

# NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

---

STRUCTURES DES FONCTIONS D'AMBIGUITE ET THEORIE  
DES CATASTROPHES : APPLICATION AUX SIGNAUX DE PAROLE

M. BOULOGNE

P. FLANDRIN B. ESCUDIE

Institut de la Communication Parlée  
L.A. 368 - 23 rue des Martyrs - GRENOBLE

Laboratoire Traitement du Signal  
I.C.P.I. - 25 rue du Plat - LYON

---

## RESUME

En analyse de la parole, l'étude de l'évolution temporelle des différentes composantes des signaux nécessite l'emploi de représentations en temps et fréquence. Une des descriptions possibles repose sur la notion de fonction d'ambiguïté. La théorie des catastrophes permet de décrire commodément le comportement des fonctions d'ambiguïté des signaux multicomposantes : leur structure se déduit géométriquement des lois temps-fréquence de chacune des composantes. Quelques exemples ayant trait à l'analyse des signaux de type "parole" illustrent la technique décrite. Ces résultats permettent d'isoler la composante fondamentale et les fréquences les plus proches des formants. L'existence dans la surface d'ambiguïté de "termes fantômes" d'amplitude non constante révèle la composante fondamentale par leur position fréquentielle et leur périodicité temporelle.

## SUMMARY

In speech processing, the analysis of multicomponent time-varying signals needs the use of joint time-frequency representations. One among the various possibilities is based upon the ambiguity function. Catastrophe theory allows an easy description of the behaviour of such ambiguity functions in the multicomponent cases : the geometry of their structure depends on the time-frequency laws of each component. The proposed technique is illustrated on a few speech signals. These results may be used to detect the pitch and the frequencies close to the formants. Such a detection relies on the frequency position and the time periodicity of "ghost" cross-terms of the ambiguity surface.



A. FONCTIONS D'AMBIGUITE, INTEGRALES OSCILLANTES ET THEORIE DES CATASTROPHES.

1°) Définitions.

On appellera fonction d'interambiguïté de deux signaux  $Z_k(t)$  et  $Z_l(t)$  la quantité :

$$\chi_{Z_k Z_l}(v, \tau) = \int_{\mathbb{R}} Z_k(t) Z_l^*(t - \tau) e^{-i2\pi v \tau} dt \quad (1)$$

Le cas  $Z_l = Z_k$  correspond à la fonction d'(auto)ambiguïté (f.d.a.), au sens de Woodward, dont les propriétés sont bien connues [1]. On sait en particulier que son caractère sesquilineaire implique que si  $Z = \sum_k Z_k$  :  $\chi_{ZZ}(v, \tau) = \sum_k \sum_l \chi_{Z_k Z_l}(v, \tau)$ . (2)

En outre, son module est maximal à l'origine et, par transformation de Fourier, un lien étroit s'établit entre f.d.a. et représentations conjointes (en particulier au sens de Wigner-Ville [2]).

2°) Intégrales oscillantes et théorie des catastrophes.

Dans le cas le plus général des signaux complexes du type :

$$Z_k(t) = A_k(t) e^{i\phi_k(t)},$$

on peut réécrire (1) suivant :

$$\chi_{Z_k Z_l}(v, \tau) = \left( \chi_{A_k A_l}(m, \tau) * I_{k1}(m, \tau) \right)_m = v \quad (3-a)$$

$$\text{avec : } I_{k1}(v, \tau) = \int_{\mathbb{R}} e^{iL_{k1}(t; v, \tau)} dt \quad (3-b)$$

$$\text{et : } L_{k1}(t; v, \tau) = \phi_k(t) - \phi_k(t - \tau) - 2\pi v \tau. \quad (3-c)$$

