

200

ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT **d'electronique**

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

PROJET DE FIN D'ETUDES

S U J E T

**ETUDE
D'UN VOCODEUR A CANAUX
REALISATION
D'UN DETECTEUR DE PITCH**

Proposé par : **M:**
Bousseksou

Etudié par : **MM:**
**Mansouri
Laib**

Dirigé par : **M:**
Bosseksou

PROMOTION **Juin 86**

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

DEDICACES

A la mémoire de mon père.

A ma chère mère.

A mes frères et sœurs .

A tous mes amis.

Je dédie ce modeste travail.

R.MANSOURI

05/06/1986

Je dédie ce travail:

à mon père.

à ma mère.

à mes frères.

à mes sœurs.

et à tous les amis...

L.AMAR

R E M E R C I E M E N T S

Nous remercions profondément notre promoteur B. BOUSSEKSOU pour sa large aide envers nous et pour ses précieuses suggestions.

Nous tenons à remercier ceux qui nous ont aidé à écrire ce polycopie notamment M^{lle} BOURENANE et M^r BENHALIMA.

Nos remerciements vont aussi à ceux qui ont contribué de près ou de loin à notre formation.

T A B L E D E M A T I E R E S
 ----- o ----- o -----

INTRODUCTION I
 ----- o -----

PREMIERE PARTIE

Chapitre I- FILTRES ANALOGIQUES 3
 I-1- Caractéristiques des filtres 3
 I-2- Filtres usuels 4

Chapitre II- NUMERISATION DES SIGNAUX 6
 II-1- Echantillonnage 6
 II-2- Quantification 7

Chapitre III - ANALYSE ET SYNTHESE DE LA PAROLE . . . 9
 III-1 Caractéristiques acoustiques du signal
 de la parole 9
 III-2- Analyse de la parole 9
 III-3- Détermination du fondamental 12
 III-4- Synthèse de la parole 13
 ----- o -----

DEUXIEME PARTIE

Chapitre IV- GENERALITES SUR LES VOCODEURS 16
 IV-1- Principe d'un vocodeur 16
 IV-2- Vocodeur à canaux 17
 IV-3- Vocodeur à formants 1
 IV-4- Vocodeur à prédiction linéaire 19

Chapitre V- ETUDE D'UN VOCODEUR A CANAUX 21
 V-1- Introduction 21
 V-2- Analyseur du vocodeur 22
 V-3- Synthétiseur du vocodeur 138

TROISIEME PARTIE

Chapitre VI- REALISATION D'UN DETECUR DE MELODIE . . .	42
VI-1- Introduction	42
VI-2- Calcul et choix des éléments	42
VI-3- Programme d'extraction et de mesure . . .	47

-----o-----o-----o-----

Le domaine des transmissions est vaste dans lequel la téléphonie occupe une place très importante.

Les recherches entreprises dans la téléphonie visent la réduction de la bande passante et la transmission à faible débit binaire.

En 1939, DUDLEY a inventé un appareil qui permet de réduire la bande téléphonique dans un rapport de 10 et fonctionne par conséquent à faible débit, il s'agit du vocodeur à canaux. Depuis lors, plusieurs générations de vocodeurs ont succédé présentant des améliorations et des optimisations sur le vocodeur original.

Dans le présent travail, on essaye de faire l'étude d'une version de vocodeur à canaux. L'exposé comprend 3 parties.

Dans la première partie, on donne les principales fonctions de traitement du signal tout en faisant appel aux relations mathématiques. Un chapitre est réservé aux méthodes d'analyse et de synthèse de la parole. Ces notions nous serviront de base pour l'étude de notre vocodeur.

Dans la deuxième partie, on fait l'étude du vocodeur proprement dite. Après un bref rappel sur les différents types de vocodeurs, on abordera la description détaillée de l'appareil. Une bonne partie sera réservée à l'étude d'un détecteur de "pitch": l'élément le plus sensible tout vocodeur.

En fin, dans la troisième partie on essayera de décrire la réalisation d'un détecteur de pitch tout en précisant le calcul des éléments et l'algorithme utilisé pour extraire ce pitch.

1^e Partie

« En toute chose, il n'y a qu'une manière
de commencer quand on veut discuter
convenablement : il faut bien comprendre
l'objet de la discussion. »

“PLATON”

CHAPITRE I

LES FILTRES ANALOGIQUES

I-I- CARACTERISTIQUES DES FILTRES

I-I-I- LE GABARIT

Le filtre passe-bas idéal suppose une atténuation nulle en bande passante et infinie en dehors de la bande, or un tel filtre n'est pas réalisable.

Pour un filtre réel réalisable, l'atténuation en bande passante n'est pas nulle, elle est inférieure à une valeur A_{max} . En bande coupée elle n'est pas infinie, mais elle est supérieure à une valeur A_{min} .

La transition entre A_{max} et A_{min} définit la raideur du filtre elle est traduite par la sélectivité k :

$$k = \frac{f_p}{f_a}$$

où f_p est la fréquence passante, f_a est la fréquence atténuée.

Ainsi le gabarit se définit par 4 paramètres: A_{max} , A_{min} , f_p et f_a . Les figures (I-1) et (I-2) montrent respectivement le gabarit d'un filtre passe-bas et celui d'un filtre passe-bande

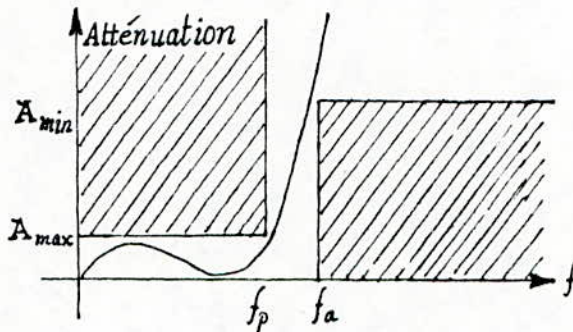


Fig I-1

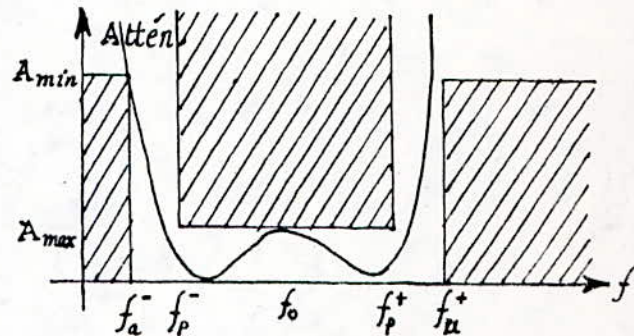


Fig I-2

Pour un filtre passe-bande, 6 paramètres sont nécessaires pour définir le gabarit. On définit ainsi la fréquence centrale f_0 , la sélectivité k et la bande B .

$$f_0^2 = f_p^+ \cdot f_p^- ; \quad k = \frac{f_p^+ - f_p^-}{f_a^+ - f_a^-} ; \quad B = \frac{f_p^+ - f_p^-}{f_0}$$

I-I-2- LE TEMPS DE PROPAGATION

Le gabarit ne donne qu'une information sur le module, or un filtre réel introduit un déphasage dans le signal à traiter.

Pour un filtre linéaire, la fonction retard est donnée par:

$$R(p) = A \cdot \exp(-Tp)$$

Ainsi le filtre introduit un retard du signal de durée T .

Le temps de propagation est donné par:

$$T = - \frac{d\phi(\omega)}{d\omega}$$

Où ϕ désigne la phase du filtre réel.

I-1-3- LA FONCTION CARACTERISTIQUE

Soient $F(w)$ la fonction de transfert du filtre et $A(w)$ son atténuation. On a :

$$F(w) = M(w) \exp(j\phi(w))$$

Où $M(w)$ désigne le module de $F(w)$ et $\phi(w)$ sa phase.

On a ainsi

$$A(w) = \frac{1}{M(w)}$$

qui peut s'écrire sous la forme :

$$A(w) = A^2(0) \cdot (1 + K(w^2))$$

La fonction K est appelée fonction caractéristique du filtre, la qualité d'un filtre est d'autant plus bonne que sa fonction caractéristique est voisine de zéro dans la bande passante et grande à l'extérieur.

I-1-4- LA TRANSPOSITION

On montre que la plupart des filtres peuvent être construits à partir du filtre passe-bas à l'aide d'une simple transposition, cette transformation exige que le filtre soit normalisé, c-à-d que sa fréquence de coupure soit égale à l'unité. Le passage aux filtres passe-haut, passe-bande et coupe-bande s'effectue respectivement par les transformations suivantes :

$$p \text{ ----- } \frac{1}{p}$$

$$p \text{ ----- } B(p + \frac{1}{p})$$

$$p \text{ ----- } B \cdot \frac{1}{p + \frac{1}{p}}$$

I-1-5- LES FILTRES DU 1^{er} ET DU 2^{ème} ORDRE

On montre que n'importe quel filtre peut être synthétisé à partir de filtres du 1^{er} et du 2^{ème} ordres regroupés en cascade.

La fonction de transfert d'un filtre du 1^{er} ordre est donnée par :

$$F(p) = F(0) \cdot \frac{1}{1 + TP}$$

Celle d'un filtre du 2^{ème} ordre est donnée par :

$$F(p) = F(0) \cdot \frac{w_0^2}{w_0^2 + 2aw_0 p + p^2}$$

I-2- FILTRES ANALOGIQUES USUELS

D'une façon générale la fonction de transfert d'un filtre linéaire est donnée par :

$$F(p) = \frac{1}{1 + a_1 p + a_2 p^2 + \dots + a_n p^n}$$

Le tableau ci-après donne les différentes valeurs des a_i pour les filtres usuels :

TYPE	ORDRE	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7
TCHÉBYCHEFF	2	1	0,995	0,907	-	-	-	-	-
	3	1	2,520	2,011	2,535	-	-	-	-
	4	1	2,694	2,749	3,456	3,62	-	-	-
	5	1	4,725	7,933	13,75	7,627	8,141	-	-
	6	1	4,456	13,63	147,4	28,02	13,47	14,51	-
	7	1	9,300	19,16	66,86	53,06	121,4	36,26	63,45
BUTTERWORTH	2	1	1,414	1	-	-	-	-	-
	3	1	2	2	1	-	-	-	-
	4	1	2,613	3,414	2,613	1	-	-	-
	5	1	3,236	5,236	5,236	3,236	1	-	-
	6	1	3,863	7,464	9,141	7,464	3,863	1	-
BESSEL	2	3	3	1	-	-	-	-	-
	3	15	15	6	1	-	-	-	-
	4	105	105	45	70	1	-	-	-
	5	945	945	420	105	15	1	-	-
	6	10395	10395	4725	1260	820	21	1	-
LEGENDRE	2	1	1,414	1	-	-	-	-	-
	3	1	2,353	2,270	1,731	-	-	-	-
	4	1	3,041	4,625	3,82	2,449	-	-	-
	5	1	4,017	7,56	9,852	6,936	4,471	-	-
	6	1	4,885	11,54	14,20	19,01	12,20	1,070	-

Tableau I-1- Paramètres des filtres usuels.

CHAPITRE II

NUMERISATION DES SIGNAUX

II-1-ECHANTILLONNAGE

II-1-1- DEFINITION

L'échantillonnage consiste à représenter un signal $x(t)$ fonction du temps par ses valeurs $x(nT)$ à des instants multiples entiers d'une durée T appelée période d'échantillonnage.

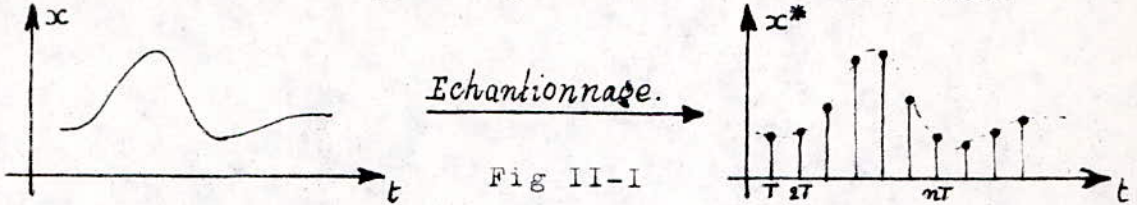


Fig II-1

II-1-2- EXPRESSION DU SIGNAL ECHANTILLONNE

L'échantillonnage est réalisé par une suite d'impulsions infiniment brèves appelée fonction peigne: $\Delta_T(t)$

$$x^*(t) = x(t) \cdot \Delta_T(t) \quad , \quad \Delta_T(t) = \sum_{-\infty}^{\infty} \delta(t - nT)$$

$$x^*(t) = \sum_{-\infty}^{\infty} x(nT) \cdot \delta(t - nT)$$

II-1-3- SPECTRE DU SIGNAL ECHANTILLONNE

On montre que la transformée de FOURIER du signal échantillonné est donnée par:

$$X^*(f) = \frac{1}{T} \sum_{-\infty}^{\infty} X(f - \frac{n}{T})$$

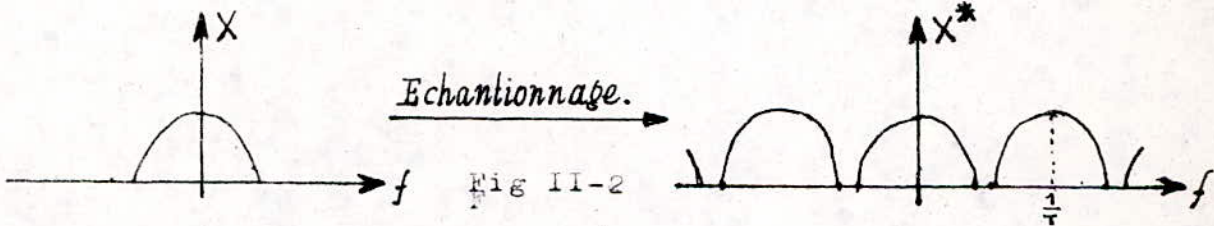


Fig II-2

REMARQUE: L'échantillonnage introduit une discrétisation du signal et une périodisation de son spectre.

II-1-4- RECONSTITUTION DU SIGNAL INITIAL - THEOREME DE SHANON

Pour que le spectre $X(f)$ ait les informations que contient $X(f)$, il faut que la fréquence maximale du spectre $X(f)$ soit au plus égale à la moitié de la fréquence d'échantillonnage, comme le montre la figure suivante:

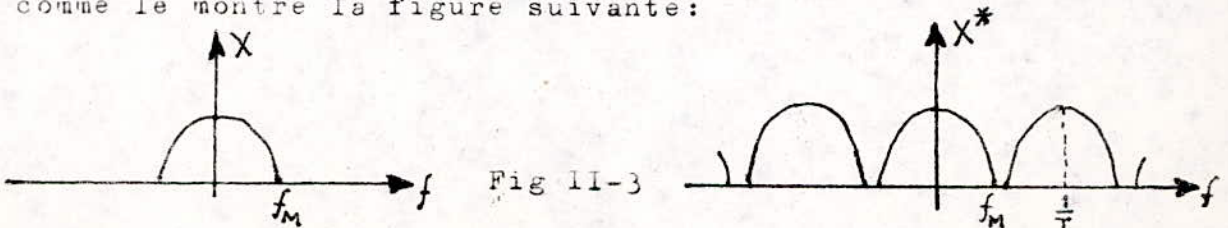
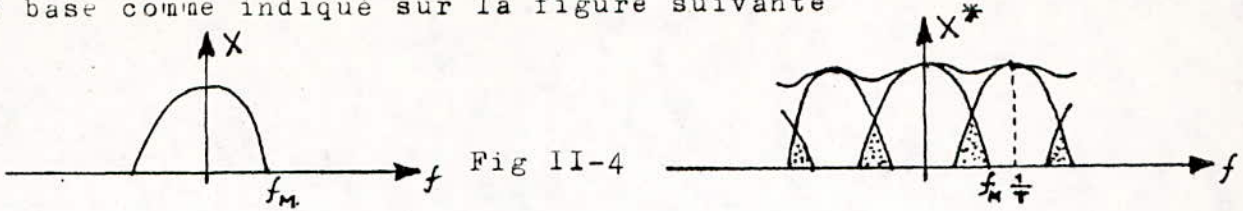


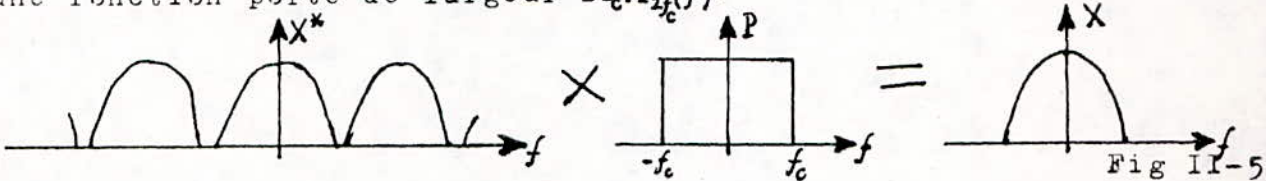
Fig II-3

Si ce n'est pas le cas, les bandes images chevauchent la bande de base comme indiqué sur la figure suivante



Pour reconstituer le signal $x(t)$ à partir des échantillons $x(nT)$, il suffit de le faire passer à travers un filtre passe-bas de fréquence de coupure f inférieure ou égale à f_M .

