



HAL
open science

Mesures de la spatialisation pour la sonorisation automobile

Alexandre Fénières

► **To cite this version:**

Alexandre Fénières. Mesures de la spatialisation pour la sonorisation automobile. Autre. 2016. dumas-01835028

HAL Id: dumas-01835028

<https://dumas.ccsd.cnrs.fr/dumas-01835028>

Submitted on 11 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

PARIS

MEMOIRE

présenté en vue d'obtenir

le DIPLOME d'INGENIEUR du CNAM

SPECIALITE : MECANIQUE

OPTION : ACOUSTIQUE

Par

Alexandre FENIERES

Sous la direction de Delphine DEVALLEZ (Arkamys) et Éric BAVU (CNAM)

Mesures de la spatialisation pour la sonorisation automobile

Soutenu le 29 janvier 2016

JURY

PRESIDENT : Alexandre GARCIA (CNAM)

MEMBRES : Frédéric AMADU (Arkamys)
Jean-Baptiste DOC (UPMC – Institut d'Alembert)
Nicolas LOPEZ (Arkamys)

Sommaire

Remerciements.....	5
Symboles, acronymes et abréviations utilisés.....	6
0. Introduction.....	7
1. Audition spatiale: modèles de localisation auditive	9
1.1. Théorie Duplex - ITD - ILD	9
1.1.1. Différence de temp de parcours inter-aural - ITD	9
1.1.2. Différence de niveau inter-aurale – ILD	13
1.1.3. Analyse conjointe ITD/ILD - Plausibilité.....	15
1.1.4. Limites du modèle	16
1.2. Découpage fréquentiel	17
1.3. Filtrage par le corps sur tous les axes - Indices monauraux - HRTF.....	18
1.4. Effet de précedence – Fusion - Influence du local d’écoute.....	20
1.5. Corrélation inter-aurale	24
1.6. Corrélation inter-aurale dans le domaine fréquentiel	28
1.7. Modèle de Jeffress et extensions	29
1.8. Boite à outils de Faller et Merimaa.....	32
1.9. Précision de localisation	32
1.10. Représentation de la scène sonore - Ségrégation	33
1.11. Résumé - Conclusion partielle - Discussion.....	36
2. Reproduction stéréophonique en véhicule.....	38
2.1. Stéréophonie.....	38
2.1.1. Principe	38
2.1.2. Recommandations et système de référence.....	38
2.2. Caractérisation acoustique et électro-acoustique des habitacles automobiles	40

2.2.1.	Local d'écoute.....	40
2.2.2.	Position des passagers.....	41
2.2.3.	Caractéristiques et implantation des haut-parleurs	42
2.2.4.	Réverbération, clarté et distance critique.....	45
2.2.5.	Analyse de la propagation	48
2.3.	Traitement du signal numérique embarqué.....	51
2.3.1.	Généralités et définitions	51
2.3.2.	Alignement temporel.....	53
2.3.3.	Représentation et égalisation phase(fréquence).....	57
2.3.4.	Filtrage par inversion	62
2.4.	Conclusion partielle	64
3.	Mesures et scène sonore	66
3.1.	Buts	66
3.2.	Résultats de mesure dans la littérature.....	66
3.2.1.	Evaluation conjointe des ITD et ILD.....	66
3.2.2.	Mesure d'IACC.....	67
3.2.3.	Modèle de Mason.....	74
3.2.4.	Modèles de type Jeffress - Lindemann – Gaik.....	76
3.2.5.	Boîte à outils de Faller et Merimaa	77
3.2.6.	Performance des modèles de latéralisation en milieu réverbéré.....	78
3.3.	Essais réalisés chez Arkamys.....	79
3.3.1.	Procédure test	79
3.3.2.	Mesure de la phase binaurale en fonction de la fréquence	80
3.3.3.	Corrélation inter-aurale dans le domaine fréquentiel	88
3.3.4.	Mesure de la phase binaurale, de la corrélation binaurale et du diagramme de Lissajous	93

3.3.5.	Boite à outils logicielle de Faller et Merimaa	95
3.3.6.	Modèle de Lindemann.....	100
3.3.7.	Conclusion partielle	115
4.	Conclusion	117
	Annexes	119
	Annexe A – Simulation d’écoute hors axe (code Matlab « ecoutehorsaxe.m »)	119
	Annexe B - Calcul de FIACC (code Matlab « FIACC.m »)	120
	Annexe C - Calculs de clarté C50 et C80 (code Matlab « clarte.m »)	121
	Annexe D - Comparaison entre écoutes critiques et résultats de la boîte à outils de Faller et Merimaa appliquée aux mesures dans le véhicule A	123
	Annexe E - Références des véhicules utilisés, des techniques de mesures et des fichiers de mesure résultants	126
	Bibliographie.....	127
	Table des illustrations.....	135
	Table des tableaux.....	140

Remerciements

Je tiens à remercier :

