

5/88

الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم و البحث العلمي
Ministère de l'Enseignement et de la Recherche Scientifique

ECOLE NATIONALE POLYTECHNIQUE

المكتبة
BIBLIOTHEQUE
Ecole Nationale Polytechnique

DEPARTEMENT : ELECTRONIQUE

PROJET DE FIN D'ETUDES

SUJET

CONTRIBUTION A LA
REALISATION D'UN
ANALYSEUR A BARC
DE FILTRES

2 PLANCHES

Proposé par :

B. BOUSSEKSOU

Etudié par :

S. AZZOUK
A. YATAGHENE

Dirigé par :

B. BOUSSEKSOU

PROMOTION JAN. 88

المدسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE = المكتبة
Ecole Nationale Polytechnique

DEDICACES

A la mémoire de ma grand-mère

Je dedie ce modeste travail

A mon père

A ma mère

A toute ma famille.

Arezki YATAGHENE

je dedie ce modeste travail

A: mon père

A ma mère

A ma femme

A mes frères et sœurs

A mes beaux frères

A mes beaux parents

A mes tantes et leurs enfants

A toute ma famille

A tous mes ami(e)s

Slimane Azzouk

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

REMERCIEMENTS

Nos vifs remerciements vont à monsieur BOUSSEKSOU, pour son aide, ses conseils et l'intérêt qu'il a porté à notre projet.

Nous remercions sincèrement celui qui nous a aidés pour le tirage de cette thèse, celui qui nous a prêté son matériel, ainsi que tous ceux qui ont participés de près ou de loin à la mise en oeuvre de ce travail.

* * * * *
* * * * *
* * * * *
* * * * *
* * * * *

INTRODUCTION

<u>CHAPITRE I - GENERALITES SUR LE SIGNAL VOCAL</u>	1
I- LE SYSTEME VOCAL ET SON FONCTIONNEMENT	2
I-1. <i>Les sources des sons</i>	
II- STRUCTURE PHONETIQUE DE LA PAROLE	3
II-1. <i>Les phonèmes</i>	
II-2. <i>La notation phonétique</i>	4
II-3. <i>Les classes de phonèmes</i>	
A/ Les voyelles	
B/ Les fricatives	5
C/ Les plosives	
III- PERCEPTION DES SONS	6
1. L'intensité	
2. La fréquence	
3. Le temps	
 <u>CHAPITRE II- ANALYSE DE LA PAROLE</u>	
Introduction	7
LES TECHNIQUES D'ANALYSE DU SIGNAL VOCAL	8
1- <i>Analyse Temporelle</i>	
1-1. Densité de passage par zéro	
1-2. Codage des extrema	
2- <i>Analyse Spectrale</i>	
2-1. Transformée de fourier rapide	9
2-2. Vocodeur à canaux	10

2-3. Vocodeur à formants	10
3- Codage Homomorphique	
4- Codage Prédicatif	11

CHAPITRE III- LA RECONNAISSANCE DE LA PAROLE

Introduction	12
1- LA RECONNAISSANCE ANALYTIQUE	13
2- LA RECONNAISSANCE GLOBALE	14
2-1. La reconnaissance de mots isolés	
2-2. La reconnaissance de mots enchainés	15
2-3. Les systèmes mono et multilocuteur	

CHAPITRE IV- LE VOCODEUR A CANAUX

I- ELEMENTS THEORIQUE D'ANALYSE CEPSTRALE	16
II- PRINCIPE DE REALISATION D'UN VOCODEUR A CANAUX	19

CHAPITRE V - ETUDE ET REALISATION D'UN ANALYSEUR A BANC DE FILTRES

Introduction	20
1- LE PREAMPLI ET LA FREACCENTUATION	21
2- LES FILTRES PASSE-BANDE	23
a/ Choix du type de filtres	
b/ choix de l'echellonnement	25
3- LA DETECTION D'ENVELOPPE	27
4- LE LISSAGE	
5- ECHANTILLONNAGE-BLOCAGE	29
6- LE MULTIPLEXAGE	32

7- LA CONVERSION ANALOGIQUE-NUMERIQUE	33
a/ Codage de la parole	35
b/ Calcul du débit binaire du banc de filtre	
8- INTERFACAGE ET STOCKAGE DES DONNEES	36
- Organigramme	38
- Programme	39
9- PROBLEMES RENCONTRES LORS DE LA REALISATION	40
LISTE DES COMPOSANTS	42

SCHEMAS DU CIRCUIT IMPRIME

CONCLUSIONS GENERALES

ANNEXE I : Brochage des circuits intégrés

ANNEXE II : Formules mathématiques du traitement du signal

ANNEXE III : Applications de la reconnaissance de la parole

Bibliographie.

INTRODUCTION

Les moyens de communication mis à la disposition de l'homme sont nombreux, mais la parole demeure le mode le plus instinctif. Ainsi, après avoir éxaucer son rêve de voler dans les airs, voila que l'homme rêve à present de dialoguer avec la machine . Ce rêve est encore loin d'être réaliser , néanmoins des étapes non négligeable ont déjà été franchies, grace aux récent progrès de la micro-électronique et de l'informatique .

Pour que le dialogue soit possible avec la machine il faudrait que celle-ci puisse entendre les sons , puis qu'elle en saisisse le sens et enfin qu'elle puisse donner une réplique cohérente.

La difficulté du problème vient tout d'abord de la complexité et de la variabilité du signal vocal, qui peut se présenter sous un grand nombre de formes , selon le type de voix, l'intonation , la vitesse d'élocution, le rythme , le niveau sonore , le locuteur , etc ...

Cependant, si l'on se fixe un objectif limité, par exemple la reconnaissance d'une centaine de mots isolés pour un seul locuteur , on peut considérer que le problème est soluble .

Divers système ont été mis au point pour permettre l'analyse et la synthèse de la parole . Le vocodeur canaux est l'un d'entre eux . Son principal avantage est de réduire le débit binaire tout en gardant une intelligibilité à la parole .

Pour notre part , nous avons essayé d'apporter une modeste contribution à la réalisation d'un vocodeur à canaux . La réalisation de l'analyseur du vocodeur a été entamer le semestre précédent avec le détecteur de mélodie. Le banc de filtres constitue donc la 2^{ème} partie de l'analyseur. Au cours de notre étude nous passerons en revu quelques techniques d'analyse du signal vocal, non sans avoir au préalable définie quelques unes de ces caractéristiques. Le vocodeur à canaux peut être utiliser pour la reconnaissance de la parole; nous citerons donc les méthodes les plus utilisées. Puis nous aborderons le vocodeur à canaux et la théorie sur laquelle il est basé. Enfin nous détaillerons les différentes parties constituant le banc de filtre, ainsi que leur fonctionnement .

CHAPITRE I

GENERALITES SUR
LE SIGNAL VOCAL

La complexité du signal de la parole à toujours représenter un obstacle pour la recherche dans ce domaine. La connaissance des modes de production des sons est donc non négligeable pour l'étude de la parole .

I - Le Système Vocal et son Fonctionnement

Globalement, le système vocal humain comprend :

- _ Les poumons et le conduit trachéo-bronchique
- _ Le larynx
- _ Le conduit vocal (cavité buccale ,cavités nasales et la cavité du pharynx) .

L'onde de pression reçue par l'oreille de l'auditeur est créée grâce à l'énergie fournie par le diaphragme et les poumons.

L'air arrive au larynx où les cordes vocales introduisent une modulation du flux d'air .

Les cordes vocales ont une course telle qu'elles peuvent obturer le larynx et donc s'opposer au passage du flux d'air , ou se séparer de façon brusque pour laisser l'air s'échapper .

Un système acoustique est décomposable en deux parties fonctionnelles , que l'on appelle " excitateur " et " résonateur " . Cette décomposition est vraie lorsque le second ne réagit pas sur le premier .

L'excitateur délivre un signal qui sera modifié par la fonction de transfert du résonateur .

Il faut , cependant , distinguer plusieurs cas, selon l'emplacement et la nature de l'excitation .

C'est le mouvement des lèvres , de la langue, du palais qui va définir la forme du conduit vocal .

I .1 - Les sources des sons

A son émission , le signal vocal naît d'un mouvement quasi-continu de l'appareil phonatoire tel que cet appareil , en cours d'articulation ne reste qu'un très court instant dans l'état caractéristique d'un son donné.

On distinguera trois sources de sons :

a/ La première sera les cordes vocales. Leurs vibrations quasi-périodique produisent le son fondamental ou/pitch, intervenant à l'émission des sons dits " voisés ". La fréquence de vibration définit la mélodie ou la hauteur d'une voie grave ou aigue et varie suivant l'âge et le sexe du locuteur . Ainsi chez les hommes elle est aux environs de 150 Hz , et chez les femmes et les enfants de l'ordre de 250 Hz .

- b/ Une seconde source provient du rétrécissement du canal vocal en un de ses points. Il s'y crée un flux d'air turbulent qui engendre une excitation incohérente du système vocal.
Des mesures indirectes ont montré qu'en ce point, le spectre de bruit créé est pratiquement uniforme sur une large bande. Par conséquent le spectre du son finalement émis est déterminé par l'ensemble des cavités vocales ultérieures.
Cette source génère les sons sourds continus.
- c/ La dernière source provient de l'occlusion du canal en un de ses points. Ces occlusions ont de multiples localisations, par exemple: labiale, dentale, palatale.
Une certaine pression est maintenue en ce point et l'ouverture brusque du canal produit une excitation transitoire des cavités vocales.
Si l'excitation est schématisée par une fonction de pression rectangulaire, le spectre décroît rapidement en fonction de la fréquence.
Cette source intervient dans la formation des sons sourds transitoires.

Les deux dernière sources décrites peuvent être respectivement activées avec la première pour donner les sons sonores continus et transitoires. On remarquera toutefois que, si un son sonore nécessite l'intervention de deux sources, il faut se garder de définir acoustiquement le son résultant comme la somme des composantes: le spectre instantané d'un son sonore continu (J., V ou Z par exemple) n'est pas le spectre du son sourd analogue (CH, F, S) additionné de voisement.

Dans le cas de la parole chuchotée, une source de bruit remplace les cordes vocales. Elle est créée essentiellement par un flux d'air turbulent au niveau de la glotte. Sur le sonogramme, le spectre n'est plus un spectre de raies (harmoniques du fondamental) mais un spectre continu. De plus, sa dynamique et son intensité sont faibles comparés à celle de la parole voisée. On remarque donc que l'ouverture de la glotte détermine la proportion de bruit dans le spectre.

II - Structure Phonétique de la Parole

II-1 Les Phonèmes

L'introduction du phonème comme unité phonétique répond au besoin d'établir un code pour représenter le signal vocal analogique. On peut le définir acoustiquement comme l'unité repérable sur un sonogramme par ses caractéristiques propres, ou linguistiquement, comme la plus petite séquence qui remplacée par un élément de la même classe (voyelle, fricative, plosive, nasale ou liquide) change le sens du mot.

Il est alors important de rechercher pour chaque phonème les critères qui le définissent et d'évaluer les paramètres employés pour le décrire.

II.2 - La notation phonétique

Les sons sont classés selon la partie de l'appareil vocal qui les a émis et selon le mode d'articulation de tout l'appareil . Ainsi on distingue :

- les sons sonores et les sons sourds selon que les cordes vocales vibrent ou non .
- les sons continus et les sons transitoires: les premiers peuvent être émis en l'absence de tout mouvement de l'appareil vocal ; les seconds impliquent par contre un mouvement de cet appareil .

La phonétique s'est essentiellement intéressée aux positions statiques du système phonatoire qui correspondent à l'émission des sons quasi-stationnaires comme les voyelles et les fricatives .

La notation phonétique, basée sur la physiologie de l'appareil vocal implique que la parole peut être décrite comme une suite d'unités phonétiques associées à des événements physiques observables .

Les avantages de la notation phonétique sont :

- de préserver les informations sémantique et esthétique en faisant une transcription phonétique avec des indications de rythme et de mélodie .
- d'associer à ces unités phonétiques un ensemble d'événements physiologiques . Dans un système de reconnaissance, on peut envisager de relever les événements acoustiques , d'en déduire les événements articulatoires et de déterminer la suite de phonèmes par le lien qui associe de manière unique un son au mouvement qui la produit .

Pourtant , les phonèmes ne se produisent généralement pas comme des séquences isolées ou simples , dans des états stationnaires qu'il suffit de concaténer pour former le signal vocal . Ainsi , une synthèse du mot " OUI " par juxtaposition des sons " OU " et " I " donnera deux sons successifs distincts alors qu'une synthèse de la zone de transition entre " OU " et " I " dans " OUI " suffit à la perception du mot .

Finalement , cette discrétisation phonétique du signal vocal conduit à une perte importante d'information en raison de la nature essentiellement transitoire de la parole .

II.3 - Les classes de phonèmes

A / Les voyelles :

Dans une transcription phonétique un phonème sur deux est une voyelle . Dans le temps , les voyelles représentent un pourcentage important de la durée du mot . On peut de plus remarquer que pour la plupart des mots d'un vocabulaire limité, la suite correcte des voyelles suffit souvent à l'identification du mot .

Pour les sons continus et sonores que sont les voyelles, les cavités vocales reste dans une configuration relativement stable. Le signal peut alors être considéré comme résultant du filtrage linéaire d'un signal d'excitation produit par les cordes vocales, par le filtre des cavités laryngo-bucco-nasales

Les voyelles de la langue française sont :

alphabet phonétique	terminologie pratique	exemples
[i]	i	lit
[e]	é fermé	été
[ɛ]	é ouvert	forêt
[a]	a antérieur	papa
[ɑ]	a postérieur	ane
[ɔ]	o ouvert	port
[o]	o fermé	peau
[u]	ou	loup
[y]	u	tu
[ø]	eu fermé	creuse
[œ]	eu ouvert	peur
[ɛ̃]	e sourd	petit
[ɛ̃]	in	brin
[õ]	un	brun
[ɑ̃]	an	blanc
[õ]	on	blond

Elles peuvent être définies et classées selon divers méthodes. Une méthode de définition utilise certaines positions articulatoires :

- le degré d'aperture de la bouche : ouvert (ɑ, ɛ, o)
- la position de la langue : antérieur (ɛ, e, i), médiane (œ) ou postérieure (ɑ, ɔ, c, o, u).

III - Perception des Sons

En plus du rôle de perception que joue l'oreille chez l'auditeur (sans lequel la parole n'aurait pas de raison d'être), elle permet également au locuteur de contrôler la hauteur tonale grâce à l'audition des sons émis.

Un son est entièrement définie par trois dimensions: l'intensité, la fréquence et le temps.

Le système auditif est sensible à ces trois paramètres.

Dans ce qui suit, nous nous limiterons à l'essentielles pour avoir une idée globale sur la perception des sons.

1- L'intensité

L'oreille est dotée d'une excellente sensibilité différentielle, c'est à dire qu'elle peut comparer deux sons qui ne diffèrent entre eux que par leur pressions acoustiques.

La loi de Weber-Fechner montre que la sensibilité de l'oreille est proportionnelle au logarithme de l'excitation.

Ainsi si pour un son I_1 la sensibilité de l'oreille est S_1 et pour un son I_2 la sensibilité est S_2 alors on a :

$$S_2 - S_1 = \log_{10} (I_2) - \log_{10} (I_1) = \log I_2/I_1$$

2- La fréquence

La plage de fréquence que l'homme peut percevoir va de 20 Hz jusqu'à 20 KHz.

Dans ce domaine également la sensibilité de l'oreille est proportionnelle au logarithme de la fréquence d'excitation. Mais il apparait des écarts à la loi logarithmique qui augmentent avec la fréquence.

3- Le temps

C'est l'audition qui permet le mieux de percevoir le temps qui s'écoule. En effet, nous percevons les événements qui marquent chaque instant, et ainsi nous pouvons savoir que tel événement a eu lieu avant ou après tel autre événement.

Notons également que le seuil de séparation temporelle varie entre 1 et 10 ms, selon l'intensité et les conditions expérimentales.

CHAPITRE II

ANALYSE DE LA PAROLE

ANALYSE DE LA PAROLE

Introduction

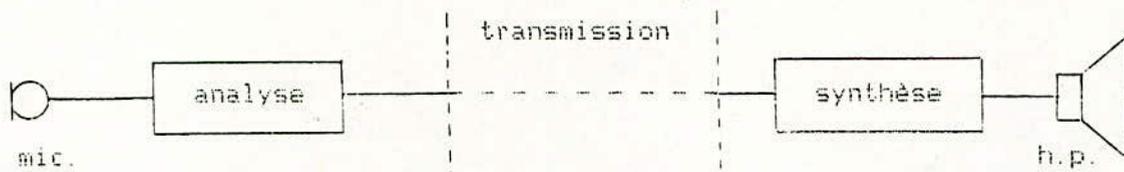
L'analyse et la paramétrisation du signal vocal font parties du premier étage d'un système de reconnaissance de la parole. Le rôle de l'analyse est d'extraire du signal vocal l'information utile au système de reconnaissance (analyse spectrale, extraction du fondamental...).

La paramétrisation consiste à déterminer les grandeurs effectivement retenue à l'entrée du système de reconnaissance. Ainsi N paramètres vont caractériser des segments de signal, d'une durée de 10 à 40 ms chacun.

Il existe plusieurs techniques d'analyse de la parole, dont la plupart dérivent des méthodes de codage de la parole en vue de sa transmission, les vocodeurs en sont un exemple.

Le signal de la parole est très dense, c'est à dire qu'il ne transporte pas uniquement l'information subjective ; il véhicule en même temps des données sur l'accent, le rythme, et l'intonation du locuteur. Il est très redondant. La transmission intégrale d'un message parlé nécessite un débit d'information de 300 Kbit/s. Il faudrait disposer d'une mémoire d'une dizaine de méga-octet pour stocker 4 à 5 minutes de parole.

Le choix de paramètres pertinents caractérisant la parole permettra une diminution du débit binaire.



synoptique du vocodeur.