Séparant par (3-a) l'influence de l'amplitude et de la phase, il a été vu précédemment [2] que  $I_{k1}$  prenait ainsi la forme d'une intégrale oscillante pour laquelle diverses méthodes d'approximation pouvaient être utilisées. De plus, la dimension du problème (une variable d'intégration (ou variable d'état)  $t$  pour deux paramètres externes (ou variables de contrôle)  $v$  et  $\tau$ ) assure que les structures possibles de  $I_{k1}$  dans le plan  $(v, \tau)$  se réduisent à un petit nombre de cas typiques. Ce résultat est directement issu de l'interprétation d'une intégrale oscillante en termes de théorie des catastrophes [3], le comportement de  $I_{k1}$  étant essentiellement décrit par l'étude des singularités de la variété critique  $\frac{\partial L_{k1}}{\partial t}$ . Renvoyant pour plus de détails à [2], nous pouvons résumer de la façon suivante la construction schématique de  $\chi_{Z_k Z_l}$  :

a) Le domaine essentiel d'existence de  $I_{k1}$  est donné par l'ensemble des points stationnaires de  $I_{k1}$  i.e. :

$$(v, \tau) \in \mathbb{R}^2 / \frac{\partial L_{k1}}{\partial t}(t; v, \tau) = 0, \quad \forall t \in (T)$$

(T) étant le support temporel des signaux. Ceci conduit à écrire, au sens de l'approximation faite :

$$I_{k1} \neq 0 \iff v = v_k(t) - v_l(t - \tau), \quad \forall t \in (T) \quad (4)$$

où  $v_k(t) = \frac{1}{2\pi} \frac{d\phi_k(t)}{dt}$  est la fréquence instantanée de  $Z_k$ .

b) Dans ce domaine essentiel d'existence, le comportement de  $I_{k1}$  pourra être approché par la méthode de la phase stationnaire si  $\frac{\partial^2 L_{k1}}{\partial t^2} \neq 0$ . Sinon, il y aura singularité. Plus précisément :

$$\text{i) si } \begin{cases} \frac{\partial^2 L_{k1}}{\partial t^2} = 0 \iff \frac{dv_k}{dt}(t) = \frac{dv_l}{dt}(t - \tau) \\ \frac{\partial^3 L_{k1}}{\partial t^3} \neq 0 \iff \frac{d^2 v_k}{dt^2}(t) \neq \frac{d^2 v_l}{dt^2}(t - \tau) \end{cases} \quad (5)$$

Il y a singularité "pli" et le comportement de  $I_{k1}$  au voisinage des points vérifiant simultanément (4) et (5) sera décrit par une fonction d'Airy [2], [3].

$$\text{ii) si } \begin{cases} \frac{\partial^2 L_{k1}}{\partial t^2} = \frac{\partial^3 L_{k1}}{\partial t^3} = 0 \\ \frac{\partial^4 L_{k1}}{\partial t^4} \neq 0 \end{cases} \quad (6)$$

Il y aura "fronce", décrite cette fois-ci par une fonction de Pearcey [2], [3].

iii) si  $\frac{\partial^m L_{k1}}{\partial t^m} = 0, \forall m \geq 2$  (7) / on pourra assimiler le comportement de  $I_{k1}$  à celui d'une distribution de Dirac.

c) Le comportement global de  $\chi_{Z_k Z_l}$  se déduit alors du b) par la convolution fréquentielle (3-a).

On vérifie immédiatement que, dans le cas d'une fonction d'auto-ambiguïté, (4) est toujours vérifiée en (0,0) ainsi que (7).  $|\chi_{ZZ}|$  est toujours maximal à l'origine.

Ainsi, le comportement d'une f.d.a. peut être approché de manière géométrique à partir des seules lois  $v_k(t)$  des différents motifs dans le plan temps-fréquence. L'influence de l'enveloppe des signaux et de leur durée limitée, n'intervient alors que par simple convolution fréquentielle.

3°) Exemples.

3-1) FM cubique

$$\text{Si l'on a : } v_i(t) = v_0 + ct^3, \quad -\frac{T}{2} \leq t \leq \frac{T}{2},$$

les points stationnaires vérifient :

$$v_s(\tau) = c(\tau - t)^3 + ct^3, \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

soit l'ensemble des points du plan  $(v, \tau)$  limité par les courbes :

$$\begin{cases} v + (\tau) = c\left(\tau - \frac{T}{2}\right)^3 + c\frac{T^3}{8} \\ v - (\tau) = \frac{c}{4}\tau^3 \end{cases} ; \tau \geq 0$$

. leurs symétriques / (0,0) ;  $\tau \leq 0$  .