Mes professeurs du CNAM, pour la qualité de leur enseignement et l'intérêt pour l'acoustique qu'ils ont suscité et renforcé chez moi, tout particulièrement Eric Bavu qui a supervisé ce mémoire et Alexandre Garcia pour ses indispensables corrections et suggestions ;

Mes collègues Delphine Devallez pour avoir été la tutrice de ce travail, pour sa disponibilité, ses conseils, et les pistes qu'elle m'a fait entrevoir, Nicolas Lopez pour ses explications toujours limpides et Yann « Sensei » Lecoeur pour m'avoir mis le pied à l'étrier ;

Mon épouse Lorena pour sa patience et ses encouragements lors des soirées, fins de semaines et jours de congés consacrés à la rédaction de ce mémoire ;

Mon père Frédéric Fénieres pour sa relecture attentive ;

Tous ceux qui partagent leurs publications et modèles d'audition informatisés, ainsi que les contributeurs des forums d'aide du logiciel Matlab.

Detroit, Michigan, USA, le 24/06/2015

Symboles, acronymes et abréviations utilisés

ω	Pulsation
t	Temps
f	Fréquence
n	Temps discret (numéro d'échantillon)
k	Fréquence discrète
c	Célérité du son
$\varphi(g, d)$	Phase entre deux signaux g et d
ITD	Différence inter-aurale de temps (<i>Interaural Time Difference</i>)
IPD	Différence inter-aurale de phase (<i>Interaural Phase Difference</i>)
ILD	Différence inter-aurale de niveau (<i>Interaural Level Difference</i>)
d^*	Complexe conjugué de d
DSP	Processeur dédié au traitement du signal (<i>Digital Signal Processor</i>)
IACC	Corrélation inter-aurale (<i>Inter Aural Cross Correlation</i>)
FIACC	Corrélation inter-aurale dans le domaine fréquentiel (<i>Frequency domain Inter Aural Cross Correlation</i>)
TR60	Durée de Réverbération pour une décroissance de niveau de 60dB après arrêt de la source
CASA	Analyse informatisée de scène sonore (<i>Computational Auditory Scene Analysis</i>)
ASW	Largeur de source apparente (<i>Apparent Source Width</i>)
LF	Fraction latérale (<i>Lateral Fraction</i>)
RIF	Réponse Impulsionnelle Finie
RII	Réponse Impulsionnelle Infinie

0. Introduction

Il est possible d'améliorer la fidélité d'un système de reproduction sonore en ajoutant à ce système un traitement du signal audio, un filtrage qui doit être réglé de façon spécifique pour chaque configuration de diffusion. La pertinence de ces réglages peut être évaluée soit avec un panel important d'auditeurs formés à l'écoute critique [1], soit avec une méthode de mesure objective [2]. Lors de réglages de systèmes audio embarqués, la cotation subjective d'écoutes du son reproduit dans le véhicule par un ou plusieurs auditeurs reste incontournable. Cela complique les réglages car il n'est pas possible de se déplacer instantanément du véhicule à évaluer à la pièce d'écoute de référence, ou d'une place à l'autre dans le véhicule. La mémoire auditive est mise en jeu, et elle est réputée être nettement moins bonne dans le temps que pour d'autres sens [3].

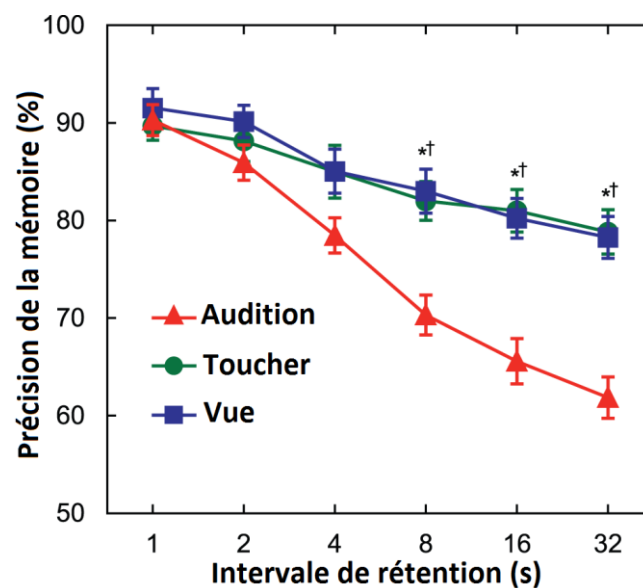


Figure 1 - Infériorité de la mémoire de l'audition par comparaison au toucher et à la vue, traduit de [3]

Le réglage du traitement audio embarqué afin d'obtenir la meilleure fidélité spectrale et spatiale est donc la plupart du temps un processus d'essais/erreurs, impliquant un auditeur formé et rarement plusieurs auditeurs présents en même temps dans le même véhicule.

Quand elle est bien faite, la mesure de l'amplitude de la pression sonore en fonction de la fréquence permet de comparer la fidélité *spectrale* à différentes places, et avec un système

de référence. Une mesure comparable pour la fidélité *spatiale*¹ permettrait d'utiliser des microphones et des modèles d'audition plutôt que des auditeurs. On en vient à se demander quelles mesures acoustiques peuvent être mises en relation avec la perception spatiale du son pour un auditeur moyen.

