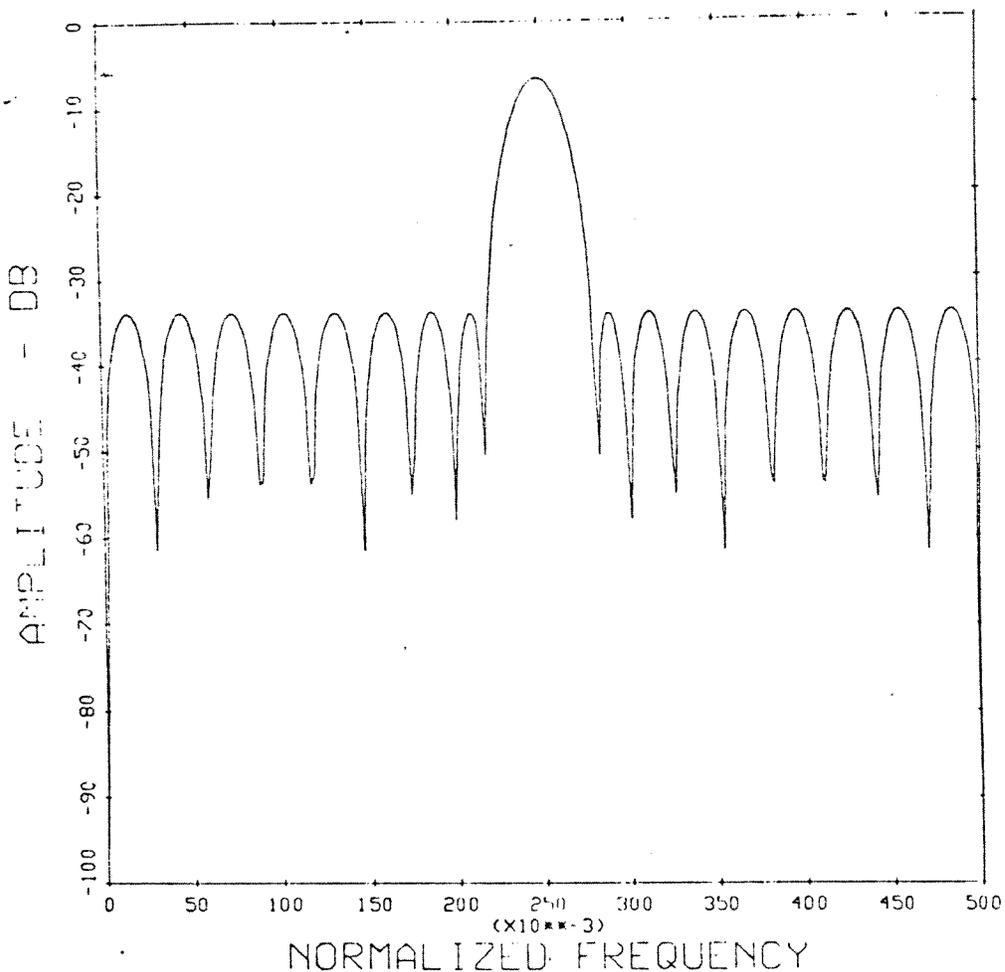
Le filtre retenu:

Le filtre retenu pour les premiers essais est un filtre à 36 coefficients d'affaiblissement 33 DB donnant une courbe de réponse du système global dont les ondulations sont inférieures à 1 DB.

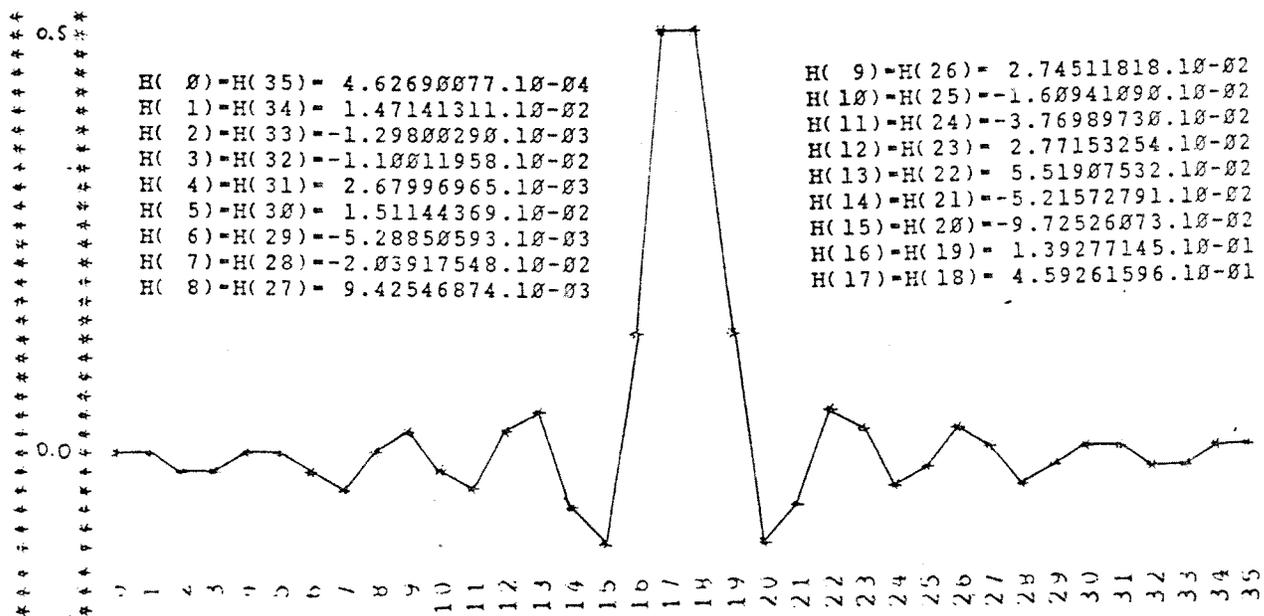
FONCTION DE TRANSFERT DU FILTRE DE BASE



FONCTION DE TRANSFERT DE $H(Z) \cdot H(-Z)$



REPOSE IMPULSIONNELLE DU FILTRE DE BASE

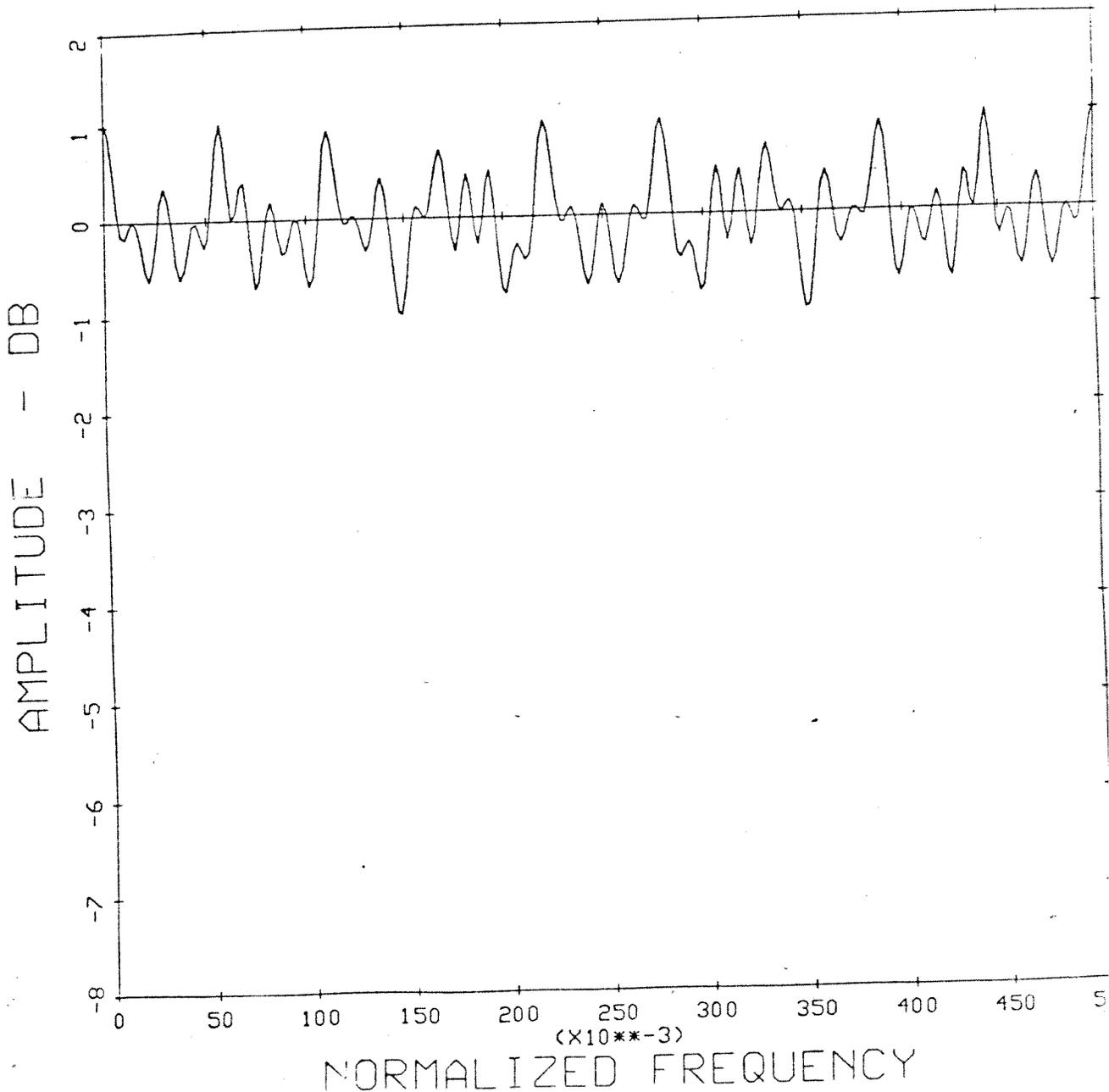


2**Le système à 8 bandes avec ce filtre de base:

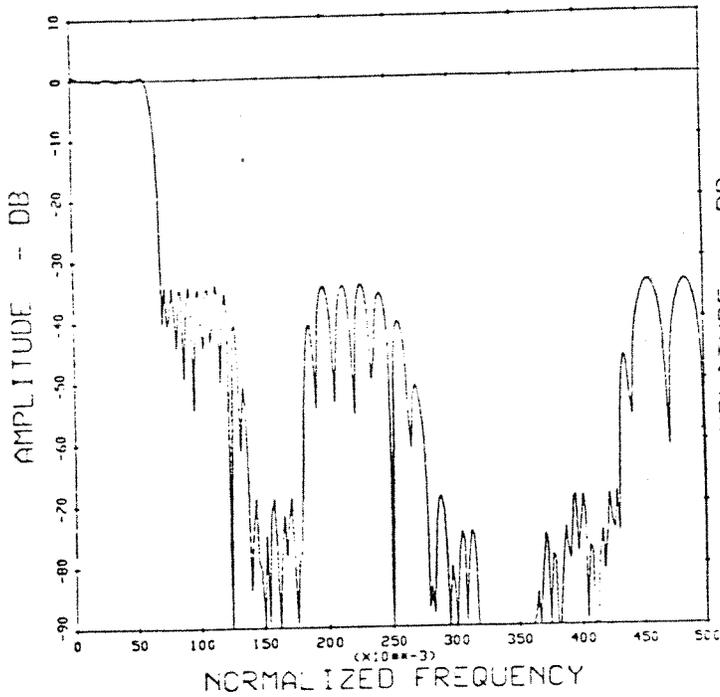
Les filtres de chaque bande dans la structure parallèle ont 246 coefficients chacun. Etant 2 à 2 symétriques seulement les 4 premiers ont été représentés. Le premier filtre est très correct avec une coupure très raide à 0.0625; le quatrième est moins bien, sa coupure à 0.25 étant plus arrondie.

Pour essayer d'obtenir 8 filtres de pentes sensiblement identiques on peut prendre un système à 3 étages dont les filtres de base de chaque étage sont différents: un d'ordre élevé pour le premier, puis deux autres d'ordre décroissant.

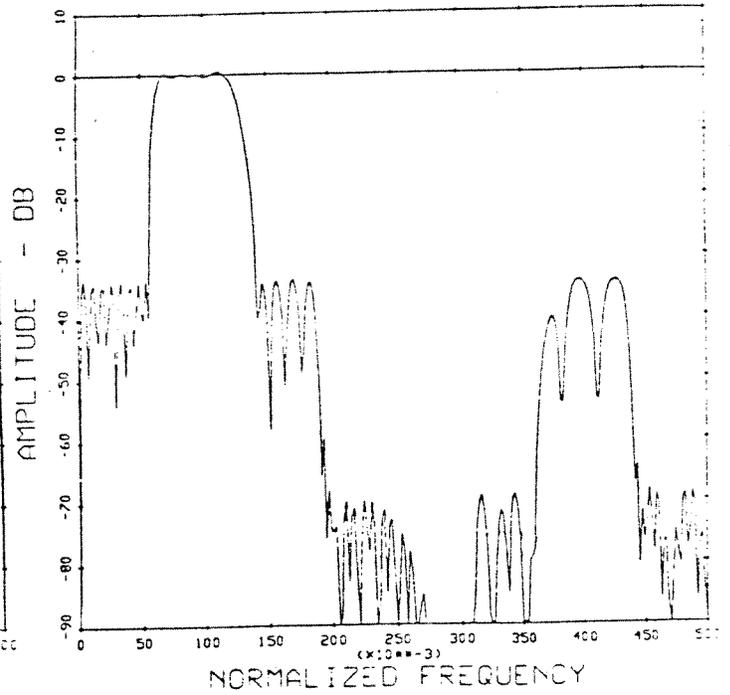
REPOSE DU SYSTEME



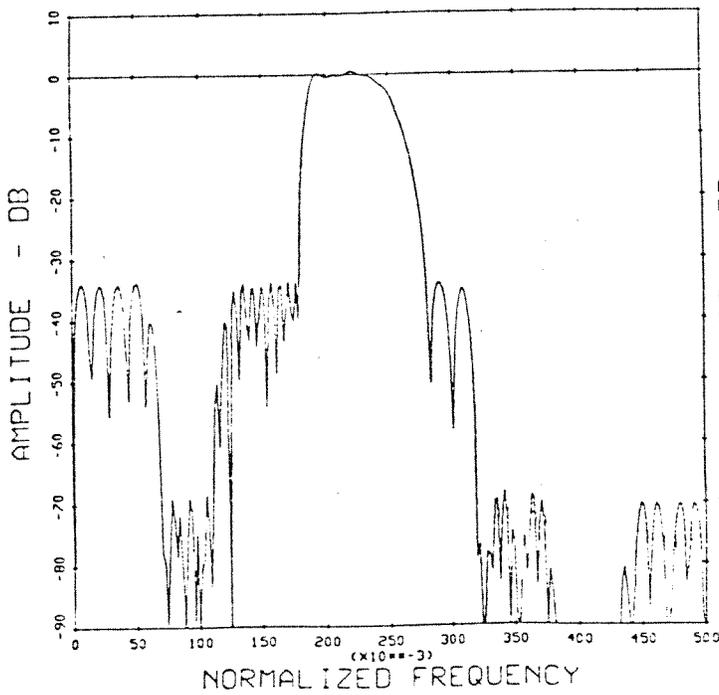
FILTRE H₄₄₄



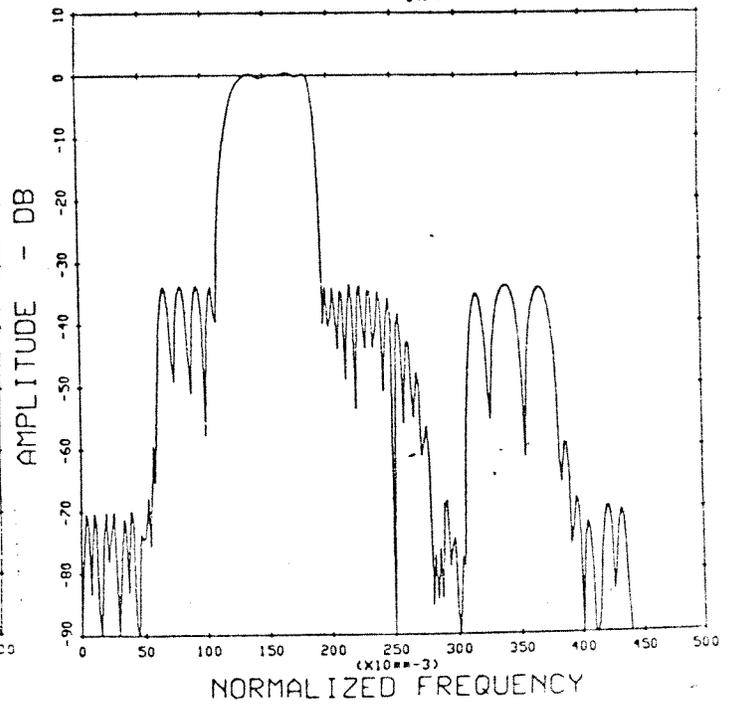
FILTRE H₄₁₂



FILTRE H₄₂₁



FILTRE H₄₂₂



CHAPITRE 5

LA DETERMINATION DU RAPPORT SIGNAL SUR BRUIT SUBJECTIF

Pour permettre une évaluation de la qualité des systèmes de transmission, le laboratoire a développé une méthode de détermination du rapport signal à bruit par des estimations subjectives. Diverses phrases sont traitées par le système à évaluer et par un générateur de bruit multiplicatif dont on peut régler le rapport signal à bruit. Elles sont ensuite présentées de façon aléatoire à des auditeurs qui donnent leurs préférences. Le système reçoit la valeur du rapport signal à bruit qui a donné un partage égal des voix entre les phrases traitées par le système et celles traitées par le générateur de bruit; une analyse statistique donne alors l'intervalle de confiance à 95%.

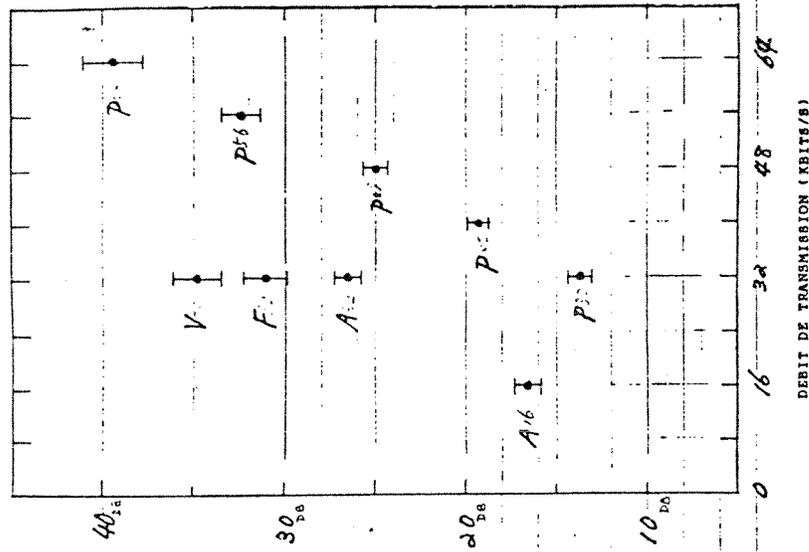
Le tableau suivant peut donner des éléments de comparaison avec des systèmes de transmission connus (voir page suivante).