Le procédé est de multiplier le spectre des échantillons par une fonction porte de largeur $2f_c$: $P_{2f_c}(f)$.



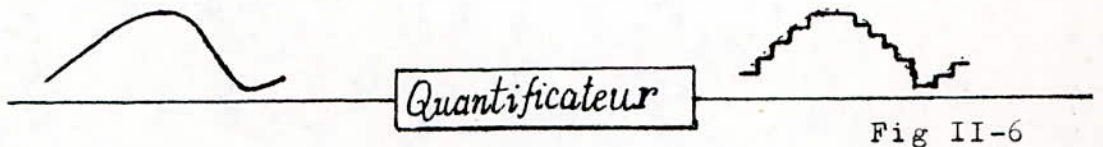
On montre alors que le signal $x(t)$ est donné par:

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin \frac{\pi}{T}(t - nT)}{\frac{\pi}{T}(t - nT)}$$

II-2- QUANTIFICATION

II-2-1- DEFINITION

La quantification est l'approximation de chaque valeur du signal par un multiple entier d'une quantité élémentaire q appelé échelon de quantification.

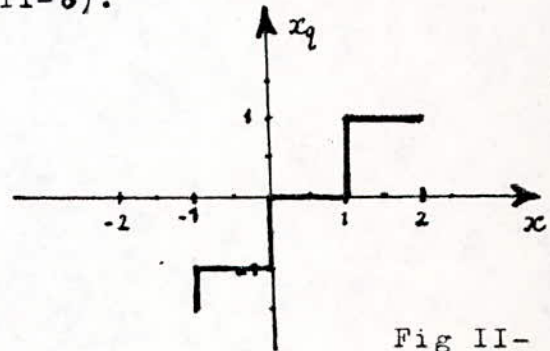
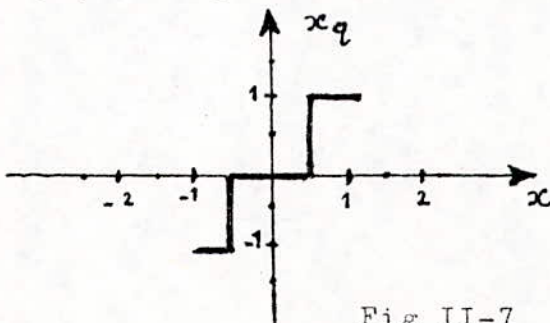


II-2-2- QUANTIFICATION UNIFORME

Si q est constant quelque soit l'amplitude du signal, la quantification est dite uniforme. Deux sortes d'approximation peuvent être envisagées:

- le cas dit arrondi, où toute valeur comprise entre $(n - \frac{1}{2})q$ et $(n + \frac{1}{2})q$ est approximée à la valeur nq (fig-II-7).

- l'approximation par défaut, où toute valeur comprise entre nq et $(n+1)q$ est arrondie à nq . (fig II-8).



II-2-3- QUANTIFICATION NON UNIFORME

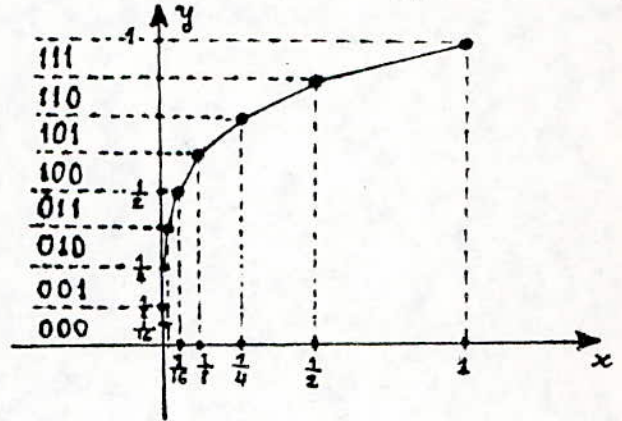
La distribution des signaux à quantifier est rarement uniforme c'est le cas du signal de la parole. Les pas de quantification ne sont plus constants. La quantification est dite non uniforme. Cette opération peut être considérée comme résultant d'un codage linéaire précédé d'une compression. Plusieurs lois de compression ont été proposées:

II-2-3-1- LA LOI A-LOGARITHMIQUE

Les amplitudes positives et négatives à coder sont divisées en 7 plages à chacune desquelles est associé un échelon de quantification dont la grandeur résulte de la multiplication d'un échelon q par une puissance de 2. La loi de compression est donnée par:

Selon CCITT, la valeur de A est optimisée à 7,6.

Fig II-9



II-2-3-2- LA LOI μ

Il existe une autre loi utilisée en télécommunication dite loi μ à 15 segments. Elle correspond à la relation de compression suivante:

La valeur de μ est optimisée à 255.

CHAPITRE III

ANALYSE ET SYNTHÈSE DE LA PAROLE

III-1- CARACTERISTIQUES ACOUSTIQUES DU SIGNAL DE LA PAROLE

- Le signal de la parole possède les caractéristiques suivantes
- Le spectre de la parole s'étend de 50 à 3000 Hz pour les sons voisés et jusqu'à 6000 Hz pour les sons non voisés.
 - La dynamique de la parole est inférieure à 50 dB.
 - Le signal issu de la glotte est périodique, il a un spectre de raies (fig III-1). Sa fréquence est appelée FREQUENCE FONDAMENTALE ou PITCH. Son spectre décroît en moyenne de 12 dB/octave
 - La fréquence fondamentale est comprise entre 70 et 500 Hz.
 - Le signal glottique est modulé par le conduit vocal dont la fonction de transfert présente un certain nombre de pôles appelés FORMANTS (fig III-2).

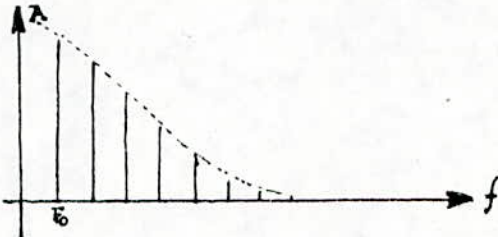


Fig III-1

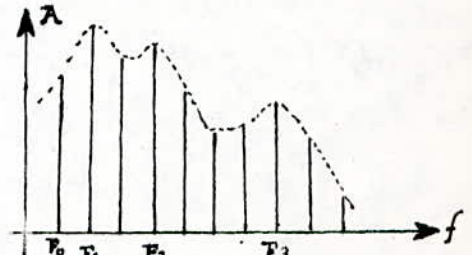


Fig III-2

III-2- ANALYSE DE LA PAROLE

Les analyseurs de la parole ont pour rôle de décomposer le signal issu du microphone en ses constituants les plus simples (pitch, formants, enveloppe spectrale; ..). Ces éléments destinés au stockage ou à la transmission sont souvent convertis sous forme numérique. Plusieurs analyseurs ont été proposés aussi bien dans le domaine temporel que dans le domaine fréquentiel, on peut citer:

III-2-1- ANALYSE SPECTRALE

III-2-1-1- ANALYSEURS A FILTRE UNIQUE, LE SONAGRAPHE

Dans ce type d'analyseurs on exploite le principe de l'hétérodynage: le signal à analyser est multiplié par un signal produit par un oscillateur local dont on peut faire varier

la fréquence. Le signal de sortie est filtré par un filtre passe-bas, on ne conserve donc que la composante dont la fréquence est la différence des fréquences des deux signaux. On peut donc explo-
rer tout le spectre du signal de la parole en faisant varier la fréquence de l'oscillateur local, mais il faut à chaque opération réinjecter le signal.

Une application de ce type d'analyseurs est le sonographe (fig III-3), c'est un appareil qui permet une représentation tridimensionnelle (Amplitude-Fréquence-Temps). Le signal issu d'une tête de lecture (une tranche de signal de 2 secondes environ étant déjà enregistrée sur un disque magnétique) est appliqué à l'entrée d'un filtre passe-bande dont la fréquence centrale est accordée à la position verticale d'un stylet. Ce stylet est en contact avec un papier sensible au courant électrique, son déplacement vertical s'effectue par l'intermédiaire d'une vis sans fin, il provoque une variation linéaire de la fréquence centrale dans la bande 0-8000 Hz.

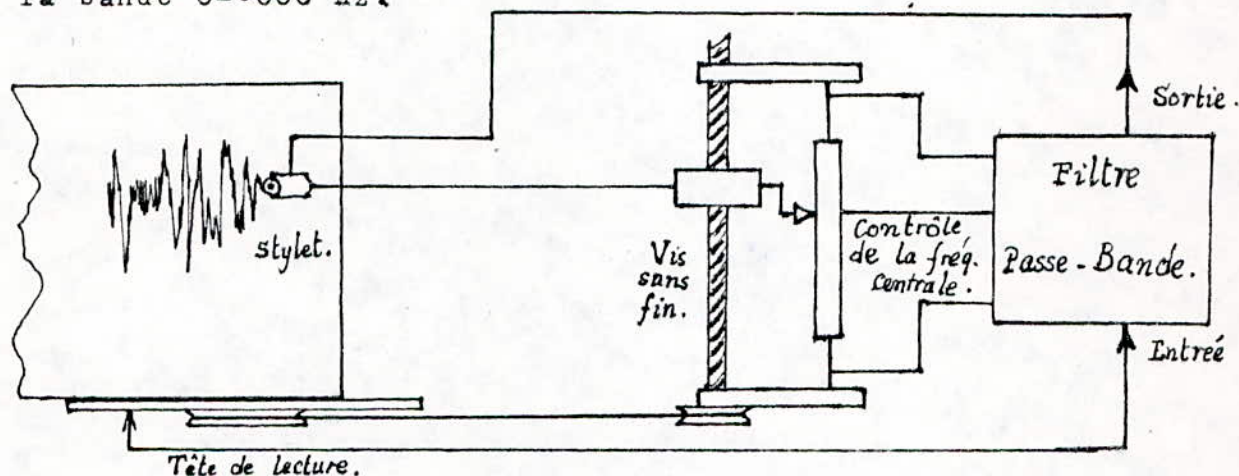


Fig III-3

III-2-1-2- ANALYSEURS A FILTRES MULTIPLES EN PARALLELE, LE VOCCO-
DEUR A CANNAUX

Les analyseurs à filtre unique sont très lents (il faut 5 ans pour analyser 2s de signal), l'analyse sera pratiquement instantannée si l'on applique le signal d'entrée à un grand nombre de filtres passe-bande montés en parallèle.

Une application de ce type d'analyseurs est le vocodeur à

cannaux dont une description détaillée est donnée dans les chapitres qui suivent.

III-2-2- ANALYSE TEMPORELLE

Dans certains cas (production d'une plosive par exemple), le signal est mieux caractérisé par l'évolution temporelle que son spectre. Les principales techniques proposées dans ce sujet sont les suivantes:

III-2-2-I- METHODE D'AUTOCORRELATION

Elle est basée sur les propriétés de la fonction d'autocorrélation du signal. Cette méthode est notamment utilisée pour la détermination du fondamental.

III-2-2-2- METHODE DE DETECTION D'EVENEMENTS

Passage par zéro, passage par un seuil ou détection de crêtes

III-2-3- ANALYSE PREDICTIVE

A intervalles de temps réguliers (10ms environ), le spectre du signal est exprimé sous la forme d'une réponse d'un filtre récursif. La réponse d'un tel filtre est de forme:

$$x(n) = \sum_{k=1}^p a_k x(n-k) + e(n)$$

où $e(n)$ représente le signal d'excitation.

L'information est stockée sous la forme des paramètres caractéristiques de ce filtre. Le problème est donc de déterminer ces coefficients.

III-2-4- ANALYSE CEPSTRALE

Le cepstre est défini comme étant le carré de la transformée du logarithme du spectre de puissance, il représente donc le spectre de puissance du logarithme du spectre de puissance du signal de la parole.

On verra par la suite que cette méthode est bien adaptée à la détermination du fondamental.

On a représenté aux figures (III-4) ET (III-5) les cepstres d'un signal voisé et d'un signal non voisé.

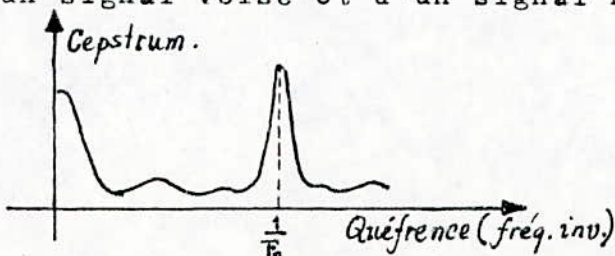


Fig III-4

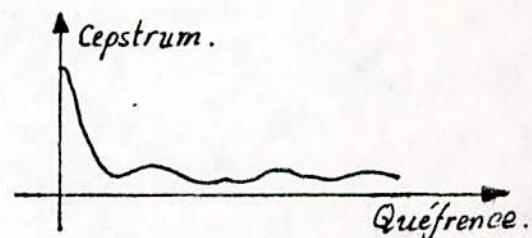


Fig III-5

III-3- DETERMINATION DU FONDAMENTAL

Le naturel de la parole synthétique dépend beaucoup de la justesse et de la finesse de l'analyse de F_0 , l'oreille humaine est plus sensible aux variations de F_0 qu'à l'emplacement et la largeur des formants.

La détermination de F_0 pour les sons voisés est une opération très délicate, de nombreuses méthodes ont été proposées. D'une manière générale et quelque soit la méthode utilisée, le traitement se fait en 3 étapes:

-Un prétraitement (filtrage passe-bas, normalisation d'amplitude de....)

-Le traitement proprement dit (extraction et mesure du fondamental)

-Un posttraitement (corrections)

III-3-1- METHODES SPECTRALES

III-3-1-1- METHODE CEPSTRALE

Le cepstre d'un signal a été défini au paragraphe III-2-4.

Si le son est voisé on observe une raie très marquée centrée sur la valeur T de la fréquence (correspondant à F_0) et une courbe plus étalée indiquant les oscillations lentes et apériodiques.

Pour les sons non voisés, il n'apparaît pas de raie caractéristique (voir fig III-4 et fig III-5). Le cepstre est donc un moyen pour déterminer si le son est voisé ou non.

Cette méthode exige que l'analyse spectrale soit faite avec une très bonne résolution, or cela est très difficile pour les F_0 faibles car les harmoniques sont très rapprochées.

III-3-1-2- INTERCORRELATION AVEC UNE FONCTION PEIGNE

Cette méthode est caractérisée par la recherche du maximum d'une fonction d'intercorrélation entre le spectre instantané du signal et une fonction peigne.

Le maximum a lieu lorsque l'emplacement des dents de la fonction peigne coïncide avec l'emplacement des raies harmoniques de F_0 est alors la fréquence de répétition des dents.

III-3-2- METHODES TEMPORELLES

III-3-2-1- LES METHODES DE DECALAGE

Dans ces méthodes, on exploite les propriétés de la fonction de corrélation qui existe généralement entre deux périodes fondamentales successives d'un signal voisé.

Le signal est découpé en fenêtres temporelles (de 10 à 30ms) dans lesquelles le paramètre à mesurer F_0 est considéré comme constant. Ensuite on cherche le maximum de la fonction:

$$f(d) = \sum x(n) \cdot x(n+d).$$

Ce maximum a lieu quand le décalage d est égal à la durée T_0 correspondant à la fréquence F_0 ;

III-3-2-2- LES METHODES DE DETECTION DES EVENEMENTS

Dans ces méthodes, le signal à traiter subit un filtrage passe-bas très sévère. Les événements à détecter peuvent être de différentes natures:

-Des passages par zéro.

-Des passages par des seuils fixes, ce genre d'information est un plus porteur d'informations sur F_0 . C'est celui-ci que nous avons utilisé pour notre détecteur de mélodie.

-Détection de crêtes.

III-4- SYNTHÈSE DE LA PAROLE

III-4-1- PRINCIPE

Le principe des synthétiseurs est de créer une analogie plus ou moins grande avec notre appareil vocal, par assemblage des constituants préenregistrés ou par simulation du conduit vocal.

Pour les appareils de la première génération, il s'agissait de créer une source avec une modification artificielle du spectre. Le vocodeur à canaux entre dans cette catégorie.

Les appareils de la seconde génération avaient pour rôle non plus de schématiser la source vocal mais aussi de simuler le processus de modulation du spectre.

La parole synthétisée est très proche de l'originale. Le vocodeur à formants est un appareil de ce type.

Pour remédier aux défauts de la première et de la deuxième génération, on a proposé des appareils dites de la troisième génération qui simulent dans la forme et les dimensions le conduit

vocal.

Les synthétiseurs à cannaux, à formants et à prédiction linéaire seront décrits respectivement aux paragraphes IV-2, IV-3 et IV-4.

III-4-2- SYNTHÈSE PAR SIMULATION DU CONDUIT VOCAL

Le principe est de décomposer le conduit vocal en tranches chacune de ces tranches peut-être simulée électriquement par une cellule électrique.