L'objet de ce mémoire est de :

- présenter les modèles habituellement utilisés pour expliquer la localisation du son, compiler et analyser comment mesures et perceptions de la scène sonore ont été reliées dans la littérature pour constituer une « boîte à outils » de mesure,
- comparer les conditions d'écoutes dans une automobile et celles recommandées pour la diffusion stéréophonique du son, puis décrire le traitement du signal audio habituellement employé dans des automobiles pour réduire l'écart entre le système embarqué et la référence,
- appliquer les modèles d'audition retenus à des mesures faites en véhicule, pour comparer l'apport d'un filtrage entre différents modèles et les résultats d'écoutes critiques.

¹ D'où vient le son reproduit ? Est-ce comparable à la perception spatiale avec le système de référence ?

1. Audition spatiale: modèles de localisation auditive

Il est généralement admis que le corps d'un auditeur produit un filtrage du son reçu par ses oreilles et que son système nerveux central étudie ce filtrage pour en déduire l'incidence des sources. Ce chapitre vise à résumer les théories les plus souvent utilisées pour modéliser l'audition spatiale humaine, plus particulièrement la localisation sur le plan azimutal (latéralisation), sans prétention d'exhaustivité mais pour mieux comprendre les modèles utilisés et les choix faits par la suite. On trouve une abondante littérature sur ce sujet, et en particulier dans Blauert [4] de nombreuses précisions.

1.1. Théorie Duplex - ITD - ILD

La théorie Duplex de Lord Rayleigh est que la localisation latérale des sources sonores est basée sur l'analyse des différences de temps et de niveaux des sons qui atteignent les deux oreilles. L'analyse des différences de temps entre les oreilles (ITD) serait prédominante pour les fréquences jusqu'à 800Hz environ, celle des différences de niveau (ILD) pour localiser les fréquences au-dessus de 2000Hz environ, et les deux analyses coexistent entre ces deux fréquences. La fréquence 1500Hz proposée par Woodworth est généralement retenue comme point de transition entre les deux analyses, soit une longueur d'onde qui correspond à une distance moyenne entre les oreilles de 22cm.

1.1.1. Différence de temp de parcours inter-aural - ITD

En dessous de 1,5kHz, on définit l'ITD comme le retard de phase inter-aural (IPD) entre les signaux $g(t)$ et $d(t)$ captés par les oreilles:

$$ITD_{f < 1.5 \text{ kHz}}(t) = IPD(t) = \frac{1}{2\pi} \frac{\varphi(d(t), g(t)) + 2\pi p}{f} \text{ où } p \text{ est un entier relatif}$$

Si on dispose de transformées de Fourier $G(f)$ et $D(f)$ des signaux temporels de pression issus des oreilles gauche et droite $g(t)$ et $d(t)$, alors :

$$ITD_{f < 1.5 \text{ kHz}}(f) = IPD(f) = \frac{1}{2\pi f} \left[\arctan\left(\frac{D(f)}{G(f)}\right) + 2\pi p \right] \text{ où } p \text{ est un entier relatif}$$

Comme Woodworth, on assimile la tête humaine à une sphère de rayon r qui occasionne un retard de propagation du son entre les deux oreilles diamétralement opposées, on suppose que les fronts d'ondes sonores sont parallèles et que l'ITD varie peu en fonction de la distance de la source, comme illustré ci-dessous.

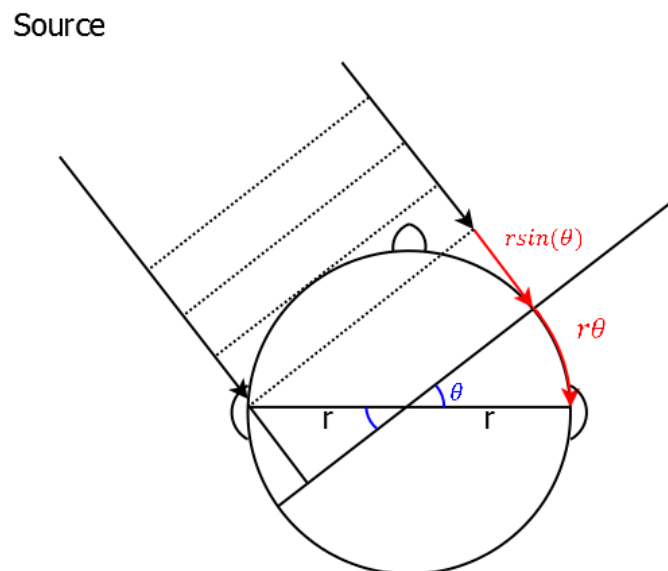


Figure 2 - Calcul de la différence de parcours avec une tête sphérique [5]