Les phrases préenregistrées ont été préalablement filtrées entre 200 Hz et 4 KHz; leur fréquence d'échantillonnage est de 8 KHz. La puissance moyenne dans chaque bande pour chaque phrase est représentée sur la figure suivante: en abscisse 8 bandes de fréquence d'une largeur de 500 hz chacune, en ordonnée le rapport entre la variance du signal dans la bande et la variance du signal global. Les enregistrements provenaient de locuteurs différents tantôt masculins tantôt féminins.

Dans l'ensemble des résultats qui suivent, le rapport signal à bruit (RSB) est le rapport variance du signal sur variance du bruit exprimé en DB. Autrement dit la qualité et la valeur du RSB varient dans le même sens.

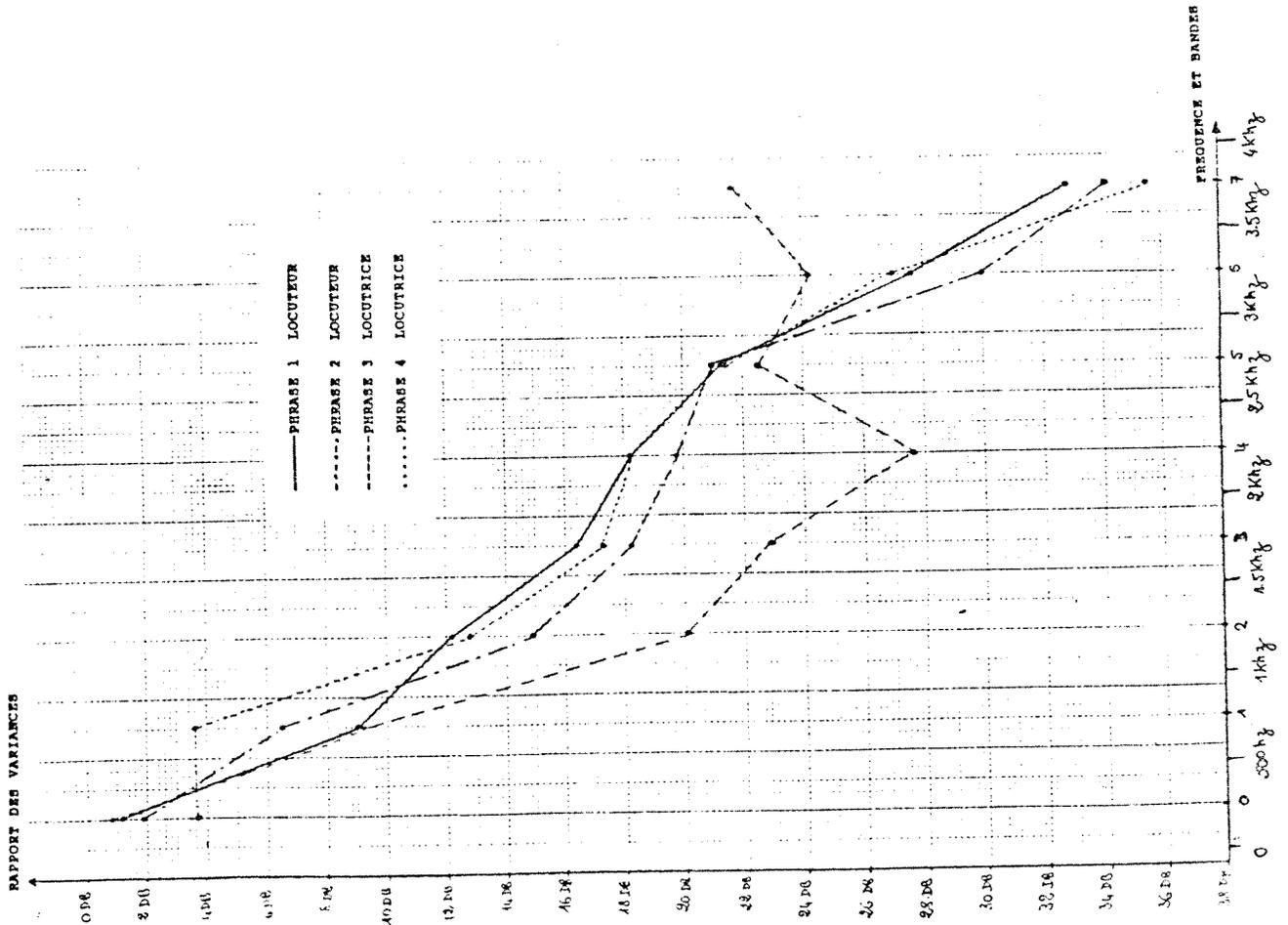
ELEMENTS DE COMPARAISON ENTRE
DES SYSTEMES DE TRANSMISSIONS

P: LOG PCM
A: ADM
V: ADPCM-PREDICTEUR VARIABLE
F: ADPCM-PREDICTEUR FIXE



RAPPORT SIGNAL A BRUIT SUBJECTIF

PUISSANCE MOYENNE DU SIGNAL PAR BANDE

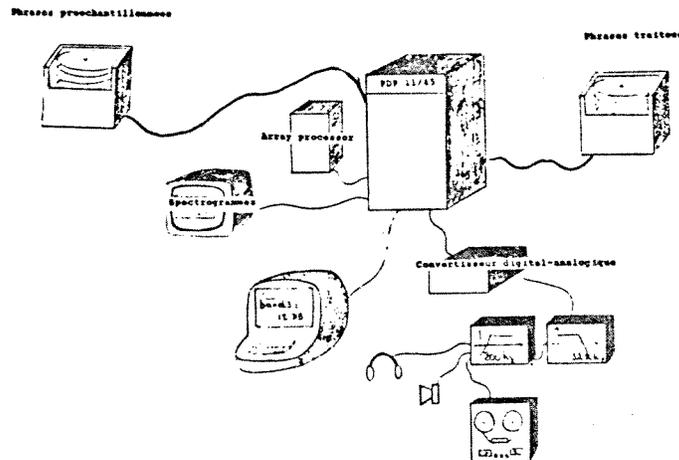


CHAPITRE 6

PRINCIPES DES EXPERIMENTATIONS

Pour tester la valeur réelle du système exposé ci-dessus les phrases prééchantillonnées mémorisées sur disque ont été traitées puis stockées de nouveau sur disque. Le résultat du traitement pouvait être écouté, être enregistré, ou faire l'objet de spectrogrammes.

SCHEMA DE PRINCIPE DES MANIPULATIONS



Les expérimentations ont été faites dans deux optiques différentes:

1-Etudier la dégradation subjective du signal par l'addition de bruit confiné dans une bande étroite du spectre.

2-Essayer de mettre à profit ces résultats pour mettre au point un procédé de codage de la parole à un débit de 16 Kb/s et 9.6 Kb/s.

En prenant le schéma du système à 8 bandes, le signal étant au départ échantillonné à 8 Khz, le signal de chaque bande est alors échantillonné à 1 Khz. Le codage de ce dernier peut se faire par différents codeurs. Pour s'affranchir des contraintes liées à l'utilisation de codeurs dont il

faudrait optimiser les divers paramètres, on s'est contenté dans une première étape de remplacer le codage par un générateur de bruit multiplicatif placé entre le sous-échantillonnage et l'interpolation.

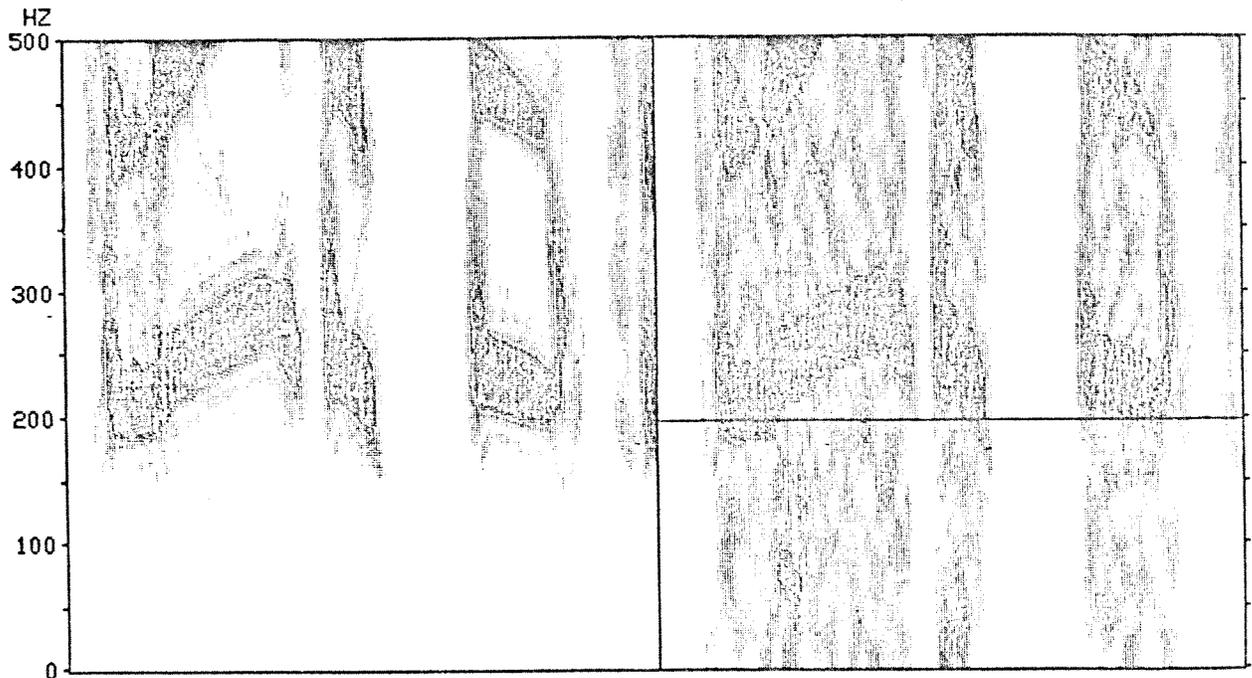
Une correspondance grossière entre le RSB du générateur et le nombre de bit de codage du codeur MIC adaptif disponible correctement réglé pouvait être établie en se reportant à diverses expériences antérieures.