Ainsi le conduit vocal se trouve représenté par une succession de cellules montées en cascade (fig III-6).

Selon leur fonctionnement, les simulateurs du conduit vocal peuvent-être classés en deux catégories :

- Les simulateurs à fonctionnement statique : ils ne peuvent synthétiser que des voyelles ou des consonnes du type S, CH, F.

- Les simulateurs à fonctionnement dynamique : dans ce cas on peut jouer automatiquement sur les paramètres de la ligne. On peut donc imiter les variations de la section du conduit vocal en utilisant une fonction d'aire. Cette fonction d'aire est très difficile à déterminer, plusieurs techniques ont été proposées : cinéradiographie, empreintes du pharynx, exploration par ultrasons etc...

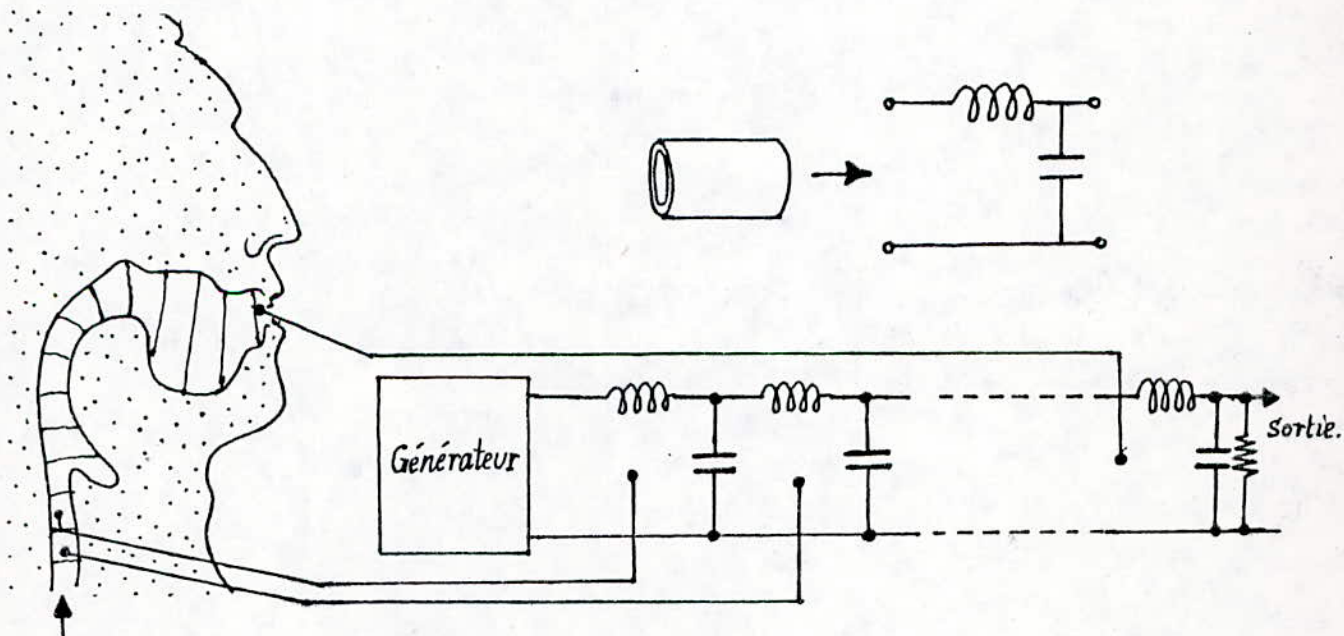


Fig III-6

2^e Partie

« Rien n'est aussi pratique
q'une bonne théorie »

"K. LEVIN"

CHAPITRE IV

GENERALITES SUR LES VOCODEURS

IV-1- PRINCIPES DE BASE

Les premières machines parlantes étaient des machines mécaniques. Dès 1791, Von Kempeler réalisa une machine qui émettait de la parole continue. La source d'excitation était formée d'une manche de cuir que l'on pouvait déformer à la main, la structure mécanique de l'appareil permet de produire la plupart des sons.

De nombreux autres machines mécaniques ont été construits mais leur réalisation et leur automatisation posaient de nombreux problèmes. Pour cette raison on s'est penché rapidement vers les simulateurs électriques ou optiques du conduit vocal.

Le premier vocodeur fût construit en 1939 par Dudley (aux Bell laboratoires-Etats unis) dans le but de réduire le débit binaire dans les transmissions téléphoniques. L'appareil fait l'analyse et la synthèse du message de la parole.

Le schéma de principe d'un vocodeur est donné (fig IV-1)

D'une façon générale, un vocodeur comporte :

- A l'analyse :
 - analyse et numérisation de l'enveloppe spectrale.
 - détection et codage du fondamental.
- A la synthèse :
 - restitution de l'enveloppe spectrale.
 - élaboration d'un signal d'excitation.

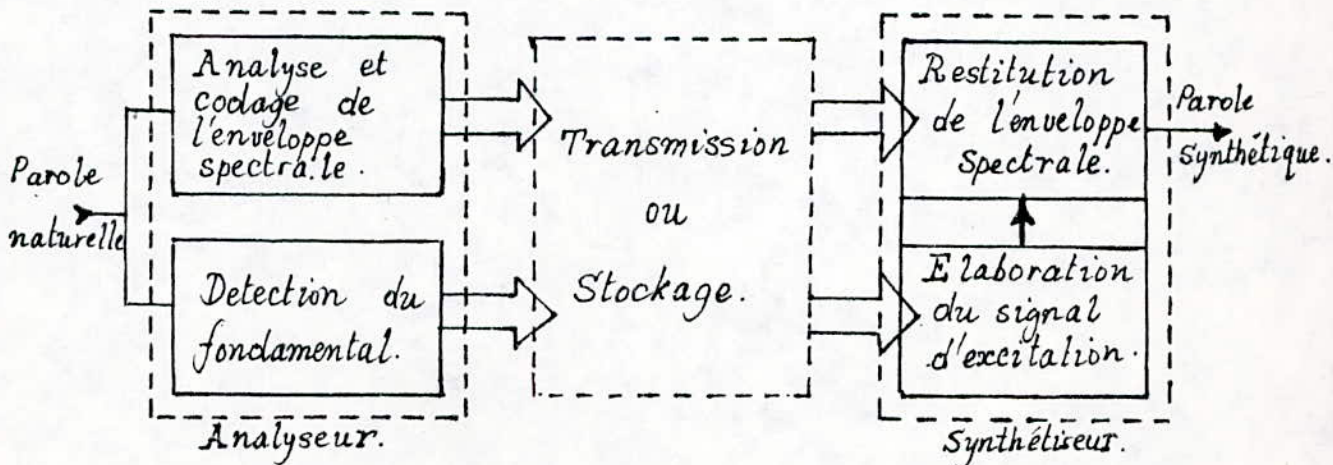


Fig IV-1

Nous examinerons dans ce qui suit les principaux vocodeurs à savoir;Levocodeur à canaux,le vocodeur à formants et le vocodeur à prédiction linéaire.

IV-2- LE VOCODEUR A CANAUX

IV-2-1- PRINCIPE

Ce vocodeur effectue l'analyse et le codage de l'enveloppe spectrale à l'aide de canaux dont le nombre peut varier de 10 à 20. Il détecte également la fréquence fondamentale à laquelle vibre la glotte.On réalise ainsi une transmission(ou un stockage)à faible débit.

Le synthétiseur réalise l'opération inverse;C'est à dire la restitution du signal de la parole.

La figure(IV-2)donne le schéma de principe d'un vocodeur à canaux.

IV-2-2- L'ANALYSEUR

Le signal issu du microphone est analysé dans différentes bandes de fréquences grâce à une batterie de filtres passe-bandes contigus.

Chaque filtre passe-bande est suivi d'un étage de détection (on n'en conserve que les parties positives)lui-même suivi d'un étage de filtrage passe-bas dont la fréquence de coupure est de l'ordre de 50Hz (on ne laisse passer que les variations lentes du signal).Le filtre passe-bas est suivi d'un étage de numérisation.

L'analyseur comporte aussi un étage de décision du caractère voisé/non voisé.Pour les sons voisés un détecteur de mélodie effectue la mesure et la détection de la fréquence fondamentale.

IV-2-3- LE SYNTHETISEUR

La synthèse est effectuée à l'aide d'un banc de filtres passe-bande identiques à ceux de l'analyse.Ces filtres sont excités par un signal d'excitation modulé par chacun des signaux issus des canaux d'analyse.Le signal d'excitation a un spectre de raies harmoniques de fondamental F_0 pour les sons voisés.Pour les sons non voisés,c'est un spectre de bruit.

Le signal de la parole synthétisée est obtenu par addition des sorties des filtres de synthèse.

Le dispositif que nous avons décrit est essentiellement intéressant dans deux cas. Dans le premier cas, les signaux sont transmis et on bénéficie ainsi d'une compression du débit de l'information.

Dans le second cas, les signaux sont stockés dans la mémoire d'un ordinateur. Ils sont ensuite utilisés pour former des phrases grâce à des programmes appropriés.

IV-3- LE VOCODEUR A FORMANTS

IV-3-1- PRINCIPE

Dans le vocodeur à formants, au lieu de transmettre des paramètres définissant l'enveloppe spectrale, on transmet des paramètres caractérisant l'emplacement, la largeur et l'amplitude des formants.

IV-3-2- L'ANALYSEUR

Dans l'analyse par vocodeur à formants, deux techniques différentes peuvent être utilisées:

La première consiste à rechercher les 3 premiers maximums du spectre instantané à l'aide d'un banc de filtres passe-bandes contigus semblables à celui de l'analyseur du vocodeur à canaux. Cette analyse en est d'autant plus lourde à réaliser, de plus elle souffre de plusieurs inconvénients: Apparition de maximum parasite le premier formant peut-être confondu avec F_0 .

La deuxième technique consiste à partager par filtrage les 3 domaines fréquentiels des 3 premiers formants. Dans chaque plage on mesure le nombre de passages par zéro et l'énergie, on en déduit ainsi la fréquence et l'amplitude dans chaque plage.

IV-3-3- LE SYNTHETISEUR

La fig(IV-3) représente le schéma de principe d'un synthétiseur à formants, on y trouve un canal pour la synthèse des sons vocaux, un deuxième pour les sons non vocaux et un troisième pour la nasalité.

La synthèse des sons voisés nécessite généralement 5 circuits de résonance dont 3 sont commandés, les 2 autres fixes. Pour les sons non voisés, une source de bruits excite un canal différent comportant 2 circuits résonants.

En tout, pour commander un synthétiseur à formants il faut une douzaine de tensions de commande. L'ensemble de ces réglages peut être pris en charge par un ordinateur.

Du point de vue structure, on trouve la structure série et la structure parallèle, la structure série comporte un nombre réduit de réglages.

IV-4- LE VOCODEUR A PREDICTION LINEAIRE

IV-4-I- PRINCIPES

Le conduit vocal peut-être assimilé à un filtre numérique récursif. Dans le vocodeur à prédiction linéaire, on calcule les paramètres définissant ce filtre numérique, on détermine ensuite la fréquence fondamentale. Ce sont les fonctions remplies par l'analyseur. A la synthèse on constitue le signal de la parole qui est la réponse d'un filtre numérique dont les paramètres sont ceux déterminés à l'analyse.

IV-4-2- L'ANALYSEUR

Le principe d'un analyseur d'un vocodeur à prédiction linéaire a été décrit au paragraphe III-2-3.

IV-4-3- LE SYNTHETISEUR

La figure (IV-4) illustre le schéma de principe d'un vocodeur à prédiction linéaire.

Selon que le son est voisé ou non et grâce à une commande binaire, le signal d'excitation est produit soit par un générateur d'impulsions dont la fréquence est celle relevée à l'analyse F_0 , soit par le générateur de bruit blanc.

Les valeurs des p coefficients relevés à l'analyse permettent de reconstituer le signal prédit $x(t)$ à partir de ses échantillons $x(n)$:

$$x(n) = \sum_{k=1}^p a_k \cdot x(n-k).$$

La valeur de $x(n)$ est combinée avec le signal d'excitation $e(n)$ lorsque ce dernier est présent. On obtient ainsi l'échantillon du signal de la parole à l'instant nT . En fin, un filtre passe-bas (5 kHz) permet la reconstitution du signal original.

La remise à jour de tous les paramètres de prédiction est assurée chaque période de pitch pour les sons voisés et chaque 10ms pour les sons non voisés.

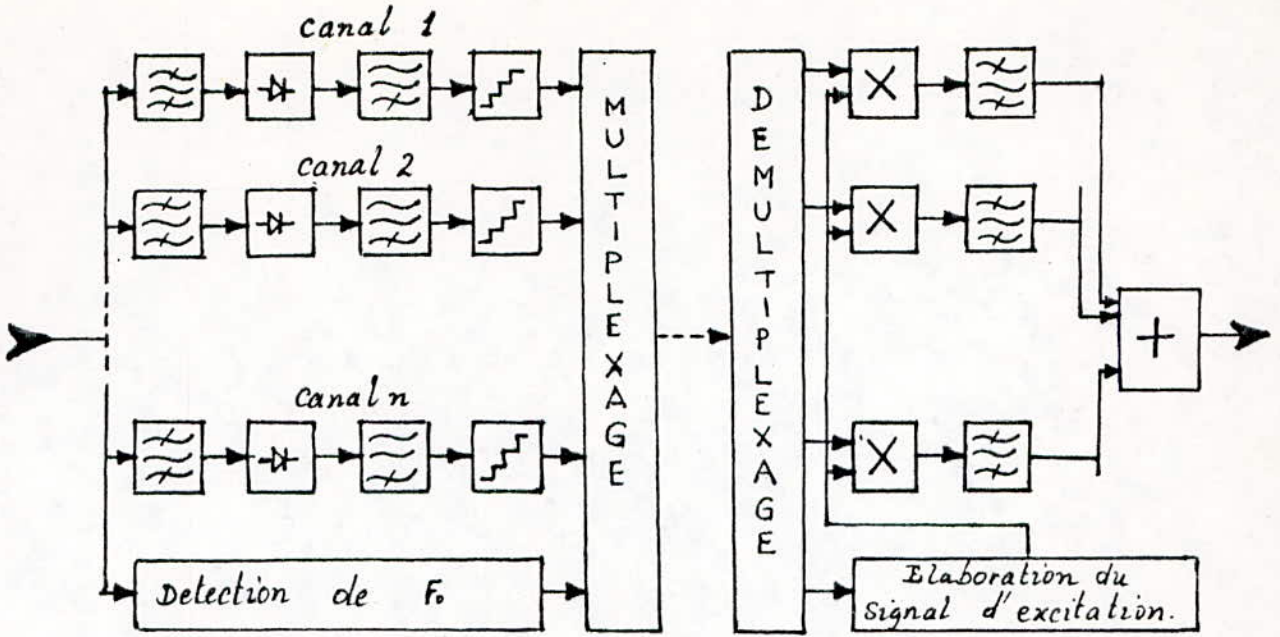


Fig IV-2

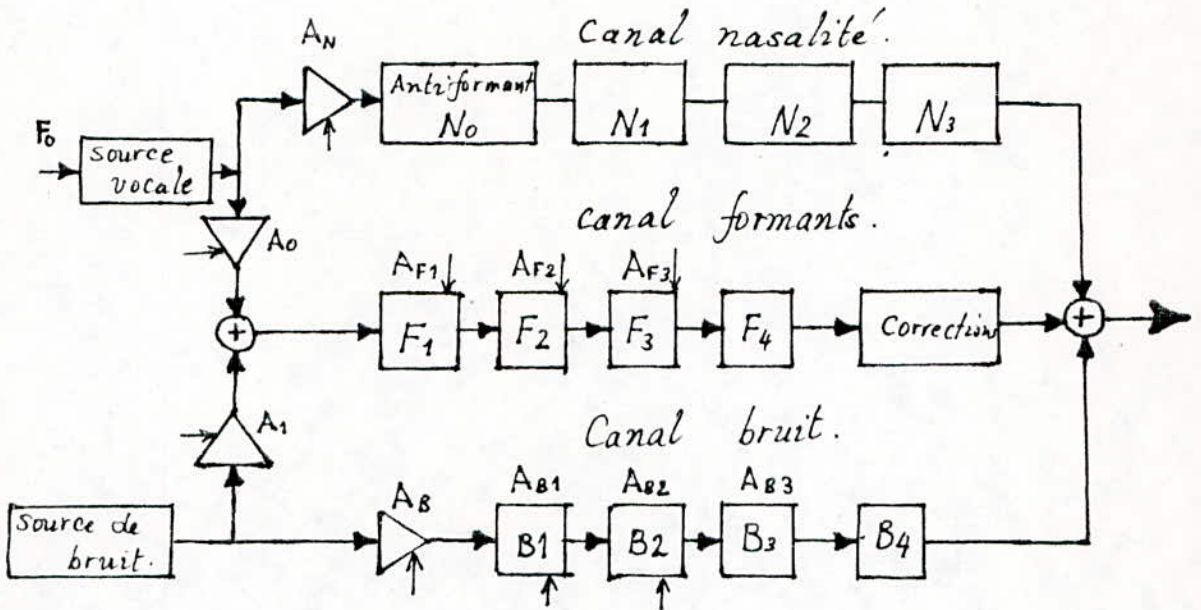


Fig IV-3

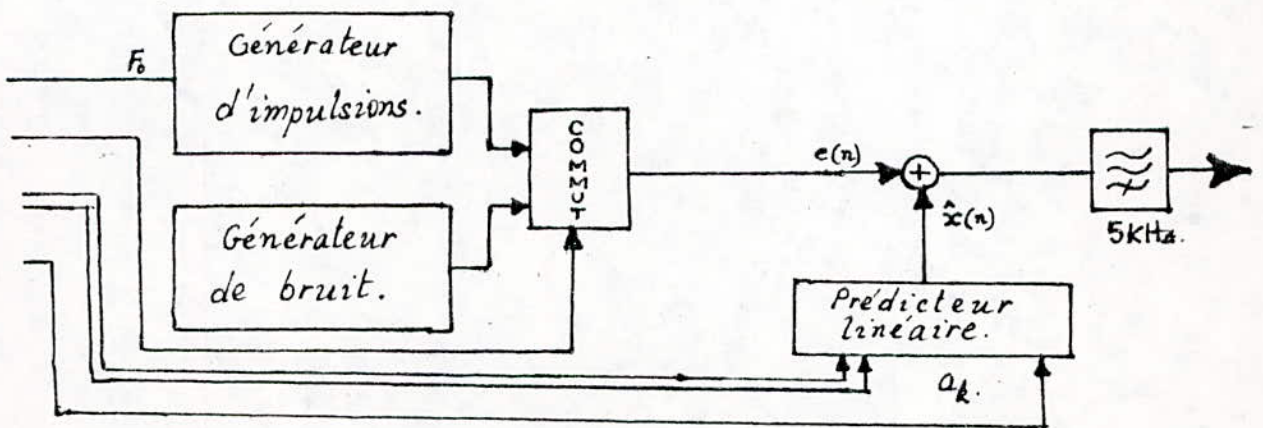


Fig IV-4

CHAPITRE V

ETUDE D'UN VOCODEUR A CANAUX

V-I- INTRODUCTION

Depuis 1939, plusieurs générations de vocodeurs à canaux ont succédé. Ces générations sont images de différentes améliorations et optimisation apportées sur le vocodeur en vu d'accroître l'intelligibilité et le naturel de la parole synthétisée.