Woodworth proposait comme relation entre ITD et latéralisation [6]:

$$ITD(\theta)_{\text{sphère}} \approx \begin{cases} \frac{r \cdot (\sin(\theta) + \theta)}{c} \text{ pour } 0 \leq \theta \leq \frac{\pi}{2}, \\ \frac{r \cdot (\sin(\theta) + \pi - \theta)}{c} \text{ pour } \frac{\pi}{2} \leq \theta \leq 3\frac{\pi}{2} \\ \frac{r \cdot (\sin(3\pi - \theta) - 2\pi + \theta)}{c} \text{ pour } 3\frac{\pi}{2} \leq \theta \leq 2\pi \end{cases}$$

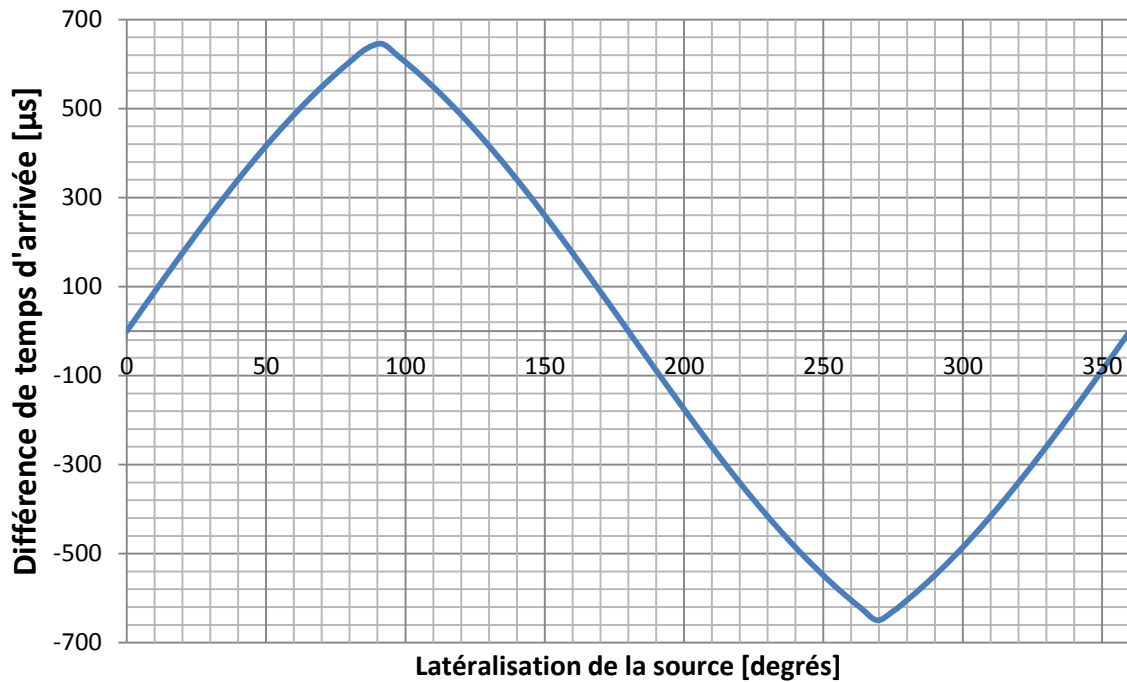


Figure 3 - ITD selon la formule de Woodworth pour une distance entre les oreilles de 18 cm

Cette formule est souvent citée dans la littérature sans les conditions sur θ , ce qui est incorrect. Aaronson et Hartmann ont confirmé les points précités avec des sujets humains [7] et ont conclu que le modèle de Woodworth fonctionne bien avec des sons large-bande, et qu'il est plus précis pour des sons à contenu spectral étroit au-dessus de 1,5kHz alors qu'en-dessous, il faudrait utiliser un modèle prenant en compte la diffraction par la tête.

De nombreux auteurs affirment que le modèle de Woodworth n'est pas valide à toutes les fréquences, et que la relation entre ITD et azimuth est plutôt du type :

$$ITD(\theta, f) = \alpha(f) \cdot r \cdot \frac{\sin(\theta) + \theta}{c}$$

Les résultats expérimentaux de Kuhn en 1977, cités par [8] et [9], indiquent comme relation :

$$ITD(\theta, f) = \frac{\alpha(f) \cdot r \cdot \sin(\theta)}{c}$$

$$\text{où } \alpha(f) = \begin{cases} 3 & f < 500 \\ 2,5 + 0,5 \cdot \cos\left(\pi \frac{\log_2\left(\frac{\sqrt{6}f}{1250}\right)}{\log_2(6)}\right) & 500 < f < 3000 \\ 2 & f > 3000 \end{cases}$$

Il paraît logique d'utiliser cette relation dorénavant. Son inversion n'est pas très complexe.

La figure suivante montre comment, selon Kuhn, les ITD varient en fonction de la fréquence si la célérité du son $c = 344\text{ms}^{-1}$ et avec un rayon effectif de la tête $r = 0.093\text{m}$.