1 1/2 bit	———	3 DB
2 bit	———	6 DB
3 bit	———	12 DB
4 bit	———	18 DB
5 bit	———	24 DB

Ainsi on constate une linéarité entre le nombre de bit et le RSB équivalent. Le débit de transmission du système s'évalue en additionnant le nombre de bit de codage de chaque bande (multiplié par la fréquence d'échantillonnage), parallèlement les systèmes dont les sommes des RSB sont égales sont déclarés avoir le même débit de transmission et peuvent ainsi être comparés.

Après cette première diminution du nombre de facteurs, il ne reste que les 8 valeurs du RSB de chaque bande. Pour ne pas avoir à envisager toutes les combinaisons possibles on a retenu que certaines configurations en se basant sur une hypothèse de 'continuité' des valeurs du RSB optimal entre les bandes.

A l'écoute les phrases étaient filtrées entre 200 Hz et 3400 Hz. Le filtrage à 200 Hz est très important. En effet l'addition de bruit dans la bande 0-500 Hz donne des fréquences basses (<200 Hz) diminuant de façon importante la qualité du signal. Ce qui pénalise injustement le système dans la mesure où le système téléphonique élimine ces fréquences. On trouvera les spectrogrammes comparant la phrase originale et la phrase après traitement par le générateur de bruit et avant le filtrage à 200 Hz.



ORIGINAL-***-APRES TRAITEMENT AVANT COUPURE DES F<200 HZ

CHAPITRE 7

PREMIERES EVALUATIONS

1**Premiere série de configurations:

Pour se rendre compte de la valeur du système les premiers essais ont consisté à attribuer la même valeur de RSB dans les 8 bandes: 6 DB, 12 DB, 18 DB.

1/ 18 DB: La valeur du RSB subjectif du système complet n'a pas pu être évaluée car aucune détérioration du signal n'est audible.

2/ 12 DB: Une détérioration est sensible le RSB subjectif mesuré est de 20 DB (± 2 DB).

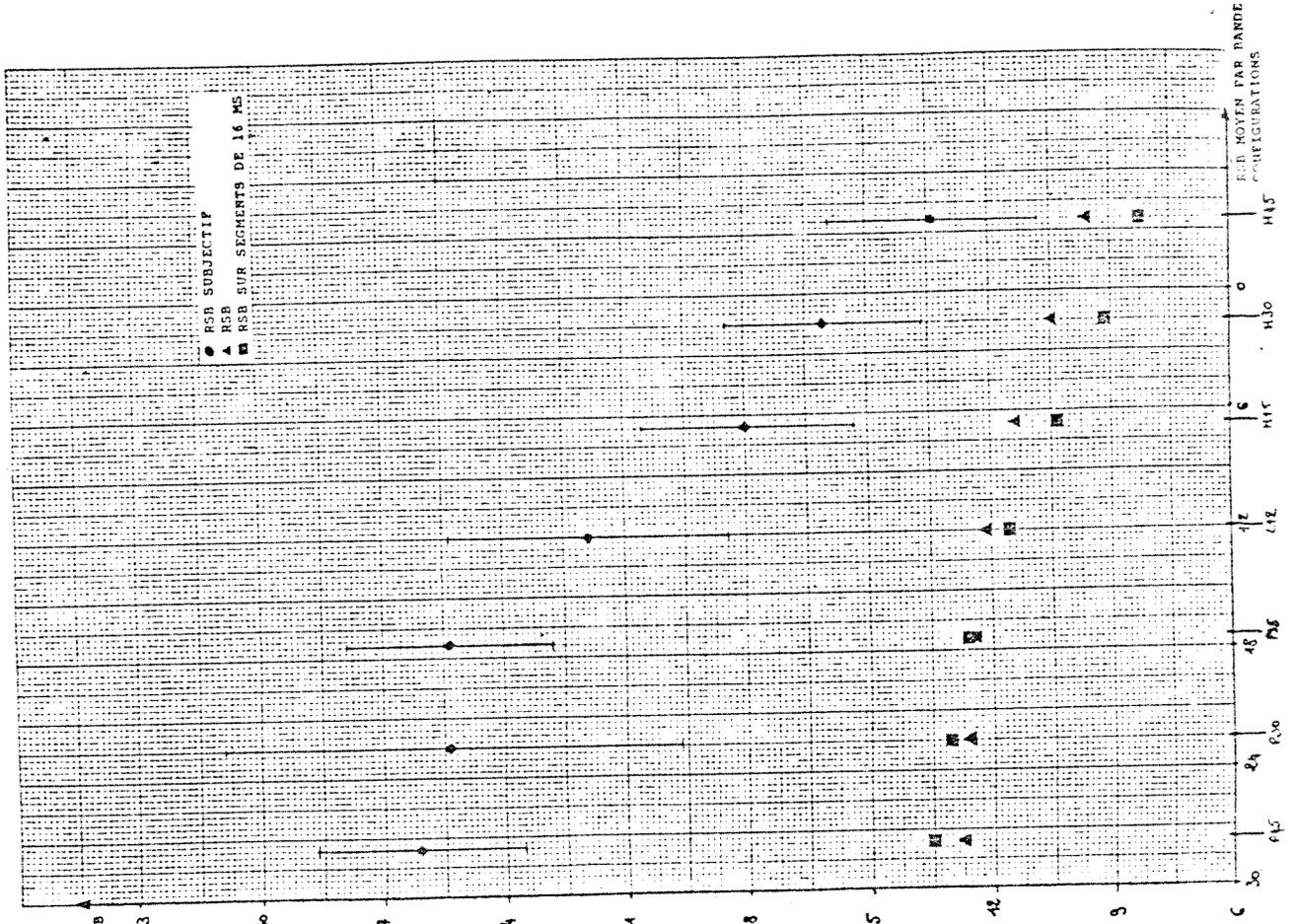
3/ 6 DB: La détérioration est encore plus marquée, le RSB subjectif se situe aux environs de 9 DB.

2**Deuxième série de tests:

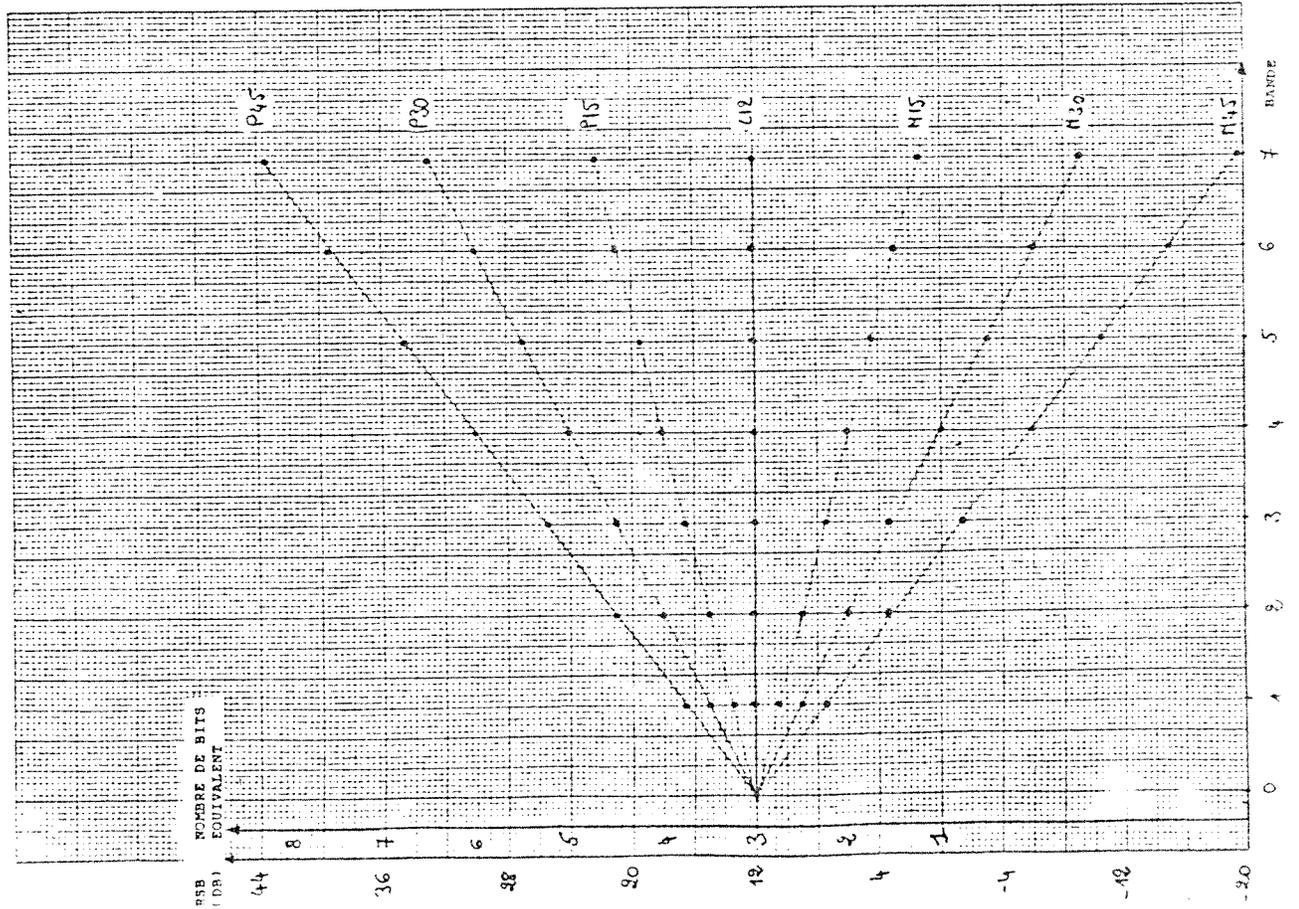
D'après des expériences précédentes (sur un découpage différent et plus réduit) une dégradation dans les fréquences aiguës n'aurait que peu d'influence sur la qualité globale. Ainsi on a essayé une série de configurations où le RSB dans la première bande était de 12 DB et où le RSB des autres bandes variait linéairement: configurations appelées P45, P30, P15, L12, M15, M30, M45. Le RSB de chaque bande dans chaque cas est indiqué sur la figure page 7-2.