A l'aide du vocodeur de DUDLEY (fig IV-2) on obtient une intelligibilité médiocre et une sonorité métallique désagréable. Ceci est dû en faite à la reconstitution non précise du spectre, au passage d'un son voisé à un son non voisé et à l'élaboration peu fidèle de la mélodie.

Pour remédier à ces défauts plusieurs structures de synthétiseurs ont été proposés, les figures V-1, V-2 ET V-3 schématisent ces différentes structures :

-La structure V-1 necessite un modulateur plus élaboré.

-La structure V-2 donne un son plus agréable.

-La structure V-3 apporte une diminution de la distorsion spectrale provoquée par les perturbations de la source d'excitation et une amélioration de la qualité de la parole.

D'autres structures ont été proposées telle que le synthétiseur à réponse vocale: c'est une structure intermédiaire entre la structure de Dudley et la structure V-3.

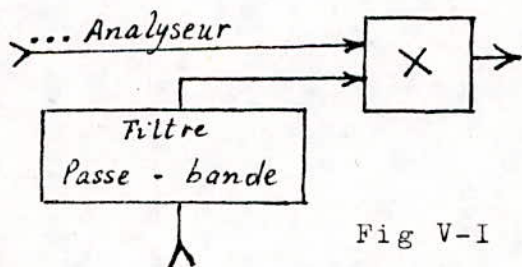


Fig V-1

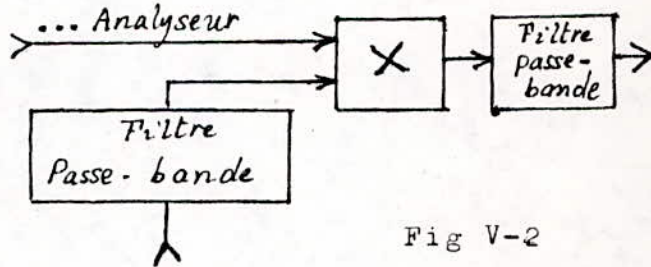


Fig V-2

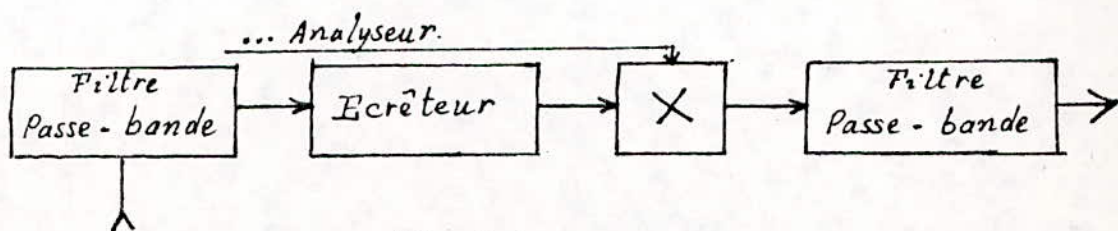


Fig V-3

La version du vocodeur que nous proposons d'étudier peut-être décrite comme suit :

V-2- L'ANALYSEUR DU VOCODEUR

V-2-I- LES FILTRES D'ANALYSE ET DE SYNTHÈSE

V-2-I-I- LES FILTRES PASSE-BANDE

Pour augmenter la fidélité de l'analyse et la netteté de la synthèse, il nous a semblé meilleur de porter le nombre de canaux à 14.

Le filtre passe-bande de chacun des canaux d'analyse et de synthèse est constitué d'une association en cascade de deux cellules de filtres passe-bande du 2^e ordre et présentant une réponse de BUTTERWORTH (courbe extra-plate en bande passante).

La fonction de transfert d'une cellule du 2^e ordre de *Butterw.* est donnée par :

$$\frac{1}{1 + \sqrt{2}p + p^2}$$

La figure V-4 schématise son circuit électrique.

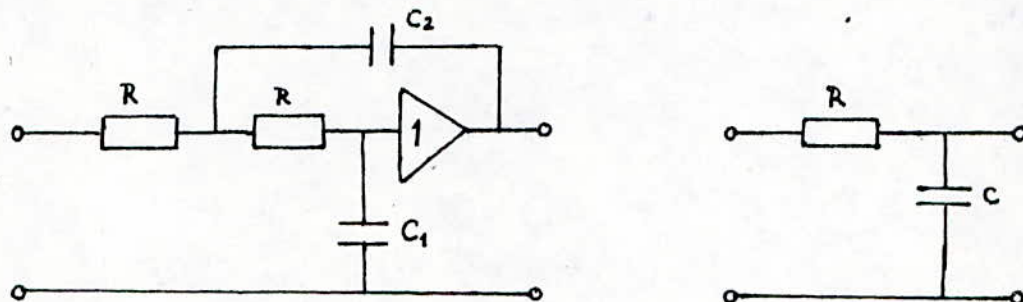


Fig V-4

Les filtres de deux canaux voisins se recoupent à -6dB, l'écart entre leurs fréquences centrales et leurs bandes sont largement discutés dans le paragraphe qui suit.

V-2-I-2- ECHELONNEMENT DES FILTRES PASSE-BANDE

A l'origine, dans le I^{er} échelonnement (échelonnement I) dit logarithmique deux fréquences centrales sont données par la relation :

$$f_{i+1} = a f_i$$

La constante *a* étant optimisée à $10^{\frac{1}{12}}$.

Le tableau (V-1) en donne le détail et la courbe I de la fig V-5 représente les variations des bandes réduites des différents canaux $\frac{\Delta f_i}{f_i}$ en fonction des valeurs de la fréquence centrale *f_i*

Cet échelonnement a pour avantage d'obtenir des coefficients

de surtension identiques pour tous les filtres d'où facilité de leur reproduction. De plus la résolution de l'analyse du bas du spectre est très bonne: le premier formant est localisé avec précision

En revanche, on constate que les filtres relatifs au bas du spectre sont trop étroits qu'en moyenne on ne trouve qu'une raie harmonique tous les 3 ou 4 canaux.

Par conséquent, une modification de F_0 (en synthèse) entraînera un déplacement des extrêmes de l'enveloppe spectrale donc une modification des phonèmes prononcés.

En 1970 un nouvel échelonnement (échelonnement II) a été proposé en remplacement de l'échelonnement I.

Le tableau V-2 en donne le détail, les variations de $-\frac{f_i}{\Delta f_i}$ en fonction de f_i sont représentées par la fig II de la fig V-5.

L'échelonnement II diffère de l'échelonnement I du fait qu'il possède des filtres à bandes beaucoup plus larges dans le bas du spectre, de plus il couvre une bande allant de 25 à 4300 Hz avec 14 canaux et présente donc un meilleur rapport bande globale sur nombre de canaux.

Par contre, on constate que les largeurs des filtres du bas du spectre sont excessives. De plus il est irrégulier, il ne permet pas donc une fonction de transfert régulière d'où une reproductivité des filtres difficile.

En 1979, un autre échelonnement est né (échelonnement III). Cet échelonnement conserve théoriquement les avantages des deux premiers tout en éliminant leurs défauts, il présente les avantages suivants:

- Les Δf_i sont en progression arithmétique.
- La progression de $-\frac{f_i}{\Delta f_i}$ est régulière.
- Bonne résolution d'analyse dans le bas du spectre.

Le tableau V-3 en donne le détail et la courbe III de la fig V-5 donne l'évolution de $-\frac{f_i}{\Delta f_i}$ en fonction de f_i .

Comparaison entre l'échelonnement II et l'échelonnement III

Des tests d'écoute de comparaison entre l'échelonnement II et l'échelonnement III ont été réalisés (5 phrases prononcées par 5 locuteurs masculins et 5 locuteurs féminins).

Les résultats des tests sont comme suit:

Canal N	Fréq. Centrale F _o	Fréq. de coupure à -6dB		F à - 6dB	$\frac{F_o}{\Delta F}$
1	274	24	301	52	5,2
2	332	301	365	64	"
3	402	365	442	77	"
4	487	442	536	94	"
5	590	536	649	113	"
6	715	649	787	138	"
7	866	777	953	166	"
8	1049	953	1155	202	"
9	1271	1155	1399	244	"
10	1540	1399	1695	296	"
11	1866	1695	2054	359	"
12	2260	2054	2488	434	"
13	2739	2488	3014	526	"
14	3318	3014	3652	638	"

Tableau V-1- Caractéristiques de l'échelonnement I

Canal N	Fréq. Centrale	Fréq. de coupure à -6dB		F à -6dB	$\frac{F_o}{\Delta F}$
1	335	250	450	200	1,7
2	541	450	650	200	2,7
3	743	650	890	200	3,7
4	945	890	1050	200	4,7
5	1168	1050	1300	250	4,7
6	1442	1300	1600	300	4,8
7	1744	1600	1900	300	5,8
8	2045	1900	2200	300	6,8
9	2345	2200	2500	300	7,8
10	2645	2500	2 00	300	8,8
11	2646	2 00	3100	300	9,8
12	3294	3100	3500	400	8,2
13	3695	3500	3900	400	9,2
14	4095	3900	4300	400	10,2

Tableau V-2- Caractéristiques de l'échelonnement II

Canal N	Fréq. Centrale F_0	Fréq. de coupure à -6dB		F à -6dB	$\frac{F_0}{\Delta F}$
1	296	250	350	100	2,95
2	410	350	480	130	3,15
3	554	480	640	160	3,45
4	729	640	830	190	3,85
5	934	830	1050	220	4,25
6	1168	1050	1300	250	4,55
7	1433	1300	1580	280	5,1
8	1728	1500	1900	310	5,60
9	2053	1900	2230	340	6,05
10	2400	2230	2600	370	6,5
11	2793	2600	3000	400	7,8
12	3200	3000	3430	430	7,45
13	3653	3430	3890	460	7,95
14	4128	3890	4380	490	8,5

Tableau V-3- Caractéristiques de l'échelonnement III

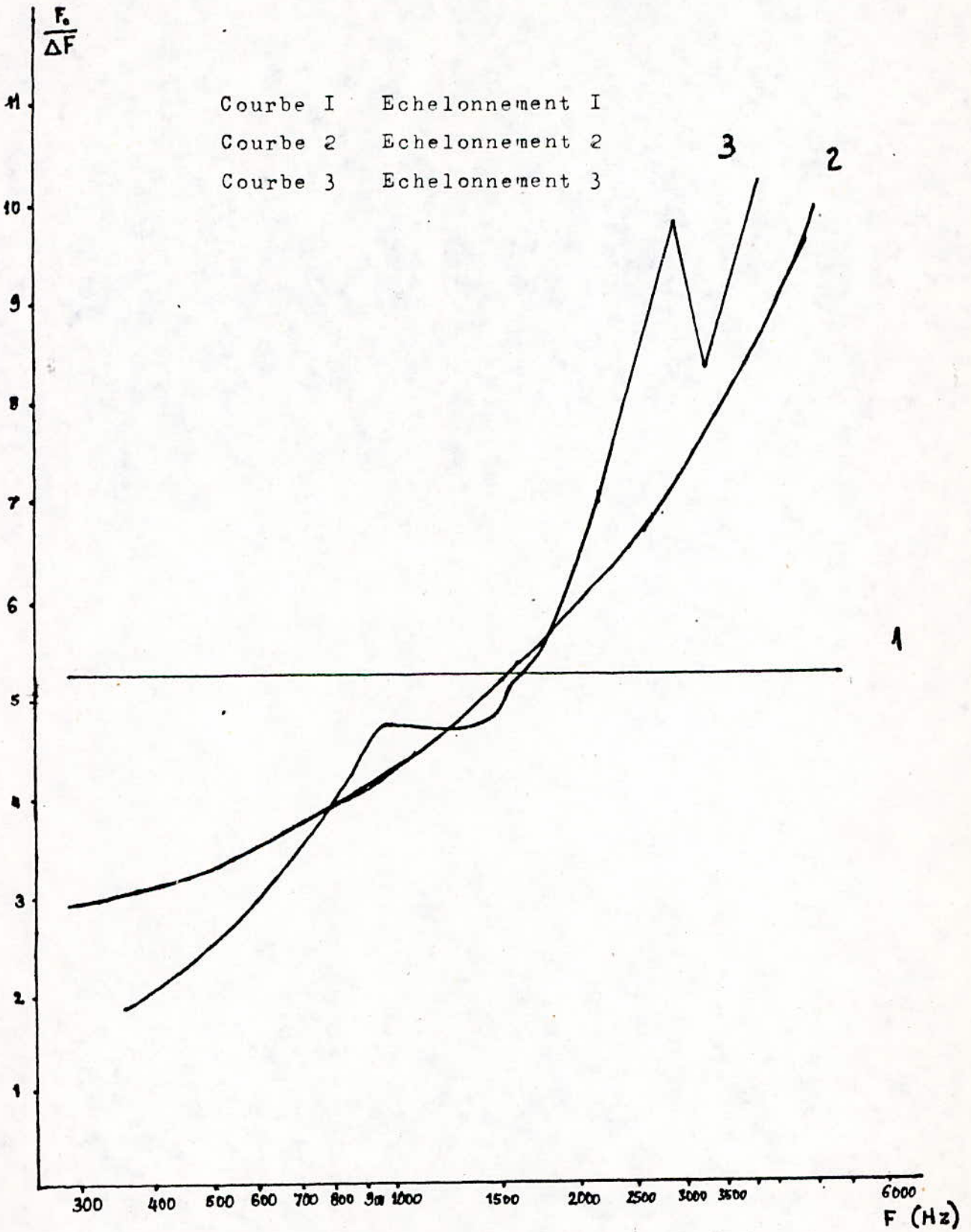


Fig V-5 Courbes représentatives de $\frac{F_0}{\Delta F}$ pour les 3 échelonnements.

	Globalement	Locuteurs mas.	Locuteurs fém.
Préférence éch. III	12%	95%	69%
Préférence éch. II	85%	5%	25%
Pas de préférence	3%	0%	6%

On constate une forte préférence de l'échelonnement III surtout pour les locuteurs masculins.

Ces constatations nous ont conduit à choisir comme échelonnement des filtres passe-bande (d'analyse et de synthèse) l'échelonnement III (tableau V-3, courbe III de la fig V-5).

V-2-1-3- LES FILTRES PASSE-BAS

Dans une région donnée du spectre, la densité spectrale de l'énergie ne varie d'une façon significative entre deux instants séparés de 40 ms, par conséquent les variations rapides de l'enveloppe correspondent à une fréquence maximale de 25Hz.

Le signal issu du détecteur subit un filtrage passe-bas dont la fréquence de coupure est de 25Hz. C'est-à-dire qu'on élimine les variations rapides du spectre.