On voit alors que les relations (5) sont vérifiées (pour  $\tau \neq 0$ ) le long de  $v = v - (\tau)$ , courbe pour laquelle il y a pli. On a alors :

$$I_{11}(v, \tau) = (6\pi c \tau)^{-\frac{1}{3}} \text{Ai} \left( \left( \frac{4\pi^2}{3c\tau} \right)^{\frac{1}{3}} (v - v_s(\tau)) \right).$$

3-2) Fréquences pures

Si l'on considère maintenant le cas d'un signal réel formé de composantes monochromatiques :

STRUCTURES DES FONCTIONS D'AMBIGUITE ET THEORIE  
DES CATASTROPHES : APPLICATION AUX SIGNAUX DE PAROLE

$$S(t) = \sum_k A_k(t) \cos(2\pi v_k t + \phi_k) \quad , t \in (T)$$

on peut écrire :

$$S(t) = \sum_k \frac{1}{k^2} A_k(t) e^{i(2\pi v_k t + \phi_k)} + \sum_k \frac{1}{2} A_k(t) e^{-i(2\pi v_k t + \phi_k)} \quad , t \in (T)$$

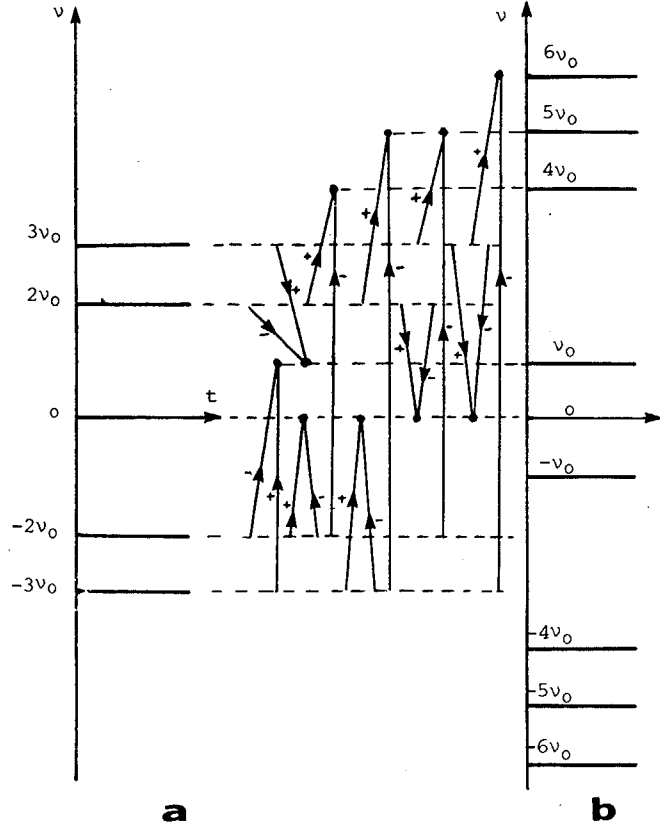
et d'après (2), la f.d.a. de S prendra en compte les contributions de toutes les interambiguïtés entre motifs du plan temps-fréquence, aussi bien en fréquences positives que négatives. Pour chacune de ces interactions, les points stationnaires seront du type :

$$v = v_k - v_l \quad (8)$$

et, en chacune de ces valeurs, il sera possible de calculer  $\chi_{Z_k Z_l}(v_k - v_l, \tau)$ . Si l'on prend par exemple le signal :

$$S(t) = A \cos 4\pi v_0 t + A \cos 6\pi v_0 t \quad , \quad (9)$$

on obtient, à partir de (8) le lien des points stationnaires : ce sont des fréquences pures en  $0, \pm v_0, \pm 4v_0, \pm 5v_0, \pm 6v_0$  (figure 1). Comme, en toutes ces valeurs, (7) est vérifiée, le comportement des différentes  $I_{kl}$  s'assimilera à celui d'un Dirac pondéré par la somme  $C_{kl}$  des contributions de chacune des intégrales (3-b) en  $v = v_k - v_l$ .