Le RSB subjectif ainsi que le RSB (rapport des variances) et le RSB sur des segments de 16 ms ont été mesurés. Les résultats sont indiqués à la page 7-2; en abscisse sont situées les configurations repérées par la somme de leurs RSB divisée par 8 (RSB moyen par bande).

RESULTATS POUR P45, P30, P15, L12, M15, M30, M45



RSB DANS CHAQUE BANDE POUR P45, P30, P15, L12, M15, M30, M45



Les conclusions:

1/les mesures subjectives:

Elles montrent une faible dégradation de la qualité entre P15 et M45 (perte de 12 DB sur le RSB subjectif) pour une moyenne des RSB par bande passant de 17,25 DB à -3,75 DB soit un 'débit de transmission' passant de 31 Kb/s à 3 Kb/s. On s'aperçoit que l'augmentation de ce débit de transmission entre P15 et P45 n'apporte que peu d'amélioration de la qualité.

Mais on ne peut tirer de ces mesures des conclusions très précises en effet les intervalles de confiance sont importants. Ceux-ci ne peuvent être que difficilement réduits par l'augmentation du nombre d'auditeurs: on se rend compte que la qualité d'une phrase est un phénomène qui est difficilement réductible à une seule dimension (l'échelle de référence du générateur de bruit multiplicatif), les dégradations dues au système et celles dues au générateur étant peu comparables. Le système apportant des modifications sur les voyelles longues, le générateur donnant un bruit de fond. Chaque auditeur juge par rapport à sa propre échelle de qualité ce qui donne beaucoup de dispersion dans les réponses. La valeur du RSB subjectif pour la configuration L12 diffère un peu de celui annoncé au paragraphe précédent car les auditeurs étaient différents.

2/Les mesures arithmétiques

Elles mettent en valeur elles aussi une faible variation de la qualité. Mais il faut se garder de tenir compte de telles mesures car seules les mesures subjectives présentent de la valeur pour l'utilisateur et de grandes différences peuvent apparaître entre le RSB subjectif et les RSB arithmétiques.

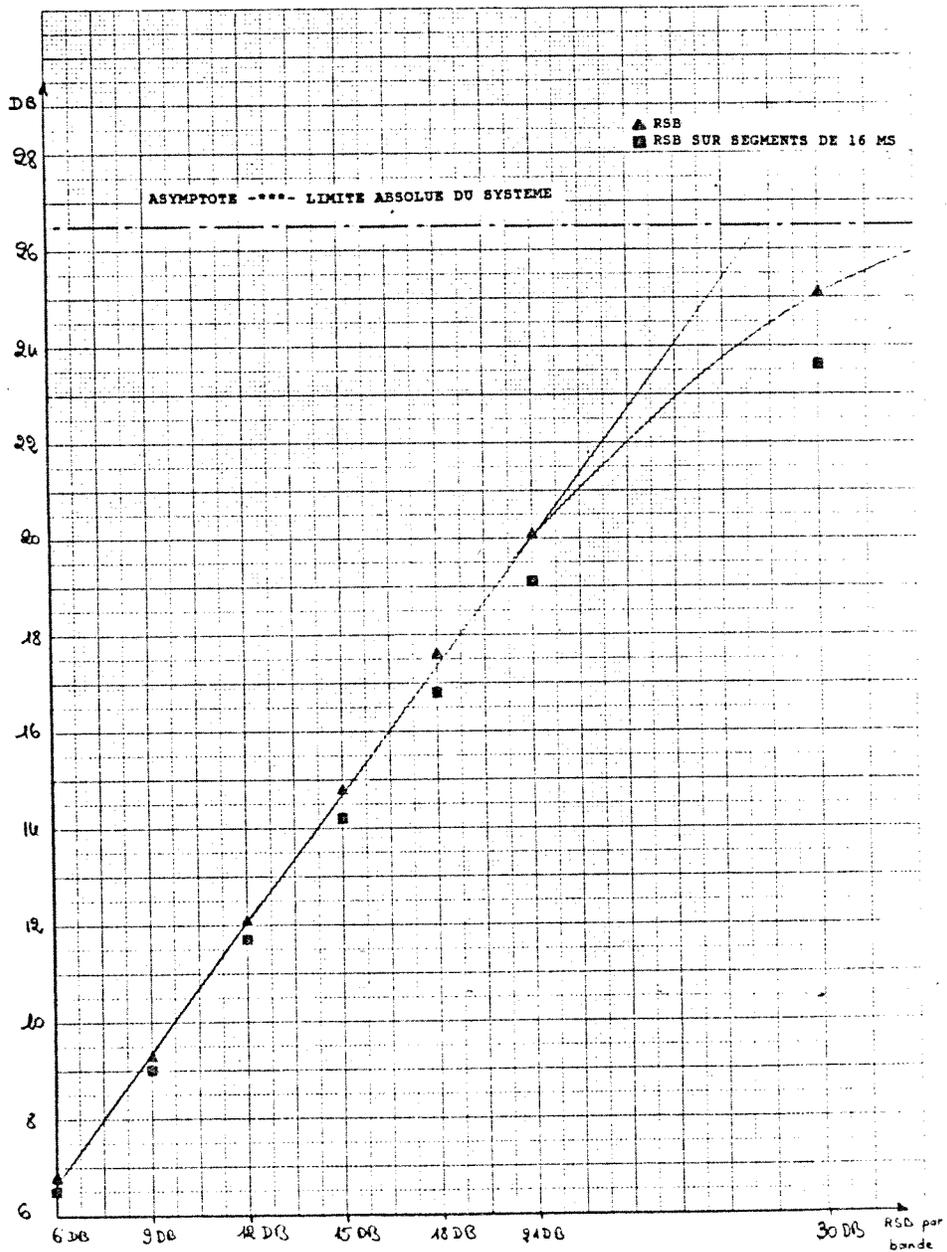
De plus si l'on cherche à attribuer un nombre de bit de codage à chaque bande, il est difficile de trouver un codeur donnant des RSB négatifs. Soit on code la bande, 1 bit pouvant donner un RSB voisin de 0 DB au minimum soit on l'écarte complètement et alors l'effet n'est pas du tout comparable à une addition de bruit même de -10 DB. Le débit de transmission de 3 Kb/s annoncé plus haut (correspondant à -3,75 DB) est irréalisable pratiquement.

3**Série de tests théoriques:

Pour connaître les limites du système avec ce filtre on a mesuré le RSB ainsi que le RSB sur des segments de 16 ms des phrases obtenues après détérioration par du bruit dans chaque bande de même RSB ainsi qu'après le traitement seul (partage et reconstruction sans addition de bruit). Le

graphe ci-dessous montre que le RSB obtenu lors du traitement sans addition de bruit (avec ce filtre 26.5 DB 24.5 DB pour le RSB sur des segments de 16 ms) est la limite supérieure de qualité du système.

RSB MESURE APRES ADDITION DE BRUIT DANS CHAQUE BANDE



CHAPITRE 8

SCHEMA DE CODEUR A 16KB/S

1**Comparaison des configurations à 16 Kb/s:

Pour comparer facilement la qualité des diverses configurations ayant un 'débit de transmission' constant, il faut fixer la somme des RSB des 8 bandes. Cherchant un modèle pour un débit de 16 Kb/s on a fixé la somme des RSB à 48 DB.