Les filtres passe-bas sont du type de BUTTERWORTH du 3^e ordre ayant une fréquence de coupure de 25Hz. La fonction de transfert de ce filtre est de la forme:

$$F(p) = \frac{1}{1 + 2p + 2p^2 + p^3}$$

On peut réaliser ce filtre par mise en cascade de deux cellules l'une du 2^e ordre (fig V-6) l'autre du 1^{er} ordre (fig V-7).

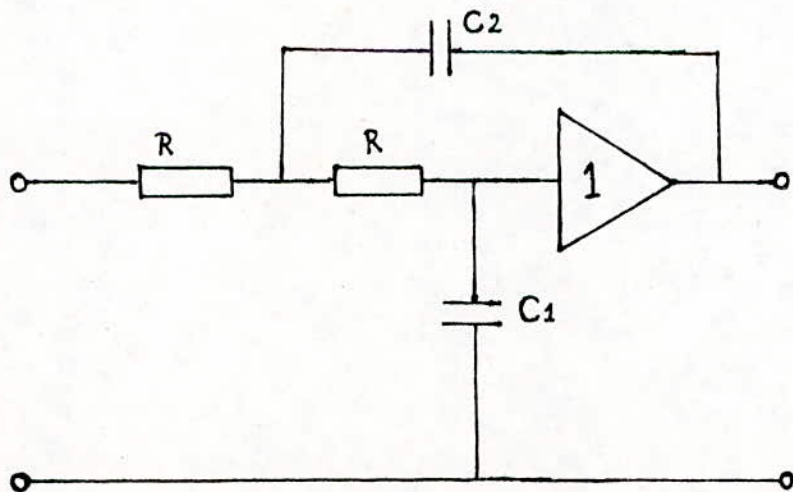


Fig V-6

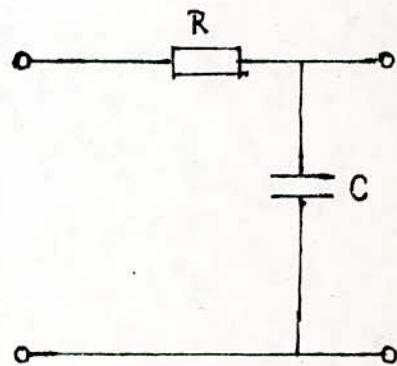


Fig V-7

V-2-2- DETECTION DES SIGNAUX

On ne traite numériquement que les signaux positifs, on est donc amené à éliminer les parties négatives. Pour ce faire on utilise des détecteurs qui sont des diodes simples (fig V-8)

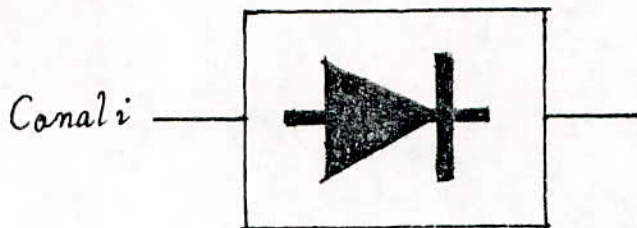


Fig V-8

V-2-3- NUMERISATION DES SIGNAUX

L'ensemble des informations recueillies sur le spectre et la fréquence fondamentale sont codées sous forme numérique.

La fréquence d'échantillonnage est de 50Hz, ce qui permet de considérer l'évolution temporelle de l'enveloppe spectrale toutes les 20 ms.

Pour coder ces informations, nous avons proposé une quantification logarithmique à 8 niveaux. Le pas de quantification est de 4dB ce qui donne 30dB comme dynamique (partie positive du signal).

On peut schématiser la numérisation comme suit (fig V-9)

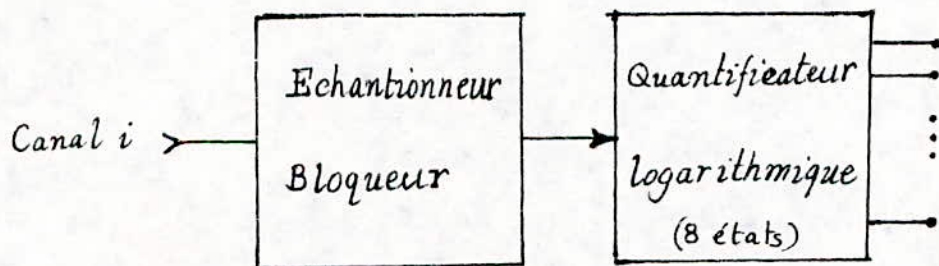


Fig V-9

V-2-4- LE DETECTEUR DE MELODIE

V-2-4-I- PRINCIPE

Le détecteur de mélodie du vocodeur étudié utilise la méthode de détection de passage par un seuil fixe S . La fig V-10 montre le schéma de principe de ce détecteur: A partir du signal de la parole $x(t)$ il donne un signal $F_o(t)$ proportionnel à la fréquence de

pitch en suivant les étapes suivantes :

- Echantillonnage en fréquence du spectre $x(t)$ par n filtres passe-bas montés en parallèle.
- Normalisation des amplitudes.
- Extraction de la fréquence fondamentale par algorithme.
- Corrections.

D'après des constatations expérimentales, les conditions de fonctionnement du détecteur sont les suivantes :

- Gamme de fréquences fondamentales à détecter de 70 à 500Hz.
- Configuration des harmoniques de la 2^e harmonique dans le cas le plus défavorable : l'amplitude de la 2^e harmonique est supérieure de 12dB à celle du fondamental.
- Fonctionnement automatique (sans réglage manuel).

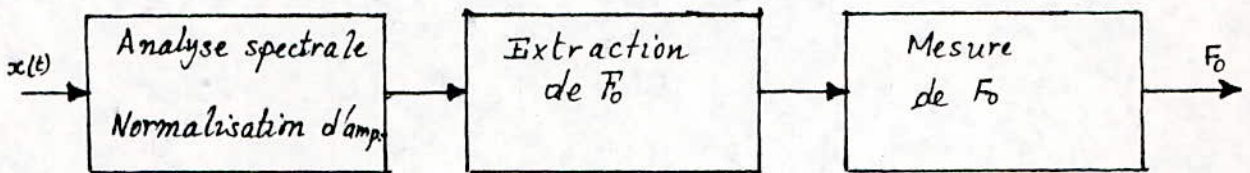


Fig V-10

V-2-4+2- CHOIX DES FILTRES

Les filtres sont disposés de sorte que la sortie d'au moins un filtre corresponde à F_0 si celle-ci existe.

Les caractéristiques de ces filtres (fréquence de coupure, courbe de réponse) sont disposées de façon que le fondamental du signal ait au moins un niveau de 15dB supérieur à la 2^e harmonique et ce quelque soit le niveau d'entrée.

La fig-V-II représente la réponse en amplitude d'un filtre passe-bas, f_{01} est la fréquence de coupure et f_{21} est la fréquence minimale d'atténuation de 30dB.

Examinons le cas le plus défavorable, soit f_{11} la fréquence fondamentale, f_{21} celle de la 2^e harmonique (l'amplitude de f_{21} est de 12dB supérieure à celle de f_{11}). On veut qu'à la sortie,

l'amplitude de la fondamentale soit d'au moins 15dB supérieure à celle de f_{21}). L'atténuation de f_{21} par rapport à f_{11} doit être donc de $15+12 = 27$ dB. On adoptera la valeur 30dB pour tenir compte des variations possible dans la bande passante (-3dB).

Conformément aux conditions imposées, f_1 représente la plus basse fréquence détectable sur le canal I et f_{c1} la plus haute fréquence. La gamme utile sera donc comprise entre f_{11} et f_{c1} (f_{11} correspond au cas le plus défavorable).

Pour des fréquences fondamentales supérieures à f_{c1} , un autre filtre passe-bas doit relayer le 1^{er} pour une extraction correcte. Pour que les filtres soient contigus, on doit avoir:

$$f_{12} = f_{c1}$$

De plus on veut que les filtres aient la même configuration de pôles et de zéros, donc:

$$\frac{f_{22}}{f_{c2}} = \frac{f_{21}}{f_{c1}} = k$$

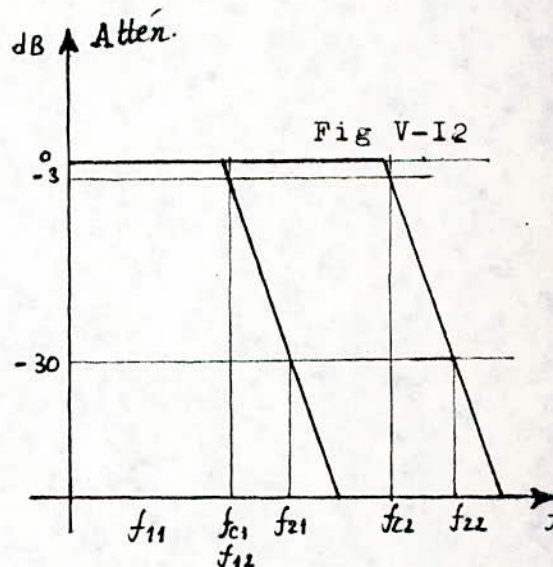
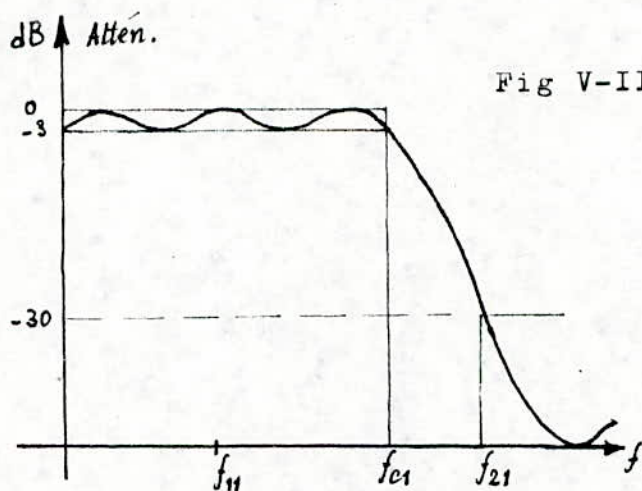
mais on a:

$$f_{21} = 2f_{11} \quad \text{et} \quad f_{22} = 2f_{12}$$

on a alors:

$$\frac{f_{22}}{f_{c2}} = \frac{2f_{12}}{f_{c2}} = \frac{2f_{11}}{f_{c1}} = k$$

La fig V-12 montre la disposition des 2 courbes d'amplitude du 1^{er} et 2^e canal.



En considérant tous les filtres jusqu'à n on aura:

$$\frac{f_{11}}{f_{cn}} \cdot 2^n = k^n$$

$$k = 2 \cdot \exp\left(-\frac{1}{n} \text{Log} \frac{f_{11}}{f_{cn}}\right)$$

avec $f_{11} = 70$ Hz , $f_{cn} = 500$ Hz.

Le nombre k constitue une mesure de performance des filtres et de leur complexité: plus le nombre k diminue, plus le filtre calculé se rapproche du filtre idéal ($k = 1$) et plus le nombre de canaux diminue (le nombre minimum est de 3 canaux).

Choix du type du filtre:

A titre de comparaison, considérons 3 types de filtres: Le filtre de Butterworth, le filtre de Tchebycheff et le filtre Elliptique. Des remarques précédentes 3 critères s'imposent:

- La valeur de k doit être aussi voisine de 1 que possible, donc un nombre minimum de canaux.

- On peut évaluer la complexité des circuits par:
 $C = \text{Nombre de canaux} \times \text{Degré de la fonction de transfert.}$
 C doit être minimal.

Ces deux critères nous conduisent aux solutions suivantes:

3 canaux ($k = 1,03$)	Elliptique 6 ^e degré
	Tchebycheff 15 ^e degré
	Butterworth 90 ^e degré
4 canaux ($k = 1,223$)	Elliptique 6 ^e degré
	Tchebycheff 7 ^e degré
	Butterworth 16 ^e degré

Les trois meilleures solutions sont:

3 canaux ($k = 1,03$) Elliptique 6^e degré.
4 canaux ($k = 1,223$) Elliptique 4^e degré.
4 canaux ($k = 1,223$) Tchebycheff 7^e degré.

- Un troisième critère s'impose: C'est la courbe de réponse. En effet la courbe de réponse en amplitude du filtre de Tchebychev présente l'avantage sur le filtre Elliptique d'avoir une atténuation croissante avec la fréquence tandis que la courbe de réponse du filtre elliptique remonte à un niveau de 30dB pour des fréquences supérieures à f_c bien qu'elle présente une chute plus rapide. Ceci viole la 2^e condition de fonctionnement du filtre à savoir que l'atténuation de f_2 par rapport à f_1 soit supérieure strictement à 30dB.

Ce dernier critère nous conduit à choisir définitivement la configuration : 4 canaux, Tchebycheff 7^e degré.

La fonction de transfert de chaque filtre est donnée par:

$$T(p) = \frac{1}{(p^2 + 0,0566 + 0,9662j)(p^2 + 0,15604p + 0,627)} \times \frac{1}{(p^2 + 0,2276p + 0,2041j)(p + 0,1263)}$$

Chaque canal peut-être réalisé par mise en cascade de 3 cellules passe-bas du 2^e ordre (fig V-6) et une cellule passe-bas du 1^{er} ordre (fig V-7).

On a représenté respectivement aux figures (V-I4) et (V-I5) la réponse en amplitude et le temps de propagation en fonction de la fréquence pour les 4 canaux.

V-2-4-3- NORMALISATION DES AMPLITUDES

Pour être appliqués à un microprocesseur, les signaux doivent avoir des valeurs normalisées. Les signaux de chacun des canaux subissent une mise en forme qui leur fait passer de la forme sinusoidale de niveau quelconque à une forme rectangulaire de niveau constant (5V).

Pour ce faire, on applique ces signaux à des comparateurs. Chaque comparateur bascule chaque fois qu'un certain seuil S est franchi. S est choisi de manière qu'il reste toujours supérieur au niveau du 2^e harmonique qui est lui-même inférieur de 30dB par rapport au niveau du fondamental.

La fig (V-I3) en donne le schéma de principe.

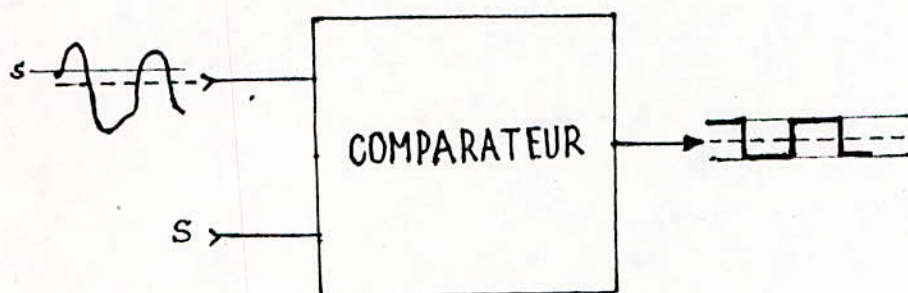


Fig V-I3

On dispose ainsi de 4 signaux rectangulaires à l'entrée du microprocesseur.

V-2-4-4- ALGORITHME DE MESURE ET D'EXTRACTION

L'algorithme a pour tâche de mesurer la période dans chaque canal par détection de deux transitions successives - la période étant le double de la durée s'écoulant entre deux passages par le seuil S- et de choisir parmi ces quatre mesures la valeur convenable. L'inversion de celle-ci donne la fréquence du fonda-

mental. Les étapes à suivre sont les suivantes :

-On cherche la 1^{ère} transition (changement d'état) éventuelle pour chaque canal. Un compteur étant mis à compter, si l'on détecte une transition on relève le nombre inscrit dans le compteur et on mémorise cette nouvelle valeur, soit C1 cette valeur.

-On cherche la 2^{ème} transition pour les signaux ayant subi une 1^{ère} transition. On relève la valeur actuelle du compteur, soit C2 cette valeur. On calcule la période qui n'est autre que

$$T_0 = (C2 - C1)T_c \times 2 = N_0 T_c$$

où T_c est la période d'incrémentatation du compteur.

On dispose aussi dans chaque canal d'une période maximale T_M .

Si $T_0 > T_M$ on rejette T_0 et on attribue à celle-ci la valeur 0

Si $T_0 < T_M$ T_0 est acceptée.

-On reboucle le programme sur lui-même pour tourner indéfiniment.

Remarques :

-La précision est d'autant meilleure que la période de comptage est courte. Mais cela entraîne un codage s'étendant sur un grand nombre de bits. Pour avoir un bon compromis entre la précision et le nombre de bits de codage, on a choisi comme période de comptage la durée du déroulement de tout le programme, soit 1mS.