Les Techniques d'Analyse du Signal Vocal

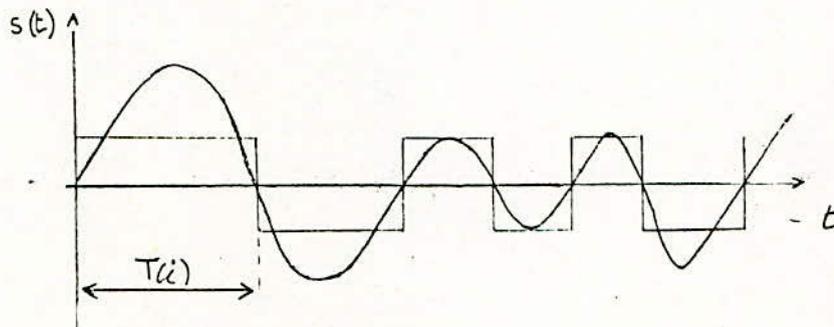
Nous décrirons brièvement, dans ce paragraphe, quelques techniques d'analyse du signal vocal.

1- Analyse Temporelle

1-1- Densité de passage par zéro

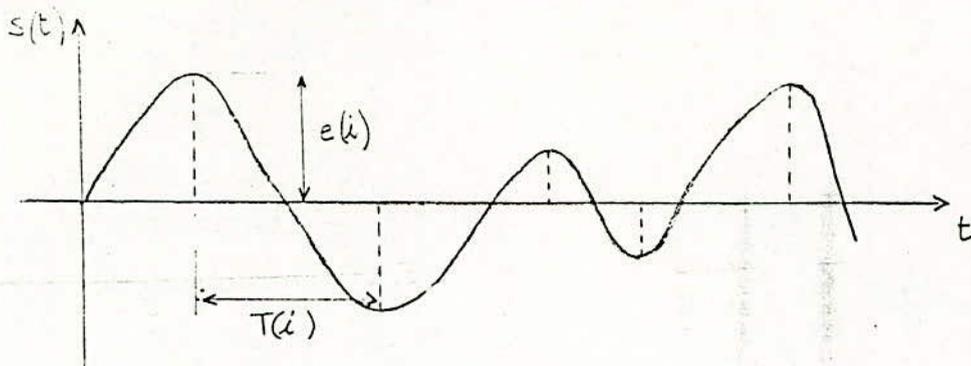
Le principe de cette méthode est basé sur la mesure de la densité de passage par zéro du signal vocal. Pour cela il suffit de compter le nombre de changement de signe du signal durant un intervalle de temps.

Cette méthode utilisée à la sortie de filtre passe-bande encadrant les trois premiers formants permet d'estimer les fréquences de ceux-ci (Schafer, Rabiner, 1974).



1-2- Codage des extrema

L'analyse par codage des extrema consiste à mesurer les distances séparant les extrema relatifs, ainsi que la valeur du signal pour chaque extremum relatif.



Si on fait précéder cette analyse par un filtre passe-bande de 0 à 600 Hz il est alors possible de mesurer la fréquence fondamentale d'un son voisé.

2 - Analyse Spectrale

2-1 - Transformée de Fourier rapide (FFT)

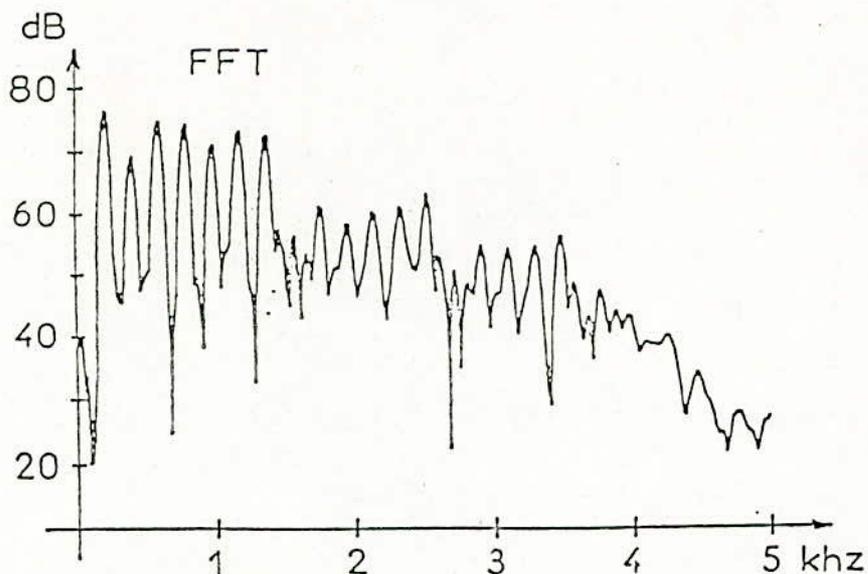
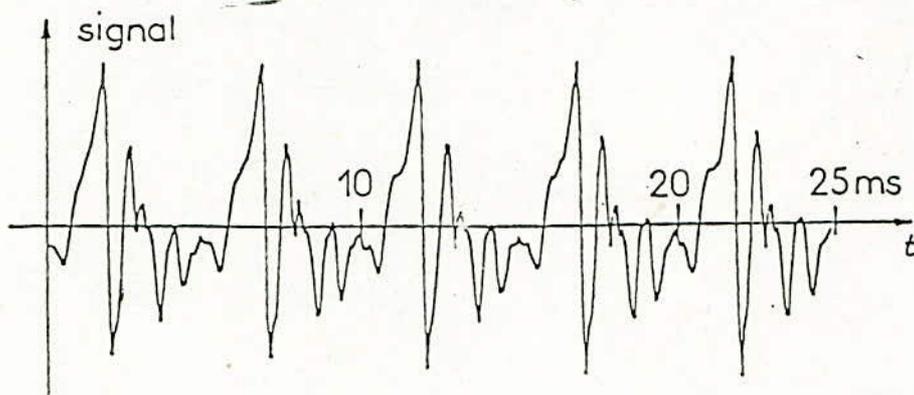
La transformée de Fourier discrète (TFD) permet le passage du domaine temporel au domaine fréquentiel. Elle est définie par l'expression :

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-j2\pi kn/N}$$

Mais elle nécessite un nombre de calcul important ce qui la désavantage du point de vue rapidité.

En 1965, Cooley et Turkey présentent un algorithme de calcul très efficace qui permet d'évaluer la TFD en un nombre moindre de calcul. Ainsi la transformée d'un signal vocal est calculée sur des segments de N échantillons pondérés par une fenêtre de Hamming.

Le spectre de puissance calculé sur un segment est alors constitué de $N/2$ échantillons répartis linéairement sur la bande de fréquence 0 à $F_e/2$, F_e étant la fréquence d'échantillonnage (de 8 à 20 KHz).



FFT sur un segment de 25 ms.

2-2- Vocodeur à canaux

L'étude du vocodeur à canaux sera détailler au chapitre IV

2-3- Vocodeur à formants

Le vocodeur à formant réalise une analyse spectrale basée sur la détection et la détermination des fréquences et amplitudes des trois premiers formants.

La recherche des maxima du spectre de puissance, obtenu au moyen d'un banc de filtres ou d'une FFT, fournit une bonne approximation de la position de ces formants

L'avantage de cette technique d'analyse, fondée sur un modèle de phonation, est qu'elle fournit un nombre réduit de paramètres caractérisant convenablement le signal vocal.

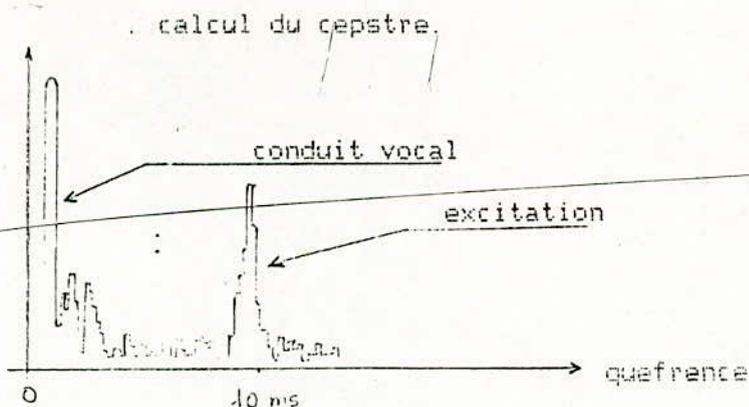
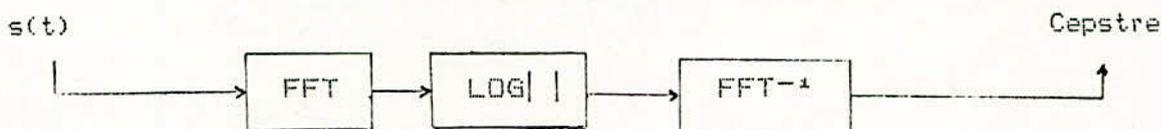
Cependant la détection automatique des trajectoires formantiques est incertaine (maxima parasites, formants rapprochés...) et rend la méthode imprécise.

3- Codage Homomorphique

L'analyse homomorphique repose sur une technique de filtrage non linéaire qui permet de séparer des signaux combinés par multiplication ou par convolution.

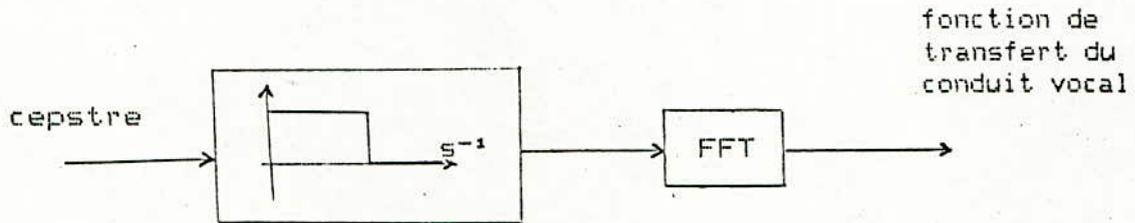
Le signal vocal peut être considéré comme le résultat de la convolution du signal d'excitation avec la réponse impulsionnelle du conduit vocal; d'où l'intérêt de cette méthode.

Il faut d'abord déterminer le cepstre défini comme étant la transformée de Fourier inverse du logarithme de la transformée de Fourier du signal vocal.



Cepstre d'une voyelle.

Dans le cepstre, on peut nettement distinguer les deux composantes du signal vocal, ainsi que le montre la figure précédente. Il suffit alors de les séparer en utilisant une fenêtre temporelle.



Les composantes dues à la source et au conduit vocal sont ensuite codées séparément pour former un ensemble de paramètres utilisables pour la transmission à faible débit ou pour la reconnaissance de la parole.

L'inconvénient de cette méthode d'analyse est l'importance du temps de calcul. Cet inconvénient est due à la présence de deux transformée de Fourier à caculer; ce qui allonge le temps de traitement.

4- Codage Prédicatif (LPC)

La prédiction linéaire est une méthode très utilisé aujourd'hui pour estimer un ensemble de paramètres d'un signal et identifier le système générateur du signal.

L'idée de base du codage par prédiction linéaire est qu'un échantillon de signal peut être approximé par une combinaison linéaire des P échantillons précédent (P=8 à 14). La valeur du signal à l'instant n est donc:

$$S(n) = \left(\sum_{k=1}^P a_k \cdot S(n-k) \right) + e(n)$$

où les constantes a_k sont les coefficients de prédiction, P est l'ordre du prédicteur, et $e(n)$ l'erreur de prédiction.

Il existe plusieurs formulations du problème du calcul des coefficients de prédiction, mais la plupart mènent à la résolution d'un système de P équations à P inconnues.

À partir des coefficients a_k , on peut directement obtenir une bonne approximation de la fonction de transfert du conduit vocal, sur des intervalles de temps donné de 10 à 25 ms.

CHAPITRE III

LA RECONNAISSANCE
DE LA PAROLE

LA RECONNAISSANCE DE LA PAROLE

Introduction

La reconnaissance de la parole entre dans le cadre général de la reconnaissance des formes. Reconnaître une forme revient à décider qu'elle représente un objet appartenant à une famille donnée.

Un système de reconnaissance comprend trois organes distincts: un capteur fournit les "images" à reconnaître; un organe de prétraitement effectue les cadrages, lissages, normalisation, extraction des paramètres; et un dispositif de décision et de classification mémorise les critères de classification des formes et les règles de parenté des "images" déformées.

Cependant, les problèmes rencontrés dans la reconnaissance de la parole sont à caractères très spécifiques.

Ainsi il faut noter le caractère continu du discours oral, qui se présente comme une suite de mots ne contenant ni pause, ni marque de séparation. De même on notera qu'à contenu phonétique égal, le signal vocal est très variable. On peut distinguer une variabilité intra-locuteur (suivant l'état émotionnel, l'intensité, la vitesse d'élocution) et une variabilité inter-locuteur (en fonction de l'âge, du sexe, des habitudes articulatoires de chacun). Enfin, à ces facteurs il faut ajouter les altérations dues aux conditions d'enregistrement, bruit de fond, réverbération, filtrages, ect., qui existent toujours dans notre environnement sonore sans que nous en soyons conscients.

Les premiers essais de reconnaissance date de 1950 et peuvent être attribués à Dreyfus-Graf. Son souhait était de réaliser un système de reconnaissance indépendant du locuteur (phonétographe). Il nota qu'il était difficile de reconnaître des classes de phonèmes.

En 1968, Vintsjuk présente un système de reconnaissance de mots utilisant pour la première fois le principe de la "programmation dynamique" mis au point par R. Bellman (1957) et Slucker (1968). Des développements seront apportés à ce principe et il sera adapté à la reconnaissance de la parole continue par Sakoe et Chiba au Japon (1971).

Divers systèmes ont été élaborés depuis et la commercialisation de certains d'entre eux a déjà débuté.

Nous donnerons, dans ce qui suit, un aperçu des méthodes de reconnaissance de la parole les plus utilisées.

On peut distinguer deux façons d'aborder la reconnaissance de la parole, qui sont :

- la reconnaissance analytique
- la reconnaissance globale

1 - La reconnaissance analytique (ou phonétique)

Cette approche consiste à localiser et à identifier les composants élémentaires du message vocal. Le problème de la reconnaissance phonétique est le fait que les organes de production sont en mouvement quasi-permanent lors de l'articulation des phonèmes. Le fait que l'articulation de deux sons successifs modifie l'articulation de l'un d'eux ou des deux est appelé effet de coarticulation. Par extension, on appelle effet de coarticulation tous phonèmes acoustiques liés à la succession de deux phonèmes.

Pour représenter le phonème il est donc nécessaire de prendre en compte ses caractéristiques stables, mais aussi les règles qui permettent de le relier aux phonèmes du contexte proche ou éloigné, et ceci aussi bien pour la synthèse que pour la reconnaissance.

Généralement cette approche fait intervenir divers niveaux de reconnaissance :

- le niveau acoustique qui effectue une paramétrisation et une segmentation du signal vocal,
- le niveau phonétique qui associe un phonème à chaque segment de signal en utilisant un dictionnaire de phonèmes et des règles contextuelles,
- le niveau phonologique et lexical qui procède à la recherche des mots contenus dans le message prononcé en utilisant un lexique et des règles phonologiques. Il est difficile de différencier Phonétique et Phonologie (étude de la forme). Ainsi si l'on place le problème de la coarticulation dans le niveau phonétique, le problème de la gémination (la ville limite / la ville imite : lavillimit) par exemple, sera délégué au niveau phonologique.
- le niveau syntaxique et sémantique qui permet de réduire le nombre de solutions en prenant en compte la grammaire du langage utilisé, et en rejetant les solutions sémantiquement incorrectes. La syntaxe traite des possibilités d'organisation des mots dans une phrase. Le niveau sémantique traite du sens général des mots et des relations entre eux.
- le niveau pragmatique qui utilise l'information relative à la tâche et au déroulement du dialogue pour restreindre les choix possibles. La pragmatique permet d'exploiter les informations accumulées au cours de la reconnaissance et de prévoir le comportement de l'interlocuteur d'après le contexte général de la conversation.

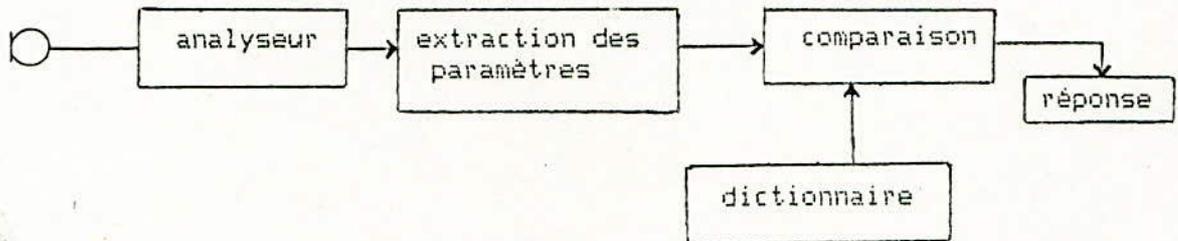
En théorie l'approche analytique permet de traiter des vocabulaires illimités. Mais, en pratique, les taux de reconnaissance phonétique sont actuellement assez faibles. Ils atteignent généralement moins de 70% de reconnaissance exacte pour un locuteur.

On peut se demander à juste titre quel est le taux de reconnaissance pour l'être humain dans son décodage phonétique. Des expériences menées en utilisant la synthèse de la parole (LIENARD), où l'on plaçait des erreurs sur des chaînes de phonèmes synthétisées (erreurs telles que la substitution de deux phonèmes différant d'un trait distinctif ou la confusion voyelle-consonne), ont montré que plus de 15% d'erreurs ne permettaient pas à l'auditeur de comprendre une phrase.

2 - La Reconnaissance Globale

A l'opposé de la reconnaissance analytique, la reconnaissance globale ne tient aucun compte de la structure et des propriétés phonétique des mots. Il s'agit d'une approche purement acoustique. L'unité de base est l'entité lexical, qui le plus souvent est un mot, mais qui peut être un groupe de mots, une phrase complète, un sifflement ou un bruit.

Cette méthode nécessite une phase d'apprentissage pour construire un dictionnaire de références acoustiques contenant au moins une référence pour chaque mot du lexique. Ce dictionnaire peut être figé ou évolutif. La reconnaissance se fait par comparaison entre un mot prononcé et l'ensemble des mots du lexique, avec choix du plus voisin, et rejet en cas de ressemblance trop incertaine. Les méthodes de comparaison sont soit de matrice binaire à matrice binaire de même dimension, soit par comparaison dynamique (ou alignement temporel dynamique, qui est un algorithme dérivé des méthodes de la programmation dynamique de BELLMAN).