La dernière bande (3500 Hz-4000 Hz) a été éliminée permettant une économie de bits de codage sans perte de qualité (cette bande étant coupée par le reste du système de communication). Pour rester proche des possibilités du codeur disponible le plus bas RSB admissible est 3 DB. On a alors envisagé diverses configurations H00, H02, H06, H11, H27. (voir page 8-2).

Du fait de l'abandon de la dernière bande la somme des bits correspondant aux configurations envisagées ne donne pas exactement 16 bits. Le ou le demi-bit restant pourra être attribué à l'une ou l'autre bande. Comme la comparaison des phrases traitées avec les références données par le générateur de bruit multiplicatif pose des problèmes, le procédé a été modifié. L'évaluation subjective de la qualité de ces configurations a pu être faite selon deux procédés:

1/ comparaisons entre elles des différentes configurations.

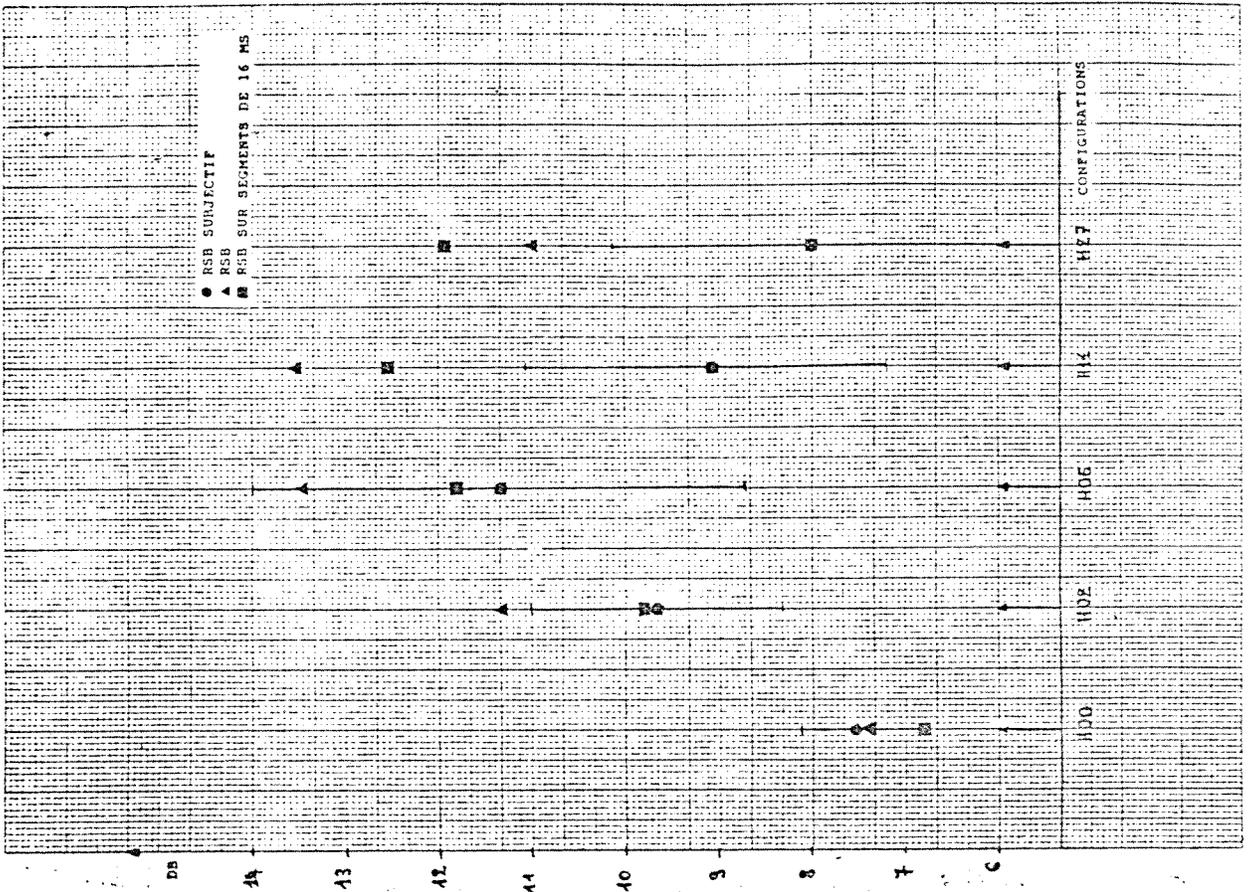
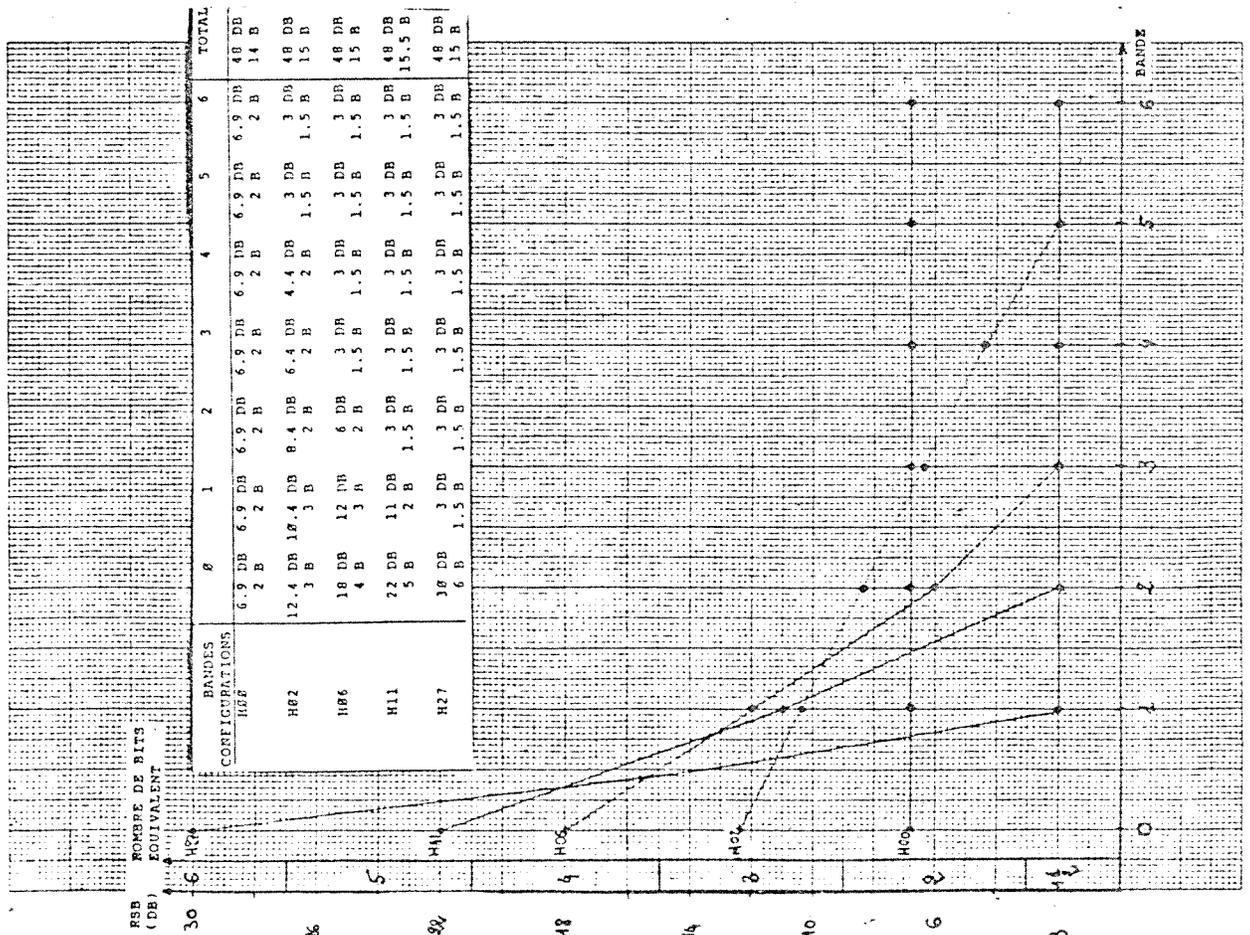
2/ comparaisons avec des phrases traitées par le système à 8 bandes dont le RSB est le même dans chaque bande (le RSB de ces phrases ainsi traitées servait d'échelle de référence).

Les résultats:

1/ Tout d'abord les résultats des comparaisons des configurations entre elles. On lira dans le tableau les pourcentages de préférence de chaque configuration sur une autre configuration. (test portant sur 2 phrases et 3 auditeurs).

COMPARAISON DES CONFIGURATIONS A '16 KB/S'

RSB DANS CHACUNE DES BANDES POUR H00, H02, H06, H11, H27



pourcentages de préférence de a sur b

a	H00	H02	H06	H11	H27
H00		100%	100%	100%	100%
H02	0%		75%	67%	8%
H06	0%	25%		42%	0%
H11	0%	33%	58%		0%
H27	0%	92%	100%	100%	
moyenne	0%	63%	83%	77%	27%

2/ L'échelle de référence est le RSB de chacune des 8 bandes. Sur le même graphique on trouvera le RSB et le RSB sur des segments de 16 ms (l'abscisse n'a aucune signification particulière) (voir page 8-2).

Conclusions:

La configuration H06: 18 DB;12 DB;6 DB;3 DB;3 DB;3 DB;3 DB;3 DB apparait comme celle donnant la meilleure qualité. Les mesures mathématiques ne coïncident pas avec ce résultat. Entre H06 et H11 le RSB augmente car le RSB de la première bande augmente (la diminution du RSB de la troisième bande ne compense pas car la puissance du signal dans cette bande est plus faible que dans la première) alors que la qualité subjective diminue.