La plus faible période se trouve représentée ainsi par : $N_{\min} = 1$

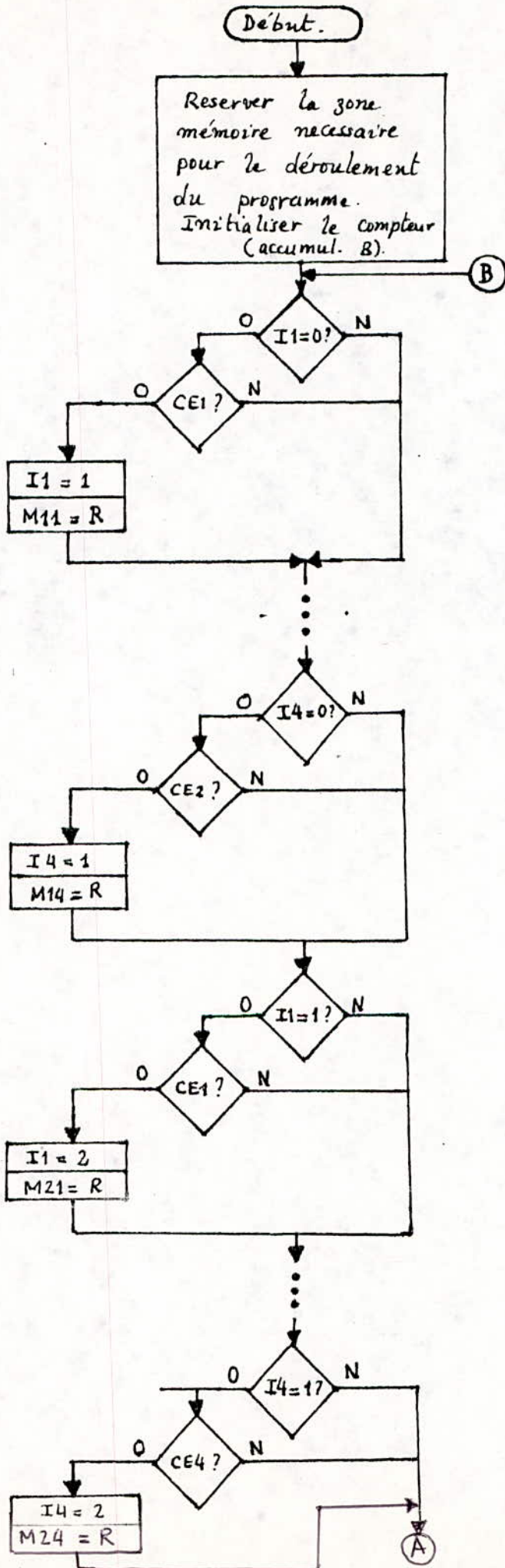
La plus forte par : $N_{\max} = 15$

-Pour identifier les différentes transitions, on a attribué à chaque canal un indice I. I étant initialement à zéro. Dès la détection de la 1^{ère} transition I se met à 1. Après la 2^{ème} transition, il se met à 2, on calcule la période et on le remet à 0. Ainsi on ne cherche la 2^e transition que si I=1 et on ne calcule la période que si I=2.

-La dernière étape consiste à choisir parmi les 4 périodes calculées la vraie valeur de la période. Le programme choisit comme période instantannée celle non nulle correspondant au filtre de fréquence de coupure la plus basse.

Si la période est nulle, le son est non voisé.

La fréquence instantannée s'obtient par inversion de la période instantannée.



I1, I2, I3 et I4 indiquent respectivement les transitions possibles des signaux issus des canaux 1, 2, 3 et 4.

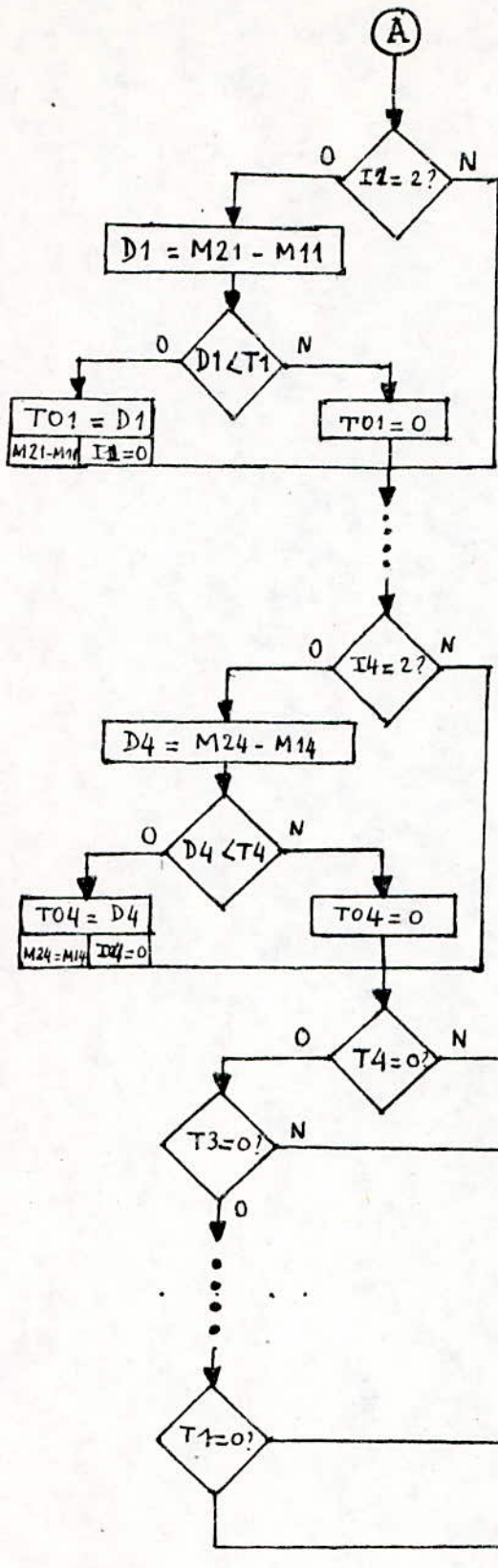
I1 = 0 initialisation après la 2^e transition.

I1 = 1 première transition.

I1 = 2 deuxième "

CE1? : tester si le signal 1 a changé d'état (haut → bas ou bas → haut).

R : valeur instantanée du Compteur.



T1: valeur maximale de la période admissible dans le canal 1 et pas dans le canal 2.

T0: valeur instantanée de la période fondamentale.

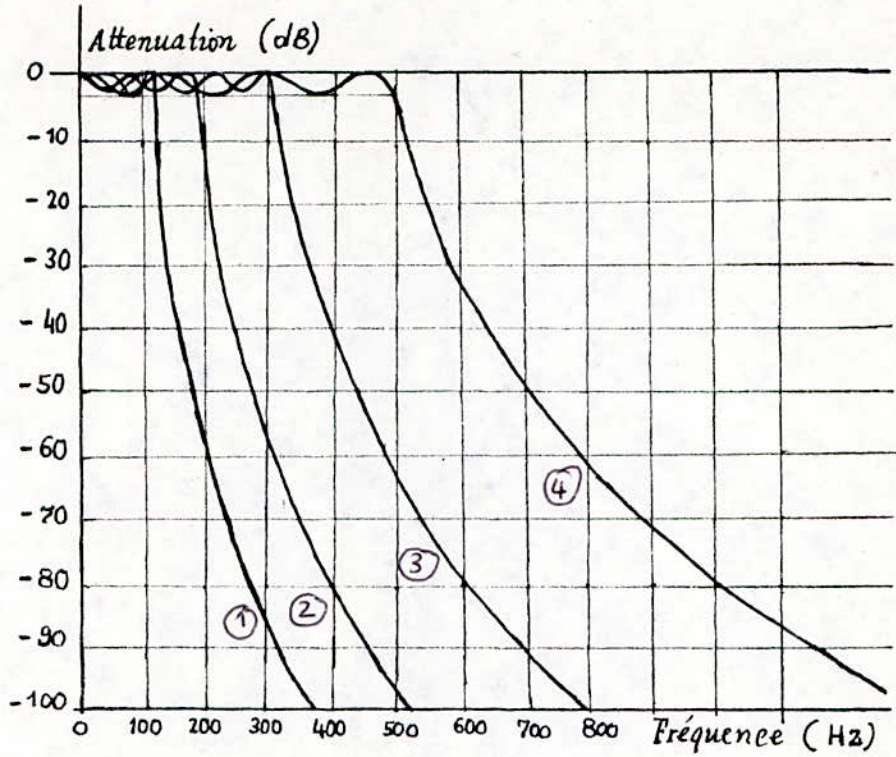


Fig V-I4 Courbes de réponses en amplitude des 4 canaux.

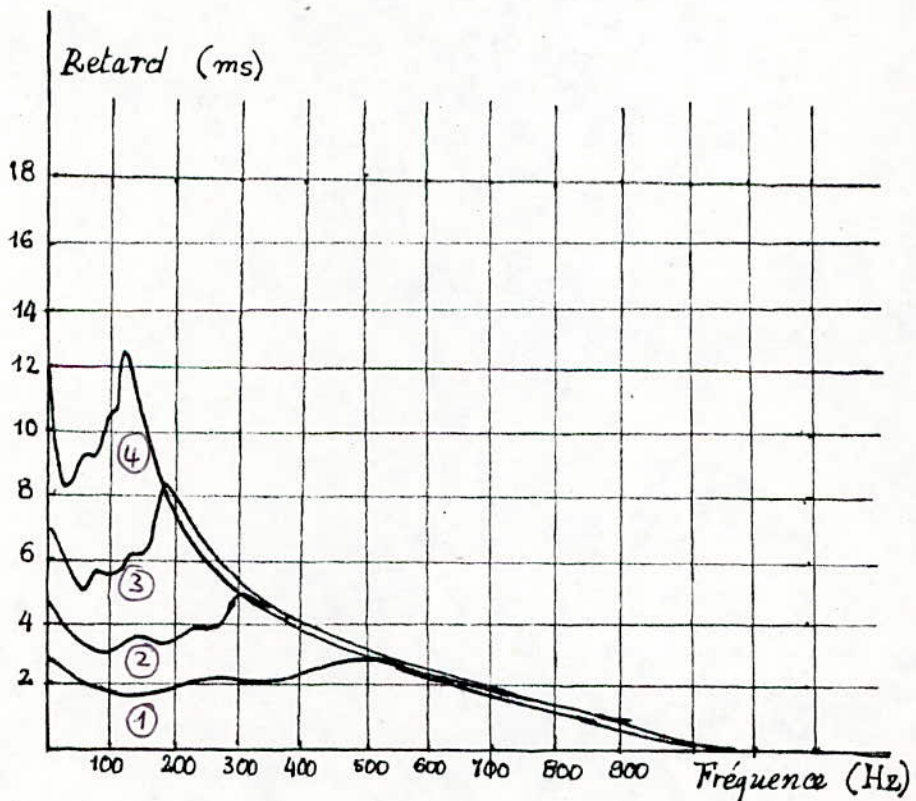


Fig V-I5 Courbes représentatives du retard en fonction de la fréquence pour les 4 canaux.

V-2-4-5- CORRECTIONS

L'analyseur de mélodie est conçu pour travailler parfaitement en temps réel dans la gamme de 70 à 500Hz. Or cette hypothèse n'est vérifiée que si le bruit de fond de l'analyseur et les temps de retard des filtres étaient nuls.

En pratique, ces deux conditions ne sont pas satisfaites. En conséquence, plusieurs défauts interviennent:

-Les tremblements: ils sont dus au bruit de fond et ils sont d'autant plus importants que le signal ait une amplitude voisine du seuil S. Pour pallier à ce défaut, on a introduit un filtre passe-haut (Tchebycheff, 7^e ordre) ayant une fréquence de coupure de 70Hz.

L'introduction de ce filtre réduit les bandes passantes échantillonnées et par conséquent il y a diminution du bruit.

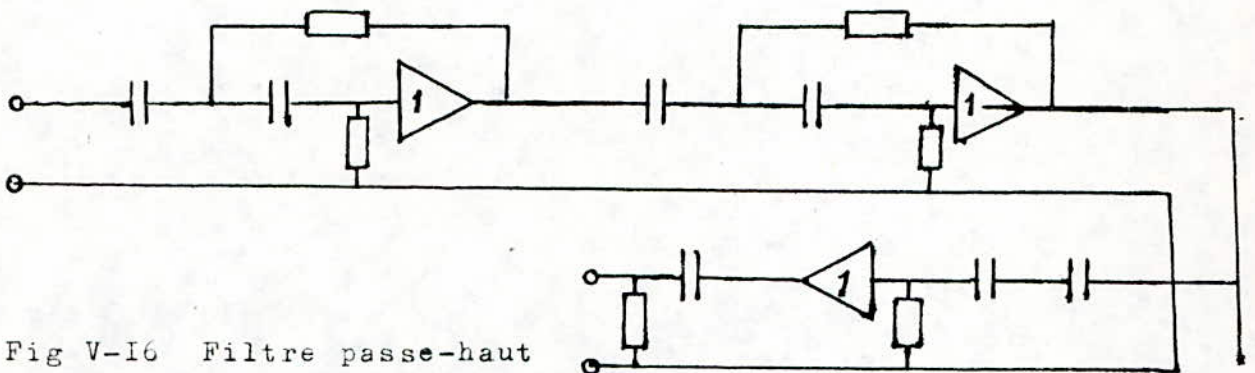


Fig V-16 Filtre passe-haut du 7^e ordre.

-La réponse aux transitoires: Ce défaut est dû principalement aux retards des filtres (fig V-15)

Pour remédier à ce défaut, on a proposé de retarder chaque canal d'une durée égale à son retard moyen dans la bande, et cela au moyen d'une ligne à retard.

Cette ligne à retard peut-être réalisée "matériellement" comme elle peut-être simulée "logiciellement" par programme. Cette dernière est plus souple (il suffit d'utiliser des boucle de temporisation) mais le programme d'extraction devient plus lourd (fig V-17).

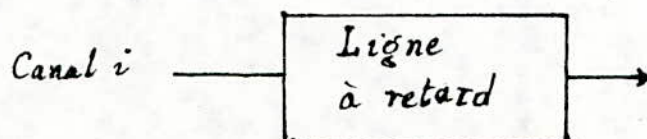


Fig V-17

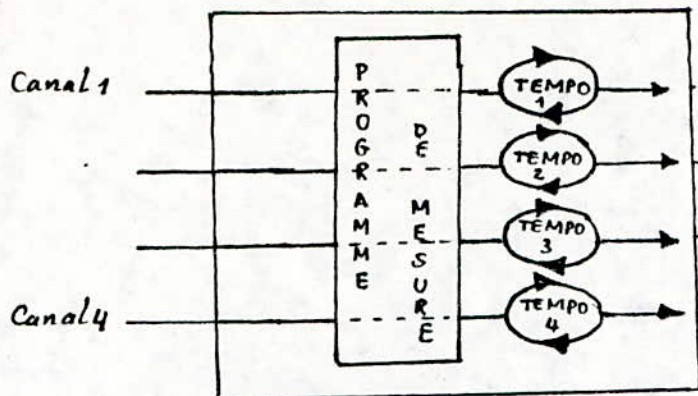


Fig V-17 bis

V-3- SYNTHETISEUR DU VOCODEUR

V-3-1- ELABORATION DU SIGNAL D'EXCITATION

Selon que le son à synthétiser est reconnu voisé ou non voisé le signal d'excitation est soit un signal périodique (séquences périodiques) soit un bruit blanc. Ces deux signaux sont commutés par la fonction de détection qui elle-même est interprétée par l'état d'un bit de sortie du détecteur de mélodie (bit de voisement).

Si ce bit est à 1, le son est voisé et c'est le générateur d'impulsions qui est sélectionné.

Dans le cas contraire, le son est non voisé et c'est le générateur de bruit qui est actionné. (fig-V-18)

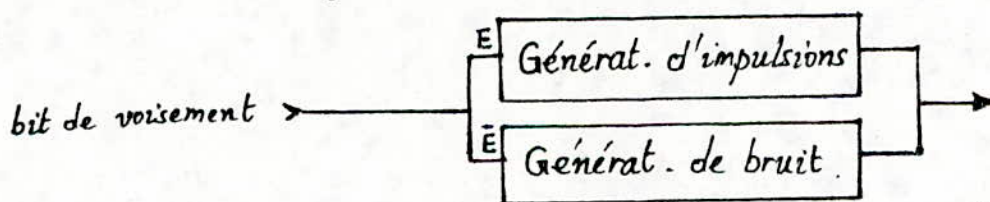


Fig V-18

V-3-1-1- GENERATEUR DE SEQUENCES PERIODIQUES

Le signal d'excitation pour les sons voisés peut-être un signal identique au signal glottique (triangulaire) ou bien une séquence d'impulsions pseudo-aléatoires.

Nous avons choisi la deuxième, du fait qu'elle a la même propriété que la première et sa puissance est plus importante d'un rapport d'environ 30, d'où une amélioration importante du rapport signal sur bruit.

La séquence est obtenue à l'aide d'un registre à décalage rebouclé sur lui-même. La période d'horloge du registre est de $\frac{T_0}{127}$ où T_0 représente la valeur du fondamental du dernier échantillon transmis. La séquence comprend alors 2^6 "1" et $2^6 - 1$ "0" répartis aléatoirement.

V-3-1-2- GENERATEUR DE BRUIT

Dans le cas des sons non voisés, l'excitation est un bruit blanc: c'est à dire un signal dont le spectre est plat.

V-3-2- MODULATION DES SIGNAUX

Le signal d'excitation est modulé en amplitude par chacun des signaux dérivant de chacun des canaux d'analyse.