- Reconnaissance globale -

2 - 1 - La reconnaissance de mots isolés

La méthode globale s'applique très bien à la reconnaissance de mots isolés.

On suppose que les mots prononcés sont séparés par des silences de durées supérieures à 200 ms, de ce fait il n'y a pas de problème de segmentation. La segmentation consiste à diviser le message en plusieurs éléments (éléments stables, transitoires, syllabes). Ainsi les entités à reconnaître sont parfaitement définies.

Les avantages de la reconnaissance acoustique des mots isolés sont que les résultats sont actuellement de bonne qualité (de l'ordre de 99% sur les chiffres de 0 à 9 par exemple), que ces systèmes sont valables pour n'importe quelle langue, et n'importe quel signal acoustique.

Pour les inconvénients, on notera l'obligation de faire un apprentissage préalable, on notera également que le système ne reconnaîtra que la personne qui a fait l'apprentissage. Un autre désavantage est que le nombre de mots du vocabulaire est limité (l'occupation mémoire pour chaque mot varie de 200 bits à 500 octets, ce qui fait que les systèmes acceptent de 40 à 120 mots environ).

2 - 2 - Reconnaissance de mots enchainés

L'approche globale et les techniques utilisées en reconnaissance de mots isolés peuvent s'appliquer à la reconnaissance de mots "enchainés", c'est à dire prononcés sans qu'il y ait de silence entre eux. Il peut s'agir de lettres ou de chiffres représentant un numéro minéralogique ou un code postal.

Les problèmes que cela pose sont que la coarticulation entre les phonèmes déforme les frontières des mots qui se succèdent, qu'il faut segmenter en mots le signal acoustique à reconnaître, et que le rythme de prononciation déforme le mot par rapport à sa prononciation en mode isolé.

La segmentation peut être obtenue en essayant de faire coïncider, par comparaison dynamique, chaque référence du dictionnaire avec les différents segments de la phrase correspondant à tous les découpages possibles. Une telle méthode est réalisable pour des systèmes disposant de très grandes puissances de calcul.

2 - 3 - Les systèmes mono- et multilocuteurs

On appelle système monolocuteur, un système de reconnaissance destiné à être utilisé par un seul locuteur. Un système multilocuteur est par contre destiné à reconnaître la voix de plusieurs locuteurs.

Dans les systèmes utilisant une approche globale, l'indépendance du locuteur est généralement obtenue par un accroissement des références en recherchant les éléments les plus représentatifs d'un ensemble de mots prononcés par un échantillon de 100 à 200 individus.

Cette méthode a l'inconvénient de diviser par 10 ou 12 le nombre de mots reconnaissables lorsque l'on passe d'un système monolocuteur à un système multilocuteur.

		Type d'élocution	
		Mots isolés	Parole continue
a p p r o c h e	Globale	reconnaissance de mots (petits vocabulaires)	localisation de mots dans les phrases
	Analytique	reconnaissance de mots (grand vocabulaire)	localisation de mots. Reconnaissance et compréhension de phrases

- tableau récapitulatif -

CHAPITRE IV

LE VOCODEUR A CANAUX

LE VOCODEUR A CANAUX

I- Eléments théorique d'analyse spectrale

Le vocodeur à canaux réalise une analyse spectrale du signal vocal. Les éléments qui plaident en faveur d'une analyse fréquentielle sont:

- 1- le concept de fréquence permet une description concise des séquences vocales,
- 2- l'oreille, en tant qu'appareil de perception de la parole, effectue entre autre au niveau de la cochlée une analyse fréquentielle du signal.

Cependant, une analyse fréquentielle par transformée de Fourier n'est pas immédiatement utilisable. En effet, pour que la transformée existe il faut que :

si $f(t)$ est le signal, alors $\int_{-\infty}^{+\infty} |f(t)| dt$ soit finie.

Or en général, le signal vocal ne satisfait pas à cette condition et de plus n'est pas connu de $-\infty$ à $+\infty$.

En outre, le signal vocal résulte de la convolution de l'excitation (au niveau de la glotte) et de la fonction de transfert du conduit vocal. Cette fonction de transfert évolue au cours du temps et la méthode d'analyse spectrale utilisée doit prendre en compte cette évolution.

Donc le traitement doit pouvoir:

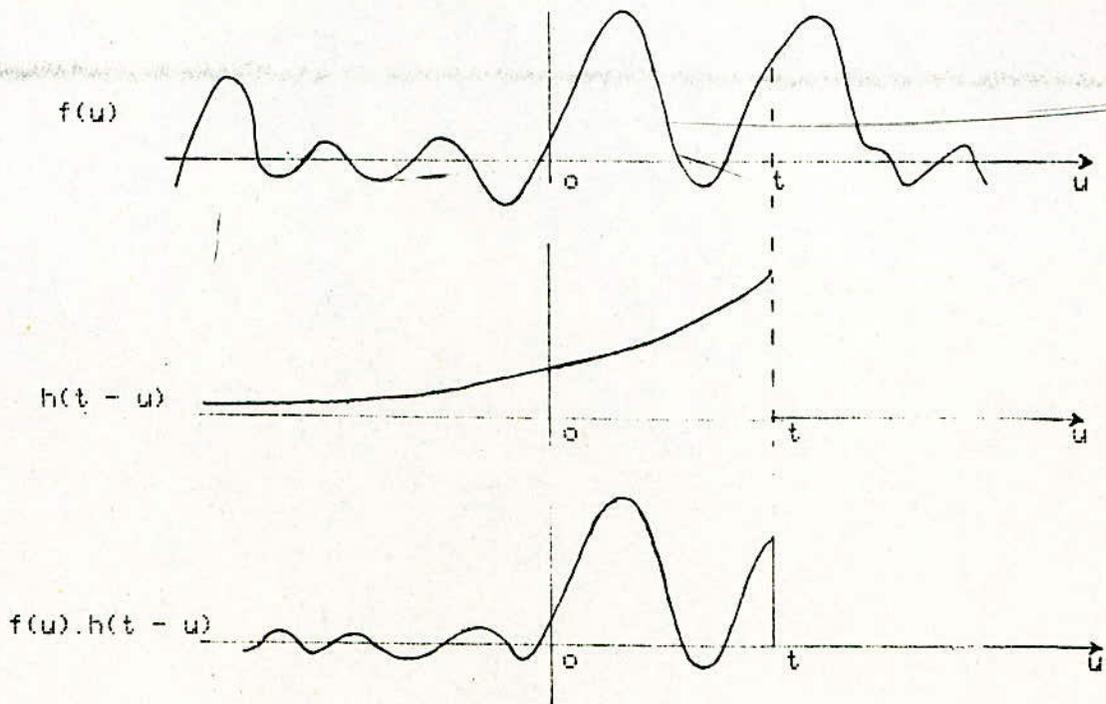
- 1 - modifier le signal pour qu'il soit sommable sur des valeurs connues du passé
- 2 - intégrer le signal dans une fenêtre de temps dans laquelle il est quasi-stable.

On arrive ainsi à la notion de transformée de Fourier à court terme (TFCT). Celle-ci permet une transformation du signal $f(t)$ de tel façon que l'intégration à un instant donné s'effectue seulement sur les valeurs du passé, et pour mettre en évidence le phénomène d'évolutivité on ne tient compte que du passé récent du signal $f(t)$.

Pour satisfaire à ces exigences on multiplie le signal $f(t)$ par une fenêtre temporelle $h(t - u)$. (voir figure page suivante).

La fenêtre temporelle $h(t - u)$ est telle que:

$$h(t - u) = 0 \quad \text{si} \quad u > t$$



La transformée de Fourier du signal vu à l'instant t à travers la fenêtre temporelle $h(u)$ a pour expression :

$$F(\omega, t) = \int_{-\infty}^t f(u) h(t-u) e^{-j\omega u} du$$

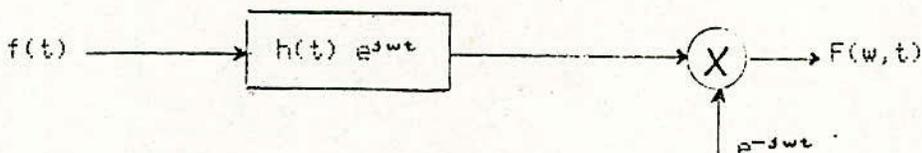
Dans l'expression de la transformée $F(\omega, t)$, si on pose $p = t - u$ alors,

$$\begin{aligned} F(\omega, t) &= \int_{+\infty}^0 f(t-p) h(p) e^{-j\omega(t-p)} (-dp) \\ &= e^{-j\omega t} \int_0^{+\infty} f(t-p) h(p) e^{j\omega p} dp \end{aligned}$$

En introduisant le produit de convolution $f(t) * h(t) e^{j\omega t}$, la relation précédente devient :

$$F(\omega, t) = e^{-j\omega t} (f(t) * h(t) e^{j\omega t})$$

Cette équation peut être schématiser par :



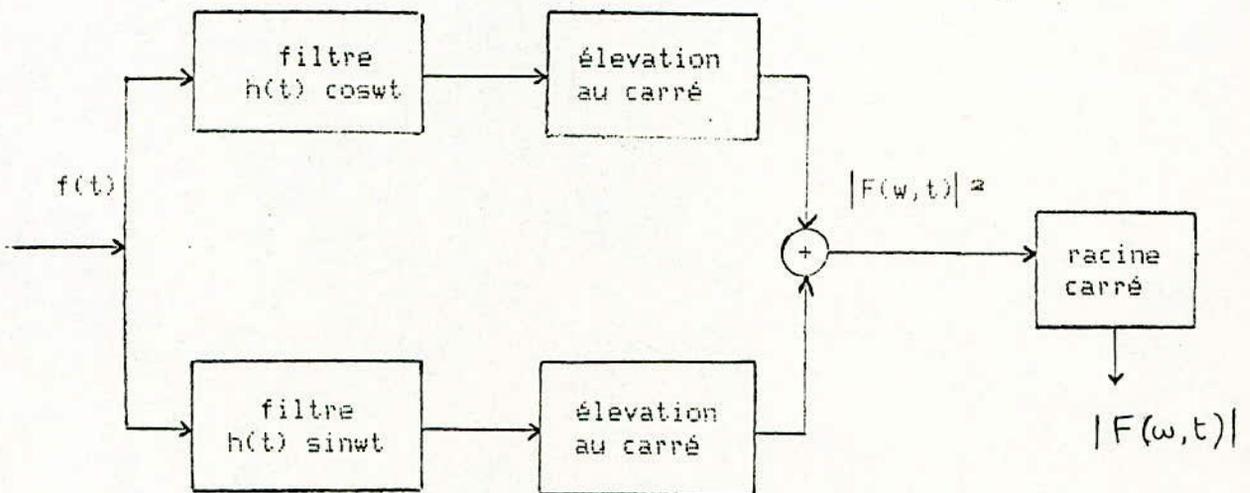
En développant l'exponentielle dans la parenthèse on aura

$$F(\omega, t) = e^{-j\omega t} (f(t) * h(t) \cos \omega t + j f(t) * h(t) \sin \omega t)$$

Le module de la TFCT sera:

$$|F(\omega, t)| = ((f(t) * h(t) \cos \omega t)^2 + (f(t) * h(t) \sin \omega t)^2)^{1/2}$$

La transformation peut être schématisé par:



Or l'enveloppe $e(t)$, d'une fonction $u(t)$ admettant une transformée de Fourier est définie par:

$$e(t) = (u(t)^2 + \hat{u}(t)^2)^{1/2}$$

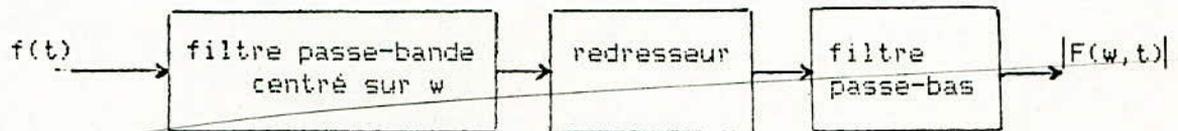
où $\hat{u}(t) = u(t) * \frac{1}{\pi t}$ est la transformée de Hilbert de $u(t)$.

Dans notre cas, et lorsque les spectres de $h(t)$ et de $\cos \omega t$ ne se chevauchent pas, la transformée de Hilbert est donnée par

$$\begin{aligned} (f(t) * h(t) \cos \omega t) * \frac{1}{\pi t} &= f(t) * (h(t) \cos \omega t * \frac{1}{\pi t}) \\ &= f(t) * h(t) \sin \omega t \end{aligned}$$

Ainsi l'amplitude $|F(\omega, t)|$ du spectre instantané que l'on recherche, sera donnée par l'enveloppe de $f(t) * h(t) \cos \omega t$.

Le schéma précédent peut être réduit à une seule branche, qui sera constituée d'un redresseur et d'un filtre passe-bas pour donner une bonne approximation de $F(\omega, t)$.



II- Principe de réalisation d'un vocodeur à canaux

Le vocodeur à canaux est composé de deux parties principales; l'analyseur et le synthétiseur.

Pour ce qui concerne l'analyseur, celui-ci comprend:

- un banc de filtres analogiques, qui réalise l'analyse spectrale que l'on a précédemment détaillé. Le domaine de fréquences analysé est situé entre 200 Hz et 7000 Hz, tandis que le nombre de voies varie entre 4 et 96.
- un détecteur de mélodie (ou du pitch). Celui-ci a pour rôle la mesure de la fréquence fondamentale des sons .
- un détecteur de voisement, qui a pour rôle d'indiquer si le message est voisé ou non.

Le synthétiseur est destiné à reproduire la parole à partir des signaux déduit de l'analyse du message.

La synthèse à canaux consiste à exciter un banc de filtre similaire à ceux de l'analyse à l'aide de l'une ou de l'autre des sources d'excitations (impulsions périodiques ou bruit blanc) possible selon que les sons sont voisés ou non; on dose l'énergie dans chaque canal en fonction de celle observée dans le même canal de l'analyse; enfin on fait correspondre la fréquence des impulsions à celle de la mélodie.

Le signal de sortie est obtenu par addition des sorties des filtres de synthèse.

CHAPITRE

V

ETUDE ET REALISATION D'UN
ANLYSEUR A BANC DE FILTRES

Introduction

Dans ce chapitre nous aborderons l'étude pratique de l'analyseur par banc de filtres analogiques à 14 canaux.

Ce banc de filtres réalisera une analyse spectrale à court terme du signal de la parole.

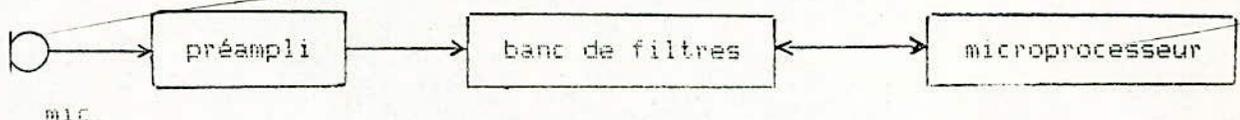
Cette méthode exploite l'hypothèse fondamentale; le signal à analyser est stationnaire pendant toute la durée de la fenêtre d'analyse, c'est à dire que ces propriétés (spectrales en particulier) n'évoluent que très lentement.

La durée de stabilité des sons est variable. Certaines voyelles ont des durées très longues (100 ms), tandis que les consonnes occlusives comme "be", "de", "gue" ne durent que 5 à 6 ms. La valeur courante de 10 ms est utilisée pour des synthèses de qualités moyennes.

Le banc de filtres peut être décomposer en plusieurs parties:

- 1 - la préamplification et la préaccentuation
- 2 - le filtrage passe-bande
- 3 - la détection d'enveloppe
- 4 - le lissage
- 5 - l'échantillonnage-blocage
- 6 - le multiplexage
- 7 - la conversion analogique-numérique
- 8 - le stockage des données

Le contrôle des opérations de multiplexage, de conversion analogique-numérique et de stockage des données sera effectuer par une carte Motorola 6802.



1- Le préampli et la préaccentuation

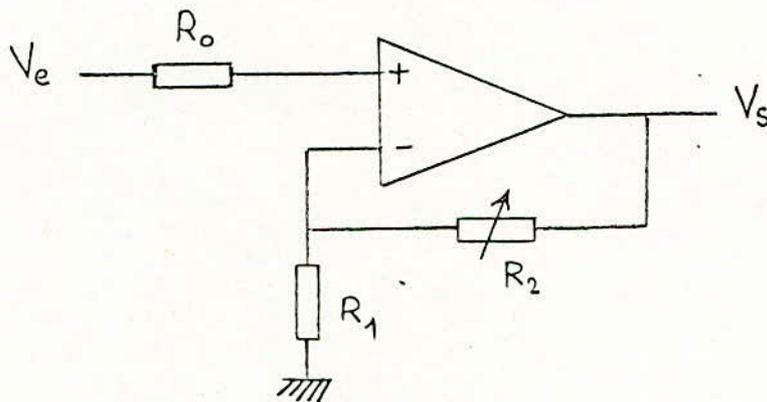
Après avoir été capté par le microphone, il est nécessaire d'amplifier l'amplitude du signal vocal, faute de quoi, celle-ci sera insuffisante pour pouvoir être traité par le banc de filtres.

De plus ce signal a subi lors de son passage dans le conduit vocal une atténuation de -6 dB/octave, particulièrement pour les hautes fréquences.

Donc pour que le signal soit exploitable il faut le préamplifier, puis renforcer l'énergie des hautes fréquences en le faisant passer par un filtre de préaccentuation.

Nous avons donc réalisé un préampli de gain constant sur une large bande et un accentuateur dont la réponse en amplitude est constante jusqu'à 500 Hz, puis présente une pente positive de 6 dB/octave jusqu'à 24 KHz, et une pente négative au delà.