Ceci montre qu'un RSB dans la première bande supérieur à 18 DB n'apporte aucune amélioration de la qualité subjective. Le niveau de 18 DB par bande est le seuil de détection auditive du bruit confiné dans des bandes étroites de 500 Hz de largeur.

La configuration H06 ayant été retenue, les configurations voisines (obtenues en ajoutant 6 DB dans les bandes 0,1 et 2 exclusivement) ont fait l'objet de tests de comparaison. Malheureusement il est impossible de distinguer à l'oreille la meilleure des trois:

bande configuration	0	1	2	3	4	5	6
H60	24 DB	12 DB	6 DB	3 DB	3 DB	3 DB	3 DB
H61	18 DB	18 DB	6 DB	3 DB	3 DB	3 DB	3 DB
H62	18 DB	12 DB	12 DB	3 DB	3 DB	3 DB	3 DB

2**Essais avec le codeur:

Le codeur MIC adaptif est une application des principes exposés dans l'article de R.E Crochiere (On the design of sub band coders for low bit rate speech communication; The Bell System Technical Journal vol 56 no 5 may june 1977). Le rapport entre le plus grand et le plus petit pas de quantification a été fixé à 100 pour toutes les bandes. La valeur maximale de ce pas de quantification dépend de la bande considérée:

bande 0:0,6 bande 1:0,4 bande 2:0,2 bande 3:0,07
bande 4:0,05 bande 5:0,05 bande 6:0,05 bande 7:0,05

La distribution des bits de codage par bande est celle qui s'approche au maximum de la configuration de RSB donnant le meilleur résultat. On trouvera le RSB obtenu dans chaque bande (moyenne sur 15 phrases).

0- 500 Hz	-----	4 bits	-----	19.0 DB
500-1000 Hz	-----	4 bits	-----	17.0 DB
1000-1500 Hz	-----	2 bits	-----	6.7 DB
1500-2000 Hz	-----	1.5 bits	-----	4.6 DB
2000-2500 Hz	-----	1.5 bits	-----	4.7 DB
2500-3000 Hz	-----	1.5 bits	-----	4.3 DB
3000-3500 Hz	-----	1.5 bits	-----	4.1 DB

Le RSB subjectif a été évalué à 14.7 DB (intervalle de confiance a 95 % :16.4 DB a 12.9 DB); le test portait sur 4 phrases présentées à 3 auditeurs. Le RSB (moyenne sur 9 phrases) est 14.1 DB. Le RSB sur des segments de 16 ms (moyenne sur 9 phrases) est de 11.6 DB.

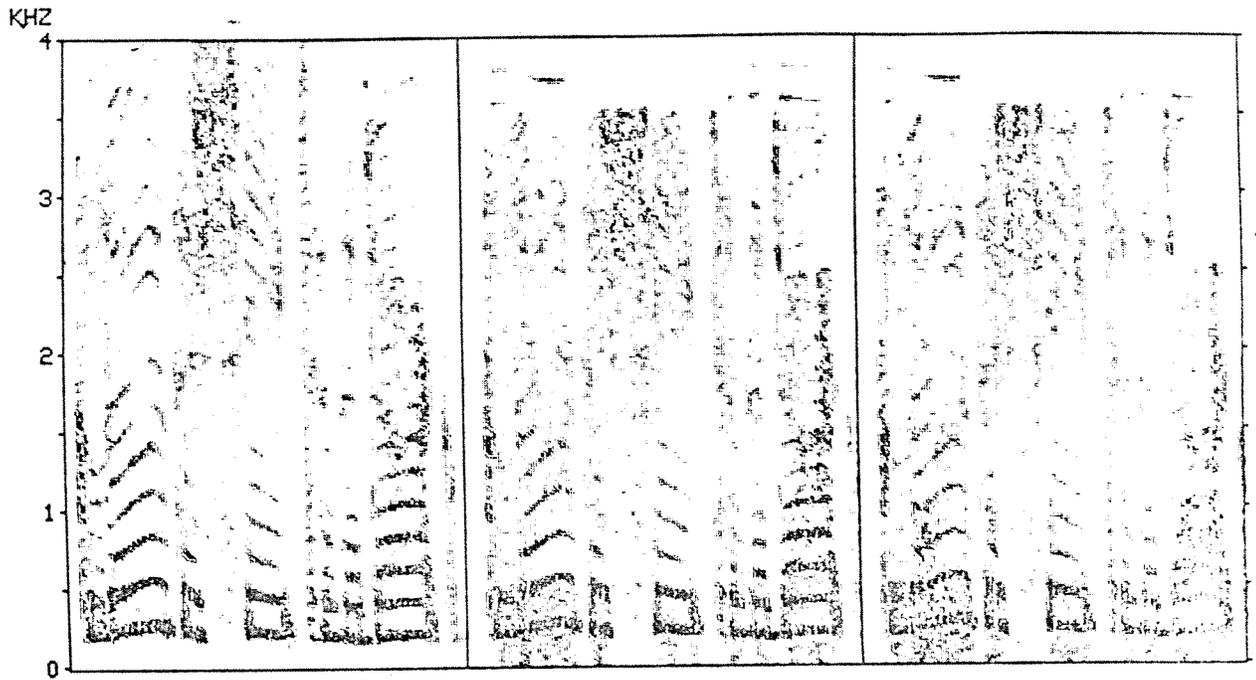
Le résultat est comparable aux résultats provenant de l' addition de bruit correspondante (voir les spectrogrammes comparant la phrase originale, la phrase codée et détériorée par du bruit).

En examinant les spectrogrammes dans chaque bande à toutes les étapes du processus, on s'est rendu compte que les filtres de bandes laissaient passer des signaux hors des limites de chaque bande (voir les spectrogrammes des bandes 0 a 3 après le dernier filtrage avant l'addition). En reprenant la fonction de transfert page 4-5 de H₁₁₁, on remarque que l'affaiblissement entre 1.5 Khz et 2 Khz n'est que de 32 DB environ. Comme l'énergie dans cette bande (ainsi d'ailleurs que dans la seconde 500-1000 Hz) est importante l'atténuation de 32 DB n'est pas suffisante et l'énergie résiduelle est d'une puissance comparable a l'énergie de la bande 3 et ainsi perturbe fortement le signal de cette bande.

Pour résoudre ce problème il faudrait refiltrer chaque bande (ou au moins les deux premières) avant de les additionner.

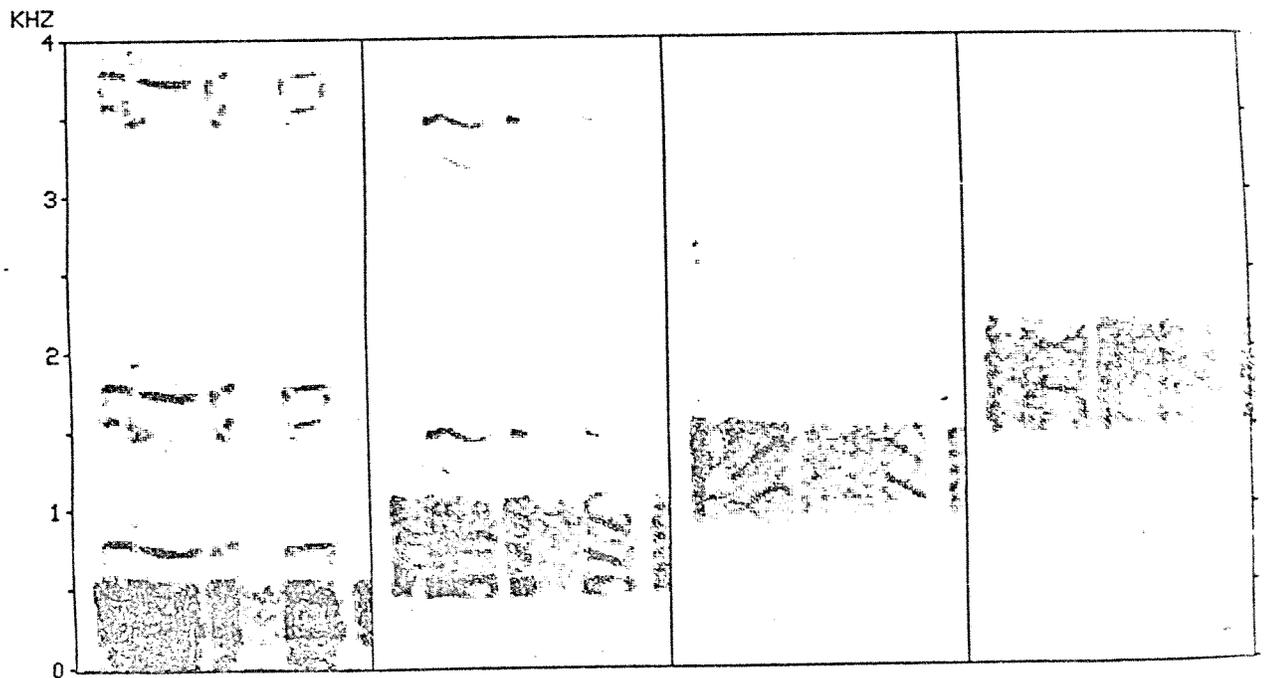
Pour des raison de facilité les essais ont été repris en refiltrant chaque bande avec les mêmes filtres donnant ainsi une première amélioration surtout sur les voyelles longues (comme 'glou').