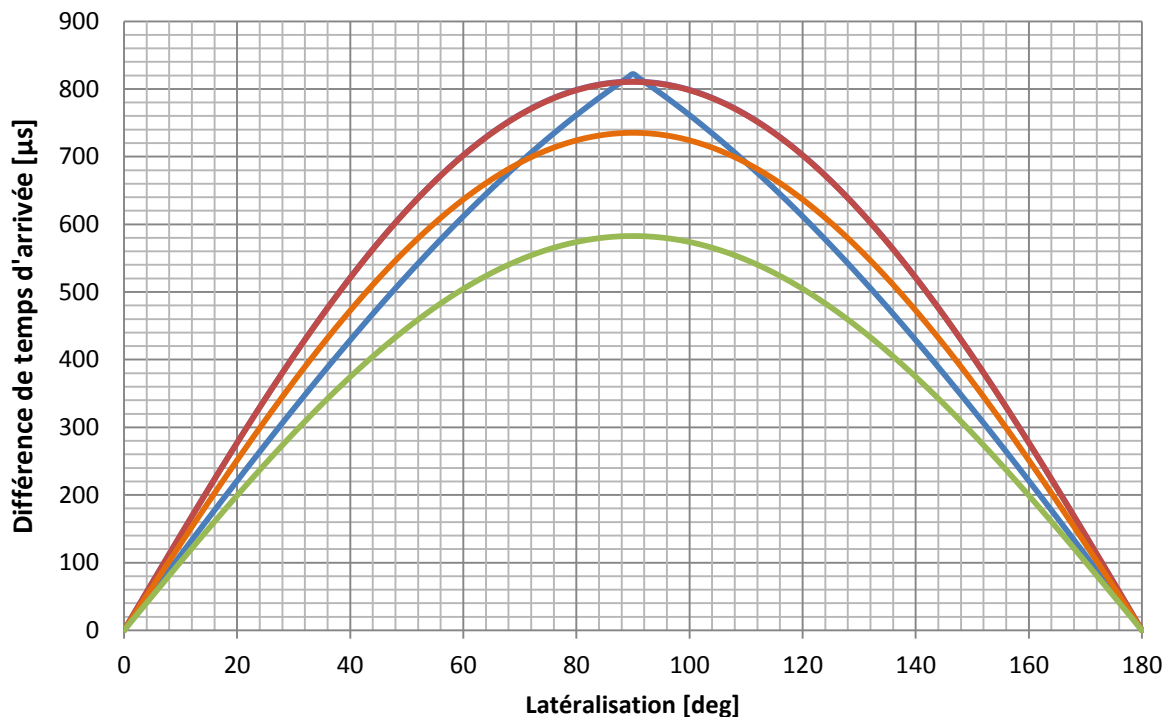


Figure 4 - Comparaison de la relation de Woodworth avec $r=11$ cm (en bleu) et de la relation de Kuhn avec $r=9,3$ cm en dessous de 500 Hz (rouge), à 1000 Hz (orange) et à 2000 Hz (vert)

A chaque fréquence considérée, les ITD sont ambigus à une période près, et la différence minimale de temps d'arrivée perceptible par un humain se situe entre 10 et 20 μs .

Au-dessus de la fréquence de raccord, il subsiste une analyse temporelle du retard d'enveloppe inter-aural (IED):

$$ITD_{f>1.5\text{kHz}} = IED = \frac{1}{2\pi} \frac{d[\varphi(d, g)(f)]}{df}$$

Les différences entre IPD et IED sont discutées plus en détails dans [10].

On trouvera plus de raffinements utiles à la compréhension du modèle dans [7] [8] et [11]. Par exemple on peut considérer que la tête est plutôt ovale, la différence de retard de parcours entre une tête ronde et ovale croît à mesure que l'angle θ est proche de $\frac{\pi}{2}$ [π]. Si on prend en compte la distance entre la source et la tête, le modèle devient plus complexe et plus précis, mais quelle distance utiliser pour les calculs si la source est virtuelle ?

1.1.2. Différence de niveau inter-aurale – ILD

Un son au-dessus de 1500 Hz a une longueur d'onde plus petite que la tête, cette dernière crée une « zone d'ombre » qui filtre le son en fonction de son incidence, ce qui est beaucoup moins marqué en basses fréquences. L'indice de latéralisation principalement analysé au-dessus de 1500 Hz est la différence d'intensité.

$$ILD(f) = 20 \log_{10} \left| \frac{D(f)}{G(f)} \right|$$

Les auteurs de [5] et [12] proposent la relation qui suit :

$$ILD(\theta, f) = \beta(f) \cdot \sin(\theta)$$

Les mesures de Gulick vérifient globalement cette relation, $\beta(f)$ étant alors une fonction croissante de la fréquence (fig. 6).