Schéma du préampli

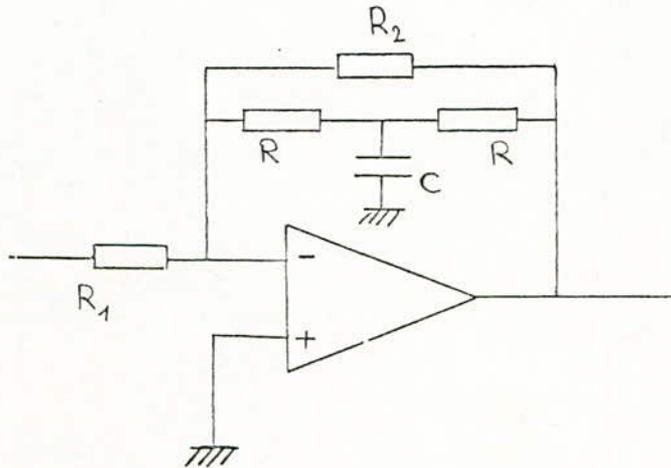


La fonction de transfert de ce montage est:

$$\frac{V_s}{V_e} = \frac{R_2 + R_1}{R_1}$$

Le gain de cet ampli est toujours supérieur à 1. En prenant pour R_2 une résistance variable on pourra ajuster le gain à la valeur désirée.

Schéma de l'accentuateur



La fonction de transfert de ce montage est de la forme:

$$F(\omega) = G_0 \frac{1 + j \omega / \omega_1}{1 + j \omega / \omega_2}$$

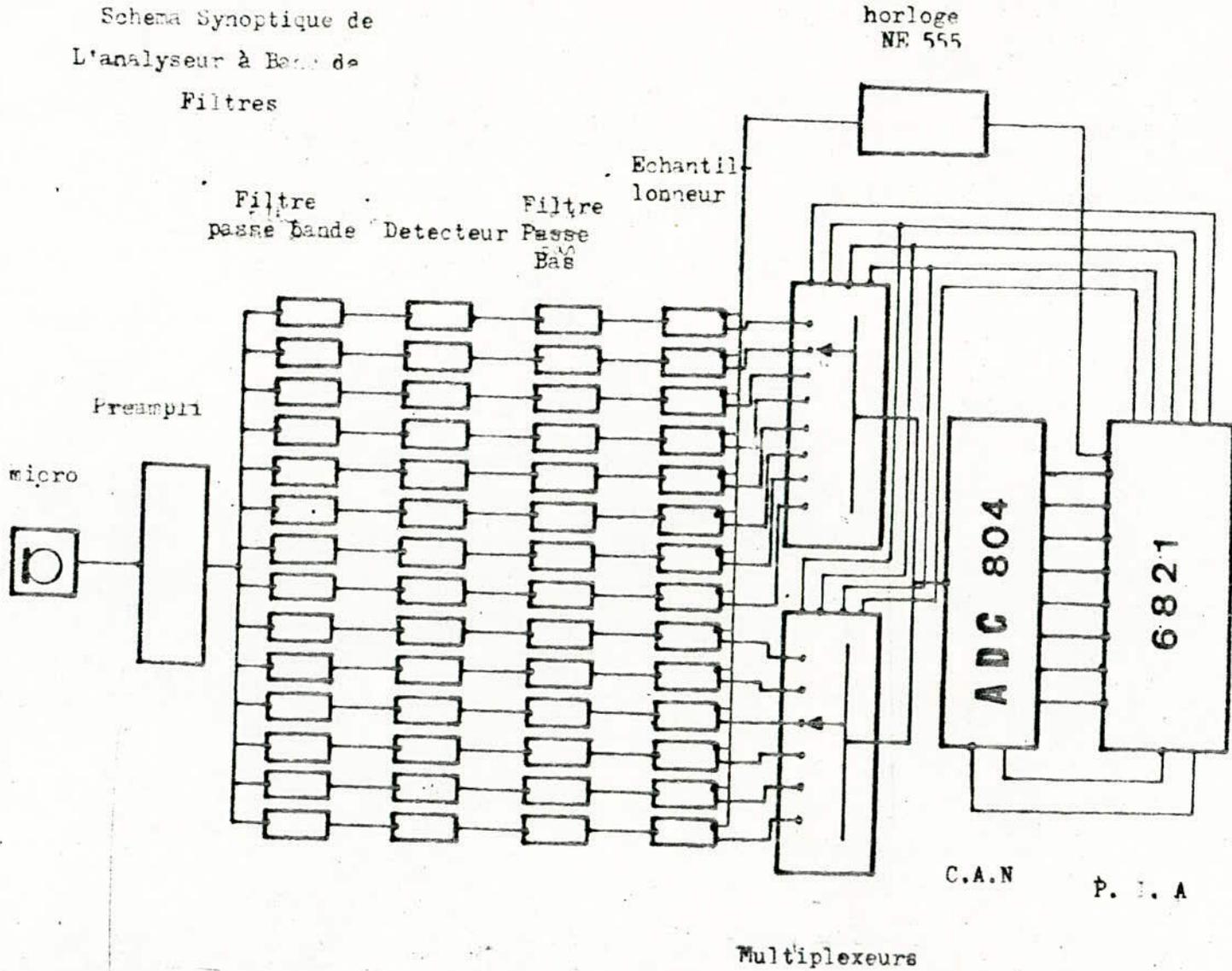
avec:

$$G_0 = - \frac{2 R R_2}{2 R R_1 + R_1 R_2}$$

$$f_1 = \frac{1}{2 \pi R C} = 500 \text{ Hz}$$

$$f_2 = \frac{R_2 + 2 R}{2 \pi R^2 C} = 24 \text{ KHz}$$

Schema Synoptique de
L'analyseur à Bande de
Filtres



2- Filtres Passe-Bande

a/ Choix du type de filtres

Le signal en provenance de la préamplification attaque les quatorze canaux montés en parallèle, chaque canal débutant par un filtre passe-bande, dont le rôle est de ne laisser passer que les signaux de fréquences voulues.

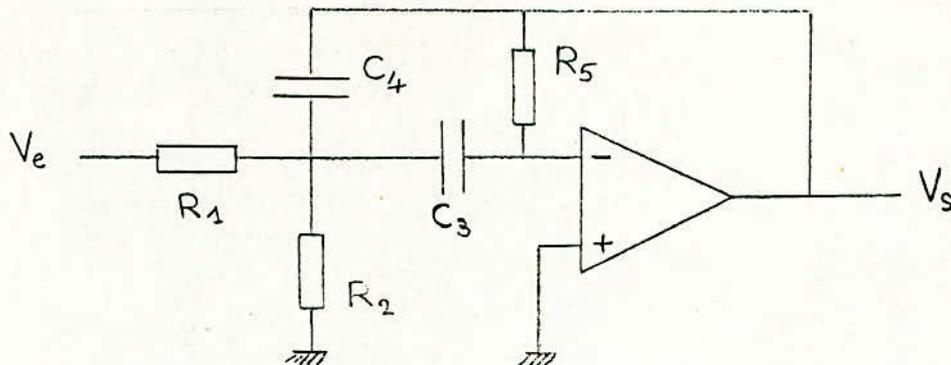
Les largeurs de bandes de ces filtres sont non disjointes, variables et choisies de manière à couvrir toute la plage de fréquence désirée, c'est à dire de 200 à 4300 Hz (voir échelonnement des filtres).

Le banc comprend donc quatorze filtres actifs passe-bande du 2^{ème} ordre à contre réaction-multiple (ou de Rauch).

Le choix des filtres actifs a été motivé par les avantages qu'ils apportent comparativement aux filtres passifs LC. Ces avantages sont:

- miniaturisation
- impédance d'entrée élevée
- faible impédance de sortie, ce qui permet des montages en cascade
- fréquences de coupure basses ou moyennes.

Le schéma du filtre passe-bande à contre réaction-multiple est donné ci-dessous:



.filtre passe-bande à contre-réaction multiple.

Les filtres passe-bandes à réaliser, doivent satisfaire à deux conditions:

- une largeur de bande précise
- une fréquence centrale contrôlable

Or parmi les divers structures possible de filtres passe-bandes, comme par exemple celle de Sallen-Key, ou la structure à contre-réaction simple, ou encore la mise en cascade d'un filtre passe-haut et d'un filtre passe-bas, il s'avère que la structure de Rauch est la plus adapté à notre réalisation.

La fonction de transfert du filtre passe-bande est donnée par :

$$F(p) = \frac{P}{ap^2 + bp + c} \quad \text{avec } p = j\omega$$

où a , b et c sont des constantes fonction des composants.
En remplaçant p par jw on aura

$$F(\omega) = \frac{j\omega}{(c - a\omega^2) + j b\omega}$$

Lorsque ω tend vers 0, la courbe de $F(\omega)$ va tendre vers 0 suivant l'asymptote d'équation:

$$g(\omega) = \frac{\omega}{c}$$

et lorsque ω tend vers l'infini, $F(\omega)$ tend vers 0 suivant l'asymptote d'équation:

$$h(\omega) = \frac{1}{a\omega}$$

Le point d'intersection de ces deux asymptotes nous donne la fréquence de coupure:

$$\omega_0 = \frac{\sqrt{c}}{\sqrt{a}} \quad ; \quad \text{qui lui correspond } g(\omega_0) = \frac{1}{\sqrt{ac}}$$

La fonction de transfert à cette pulsation a pour valeur: $F(\omega_0) = \frac{1}{b}$

Le coefficient de surtension Q est donnée par le rapport:

$$Q = \frac{F(\omega_0)}{g(\omega_0)} = \frac{\sqrt{ac}}{b}$$

On déduit de ce qui précède les constantes a, b et c

$$b = \frac{1}{G} \quad ; \quad a = bQ/\omega_0 \quad ; \quad c = bQ\omega_0$$

avec G qui est le gain nominal, et $Q = \frac{\omega_0}{\Delta\omega}$

En appliquant les lois de Kirchoff au circuit électrique du filtre et en travaillant en admittance on a:

$$\frac{V_s}{V_e} = \frac{Y_1 Y_3}{Y_5 (Y_1 + Y_2 + Y_3 + Y_4) + Y_3 Y_4}$$

avec $Y_1 = \frac{1}{R_1} \quad ; \quad Y_2 = \frac{1}{R_2} \quad ; \quad Y_3 = j\omega C_3$

$$Y_5 = \frac{1}{R_5} \quad ; \quad Y_4 = j\omega C_4$$

Après développement de cette expression et identification avec la fonction de transfert initial, les équations donnant la valeur des composants seront:

$$R_1 = a/C_4 = \frac{Q}{G \omega_0 C_4} = \frac{1}{G \Delta\omega C_4}$$

$$R_5 = \frac{R_1 (C_3 + C_4)}{b C_3} = \frac{(C_3 + C_4)}{\Delta\omega C_3 C_4}$$

$$R_2 = \frac{R_1}{c R_5 C_3 - 1}$$

Lors du calcul des composants, on se fixe la valeur de C_3 et C_4 , puis on calcule les valeurs de R_1 , R_2 et R_5 .

Les résistances R_1 et R_5 contrôlent la largeur de bande, tandis que la résistance R_2 permet de contrôler la fréquence centrale.

b/ Echelonnement des filtres

Divers échelonnements des filtres furent développés, notamment au CNET (Lannion, FRANCE).

Le premier échelonnement qui y fut proposé, appelé échelonnement 1, était logarithmique. La relation entre deux filtres successifs est définie par:

$$f(i+1) = 10^{1/12} f(i)$$

La bande globale analysée était comprise entre 250 et 3600 Hz. On reprochait à cet échelonnement d'avoir une bande trop étroite en dessous de 1000 Hz.

Le second échelonnement qui fut développé (échelonnement 2) augmentait le rapport bande globale sur nombre de canaux, grâce à un élargissement de 650 Hz de la bande globale pour le même nombre de canaux. Cependant cet échelonnement n'était pas pleinement satisfaisant. En effet les premiers canaux étaient de largeurs excessives.

D'autres échelonnements furent proposés pour éliminer les défauts des précédents.

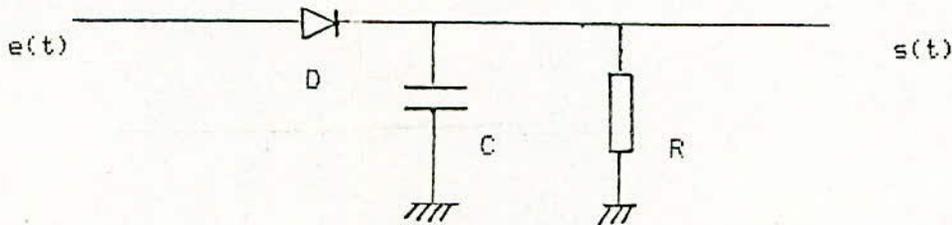
Des tests d'écoute de comparaison entre les différents échelonnements ont été réalisés. Il ressort des résultats de ces tests qu'un échelonnement optimum pour les voix masculines serait différent d'un échelonnement optimum pour les voix féminines. Autrement dit, l'échelonnement optimum ne tient pas uniquement compte du "récepteur oreille" mais aussi de "l'émetteur source - conduit vocal".

Nous utiliserons pour la réalisation de notre banc de filtres l'échelonnement 2 qui présente une bonne résolution dans le haut du spectre. Le tableau ci-dessous en donne le détail:

N°	F ₀ (Hz)	F _c à -6 dB		ΔF à -3 dB	Q = F ₀ /F
1	335	250	450	100	3.35
2	541	450	650	100	5.41
3	743	650	850	100	7.43
4	945	850	1050	100	9.45
5	1168	1050	1300	120	9.73
6	1442	1300	1600	150	9.61
7	1744	1600	1900	150	11.62
8	2045	1900	2200	150	13.63
9	2345	2200	2500	150	15.63
10	2645	2500	2800	150	17.63
11	2945	2800	3100	150	19.64
12	3245	3100	3500	200	16.47
13	3695	3500	3900	200	18.47
14	4095	3900	4300	200	20.47

3- La Détection d'Enveloppe

Un circuit de détection d'enveloppe est disposé après chaque filtre passe-bande de chaque canal. Le circuit de détection est constitué d'une diode de redressement suivie d'un filtre capacitif. Le filtrage capacitif consiste en un condensateur C placé en parallèle avec une résistance R, comme le montre la figure ci-dessous:



Pour que le filtrage soit efficace il faut que le produit RC correspondant à la constante de temps du circuit soit très supérieur à la période maximale du signal d'entrée.

$$RC \gg T_{\max} = \frac{1}{f_{\min}} \quad \text{et } f_{\min} = 250 \text{ Hz}$$

4- Filtre Passe-Bas

Les filtres passe-bas doivent effectuer un lissage de l'enveloppe détectée.

En effet, sur l'enveloppe détectée il reste des pics et des variations parasites, dont certaines sont dues aux filtres capacitifs. Pour supprimer ces variations indésirables, on utilise un filtre passe-bas du 2^{ème} ordre à réponse de Butterworth dont la fréquence de coupure est de 50 Hz. Un filtre à réponse de Butterworth offre l'avantage d'une réponse plate dans la bande passante. Le schéma du filtre est donné en page suivante. La fonction de transfert d'un tel filtre est:

$$\frac{V_s}{V_e} = \left[\frac{1}{1 + (w/w_0)^4} \right]^{1/2} \quad \text{avec } w_0 = 2 \pi f_0$$

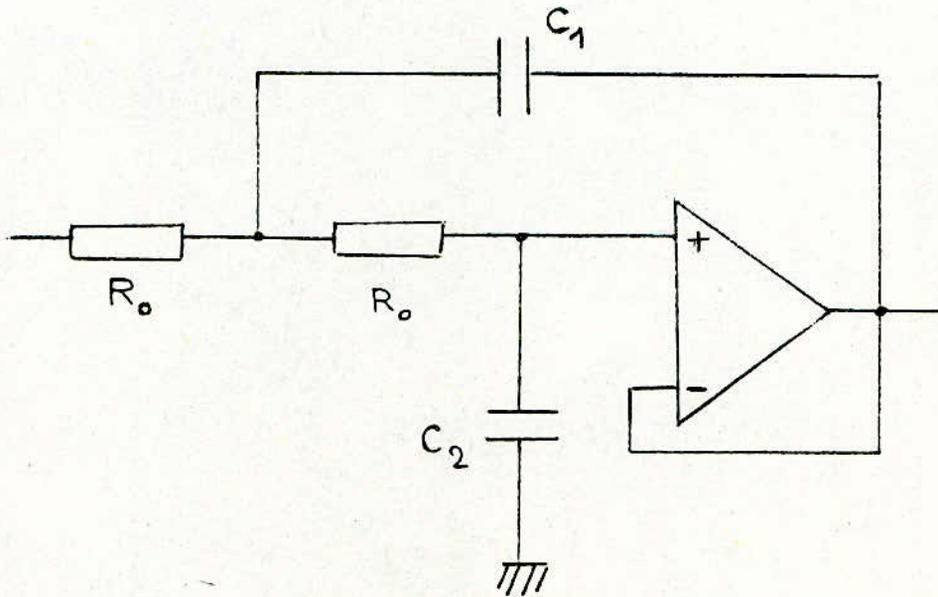
Pour le calcul des composants, on se fixe la valeur de la résistance R_0 puis on calcule la valeur de la capacité intermédiaire :