**a** **b**  
Figure 1 : Construction des points stationnaires dans le cas d'un signal à 2 composantes  $2v_0$  et  $3v_0$ .  
a : loi temps fréquence des composantes du signal  
b : lieu des points stationnaires de la fonction d'ambiguïté.  
Ex. : La composante en  $5v_0$  est obtenue par combinaison des composantes fréquentielles  $(3v_0 - (-2v_0)) (2v_0 - (-3v_0))$ .

Ainsi, si l'on considère par exemple la composante en  $4v_0$ , elle est résultante de la seule interaction des fréquences  $2v_0$  et  $-2v_0$  i.e. de  $I_{2,-2}$ . Ou alors :

$$I_{2,-2}(v, \tau) = \int_R e^{i[4\pi v_0 t + 4\pi v_0(t-\tau) - 2\pi v t]} dt \quad (10)$$

$$= e^{-i4\pi v_0 \tau} \delta(v - 4v_0)$$

soit  $|C_{2,-2}| = 1$ .

De la même façon, la composante en  $5v_0$  est résultante des interactions  $(3v_0, -2v_0)$  et  $(2v_0, -3v_0)$  (cf fig.1). On obtient alors :

$$I_{3,-2}(v, \tau) = \int_R e^{i[6\pi v_0 t + 4\pi v_0(t-\tau) - 2\pi v t]} dt \quad (11)$$

$$= e^{-i4\pi v_0 \tau} \cdot \delta(v - 5v_0)$$

$$I_{2,-3}(v, \tau) = \int_R e^{i[4\pi v_0 t + 6\pi v_0(t-\tau) - 2\pi v t]} dt \quad (12)$$

$$= e^{-i6\pi v_0 \tau} \cdot \delta(v - 5v_0)$$

$$\text{et } |C_{3,-2} + C_{2,-3}| = 2 |\cos \pi v_0 \tau| \quad (13)$$

On procéderait de même pour les trois autres composantes.

Le squelette schématique du module de  $\sum_{k,l} I_{kl}$  ainsi calculé est représenté sur la fig. (2-a). L'allure globale de  $\chi_S$  s'obtient alors par convolution fréquentielle avec les  $\chi_{A_k A_l}$ . Si l'on suppose, comme ici, que les amplitudes affectées à chaque composante sont identiques et constantes ( $=A$ ) et que le calcul effectif de  $\chi_S$  se fait suivant :

$$\int_R W(\tau) S(t) \cdot S(t-\tau) e^{-i2\pi v t} dt \quad (14)$$

où  $W(\tau)$  est une fenêtre de pondération, on obtient :

$$\chi_{A_k A_l}(v, \tau) = A^2 w(v) \frac{v}{t} W(t) \quad .$$

Le résultat conduit à remplacer (dans la fig.2-a) chaque

$$C_{kl} \delta(v - (v_k - v_l)) \text{ par } A^2 C_{kl} w(v - (v_k - v_l)) \text{ .}$$

4°) Fonction d'ambiguïté : résultats de simulation.

Le calcul de la fonction d'ambiguïté a été programmé selon l'expression (14). Plusieurs types de fenêtres de pondération peuvent être utilisées : fenêtre rectangulaire, fenêtre de Hanning ou Hamming. L'intégrale (14) est calculée par corrélation pour les différentes fréquences considérées. Le résultat est donné en partie réelle et imaginaire. Le module et la phase sont déterminés par l'algorithme de CORDIC de manière à diminuer les temps de calcul.

Les résultats présentés (figure 2-b) ont été

calculés dans les conditions suivantes :

- signal (9) avec  $v_0 = 200$  Hz
- fréquences d'échantillonnage 10 kHz
- durée analysée : 12,5 ms soit 125 échantillons
- pas d'analyse en fréquence : 12,5 Hz
- fréquence d'analyse maximale : 1250 Hz soit 100 tranches.