COMPARAISON ENTRE LE CODAGE ET L'ADDITION DE BRUIT



ORIGINAL-***-CODE 16 KB/S -***-ADDITION DE BRUIT CORRESPONDANTE

SPECTROGRAMMES DES BANDES 0 A 4



BANDES 0-***-1-***-2-***-3 BLANC:-60 DB NOIR:-10 DB

3**Deuxième configuration:

La bande 0-500 Hz est inutilement large. Seules les fréquences au dessus de 300 Hz sont transmises par le système téléphonique. De plus le codeur MIC adaptif perd de son efficacité en présence des fréquences aiguës (c'est le cas pour la bande 300-500 Hz échantillonnée a 1 Khz).

Cette bande a été redécoupée en 2 sous-bandes selon le même principe de base. Seule la bande supérieure a été conservée. L'échantillonnage étant réduit à 500 Hz, on a pu gagner quelques bits de codage et les reporter sur les autres bandes. La configuration essayée est la suivante:

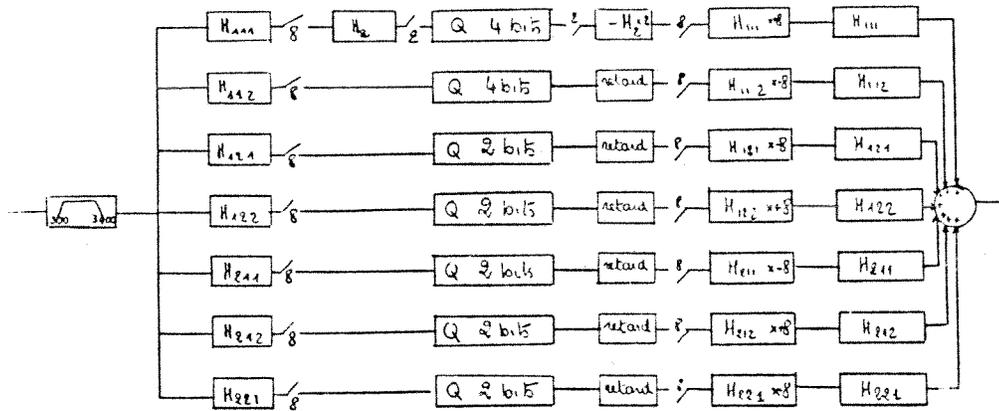
250- 500 Hz	-----	4 bits	-----	18.1 DB
500-1000 Hz	-----	4 bits	-----	17.7 DB
1000-1500 Hz	-----	2 bits	-----	6.8 DB
1500-2000 Hz	-----	2 bits	-----	6.9 DB
2000-2500 Hz	-----	2 bits	-----	7.0 DB
2500-3000 Hz	-----	2 bits	-----	7.4 DB
3000-3500 Hz	-----	2 bits	-----	7.2 DB

La qualité obtenue est supérieure à celle obtenue dans le cas précédent bien que l'on observe un changement de tonalité de la voix dû à la perte du premier formant entre 200 Hz et 250 Hz, changement qu'il faut compenser par un filtrage analogique adéquat pour pouvoir comparer l'original et la phrase traitée. Cependant un problème est soulevé: la bande 0-250 Hz étant abandonnée, lors de l'addition des bandes il ne peut plus y avoir de compensation entre les deux bandes 0-250 Hz et 250-500 Hz (l'annulation du terme $X(-z) \{ H_1(-z) \cdot H_1(z) - H_2(-z) \cdot H_2(z) \}$). On se rend compte de ce problème sur le spectrogramme suivant page 8-7. La zone A provient de la zone A' qui n'a pas pu être compensée par l'effet des filtres miroirs. Pour limiter cet inconvénient il suffit de filtrer au préalable le signal entre 300 Hz et 3400 Hz au lieu de traiter comme c'est le cas des phrases filtrées à 200 Hz seulement.

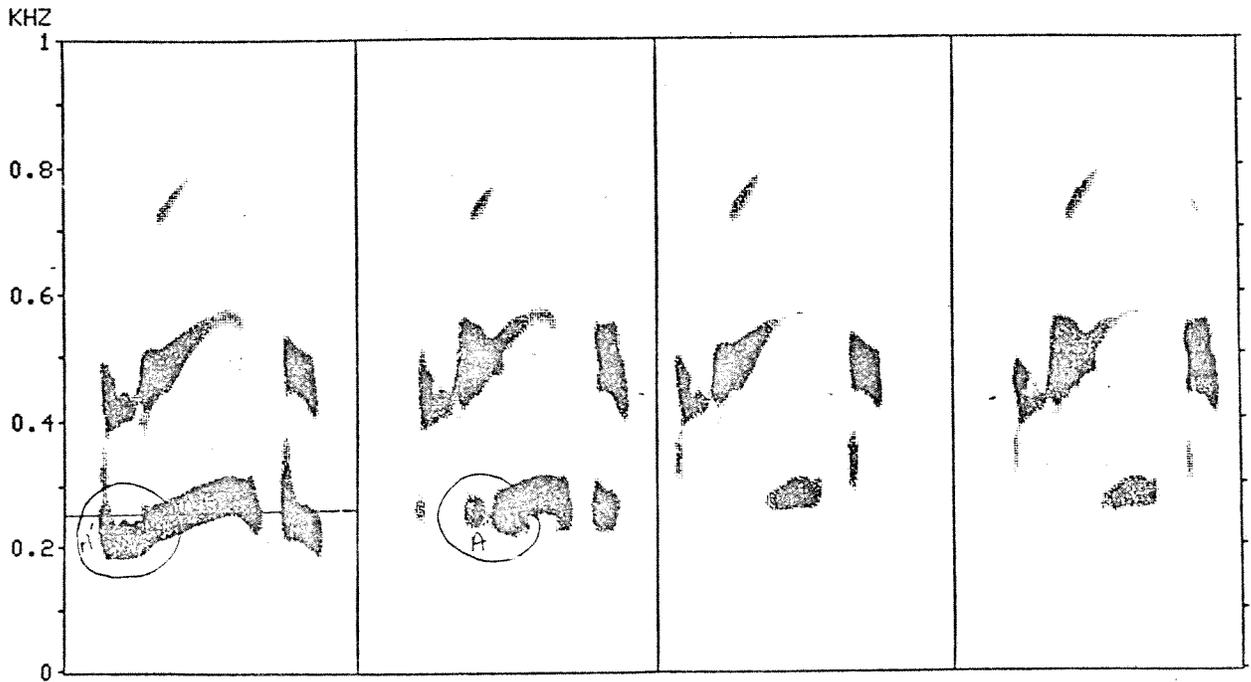
Le RSB subjectif a été mesuré : 19.2 DB (intervalle de confiance 17,7 DB à 20,7 DB),

La transmission des voyelle est nettement améliorée; le bruit est un bruit plus métallique.

NOUVEAU SCHEMA DU SYSTEME DE CODAGE



PROBLEME DE COMPENSATION ENTRE LES DEUX PREMIERES SOUS-BANDES



ORIGINAL - CODE -***- ORIGINAL FILTRE - CODE

4**Application à 9,6 Kb/s:

De très rapides expériences ont été menées pour simuler un débit de 9,6 Kb/s. Si l'on veut rester dans les limites imposées par le codeur MIC adaptif (1 1/2 bits au minimum) il est impossible de conserver le schéma précédent à 7 bandes avec refiltrage et redécomposition de la première bande. D'après l'article de Carlson et Granstrom (Vowel perception: the relative perceptual salience of selected acoustic manipulations), il est possible d'abandonner une partie du spectre (9000 Hz au maximum) sans affecter grandement la qualité du signal.

Deux configurations ont été retenues:

	configuration 1	configuration 2
250- 500 Hz -----	4 (-->2)	4 (-->2)
500-1000 Hz -----	3	3
1000-1500 Hz -----	1 1/2	1 1/2
1500-2000 Hz -----	1 1/2	1 1/2
2000-2500 Hz -----	1 1/2	Ø
2500-3000 Hz -----	Ø	1 1/2
3000-3500 Hz -----	Ø	Ø
3500-4000 Hz -----	Ø	Ø

La deuxième configuration donnerait des résultats meilleurs que la première. Les phrases sont déformées mais intelligibles. Mais les expériences doivent être poursuivies pour avoir des résultats plus fiables.

CHAPITRE 9

CONCLUSION

Cette première approche d'un codeur à 16 Kb/s par codage des sous-bandes a permis d'obtenir un résultat très convenable mais non parfait. Pour améliorer le codeur il faudrait essayer différents filtres de base et pousser l'optimisation du codeur en tenant compte de la puissance et de la dynamique de chaque bande.

Les problèmes de réalisation pratique n'ont pas été soulevés, les expérimentations étant faites par simulations sur ordinateur. Les éléments fondamentaux d'une telle réalisation seraient les filtres numériques que l'on met en cascade. Des difficultés peuvent naître de l'imprécision des calculs dans une telle réalisation pratique.

En outre les essais ont permis de situer le seuil de détection d'un bruit dans les bandes les plus graves; il se situe entre 15 DB et 18 DB. Ainsi une réalisation d'un codeur à 24 Kb/s peut donner un résultat quasi parfait.

Je tiens à remercier le Docteur P.Kabal et le Docteur P.Mermelstein qui ont aidé par leurs conseils et leurs travaux antérieurs à la réalisation de ce travail et à la rédaction de ce rapport.