$$C_0 = \frac{1}{R_0 w_0}$$

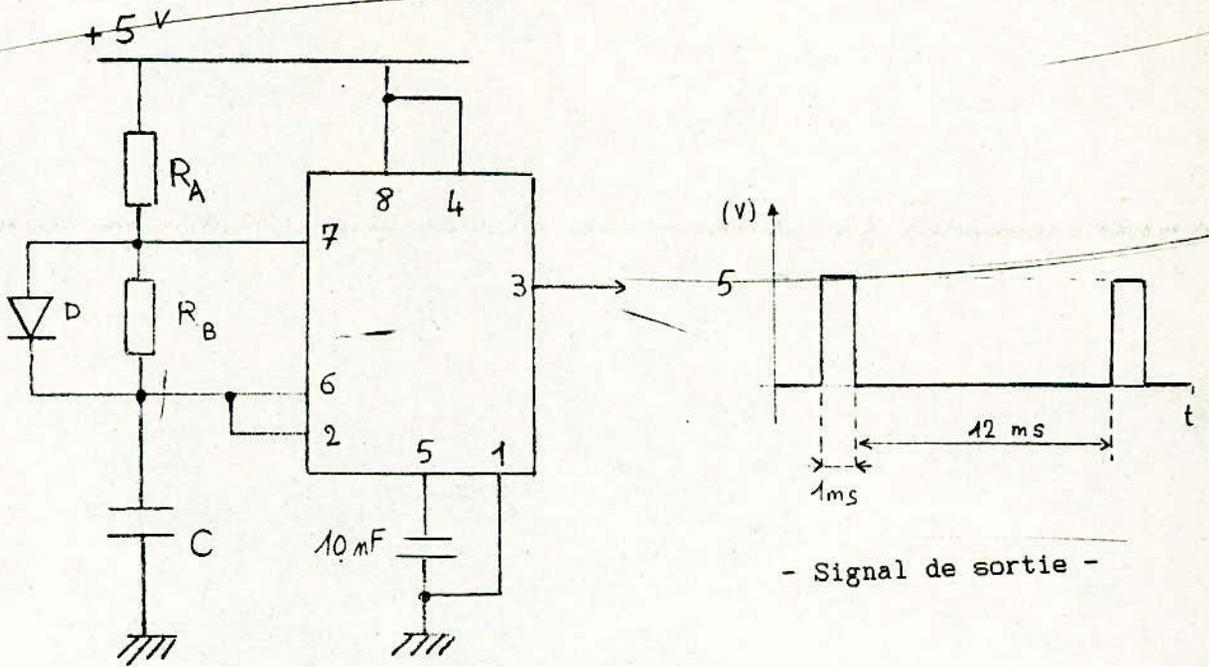
et enfin les valeurs des capacités C_1 et C_2 , donnée par

$$C_1 = 0.707 C_0$$

$$C_2 = 1.414 C_0$$



- Schéma du filtre passe-bas -



- Signal de sortie -

- Schéma de l'astable (NE 555) -

5 - Etude de l'Echantillonnage-Blocage

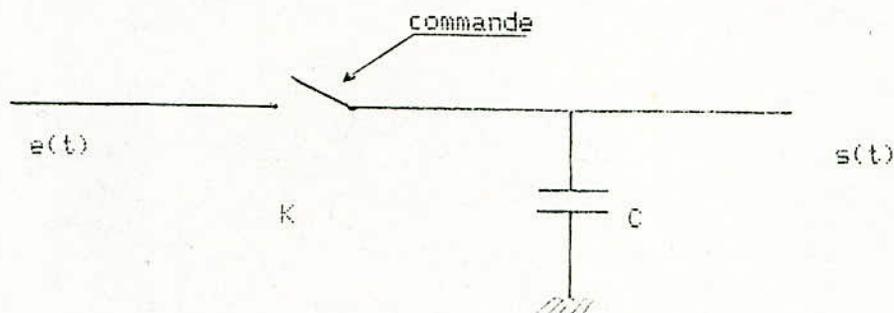
a/ Echantillonnage - Blocage

Echantillonner un signal $x(t)$ revient à prélever périodiquement une portion du signal. Dans le cas idéal la portion de signal prélevé est réduite à un point. Ce qui revient en fait à multiplier $x(t)$ par une suite d'impulsions de DIRAC périodiques.

L'échantillonneur-bloqueur réalise deux opérations successives; l'échantillonnage comme il a été décrit ci-dessus et le blocage qui consiste à stocker l'information analogique.

Les convertisseurs analogique-numérique sont souvent précédés par des échantillonneurs-bloqueurs qui permettent de maintenir la tension à coder constante pendant la durée de conversion.

Le principe de fonctionnement est simple; on considère un signal analogique $e(t)$, que l'on applique à l'entrée du circuit de la figure ci-dessous.



Celui-ci est constitué d'un interrupteur K et d'un condensateur C. L'interrupteur K est commandé par un signal logique à deux états.

Lorsque K est fermé, la tension $s(t)$ reproduit fidèlement $e(t)$; c'est la période d'échantillonnage. Le condensateur se charge.

Puis K se ^{s'ouvre} ferme. Le condensateur C isolé et dont nous admettons qu'il ne débite sur aucune charge, conserve en mémoire la tension qu'il recevait à l'instant de l'ouverture de K; c'est la phase de blocage.

Dans la réalité divers phénomènes parasites perturbent le fonctionnement théorique de l'échantillonneur-bloqueur. Ainsi il arrive que dans la phase d'échantillonnage la sortie $s(t)$ ne reproduise pas fidèlement $e(t)$; c'est le cas lorsque le gain diffère de l'unité. Dans la phase de blocage des erreurs peuvent être dues aux courants de fuite dans le condensateur qui conduisent à une décharge de C. Une autre source d'erreur provient de l'interrupteur K qui n'est jamais parfait.

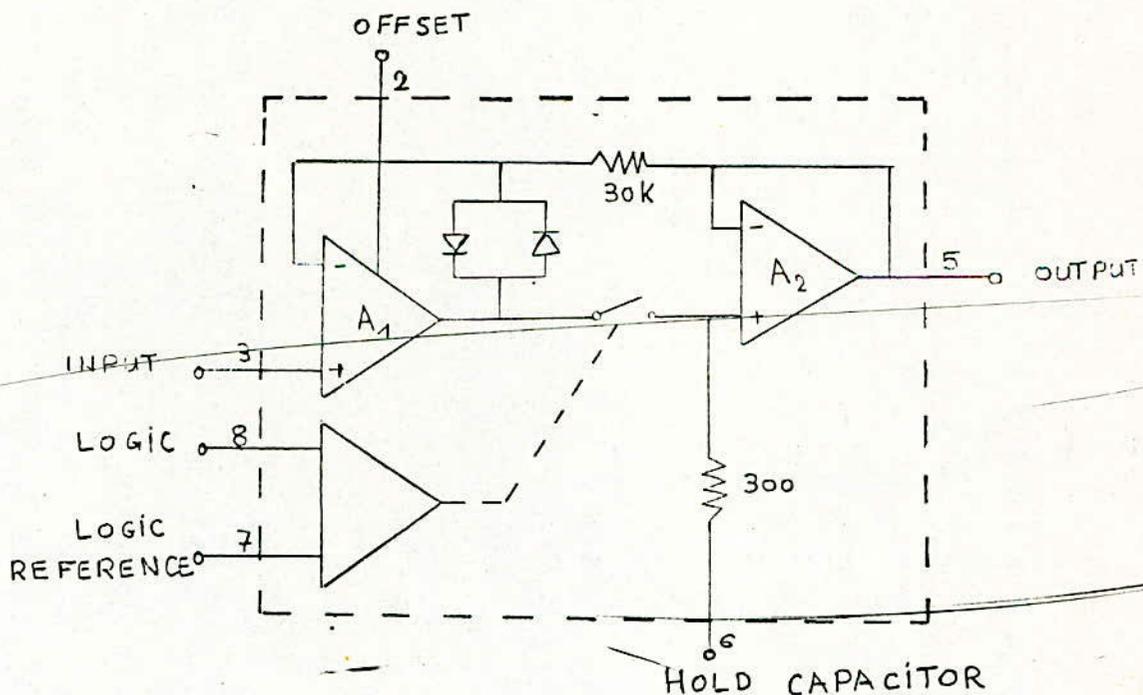
b/ L'échantillonneur bloqueur LF 398

Description générale:

Ce circuit intégré de National Semiconductor, a pour caractéristiques principales :

- tensions d'alimentation : de + ou - 5 V à + ou - 18 V
- erreur de gain (pour une résistance de charge de 10 K Ω): 0,004 %
- tension d'offset : 2 mV
- impédance d'entrée: 10⁴ M Ω
- large bande passante : 1 MHz
- la logique d'entrée est compatible TTL, PMOS, CMOS

Schéma fonctionnel interne:



Fonctionnement

L'amplificateur opérationnel A_1 est monté en suiveur, ce qui lui confère une faible impédance de sortie. De ce fait, lors de la phase d'échantillonnage le condensateur se chargera avec une faible constante de temps.

Le second amplificateur opérationnel est également monté en suiveur. Son impédance d'entrée, qui est très grande, devient alors l'impédance de charge du condensateur. Ce qui a pour effet de réduire la vitesse de décharge du condensateur et donc d'améliorer la capacité de l'échantillonneur-bloqueur à conserver en mémoire la tension appliquée.

Les diodes montées en tête-bêche ont pour rôle d'éviter la saturation de l'ampli-op A_1 .

Lors du blocage l'une des deux diodes conduit; la différence de tension $s(t)-e(t)$ entre la sortie et l'entrée est appliquée à la résistance de $30\text{ K}\Omega$. Celle-ci a pour rôle de limiter le courant qui passe entre A_1 et A_2 .

Pendant la phase d'échantillonnage les diodes sont bloquées, la tension à leurs bornes restant très inférieure au seuil de conduction. Le réglage de la tension d'offset sera réalisé en reliant la broche 2 avec un pont diviseur de tensions.

Principe de l'utilisation

Le signal de parole à analyser est considéré comme stationnaire sur des durées de 12 ms. L'échantillonneur-bloqueur sera donc utiliser pour réaliser des fenêtré d'analyse de 12 ms.

Le signal logique sera délivré par un NE 555 monté en astable de période:

$$T = (R_A + R_B) \cdot C \cdot \log 2$$

Cette horloge délivrera un signal rectangulaire dont les caractéristique sont:

$$t_{ON} = 1\text{ ms} \qquad t_{OFF} = 12\text{ ms}$$

Ce signal attaquera les quatorze échantillonneurs-bloqueurs pour que la prise d'échantillons se fasse au même instant sur tous les canaux.

6- Le Multiplexage

Chacun des quatorze canaux fournit une donnée analogique à convertir. Il aurait été possible d'utiliser quatorze convertisseurs analogiques-numériques, mais cette possibilité est déraisonnable du point de vue cout.

Le multiplexage permet une économie importante en CAN. En effet il permet de réduire leurs nombres de quatorze à un.

Le multiplexeur a pour rôle de scruter tour à tour les sorties des quatorze canaux.

Chaque échantillon étant bloquer durant 12 ms, il faut que toutes les sorties aient été scruter en, au plus, 12 ms.

L'opération de multiplexage sera réaliser par deux CD 4051, qui est multiplexeur/démultiplexeur analogique à huit canaux.

Le CD 4051 est un circuit intégré de technologie CMOS, dont les principales caractéristiques sont:

- faible impédance d'entrée
- grande impédance de sortie
- faible courant de fuite: 10 pA par canal.
- large plage de variation de la tension d'entrée:
 - entrée digitale : de 3 à 15 V
 - entrée analogique : 15 V crête à crête

Le CD 4051 possède une entrée de validation (inhibit) qui permet d'ouvrir simultanément tous les commutateurs.

La sélection d'entrée se fait par les broches 9, 10 et 11, à l'aide de tensions digitales. Dans notre cas la sélection d'entrée se fera par le biais du microprocesseur 6802. Celui-ci sera programmé pour générer les signaux de contrôle.

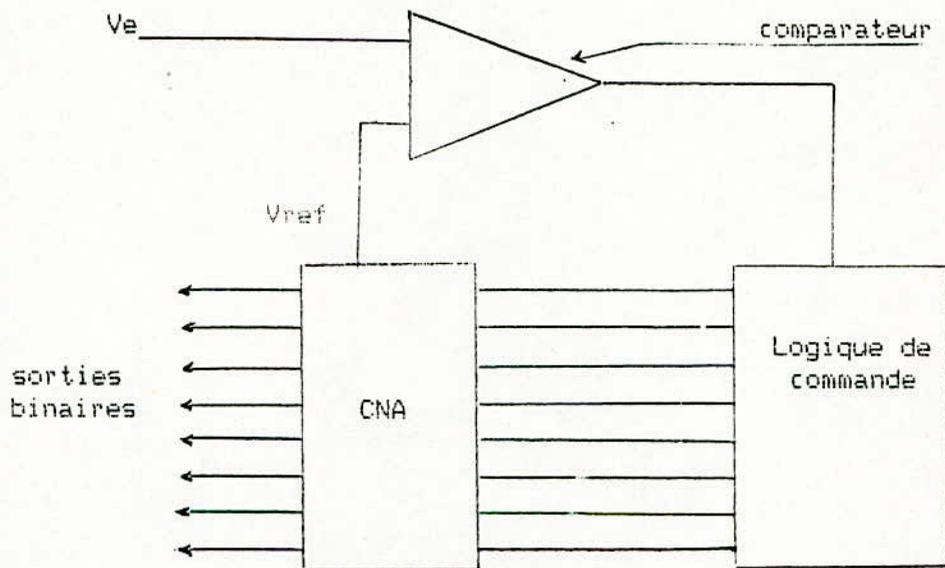
Nos deux CD 4051 nous permettrons de constituer un multiplexeur à seize entrées. De ce fait, deux entrées de l'un des deux multiplexeurs ne seront pas utilisées.

Le brochage de ce circuit intégré est donné en annexe I.

7- La Conversion Analogique Numérique

Les opérations de conversion analogique-numérique sont réalisées par l'ADC 804 qui est un convertisseur analogique numérique délivrant en sortie un signal sur 8 bits.

C'est un convertisseur à approximations successives. Rappelons que cette méthode de conversion est basée sur des comparaisons successives de la tension d'entrée avec les tensions correspondantes aux poids des différents bits.



synoptique d'un CNA à approximations successives

L'ADC 804 est circuit intégré de technologie CMOS, dont les principales caractéristiques sont:

- tension d'alimentation : + 5 V
- résolution de 8 bits
- temps de conversion : $100 \mu s$
- erreur de conversion : + ou - 1 LSB

Les signaux de commande de l'ADC 804 sont:

1- Partie digitale

- \overline{CS} : broche 1; active à l'état bas. Elle permet la validation des signaux \overline{RD} et \overline{WR} .
- \overline{RD} : broche 2; active à l'état bas. Les sorties du convertisseur sont validés lorsque \overline{RD} est activé.
- \overline{WR} : broche 3; le passage de l'état haut à l'état bas de cette broche remet à zéro en interne le CAN. La remontée de cette ligne déclenche la conversion.
- \overline{INTR} : broche 5; le passage à l'état bas de cette ligne signifie que la conversion est achevée. Il est alors possible de lire le mot binaire en sortie du CAN. Elle est remise automatiquement à 1 lors de la descente suivante de \overline{WR} .

2- Partie analogique

- V_{inc+} : broche 6; entrée de la valeur à convertir
- V_{inc-} et $V_{ref/2}$: broches 7 et 9; ces broches permettent de définir la plage de travail du convertisseur. Ainsi si $V_{ref/2}$ est laissé en l'air et V_{inc-} est à la masse le CAN convertira toute tension comprise entre 0 et 5 V. Si on relie la broche $V_{ref/2}$ à une source de tension externe de valeur V_x , celle-ci sera utilisée comme référence. Le circuit pourra alors mesurer toute tension allant de 0 à $2V_x$.

Les lignes DB0 à DB7 (broches 18, 17, 16, ..., 11) sont les sorties digitales, DB7 étant la ligne de poids fort.

Le brochage de ce circuit intégré est donné dans l'annexe I.

a/Codage de la parole

Le codage désigne le passage d'une représentation analogique du signal de parole à une représentation numérique, caractérisée par son débit, exprimé en éléments binaires par seconde (bit/s).

Par rapport à la représentation analogique la représentation numérique possède de nombreux avantages: la possibilité de stocker la parole, une grande facilité de traitement du signal et dans le réseau de télécommunications, une meilleure résistance aux perturbations apportées par les organes et les circuits de transmission, ainsi qu'une grande souplesse de manipulation des voies téléphoniques(multiplexage,...)

Cependant si le débit est inférieur à un certain seuil, la qualité subjective de la parole est dégradée. Pour obtenir un bon rapport qualité sur débit, il faut avoir recours à des équipements sophistiqués.

Le procédé de codage utilisé en télécommunication est le MIC (Modulation par Impulsions et Codage) qui nécessite un débit de 64 Kbit/s. Cette technique basée sur la statistique de la distribution des niveaux d'amplitude du signal, ne tient aucun compte de la redondance temporelle du signal.

b/Calcul du débit binaire du banc de filtres

Notre banc de filtres comprend:

- 14 canaux
- 8 bits par canal
- une fréquence d'analyse de 75 hz

Le débit sera donc donné par : $14 \times 8 \times 75 = 8400 \text{ bit/s}$

8- Interfaçage et Stockage des Données

Lors de l'acquisition des échantillons de parole, il devra exister une synchronisation entre la sélection d'une entrée d'un multiplexeur, le déclenchement de la conversion analogique-numérique et la période de fonctionnement de l'échantillonneur-bloqueur.

Chaque échantillon de parole, constitué de 14 données (chacune sur 8 bits), devra être stocker en mémoire.

Nous avons prévu d'utiliser le Kit Motorola 6802 pour ces opérations. Le programme en assembleur que nous avons établi se chargera de:

- sélectionner les entrées des multiplexeurs dans l'ordre désiré
- déclencher les conversions analogique-numérique
- stocker les résultats des conversions en RAM

Nous utiliserons donc l'interface parallèle programmable (PIA) 6821 pour relier notre analyseur avec le micro-processeur.

La PIA 6821 comporte deux ports, nommés A et B, qui peuvent être programmer en entrée ou en sortie, et cela indépendamment l'un de l'autre.

Le port A sera programmer en entrée et servira à recevoir les données en provenance du convertisseur analogique-numérique.

Le port B sera programmer en sortie, et nous n'utiliserons que cinq sorties sur huit:

- les trois premières serviront à sélectionner l'entrée d'un multiplexeur et seront donc relier aux broches 9, 10 et 11 de ces derniers,
- les deux dernières serviront à valider ~~alternativement le fonctionnement des multiplexeurs, et seront donc relier aux broches 6 (inhibit) de même multiplexeurs.~~

Ainsi, si l'on désire sélectionner, par exemple, le canal 2, il faudra que le port B soit porter à la valeur binaire 000010001.

Les échantillonneurs-bloqueurs passent en phase de blocage lorsque le timer délivre un front descendant. Ce même front descendant sera reçu sur la broche CB1 du PIA 6821, et déclenchera ~~par programme la sélection du canal 1, puis la conversion de la donnée de ce canal et enfin le stockage de celle-ci en RAM.~~ Puis le processus se poursuivra pour les autres canaux.

L'ordre de début de conversion sera donner par la broche CA1 du PIA. Elle sera donc relier à la broche WR de l'ADC 804.