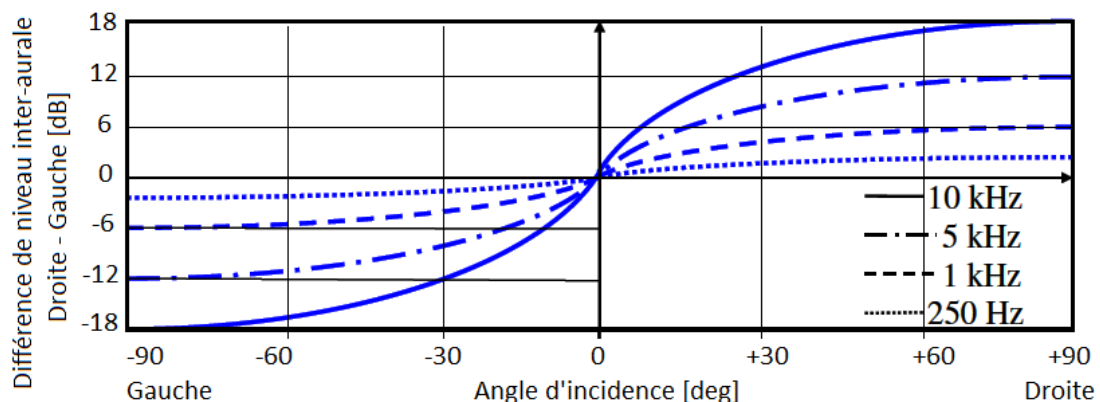


Figure 5 - ILD mesuré en fonction de l'angle d'incidence à différentes fréquences selon Gulick [12]

Une simulation plus récente de l'ILD pour différentes positions d'une source, sur la base d'un modèle de tête sphérique avec oreilles aux antipodes et prenant en compte la diffraction, montre une relation entre position et filtrage plus complexe et moins monotone, d'autant plus que la source est latéralisée sur le côté (fig. 8).

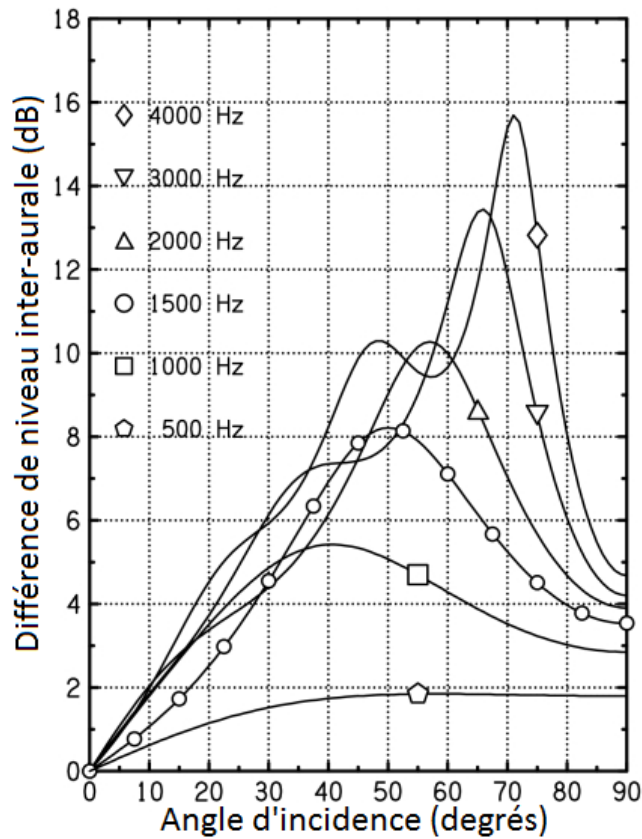


Figure 6 - ILD calculée en fonction de l'angle d'incidence à différentes fréquences de sinus pur, traduit de [11]

Dans [8] Larcher résume ainsi les travaux de Wrightman et Kistler : « l'ITD est prépondérant pour un son large bande, mais l'ILD l'emporte sur l'ITD pour un son dénué de basses fréquences ».

1.1.3. Analyse conjointe ITD/ILD - Plausibilité

Dans la zone fréquentielle de raccord (de 800 Hz à 2 kHz environ), les deux indices sont analysés conjointement, et Franssen a proposé des abaques de localisation par bande de fréquence pour des différences de temps et de niveau entre les haut-parleurs.