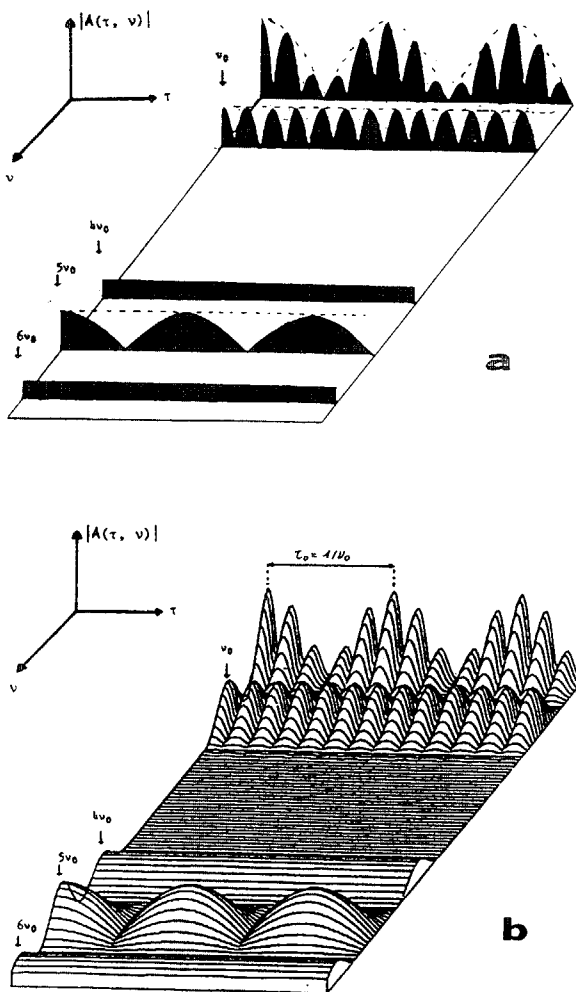


Figure 2 : Surfaces d'ambiguïté pour le signal à 2 composantes  $S(t) = A \cos 4\pi\nu_0 t + A \cos 6\pi\nu_0 t$ .  
a : squelette schématique.  
b : calcul réel avec pondération de Hanning.

- pondération utilisée : Hanning.

A titre indicatif, le calcul précédent effectué sur un calculateur PDP 11/34 nécessite un temps de calcul de l'ordre de 1 minute.

La comparaison de la surface d'ambiguïté obtenue fig. 2-b, avec la structure schématique proposée fig. 2-a, justifie donc bien la méthode approchée de construction décrite en 3-2.

## B. APPLICATION AUX SIGNAUX DE PAROLE.

1°) Interprétation de l'exemple précédent :  
(figure 2-b).

Le signal (9) représente dans le domaine de la parole, un signal voisé stationnaire réduit à 2 composantes et n'ayant pas de composante fondamentale phy-

siquement présente. On constate que sa fonction d'ambiguïté présente néanmoins une composante à la fréquence fondamentale. Celle-ci provient d'après  $\delta$ , du fait que les composantes sont 2 harmoniques successives. Ce résultat se généralise à tout signal voisé dans lequel au moins 2 harmoniques quelconques mais consécutives sont présentes.

2°) Détection de la fréquence fondamentale à l'aide de la fonction d'ambiguïté : principe ([4]).

La fonction d'ambiguïté en  $\nu = 0$ , représente la fonction de corrélation du signal. Si celui-ci est périodique, la fonction de corrélation le sera également. On peut donc envisager une procédure de détection de fréquence fondamentale de la façon suivante :

- calcul de la fonction de corrélation par  $\chi(0, \tau)$  et recherche des maximums ce qui donne plusieurs valeurs probables par la période fondamentale.
- calcul de  $\chi(\nu, 0)$ . Le premier maximum donne la valeur de la fréquence fondamentale.
- recherche de la solution commune avec deux étapes précédentes.

3°) Exemple de détection. [4].

Un exemple de résultat de détection de fréquence fondamentale utilisant cette méthode est donné à la figure 3-b.

Une mesure de fréquence fondamentale est effectuée toutes les 10 ms, ce qui donne un bon suivi de l'évolution de cette fréquence tout en satisfaisant les hypothèses de stationnarité locales implicitement nécessaires dans le modèle envisagé.

La figure (3-a) donne à titre de comparaison, la détection de fréquence fondamentale obtenue par une méthode à base de filtrage [5].