Le passage au niveau bas de la sortie INTR du convertisseur indiquera au PIA que la conversion est achevée et que la donnée est disponible sur les sorties du convertisseur.

Pour une utilisation optimal de la place mémoire disponible sur le Kit Motorola, le stockage des données ne doit être activé que lorsqu'un signal vocal se trouve effectivement à l'entrée du microphone.

Le programme de stockage des données devra donc être déclencher que lorsque l'amplitude du signal reçue est supérieure à un certain seuil à définir. Ce seuil correspond à l'amplitude du bruit ambiant qui est très variable. Ces sources de bruit sont divers et nombreuses:

- bruits de conversations
- bruits de pas
- ouvertures et fermetures des portes
- bruits extérieurs (véhicules, ect...)
- climatisation
- tubes fluorescents

Après la conversion analogique-numérique, il faudra comparer la valeur obtenue avec la valeur moyenne du bruit ambiant qu'il faut mesurer. En utilisant un microphone unidirectionnel, nos mesures effectuées dans une salle peu bruyante ont donné une valeur variant autour de 50 mV, après la préamplification.

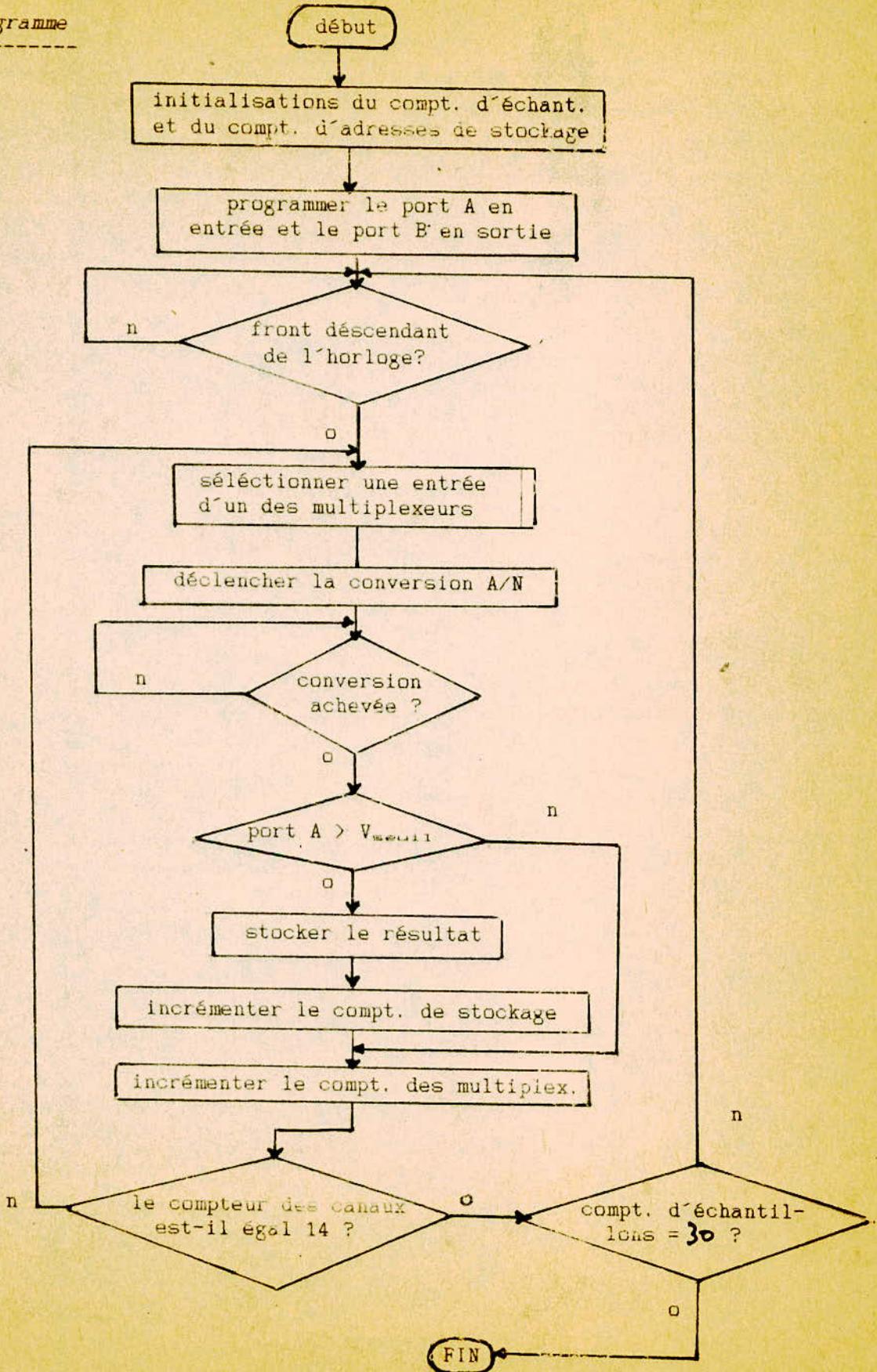
Le convertisseur analogique-numérique devant convertir toute tensions comprises entre 0 et 5V, la plus petite tension convertible est:

$$\frac{5}{256} = 19,5 \text{ mV}$$

Ce qui signifie que lorsqu'une tension de 19,5 mV se présente à l'entrée du CAN, sa sortie sera égale à 0000 0001

Le mot binaire correspondant à la tension de 50 mV est 0000 0011 ce qui en correspond à 03 en hexadécimal en arrondissant au bit supérieur.

Organigramme



Programme en assembleur pour l'acquisition des données:

LDX # \$ E000	7F 00 00
LDX # \$ E000	CE E0 00
CLR \$ E481	7F E4 81
CLR \$ E483	7F E4 83
CLR \$ E480	7F E4 80
LDA A # \$ FF	86 FF
STA A \$ E482	B7 E4 82
LDA A # \$ 04	86 04
STA A \$ E483	B7 E4 83
P4 LDA B # \$ 08	C6 08
P0 TST \$ E483	7D E4 83
BPL P0	2A FB
P2 STA B \$ E482	F7 E4 82
LDA A # \$ 34	86 34
STA A \$ E481	B7 E4 81
LDA A # \$ 3C	86 3C
STA A \$ E481	B7 E4 81
P1 TST \$ E481	7D E4 81
BPL P1	2A FB
LDA A \$ E480	B6 E4 80
CMP A # \$ 03	81 03
BLT P3	2D 03
STA A \$ 00, X	A7 00
INX	08
P3 INC B	5C
CMP B # \$ 0F	C1 0E
BNE P2	26 DF
LDX # \$ 0000	7F 00 00
LDX # \$ 0000	7F 00 00
LDX # \$ E1A4	B6 00
BNE P4	8E E1 A4
SWI	26 D0
	3F

Les données seront stockées à partir de l'adresse E000. Chaque échantillon sera composé de 14 données. L'adresse 0000 servira de compteur d'échantillons. Le programme s'achève après stockage du nombre d'échantillons désiré.

9- Problèmes rencontrés lors de la réalisation du banc de filtres

Lors des tests que nous avons effectués pour la réalisation de notre projet, divers difficultés sont apparues.

Ainsi au départ notre réalisation était basée sur le schéma électrique préparé le semestre précédent lors de la réalisation du détecteur de mélodie. C'était un schéma qui n'avait pas été testé mais qui nous à quand même servit à la fois de point de départ et de référence.

Avec les premières difficultés et après mûre réflexion, il apparut que divers modifications devaient être apportées à ce schéma.

Pour les filtres passes-bandes la structure qui consiste à mettre en cascade un filtre passe-haut avec un filtre passe-bas, avec laquelle nous avons débuté, ne répondait pas à nos exigences concernant le gain à la fréquence centrale et la largeur de bande. Pour l'échelonnement des filtres les valeurs du CNET sont données avec les fréquences de coupures à -6 dB. Nous avons donc déterminées les fréquences de coupures à -3 dB, afin de pouvoir calculer les composants, car les relations donnant les valeurs de ces derniers sont généralement liées avec la fréquence centrale et la largeur de bande à -3 dB.

La première modification que nous avons due apporter, fut donc de choisir une autre structure pour ces filtres (voir "choix des filtres" page 23). Après avoir opter pour la structure de Rauch d'autres problèmes apparurent.

Le calcul des composants se fait à partir des relations les liant aux caractéristiques de la réponse désirée. Cela revient en fait à résoudre un système d'équations où les inconnues sont les valeurs recherchées. Dans le cas où le nombre d'inconnues est supérieur au nombre d'équations on fixe la valeur d'un ou de deux composants, suivant le cas, puis en on déduit la valeur des autres. Le tout est que les valeurs de ces composants existent.

Dans le cas des filtres passes-bandes à structure de Rauch on se fixait les valeurs des capacités, puis on calculait les valeurs des résistances. Aussi, théoriquement, les 14 filtres passes-bandes peuvent être réalisés avec des capacités identiques ou dont les valeurs ne sont pas trop éloignées.

Pratiquement on constate que si les deux premiers filtres fonctionnent très bien avec des capacités de 220 nF, c'est à dire pour les fréquences comprises entre 250 et 650 Hz, les filtres suivants verront apparaitre des bruits indésirables en dehors de la bande désirée, lorsqu'ils sont réalisés avec les mêmes condensateurs.

Ces bruits disparaissent lorsque l'on utilise des filtres dont les composants ont été calculer avec des capacités de plus faibles valeurs. Ce phénomène se reproduit pour tous les autres canaux. Il faut à chaque fois tester plusieurs combinaisons de valeurs pour ne retenir que la plus efficace. Mais lorsque l'on diminue les valeurs des capacités on ne doit pas descendre en dessous d'une certaine limite, différente pour chaque canal, sinon d'autres défauts apparaissent (largeur de bande incontrôlable)

La seconde modification que nous avons apporté au premier schéma, concerne le signal de commande des échantillonneurs-bloqueurs. Le montage utilise quatorze échantillonneurs-bloqueurs à qui il faut envoyer un signal de commande logique. Nous avons préféré utiliser un seul circuit de commande (astable NE 555) relié aux quatorze échantillonneurs-bloqueurs, afin que la prise d'échantillons se fasse au même instant sur tous les canaux. Plusieurs timer, même s'ils sont identiques, pourraient générer des signaux non synchronisés, qui déclencheraient des prises d'échantillons à des instants différent sur chaque canal.

Pour l'échantillonneur-bloqueur on notera l'importance de la capacité de charge. Il faut tenir compte de deux caractéristiques:

- la valeur de la capacité, qui joue sur le temps d'acquisition
- les courants de fuites dans le condensateur

Nous utilisons des capacités au tantale présentant de faibles courants de fuites.

Un réglage de l'offset est indispensable pour le bon fonctionnement du circuit intégré.

Pour la conversion analogique-numérique nous avons également apporté une modification par rapport au schéma initial.

Tenant compte du fait que le signal à convertir est bloqué par l'échantillonneur-bloqueur pendant 12 ms , et sachant que le temps de conversion du convertisseur analogique-numérique est de 100 μ s , nous disposons donc de suffisamment de temps pour effectuer la conversion des quatorze données (une par canal) à l'aide d'un seul convertisseur analogique-numérique et de multiplexeurs. Ce qui représente une économie en composant ainsi qu'un encombrement moins important du circuit imprimé.

Notre réalisation sera partagée sur deux plaques de circuits imprimés. La première plaque contient:

- les huit premiers canaux
- le timer
- les multiplexeurs
- le convertisseur analogique-numérique

La seconde plaque contient les six derniers canaux .

Calculs des composants

Le banc de filtres

Canal 1	$F_c = 335 \text{ Hz}$	$\Delta F = 100 \text{ Hz}$	

	$C_3 = 220 \text{ nF}$	$C_4 = 220 \text{ nF}$	
	$R_1 = 7,5 \text{ K}\Omega$	$R_2 = 0,33 \text{ K}\Omega$	$R_5 = 15 \text{ K}\Omega$
Canal 2	$F = 541 \text{ Hz}$	$\Delta F = 100 \text{ Hz}$	

	$C_3 = 220 \text{ nF}$	$C_4 = 220 \text{ nF}$	
	$R_1 = 33,8 \text{ K}\Omega$	$R_2 = 0,12 \text{ K}\Omega$	$R_5 = 15 \text{ K}\Omega$
Canal 3	$F_c = 743 \text{ Hz}$	$\Delta F = 100 \text{ Hz}$	

	$C_3 = 33 \text{ nF}$	$C_4 = 47 \text{ nF}$	
	$R_1 = 33,8 \text{ K}\Omega$	$R_2 = 0,33 \text{ K}\Omega$	$R_5 = 82 \text{ K}\Omega$
Canal 4	$F_c = 945 \text{ Hz}$	$\Delta F = 100 \text{ Hz}$	

	$C_3 = 33 \text{ nF}$	$C_4 = 47 \text{ nF}$	
	$R_1 = 33,8 \text{ K}\Omega$	$R_2 = 0,202 \text{ K}\Omega$	$R_5 = 82 \text{ K}\Omega$
Canal 5	$F_c = 1168 \text{ Hz}$	$\Delta F = 120 \text{ Hz}$	

	$C_3 = 33 \text{ nF}$	$C_4 = 47 \text{ nF}$	
	$R_1 = 28,2 \text{ K}\Omega$	$R_2 = 0,162 \text{ K}\Omega$	$R_5 = 68,6 \text{ K}\Omega$
Canal 6	$F_c = 1442 \text{ Hz}$	$\Delta F = 150 \text{ Hz}$	

	$C_3 = 33 \text{ nF}$	$C_4 = 47 \text{ nF}$	
	$R_1 = 22,56 \text{ K}\Omega$	$R_2 = 0,13 \text{ K}\Omega$	$R_5 = 54,1 \text{ K}\Omega$
Canal 7	$F_c = 1744 \text{ Hz}$	$\Delta F = 150 \text{ Hz}$	

	$C_3 = 33 \text{ nF}$	$C_4 = 47 \text{ nF}$	
	$R_1 = 22,56 \text{ K}\Omega$	$R_2 = 0,092 \text{ K}\Omega$	$R_5 = 54,1 \text{ K}\Omega$
Canal 8	$F_c = 2045 \text{ Hz}$	$\Delta F = 150 \text{ Hz}$	

	$C_3 = 23 \text{ nF}$	$C_4 = 23 \text{ nF}$	
	$R_1 = 40,8 \text{ K}\Omega$	$R_2 = 0,144 \text{ K}\Omega$	$R_5 = 92,2 \text{ K}\Omega$
Canal 9	$F_c = 2345 \text{ Hz}$	$\Delta F = 150 \text{ Hz}$	

	$C_3 = 8,2 \text{ nF}$	$C_4 = 8,2 \text{ nF}$	
	$R_1 = 148 \text{ K}\Omega$	$R_2 = 0,243 \text{ K}\Omega$	$R_5 = 256 \text{ K}\Omega$

<u>Canal 10</u>	$F_c = 2645 \text{ Hz}$	$\Delta F = 150 \text{ Hz}$	
	$C_3 = 8,2 \text{ nF}$	$C_4 = 8,2 \text{ nF}$	
	$R_1 = 148 \text{ K}\Omega$	$R_2 = 0,186 \text{ K}\Omega$	$R_5 = 256 \text{ K}\Omega$
<u>Canal 11</u>	$F_c = 2496 \text{ Hz}$	$\Delta F = 150 \text{ Hz}$	
	$C_3 = 6,8 \text{ nF}$	$C_4 = 6,8 \text{ nF}$	
	$R_1 = 156 \text{ K}\Omega$	$R_2 = 0,182 \text{ K}\Omega$	$R_5 = 312 \text{ K}\Omega$
<u>Canal 12</u>	$F_c = 3294 \text{ Hz}$	$\Delta F = 200 \text{ Hz}$	
	$C_3 = 4,7 \text{ nF}$	$C_4 = 2 \text{ nF}$	
	$R_1 = 358 \text{ K}\Omega$	$R_2 = 0,353 \text{ K}\Omega$	$R_5 = 565 \text{ K}\Omega$
<u>Canal 13</u>	$F_c = 3695 \text{ Hz}$	$\Delta F = 200 \text{ Hz}$	
	$C_3 = 4,7 \text{ nF}$	$C_4 = 2 \text{ nF}$	
	$R_1 = 279 \text{ K}\Omega$	$R_2 = 0,316 \text{ K}\Omega$	$R_5 = 565 \text{ K}\Omega$
<u>Canal 14</u>	$F_c = 4095 \text{ Hz}$	$\Delta F = 200 \text{ Hz}$	
	$C_3 = 4,7 \text{ nF}$	$C_4 = 2 \text{ nF}$	
	$R_1 = 300 \text{ K}\Omega$	$R_2 = 0,24 \text{ K}\Omega$	$R_5 = 565 \text{ K}\Omega$