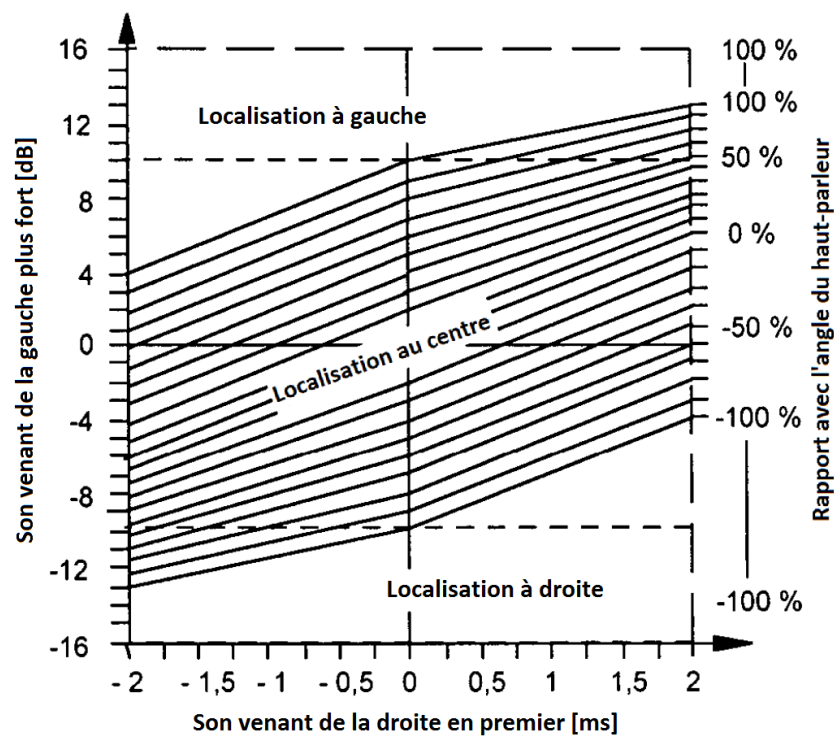


Figure 7 - Un abaque de Franssen pour une bande critique, traduit de [16]

Les expériences de Gaik indiquent qu'il existe des relations « naturelles » entre ITD et ILD [4][8]. Si ces relations naturelles ne sont pas respectées l'auditeur peut percevoir des sources multiples. En 2002 Oag et Wittek ont étudié la localisation de sources fantômes² avec des différences de temps et de niveau contradictoires entre les haut-parleurs, par exemple haut-parleur gauche plus fort mais signal venant du haut-parleur droit en premier, et concluent que la précision de localisation est alors faible sauf si la différence de niveau est :

- inférieure à 3dB, dans ce cas il est possible de compenser la différence de niveau avec un délai pour recentrer la source fantôme

² Une voix de femme parlée

- ou supérieure à 12dB, dans ce cas la source fantôme est localisée sur le haut-parleur le plus fort

Quand ils augmentent les différences de temps et de niveau, la source fantôme se sépare en deux parties, une avec les hautes fréquences et l'autre avec les basses fréquences [13]. On retrouve le concept de couple ITD-ILD « naturels ». Ces résultats ressemblent au ressenti dans les automobiles. L'analyse conjointe des ITD et ILD est donc un moyen de déterminer la plausibilité et la précision de la latéralisation perçue.

1.1.4. Limites du modèle

La théorie Duplex n'explique pas tous les attributs de localisation du son (élévation, distance³). Elle ne fonctionne que pour la localisation azimutale, c'est-à-dire dans le plan dans deux conduits auditifs, et n'explique pas la localisation de sons sur le plan zénithal. Elle n'explique donc pas la différenciation entre deux sources formant le même angle par rapport à l'axe médian entre les deux oreilles, si une source est devant l'auditeur et l'autre derrière l'auditeur.

Plusieurs positions de sources réelles donneraient des ITD égales ou trop proches pour lever l'ambiguïté, c'est-à-dire pour déterminer avec précision quelle position possible est la bonne. Pour une ITD donnée, ces positions ambiguës sont situées sur une hyperbole de rotation centrée sur l'axe des deux oreilles couramment appelée « cône de confusion ». Pour des sources proches de l'auditeur, les ILD permettraient de lever certaines ambiguïtés sur ces cônes de confusion, les positions potentielles des sources seraient alors situées sur des « tores de confusion ».[14]

³ Brungart et Rabinowitz montrent qu'en plus d'être un indice de latéralisation, l'ILD est un indice perceptif de distance de la source en champ proche [10]. Un procédé de mesure analysant cet effet permettrait donc, en plus d'analyser la latéralisation, de déduire la perception de distance des sources virtuelles et de profondeur de la scène sonore.

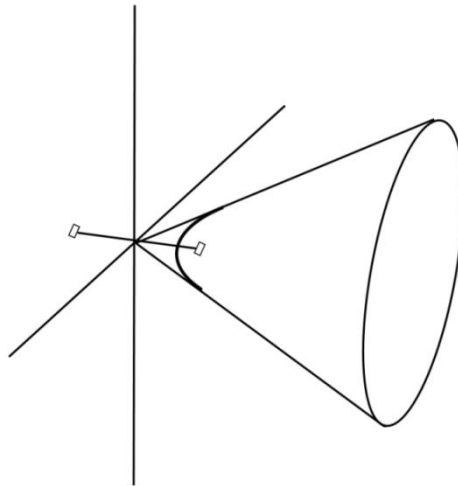


Figure 8 – Un « cône » de confusion considéré pour une oreille, adapté de [4]

Enfin la théorie Duplex n'explique pas comment se fait la perception spatiale des réflexions artificielles ou naturelles.

1.2. Découpage fréquentiel

Il est admis que l'analyse fréquentielle par le système nerveux est découpée en bandes de fréquences dites « bandes critiques », au niveau de la cochlée. La plupart des modèles incorporent donc une étape de filtrage de séparation appelé « gammatone »⁴.