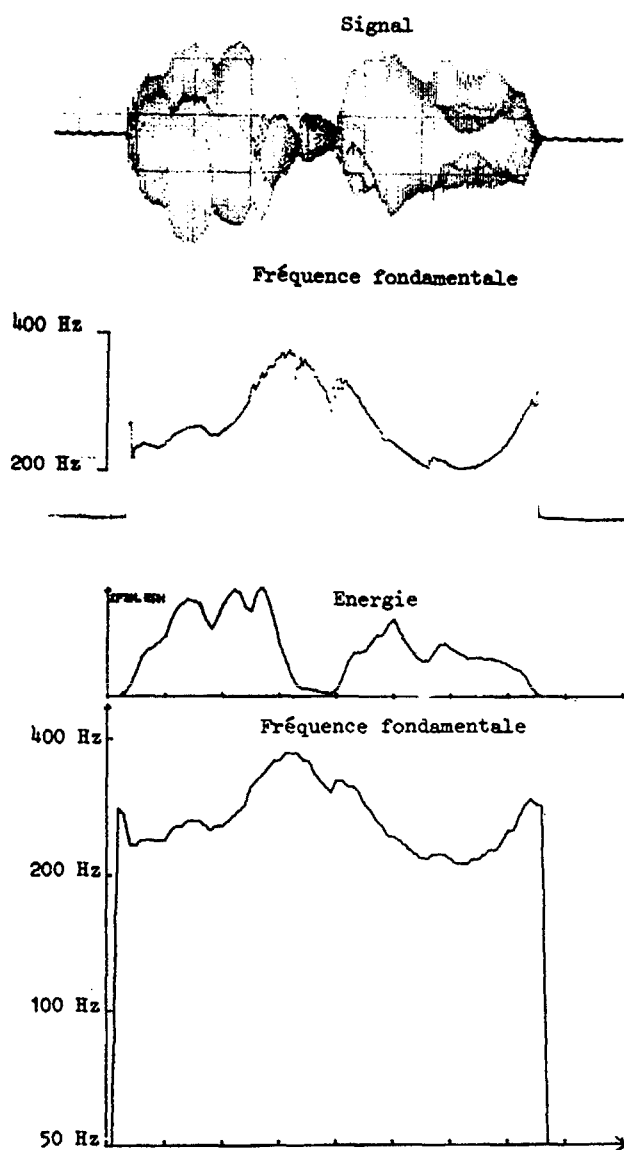


Figure 3. - Locuteur n° 3 : "le lendemain"

Evolution de la fréquence fondamentale détectée par :

- a) filtrage
- b) autocorrélation et ambiguïté

#### CONCLUSION.

L'étude théorique des fonctions d'ambiguïté a permis une meilleure compréhension de leur structure, autorisant en retour une meilleure caractérisation des signaux par une telle analyse temps-fréquence. Cette analyse a permis de proposer une méthode de détection de fréquence fondamentale reposant sur l'existence et la structure de termes dits "fantômes".

Cette méthode a été testée sur un vaste corpus et a donné entière satisfaction avec une restriction en ce qui concerne les temps de calcul.

- [1] A.W. RIHACZEK  
"Principles of high-resolution radar",  
Mc Graw Hill, 1969.
- [2] P. FLANDRIN, B. ESCUDIE  
"Géométrie des fonctions d'ambiguïté et des représentations conjointes de Ville : l'approche de la théorie des catastrophes"  
8ème Colloque GRETSI, 69-74, Nice, 1981.
- [3] T. POSTON, I. STEWART  
"Catastrophe Theory and its applications"  
Pitman, 1978.
- [4] M. BOULOGNE  
"Détection du voisement et de la fréquence fondamentale d'un signal de parole"  
Thèse Docteur-Ingénieur, I.N.P.GRENOBLE, 1979.
- [5] C. ABRY, L.J. BOE, J.F. ZURCHER  
"La détection du voisement par les propriétés physiques résultant de l'excitation du conduit vocal : comparaison statistique de trois procédés"  
6ème J.E.P. Groupe Communication Parlée du G.A.L.F. Toulouse, 1, 194-204, 1975.