Détecteur d'enveloppe

diode 1N 4148 , $C = 10 \mu\text{F}$ $R = 12 \text{ K}\Omega$

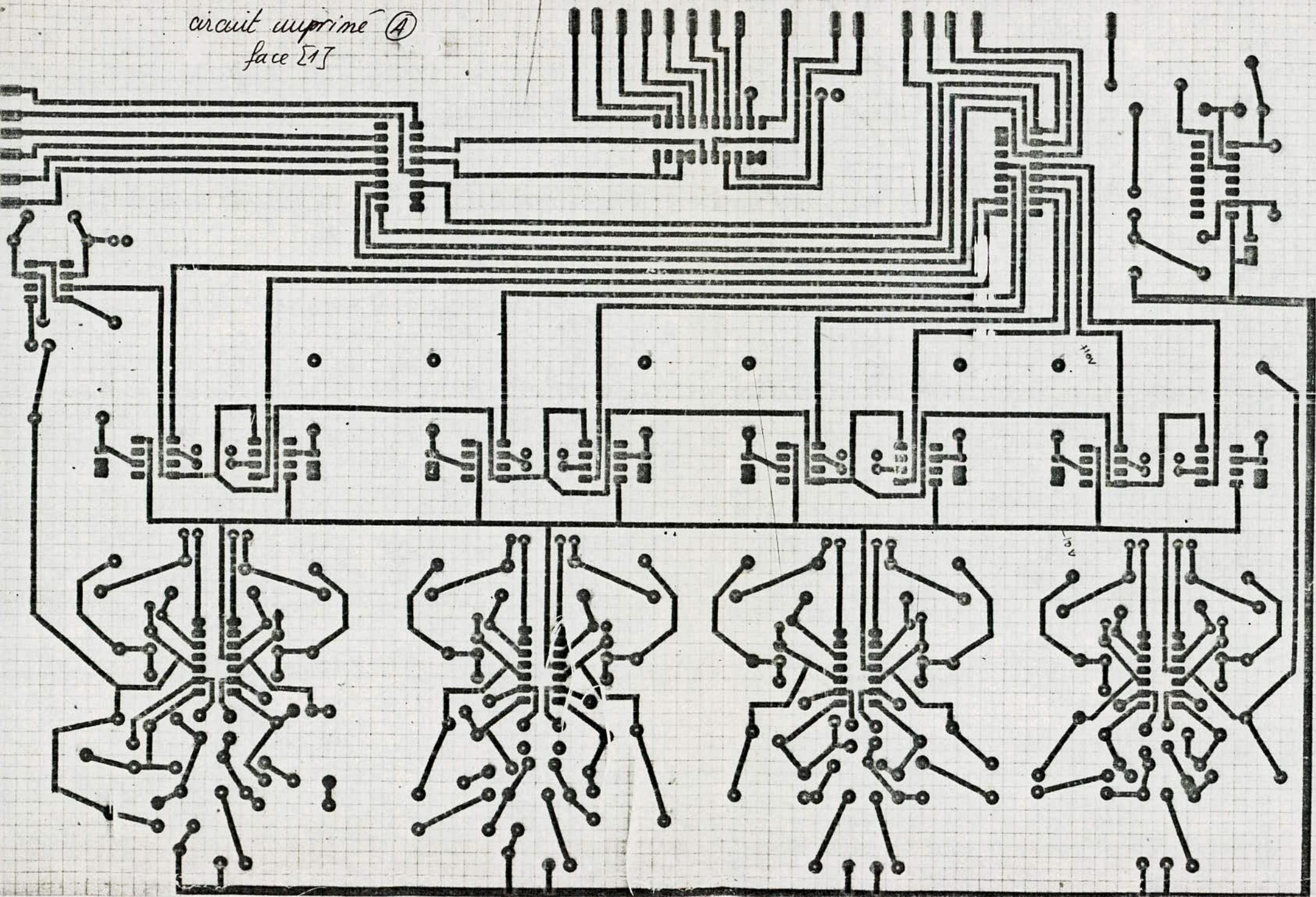
Filtre passe bas

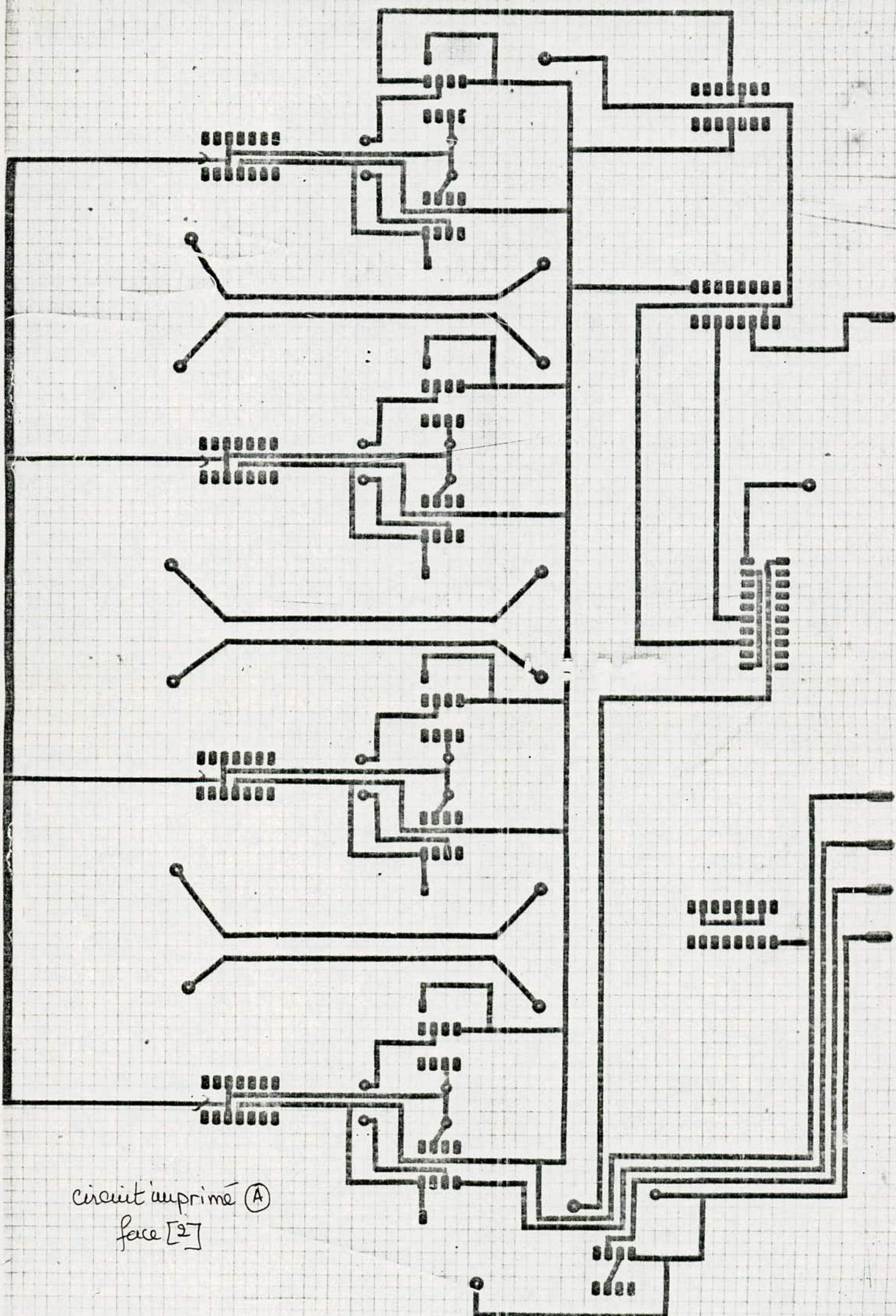
$R_0 = 10 \text{ K}\Omega$ $C_1 = 220 \text{ nF}$ $C_2 = 470 \text{ nF}$

Echantillonneur-bloqueur

ajustable de 1K ; $R = 12 \text{ K}\Omega$; $C = 220 \text{ nF}$

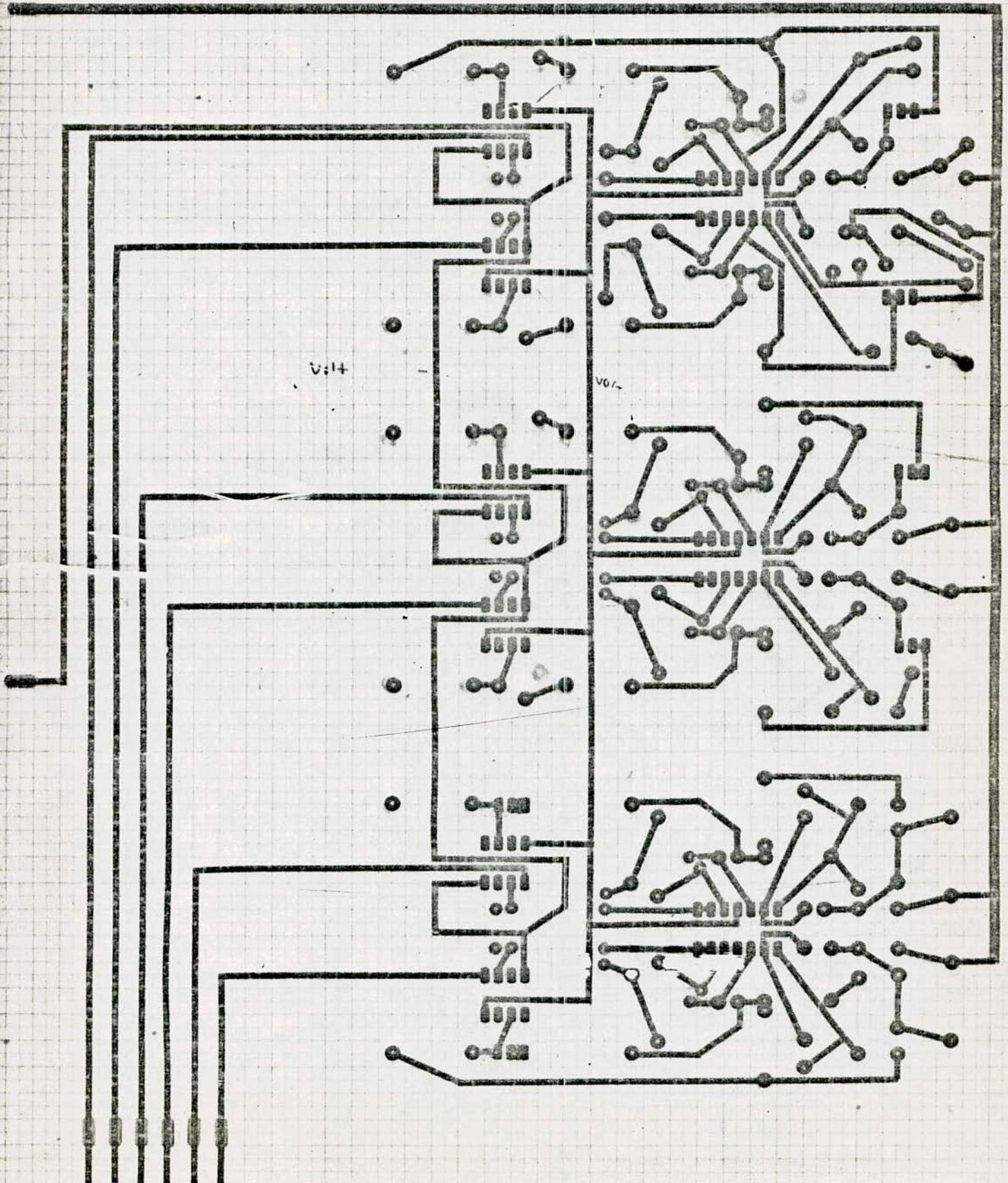
circuit imprimé (A)
face [17]



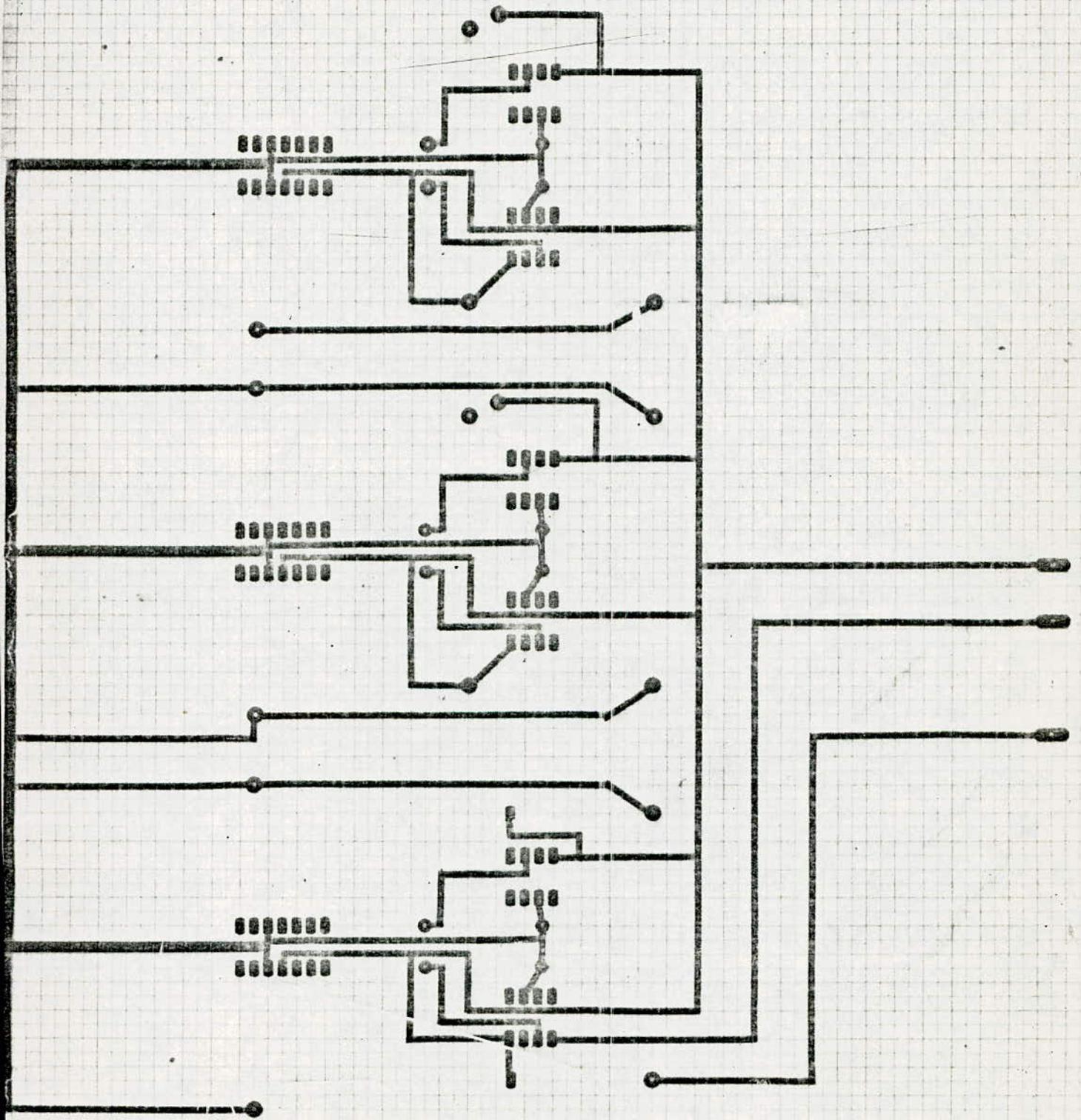


circuit imprimé (A)
face [2]

circuit imprimé (B)
face (D)



circuit imprimé (B) face [2]



CONCLUSIONS GÉNÉRALES

Comme nous l'avons déjà dit ce travail entre dans le cadre de la réalisation d'un vocodeur à canaux à l'ENP. Ainsi, après que le détecteur de mélodie ait été réalisé, ce travail constitue la 2^{ème} partie de la réalisation de ce vocodeur.

Mais, à plus long terme, il entre dans le cadre de la reconnaissance de la parole continue. Il est évident qu'il reste encore beaucoup à faire. Ainsi l'apparition de processeurs spécialisés en traitement numérique du signal sur le marché des composants (ex: INTEL 2920, TMS 320), doit mener au remplacement de l'analyseur analogique par un analyseur numérique intégré, ce qui augmenterait la qualité de la reconnaissance.

La réalisation du banc de filtre analogique nous a permis entre autre de constater l'application des théories du traitement du signal au domaine, à la fois concret et complexe, de la parole, mais aussi d'approfondir nos connaissances dans le domaine des systèmes d'acquisition de données.

Le retard pris à cause de la difficulté d'obtention de certains composants, nous a empêché d'exécuter tous les essais que l'on aurait souhaiter faire.

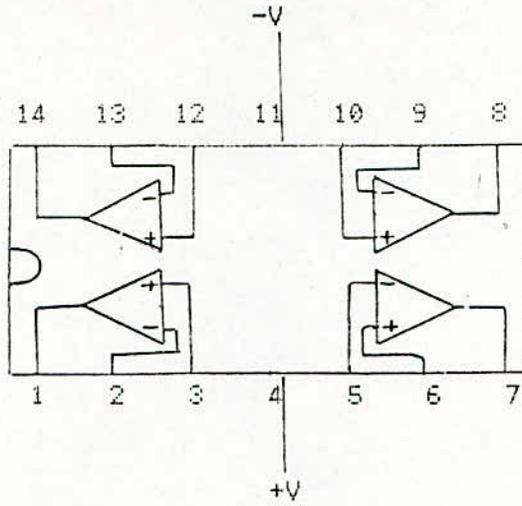
La réalisation d'un détecteur de voisement constituera la dernière étape de la réalisation de l'analyseur du vocodeur. Il restera alors à les faire fonctionner ensemble, puis de réfléchir à la réalisation d'un synthétiseur.

Les domaines d'applications futurs de la reconnaissance de la parole semblent sans limites, car déjà aujourd'hui elle touche à de nombreux secteurs de la vie active (annexe 3).