Nous avons proposé un modulateur en anneau (fig V-19) parcequ'il présente plusieurs avantages.

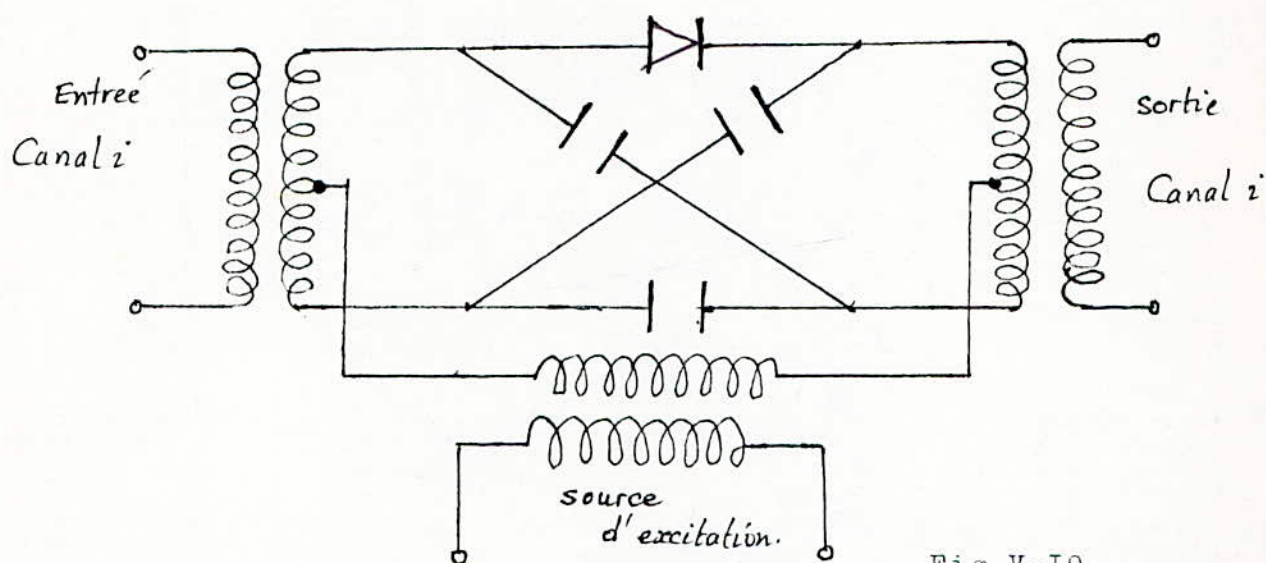


Fig V-19

V-3-3- SOMMATION DES SIGNAUX

Le signal de la parole synthétique est obtenu par addition de tous les signaux modulés.

Cette opération peut-être réalisée par le montage de la (fig V-20).

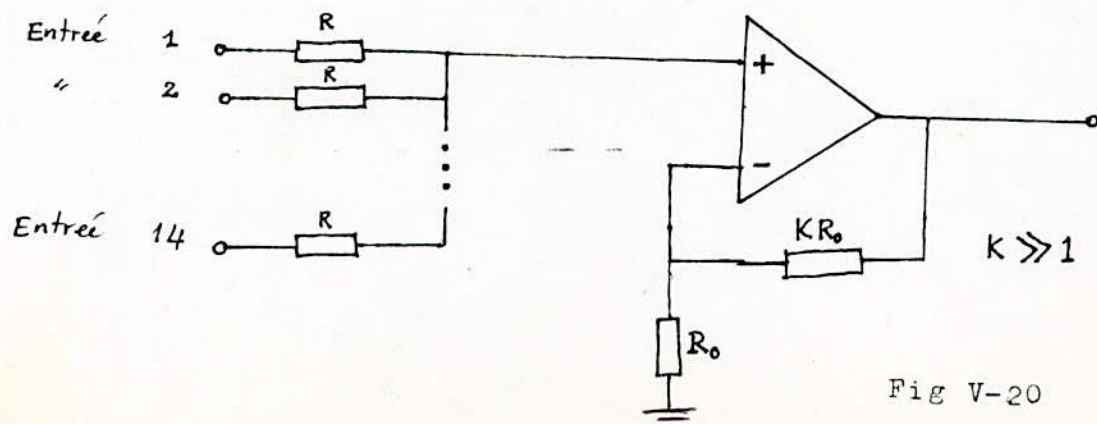


Fig V-20

V-4- PERFORMANCES DU VOCODEUR ETUDIE

-Cette version de vocodeur s'accomode bien à la bande téléphonique (300 - 3300Hz).

-Etant donné que la fréquence d'échantillonnage est de 50Hz, que le nombre de canaux est de 14 et que F_0 est codée sur 8 bits, le débit binaire est de $(14 \times 4 + 8) \times 50 = 3100 \text{ bits/S}$. Ce vocodeur réalise ainsi une transmission à faible débit (6400 bits/S pour la téléphonie normale).

-D'autre part, un test de comparaison entre ce vocodeur à canaux et un vocodeur à prédiction linéaire 9000bits/S. a donné les résultats suivants:

Préférence pour le vocodeur à canaux	4 ,4/
Préférence pour le vocodeur à prédiction linéaire.	39 /
Pas de préférence	12,6/

Ces résultats confirment la supériorité de cette version sur les anciennes et même sur le vocodeur à prédiction linéaire

V-5- SCHEMA GLOBAL

La figure (V.-7) représente le schéma détaillé global de la version du vocodeur que nous avons étudié.

3^e Partie

« Les fleurs de demain sont
Contenues dans les semences
d'aujourd'hui. »

- PROVERBE CHINOIS -

CHAPITRE VI

REALISATION D'UN DETECTEUR DE MELODIE

VI-I- INTRODUCTION

Nous nous proposons de décrire - dans cette partie - la réalisation d'un détecteur de mélodie conformément à l'étude théorique précédente.

Le schéma détaillé du détecteur est donné à la fin de cette partie fig(VI-7). Nous allons calculer les éléments de ce montage et établir le programme approprié pour l'extraction de la mélodie.

VI-2- CALCUL ET CHOIX DES ELEMENTS

VI-2-I- LES FILTRES PASSE-BAS

Chaque canal est constitué d'un filtre passe-bas de Tchebycheff du 7^e ordre. Ces filtres se recoupent à -3dB et sont disposés de manière que:

$$\frac{f_{i2}}{f_{ci}} = k \quad \text{où } k=1,233.$$

comme $f_{i2} = 2f_{(i-1)}$

et $f_{11} = 70\text{Hz}$.

On peut tirer les fréquences de coupure:

$$f_{c1} = 117 \text{ Hz}$$

$$f_{c2} = 191 \text{ Hz}$$

$$f_{c3} = 321 \text{ Hz}$$

$$f_{c4} = 539 \text{ Hz}$$

La fonction de transfert de chaque canal peut-être mise sous la forme:

$$T(p) = \frac{1}{(p^2 + 0,0566p + 0,966)(p^2 + 0,15604p + 0,627)}$$
$$\times \frac{1}{(p^2 + 0,227p + 0,2041)(p + 0,1263)}$$

On voit que le filtre peut-être réalisé à l'aide de 3 filtres passe-bas du 2^e ordre et un filtre passe-bas du 1^e ordre placés en cascade et dont les fonctions de transfert sont respectivement:

$$T_1(p) = \frac{1}{1 + 0,05 p + 1,035 p^2}$$

$$T_2(p) = \frac{1}{1 + 0,2529 p + 1,594 p^2}$$

$$T_3(p) = \frac{1}{1 + 0,9419 p + 4,149 p^2}$$

$$T_4(p) = \frac{1}{1 + 7,917 p}$$

où $p = j \frac{\omega}{\omega_{ci}}$

Les filtres du 2^e ordre peuvent être réalisés à l'aide du montage de la fig(VI-1) et celui du 1^e ordre à l'aide du montage de la fig(VI-2).

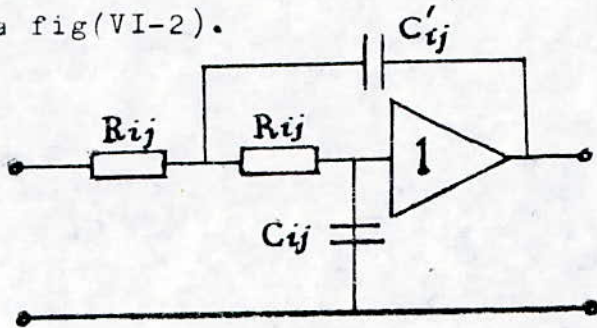


Fig VI-1

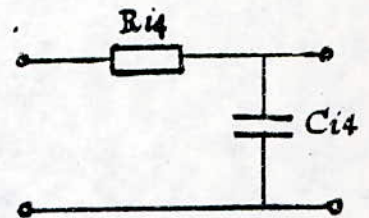


Fig VI-2

Ces deux montages ont pour fonctions de transfert respectivement:

$$T_{ij}(\omega) = \frac{1}{1 + 2R_{ij}C_{ij}p + R_{ij}^2 C_{ij} C'_{ij} p^2} \quad \text{pour } j=1,2,3.$$

$$T_{ij}(\omega) = \frac{1}{1 + R_{ij} C_{ij} p} \quad \text{pour } j=4.$$

où i désigne le rang du canal, et j le rang de l'étage dans chaque canal.

Identifions T(p) et T(ω):

On aura donc: pour j=1 $\frac{C_{i1}}{C_{i1}} = \frac{4 \times 1,035}{\pi \times (0,05)^2}$; $R_{i1} = \frac{0,05}{4\pi f C_{i1}}$

j=2 $\frac{C_{i2}}{C_{i2}} = \frac{4 \times 1,594}{\pi \times (0,2529)^2}$; $R_{i2} = \frac{0,2529}{4\pi f C_{i2}}$

j=3 $\frac{C_{i3}}{C_{i3}} = \frac{4 \times 4,149}{\pi \times (0,9419)^2}$; $R_{i3} = \frac{0,9419}{4\pi f C_{i3}}$

$$j=4, \quad R_{i4} = \frac{7,917}{2\pi f_{ci} C_{i4}}$$

Il suffit donc de choisir la valeur d'une capacité et de calculer les valeurs des autres composants.

Pour les amplificateurs à gain unité on a choisi un montage suiveur réalisé à l'aide d'un ampli Op. du type μ A741 dans lequel on a relié l'entrée inverseuse à la sortie.

La fig(VI-3) illustre le sens de brochage du μ A741.

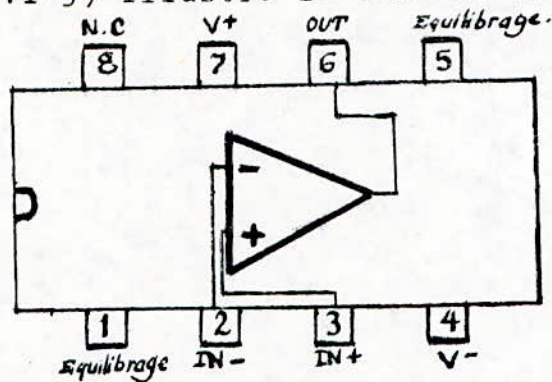


Fig VI-3

Ses principales caractéristiques sont les suivantes:

- Résistance d'entrée : 1M Ω .
- Résistance de sortie: quelques Ω .
- Fréquence de transition : 1MHz.
- Alimentation : jusqu'à $\pm 15V$.

VI-2-I-I- CALCUL DU 1^e CANAL (i=1)

On a $f_{c1} = 117Hz$.

1^e étage (j=1), choisissons $C_{11} = 1nF$

On aura $\bar{C}_{11} = 1,2 F$ et $R_{11} = 39k$

2^e étage (j=2), Choisissons $C_{12} = 3nF$

On aura $\bar{C}_{12} = 99nF$ et $R_{12} = 56k$

3^e étage (j=3), Choisissons $C_{13} = 10nF$

On aura $\bar{C}_{13} = 157,6nF$ et $R_{13} = 2k$

4^e étage (j=4), Choisissons $C_{14} = 47nF$

On aura $\bar{R}_{14} = 229k$

VI-2-I-2- CALCUL DU 2^e CANAL (i=2)

On a $f_{c2} = 191Hz$

1^e étage (j=1), Choisissons $C_{21} = 1nF$

On aura $\bar{C}_{21} = 1,2 F$ et $R_{21} = 24,4k$

2^e étage (j=2), Choisissons $C_{22} = 3nF$

On aura $\bar{C}_{22} = 299nF$ et $R_{22} = 35,1k$

3^e étage (j=3), Choisissons $C_{23} = 10nF$

On aura $\bar{C}_{23} = 157,6nF$ et $R_{23} = 46,5k$

4^e étage (j=4), Choisissons C24 = 47nF

On aura R24 = 140k

VI-2-I-3- CALCUL DU 3^e CANAL (i=3)

On a $f_{c3} = 321\text{Hz}$

1^e étage (j=1), Choisissons C31 = 1nF

On aura $C_{\bar{3}1} = 1,2 \text{ F}$ et R31 = 14,5k

2^e étage (j=2), Choisissons C32 = 3nF

On aura $C_{\bar{3}2} = 299\text{nF}$ et R32 = 20,9k

3^e étage (j=3), Choisissons C33 = 10nF

On aura $C_{\bar{3}3} = 157,6\text{nF}$ et R33 = 27,6k

4^e étage (j=4), Choisissons C34 = 47nF

On aura R34 = 3,5k

VI-2-I-4- CALCUL DU 4^e CANAL (i=4)

On a $f_{c4} = 539\text{Hz}$

1^e étage (j=1), Choisissons C41 = 1nF

On aura $C_{\bar{4}1} = 1,2 \text{ F}$ et R41 = 6k

2^e étage (j=2), Choisissons C42 = 3nF

On aura $C_{\bar{4}2} = 299 \text{ F}$ et R42 = 12,5k

3^e étage (j=3), Choisissons C43 = 10nF

On aura $C_{\bar{4}3} = 157,6 \text{ F}$ et R43 = 16,5k

4^e étage (j=4), Choisissons C44 = 47nF

On aura R44 = 49,7k

VI-2-2- LES FILTRES PASSE-HAUT

La fonction de transfert du filtre passe-haut s'obtient à partir de celle du filtre passe-bas de même fréquence de coupure par la transposition :

$$p \text{ ----- } \frac{1}{p}$$

On obtient:

$$T(p) = \frac{p^2}{(1+0,0566p+0,966p^2)} \times \frac{p^2}{(1+0,15p+0,627p^2)}$$

$$\times \frac{p^2}{(1+0,227p+0,2041p^2)} \times \frac{p}{(1+0,263p)}$$

Ce filtre peut-être réalisé par mise en cascade de trois filtres passe-haut du 2^e ordre et d'un filtre passe-haut du 1^e ordre fig(VI46).

Ces montages ont pour fonctions de transfert respectivement:

$$T_{ij} = \frac{R_i R_i C_i^2 (j\omega)^2}{1 + 2R_i C_i (j\omega) + R_i R_i C_i^2 (j\omega)^2} \quad \text{pour } j=1,2,3.$$

$$T_{ij} = \frac{C_i R_i (j\omega)}{1 + C_i R_i (j\omega)} \quad \text{pour } j=4$$

où j désigne le rang de l'étage.

En identifiant les fonctions de transfert, on obtient:

Pour j=1,	$\frac{R_1}{R_1} = \frac{4 \times 0,9662}{(0,0566)^2}$	et	$R_1 = \frac{0,0566}{4 \times 70 \times C_1}$
Pour j=2,	$\frac{R_2}{R_2} = \frac{4 \times 0,627}{(0,15604)^2}$	et	$R_2 = \frac{0,15604}{4 \times 70 \times C_2}$
Pour j=3,	$\frac{R_3}{R_3} = \frac{4 \times 0,2041}{(0,2276)^2}$	et	$R_3 = \frac{0,2276}{4 \times 70 \times C_3}$
Pour j=4,	$R_4 = \frac{0,1263}{2 \times 70 \times C_4}$		

On peut tirer les valeurs suivantes:

1^e étage; Choisissons C1 = 330nF

On aura $R_1 = 194,9 \text{ ohms.}$ et $R_1 = 23^r, 2k.$

2^e étage; Choisissons C2 = 330nF

On aura $R_2 = 516 \text{ ohms.}$ et $R_2 = 54,4k.$

3^e étage; Choisissons C3 = 330nF

On aura $R_3 = 74 \text{ ohms.}$ et $R_3 = 12,4k.$

4^e étage; Choisissons C4 = 330nF

On aura $R_4 = 70 \text{ ohms.}$

VI-2-3- LES COMPARATEURS

VI-2-3-I- CALCUL DU SEUIL S

Le seuil doit être légèrement supérieur à l'amplitude du deuxième harmonique. Il est donc de 15dB inférieur à l'amplitude du fondamental. On adoptera S = 0,2V

VI-2-3-2- CALCUL ET CHOIX DES COMPOSANTS

Pour les comparateurs, on a choisi un quadruple comparateur du type LM339N. dont les principales caractéristiques sont les suivantes:

Tension de polarisation : $\pm 15V.$

Tension de décalage à l'entrée : 50mV.

Temps de montée : 1300nS.

La fig(VI-6) illustre son schema de brochage :

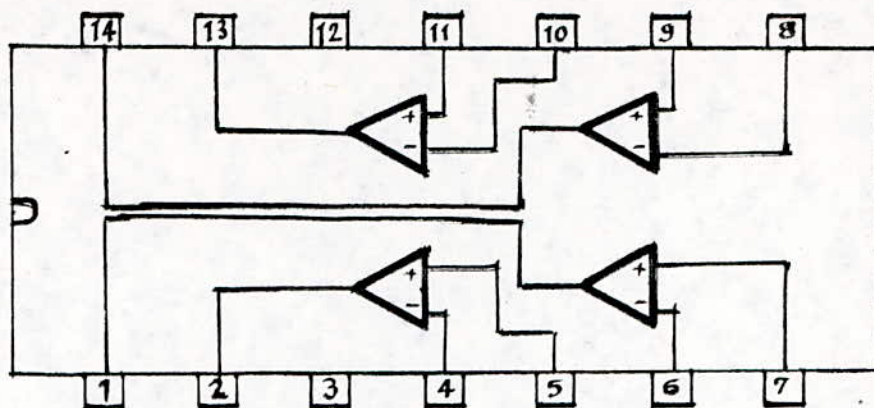


Fig V-6

On a adopté $\pm 15V$ pour l'alimentation de toute la maquette.
 Pour obtenir le seuil S on utilise un pont diviseur.

A la sortie des comparateurs, la tension doit-être normalisée à 15V, on utilise également des ponts diviseurs.

VI-3- PROGRAMME D'EXTRACTION DE LA MELODIE

On a choisi d'exécuter le programme sur un kit de MOTOROLA du type 6800. Ce dernier nous permet de bien suivre l'évolution du signal ainsi qu'il s'adapte bien aux calculs précis de temporisations. L'interface utilisé avec le microprocesseur 6800 et notre périphérique est le PIA 6821.

Langage assembleur

Commentaires

CLR A

CLR B

CLR \$31

CLR \$32

CLR \$33

CLR \$34

LDAA #2

STAA \$31

LDAA #3

STAA \$32

LDAA #5

STAA \$33

LDAA #9

STAA \$34

CLR \$E4 1

CLR \$E4 3

Initialisation préliminaire.

On écrit les nombres relatifs aux périodes maximale dans les canaux 1, 2, 3, et 4

```
CLR $E4 0
LDAA #$FF
STAA $E4 2
LDAA #$04
STAA $E4 1
STAA $E4 3
LDAA $E4 0
STAA $0000

LPO LDX #$04
LP1 LDAA $00
   ANDA #$01
   STAA $06,X
   ROR $00
   DEX
   BNE LP1
   LDX #$04
LP30 LDAA $01,X
     CMPA #$00
     BNE LP20
     LDAA $06,X
     CMPA $11,X
     BNE LP20
     LDAA $06,X
     STAA $11,X
     INC $01,X
     STAB $16,X

LP20 DEX
     BNE LP30
     LDX #$04
LP31 LDAA $01,X
     CMPA #$01
     BNE LP21
     LDAA $06,X
     CMPA $11,X
     BNE LP21
     LDAA $06,X
     STAA $11,X
     INC $01,X
     STAB $21,X
```

Initialisation du PIA

Le port A en entrée.

Recherche de la première transition

Recherche de la deuxième transtion.

LP21	DEX	
	BNE LP31	
	LDX /\$04	
LP42	LDAA /\$01,X	Calcul des périodes dans les 4 canaux.
	CMPA /\$02	
	BNE LP22	
	LDAA /\$21,X	
	SUBA /\$16,X	
	STAA /\$26,X	
	CMPA /\$31,X	
	BLT LP32	
	CLR /\$26,X	
LP32	LDAA /\$21,X	
	STAA /\$16,X	
	CLR /\$01,X	
LP22	DEX	
	BNE LP42	
	LDX /\$04	
	LDAA /\$26,X	Calcul de la période instantanée.
	CMPA /\$00	
	BEQ LP23	La période instantanée est contenue
	STAA /\$40	dans l'adresse /\$40 .
	JMP LP33	
LP23	DEX	
	BNE LP43	
LP33	INCB	
	CMPB /\$FF	
	BNE LPO	Le programme est rebouclé sur lui-même
	CLRB	
	BRA LPO	

ENTRÉE

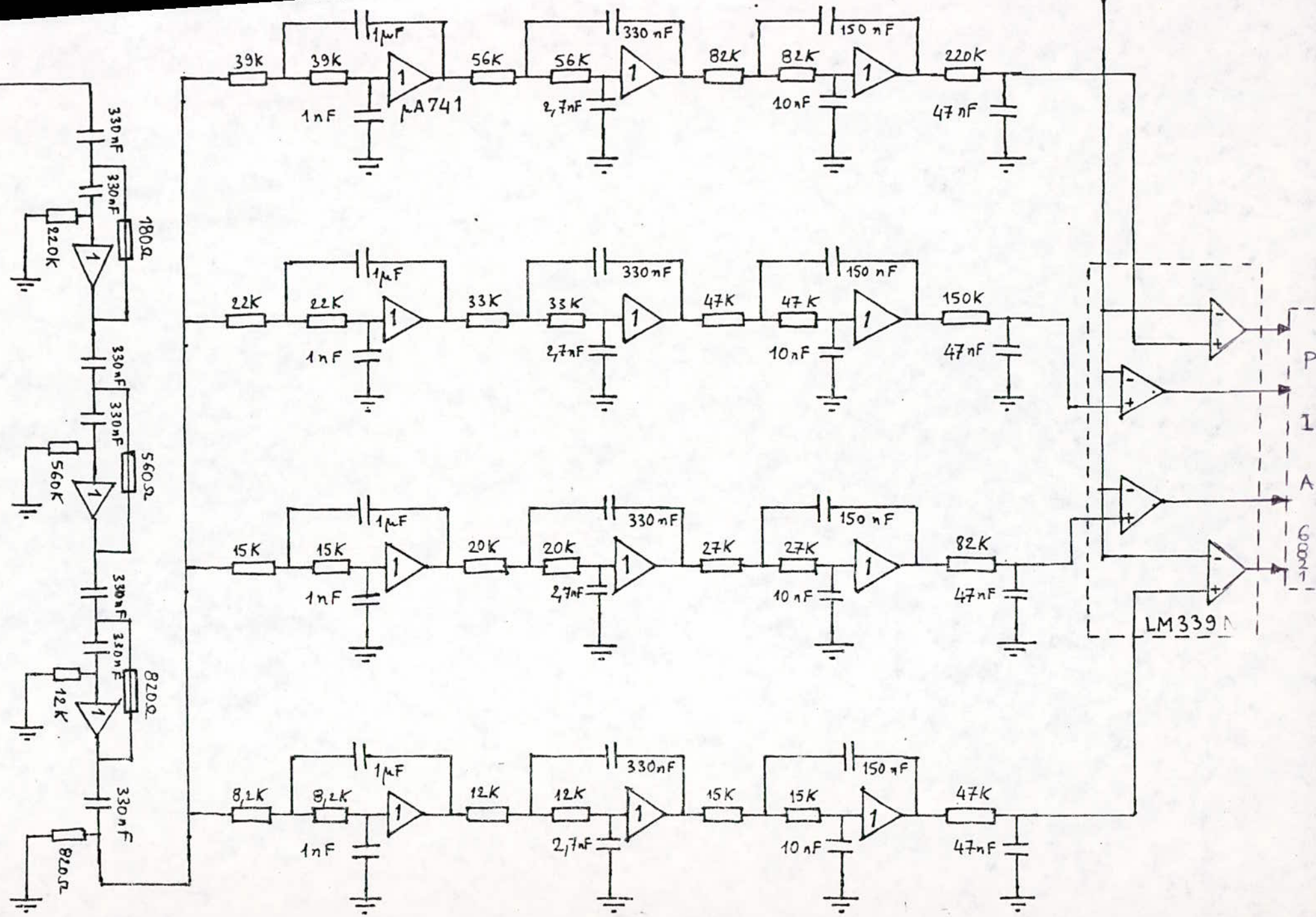


Fig VI - 7

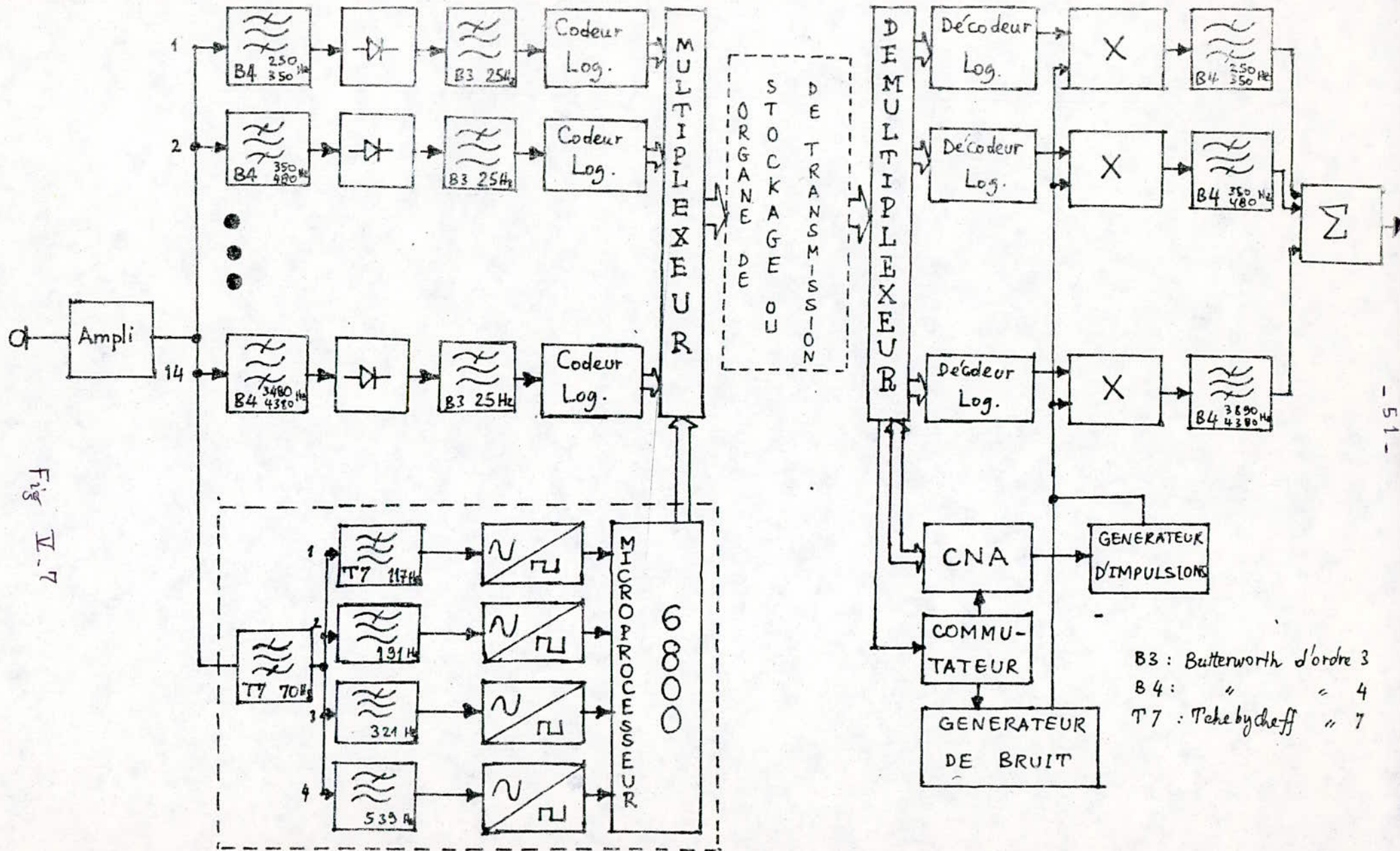
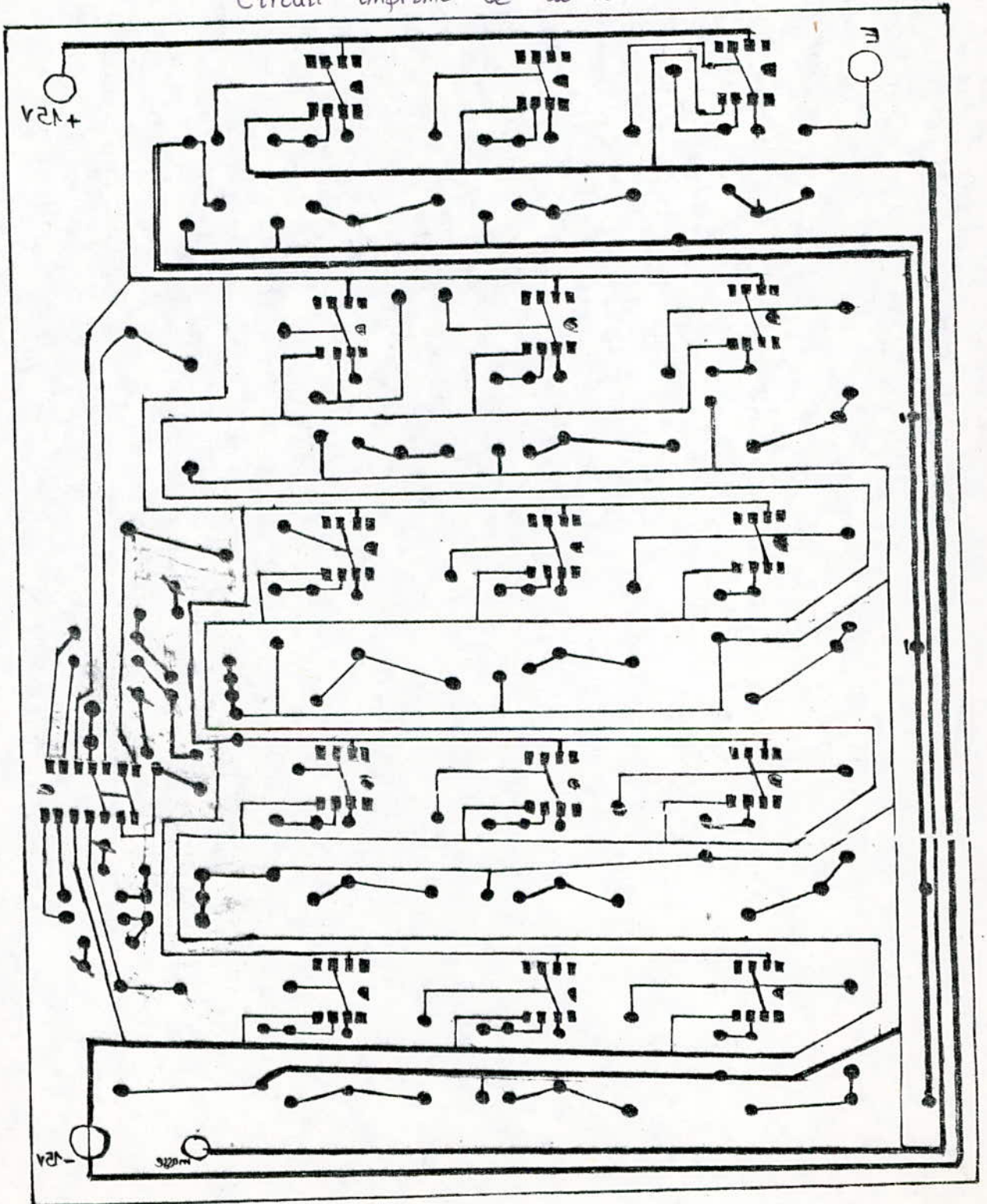


Fig. V.7

B3 : Butterworth d'ordre 3
 B4 : " " " 4
 T7 : Tchebycheff " 7

Circuit imprimé de la réalisation.



CONCLUSION

Notre travail nous a permis d'ouvrir une fenêtre sur un domaine très vaste -domaine de la parole- et sur les différentes techniques utilisées dans ce domaine.

Il nous a permis aussi d'approfondir nos connaissances sur le traitement du signal.

En conclusion, nous pouvons dire que nous avons préparé l'édifice de la réalisation d'un vocodeur à canaux complet. L'état théorique représente la plate forme de cet édifice et la réalisation du détecteur de mélodie ses piliers. Il reste aux promotions à venir d'en terminer la construction.

LAIB A.

MANSOURI R.

B I B L I O G R A P H I E

- Les méthodes rapides de transformation du signal. LIFERMAN.
- Méthodes et techniques du traitement du signal. J.MAX.
- Traitement numérique du signal. BELLANGER.
- Schémas d'électronique. J.MORNAND.
- Pratique de l'électronique:Comparateurs, filtres actifs.
M.AUMIAUX
- La parole:Compréhension et synthèse par les ordinateurs
J.GUIBERT
- Les techniques de synthèse de la parole. R.DESCOUT.
- Les appareils de synthèse et leurs applications. R.CARRE et
J.PAILLE
- Proglégomènes à l'étude des structures intonatives.
PIERRE R.LEON et PHILIPPE MART

----- o ----- o -----