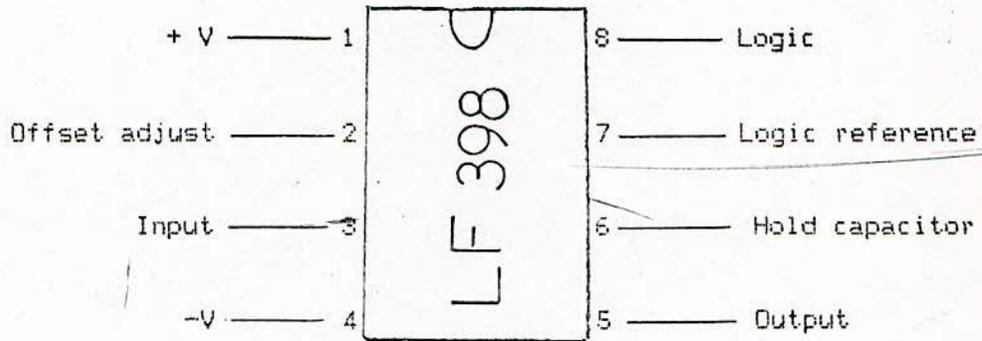
ANNEXE I

BROCHAGE DES CIRCUITS INTEGRES

1/ Le quadruple ampli-op LM 324



2/ L'échantillonneur-bloqueur LF 398



3/ Le multiplexeur/demultiplexeur CD 4051

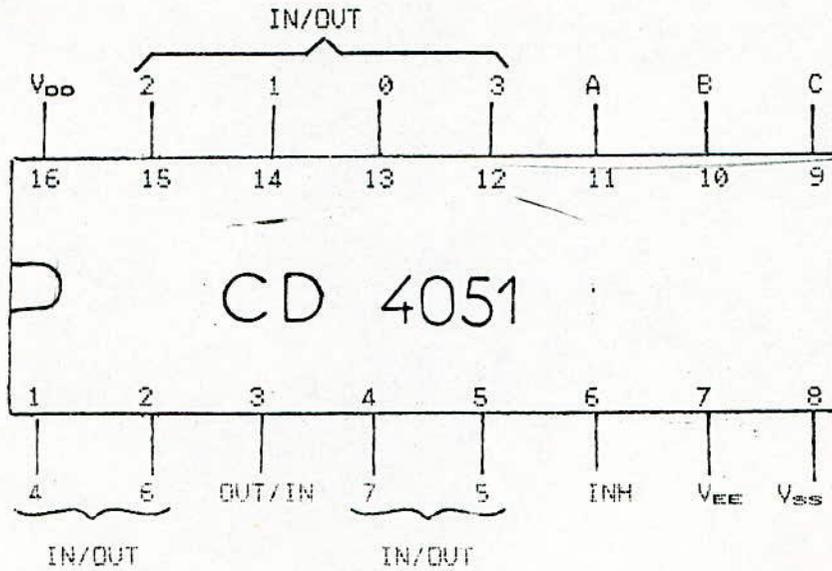
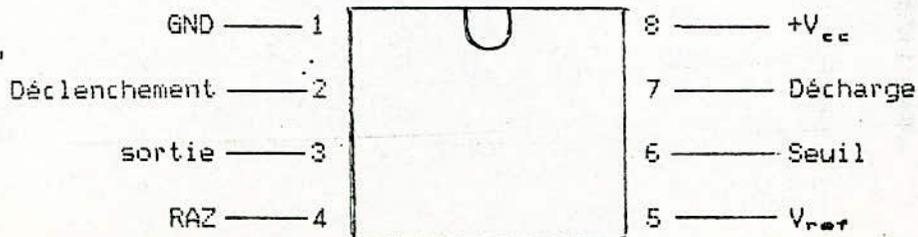


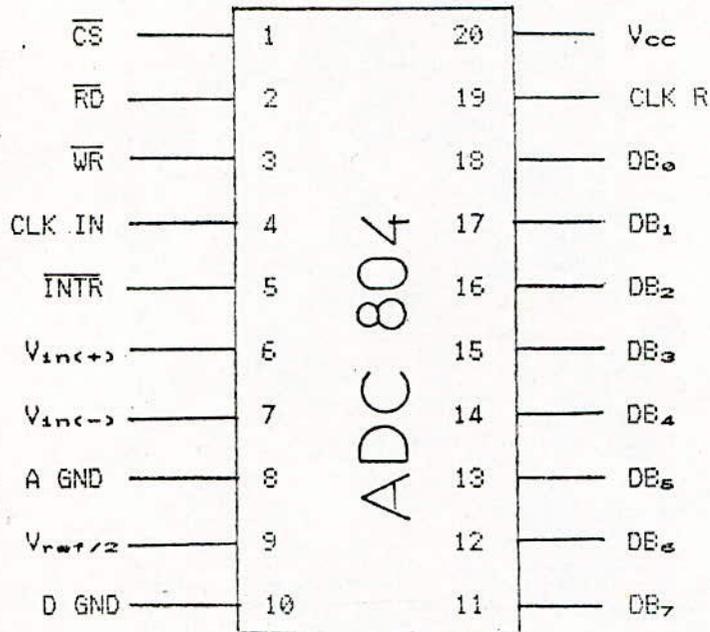
Table de vérité

Inhibit	C	B	A	Canal sélectionné
0	0	0	0	0
0	0	0	1	1
0	0	1	0	2
0	0	1	1	3
0	1	0	0	4
0	1	0	1	5
0	1	1	0	6
0	1	1	1	7
1	*	*	*	aucun

4/ Le timer NE 555



5/ Le convertisseur analogique numérique ADC 804



3/ Le fenêtrage

Pour étudier le signal temporelle sur une durée limitée, il faut le multiplier par une fenêtre temporelle. Cette limitation en temps implique une distorsion au niveau spectral: apparition de lobes secondaires et élargissement du lobe principal. Il existe plusieurs types de fenêtres: Hamming, Hanning et rectangulaire (voir J.MAX et KUNT)

Le choix d'une fenêtre plutôt qu'une autre est un compromis entre l'élargissement du lobe principal et l'énergie des lobes secondaires.

Fenêtre de Hanning

Elle est donnée par: $x(k) = \frac{1}{2} \left(1 - \cos\left(\frac{2k}{N-1}\right) \right)$ pour $0 < k < N-1$

Fenêtre de Hamming

Elle est donnée par: $x(k) = 0,54 - 0,46 \cos\left(\frac{2k}{N-1}\right)$ pour $0 < k < N-1$

4/ Moyenne temporelle

Elle est définie par:

$$M = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{t_0}^{t_0+T} x(t) dt$$

Pour des signaux de durée finie :

$$M = \frac{1}{T} \int_{t_0}^{t_0+T} x(t) dt$$

ANNEXE

II

Les notions de théorie du signal que nous avons utilisé dans notre étude seront définies dans cette annexe.

1/ Transformée de Fourier

La transformée de Fourier de la fonction $x(t)$ de la variable réelle t est une fonction de la variable réelle ω , $X(j\omega)$ telle que:

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt$$

comme on a: $\omega = 2\pi \cdot f$

on obtient alors:

$$X(j\omega) = x'(f) = \int_{-\infty}^{+\infty} x(t) e^{-2j\pi ft} dt$$

La transformée inverse de la fonction $X(j\omega)$ de la variable réelle ω est une fonction de la variable réelle t , $x(t)$ telle que:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\omega) e^{j\omega t} d\omega$$

2/ Transformée de Fourier discrète

Soit le signal $x(t)$ échantillonné à la fréquence $1/T$ et de durée $T = N \cdot t$ où N est le nombre d'échantillons.

La transformée de Fourier discrète de $x(t)$ est alors:

$$X(m) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{-2j\pi mk/N} \quad \begin{array}{l} m=0, \dots, N-1 \\ k=0, \dots, N-1 \end{array}$$

La transformée de Fourier discrète inverse s'écrit :

$$x(m) = \sum_{k=0}^{N-1} X(k) e^{+2j\pi mk/N} \quad \begin{array}{l} m=0, \dots, N-1 \\ k=0, \dots, N-1 \end{array}$$

3/ Convolution

Le produit de convolution de deux fonctions $x(t)$ et $y(t)$ est l'intégrale:

$$u(t) = \int_{-\infty}^{+\infty} x(t') \cdot y(t-t') dt'$$

On le note : $u(t) = x(t) * y(t)$

Transformée de Fourier d'une convolution

La transformée de Fourier $u(f)$ est: $u'(f) = x'(f) \cdot y'(f)$

ANNEXE III

UTILISATIONS DE LA RECONNAISSANCE DE LA PAROLE

Contrôle de qualité, inspection des chaînes de montages

- inspection d'écrans de téléviseurs (l'utilisateur peut se déplacer, manipuler les écrans, lire des mesures, les introduire vocalement)
- inspection des chaînes d'assemblages des véhicules (GENERAL MOTORS)
- inspection des chaînes d'assemblage de circuit intégrés (LOCKHEED)

Enseignement assisté par ordinateur

- pilotage (LOGICON, AFTS "automated flight training system")

Médecine

- diagnostic aidé par ordinateur
- choix de médicaments (accès aux banque de données)

Contrôle à distance

- industrie du pétrole (EXXON, ELF)
- contrôle de microscope chirurgical
- commande d'un robot agricole

Aide aux handicapés

- aide à la lecture labiale
- commande de voitures (KEMPF)

Téléphone

- carte de crédit téléphonique (Etat d'ILLINOIS pour 4000 personnes)
- transactions bancaires (VERBEX)

Machine à écrire automatique

- projet MATSUSHITA (à partir de 68 syllabes japonaises)

Calculatrice vocale

- SANYO-HITACHI

BIBLIOGRAPHIE

- 1/ J. MAX
Méthode et techniques de traitement du signal et application aux mesures
tome I et II . MASSON. (1981)
- 2/ M. KUNT
Traitement numérique des signaux. DUNOD. (1981)
- 3/ J. AUVRAY
Electronique des signaux analogiques. DUNOD
- 4/ M. AUMIAUX
Pratique de l'électronique
Tome I " L'amplification "
Tome II " Comparateurs, filtres actifs, techniques numériques "
MASSON. (1977)
- 5/ J-P. GUILLET
Réseaux actifs. Filtres MASSON. (1972)
- 6/ A. CHAMPENOIS
Electronique industrielle E. R. P. I. (1984)
- 7/ J-S. LIENARD
Les processus de la communication parlée MASSON. (1977)
- 8/ M et J. GUIBERT
Les ordinateurs qui parlent. BELFOND. (1986)
- 9/ M. FERRETTI et F. CINARE
Synthèse, reconnaissance de la parole EDITEST. (1983)
- 10/ ~~Thèse de Magister en Sciences Physiques~~
~~" Conception et réalisation d'un système d'acquisition par~~
~~microprocesseur du signal de la parole "~~
~~soutenue par M^{me} YAHIAT à l'USTHB (Juin 1986)~~

11/ M. AUMIAUX

L'emploi des microprocesseurs

MASSON. (1982)

12/ Thèse de projet de fin d'études

" Contribution à la réalisation d'un vocodeur à canaux "
soutenue par S. MEKAOUI et K. SEMMAR à l'ENP (Juin 1987)

Revue et Notes Techniques

1-Toute l'électronique:

N° 489. Dec 1983 ; N° 490. Jan 1984 ; N° 496. Sep 1984

2-L'onde électrique:

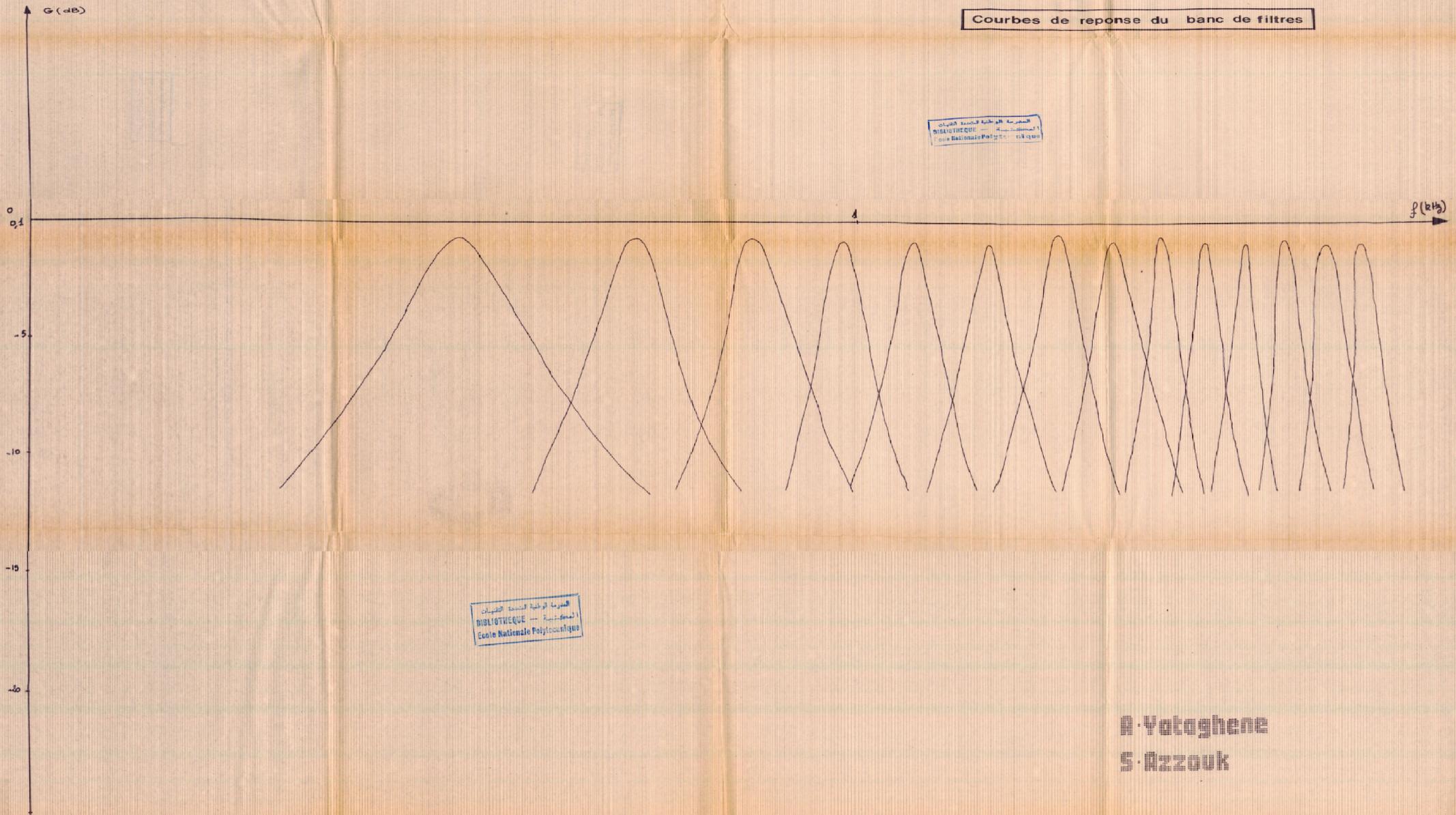
"La numérisation de la parole". vol. 61 , 1981 pages 26 à 53

3- F. ZURCHER

"Le vocodeur à canaux une nouvelle jeunesse?"

NT/LAA/TSS/43. Dec. 1983

Courbes de reponse du banc de filtres

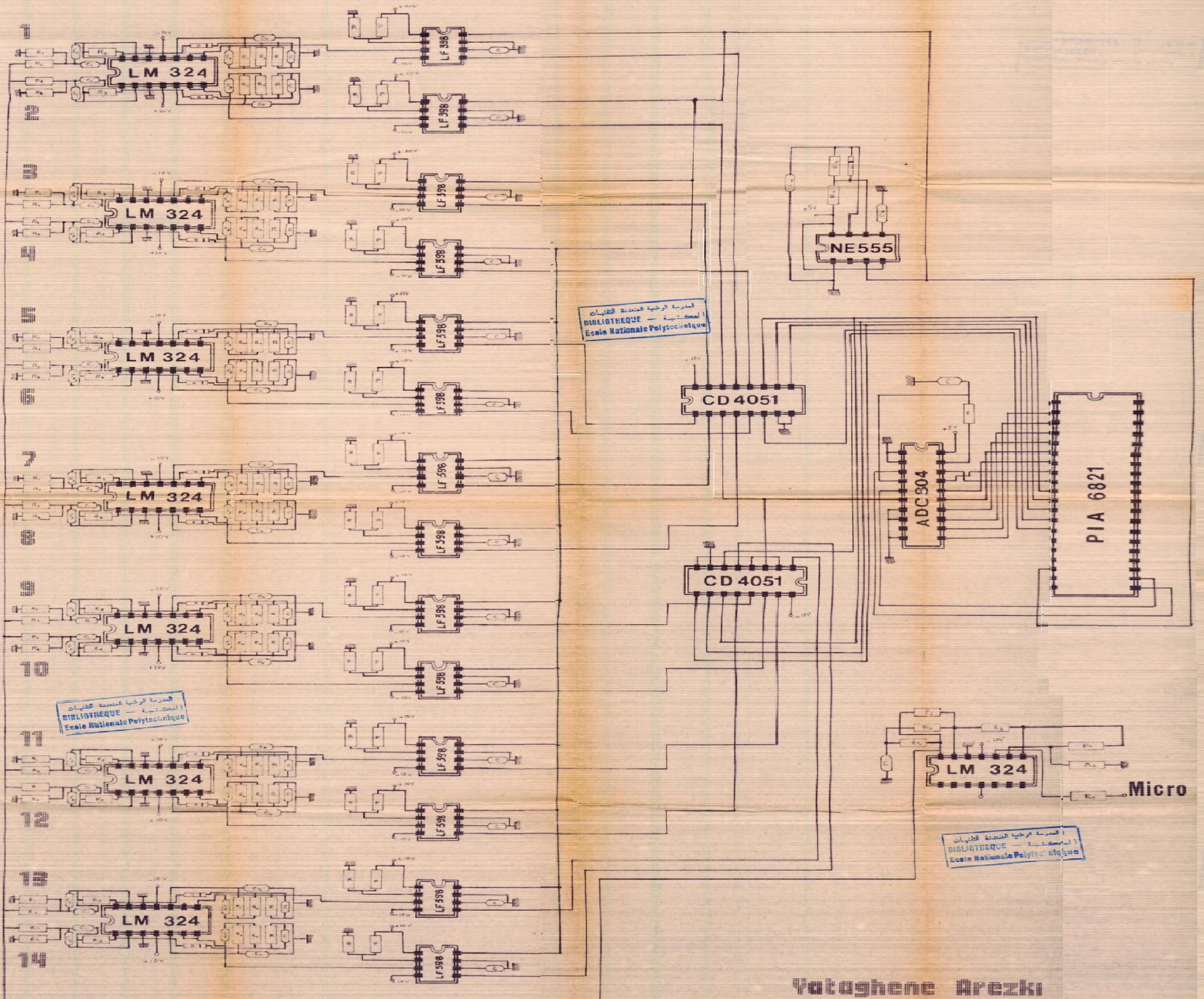


المعهد الوطني للتكنولوجيا
BIBLIOTHEQUE
Ecole Nationale Polytechnique

المعهد الوطني للتكنولوجيا
BIBLIOTHEQUE
Ecole Nationale Polytechnique

A. Yatoghene
S. Azzouk

Circuit Electrique De L'analyseur A Banc De Filtres



المدرسة الوطنية المتعددة التخصصات
BIBLIOTHEQUE
المعهد
Ecole Nationale Polytechnique

المدرسة الوطنية المتعددة التخصصات
BIBLIOTHEQUE
المعهد
Ecole Nationale Polytechnique

المدرسة الوطنية المتعددة التخصصات
BIBLIOTHEQUE
المعهد
Ecole Nationale Polytechnique

Yatoghene Avezki
Rzzouk Simone