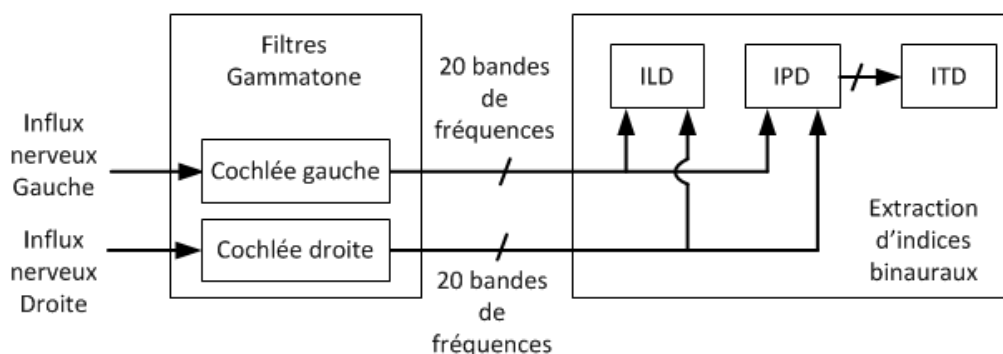


Figure 9 - Position du filtrage gammatone dans un modèle d'audition

⁴ On trouvera plus de précisions sur ce filtrage dans [12] pp 21-23

Blauert estime que ce découpage peut être modélisé de façon satisfaisante par un réseau de filtres IIR à recouvrement, de topologie Chebychev d'ordre 4 à 6 comme illustré ci-dessous.[4]

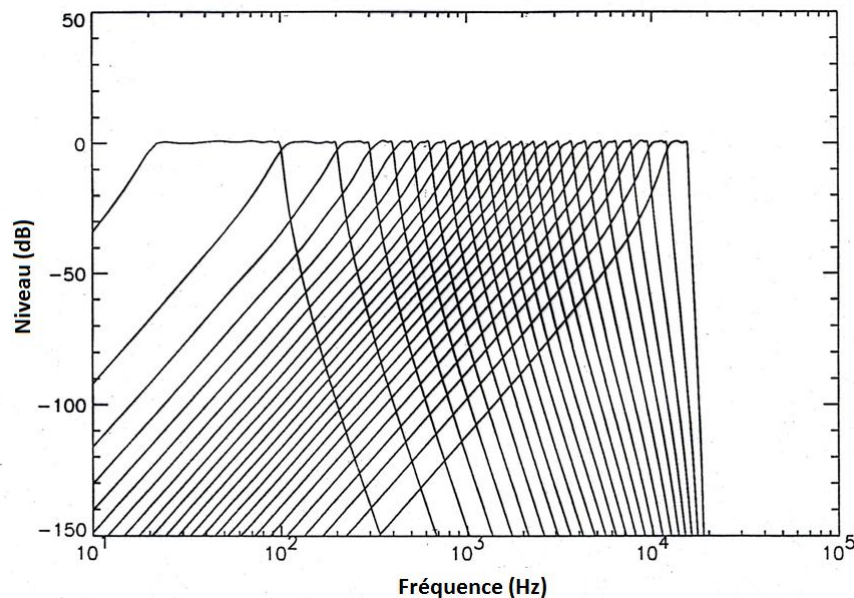


Figure 10 - Réseau de filtrage en bandes critiques, selon [4]

1.3. Filtrage par le corps sur tous les axes - Indices monauraux - HRTF

Jusqu'ici le filtrage du son par le corps de l'auditeur a été analysé de façon binaurale sur l'axe azimutal, mais ce filtrage se produit en réalité pour les sons arrivant de tout l'espace autour de l'auditeur. Certains indices de filtrage peuvent être analysés indépendamment pour chaque oreille et sont appelés indices « monoraux ». Par exemple, comme un son arrivant de derrière la tête est filtré d'une certaine façon, un son joué sur une oreillette avec le même filtrage sera souvent ressenti derrière la tête [8]. Les filtrages du son par le torse, la tête et le pavillon des oreilles peuvent être mesurés, et sont appelés HRTF (*Head Related Transfer Functions*). Il existe des bases de données de HRTF, faites en mesurant le filtrage en intensité et phase en fonction de l'incidence d'une source sonore autour d'un auditeur ou d'un simulateur tête-torse, en milieu anéchoïque. Certaines de ces bases de données sont disponibles librement, on citera celles de l'IRCAM⁵, CIPIC⁶ et celles réalisées avec le mannequin KEMAR⁷.

⁵ <http://recherche.ircam.fr/equipes/salles/listen/>

Pour mesurer la perception spatiale, on pourrait utiliser un simulateur tête-torse pour faire les mesures binaurales, et sa base de données de HRTF, classés par position de la source réelle pour laquelle ils ont été obtenus⁸. Le filtrage mesuré en véhicule serait comparé à la base de données pour estimer la localisation de la source fantôme. Un avantage de cette technique dans le cadre qui nous intéresse est qu'elle permettrait de lever les ambiguïtés de latéralisations avant/arrière et les ambiguïtés sur les tores de confusion.

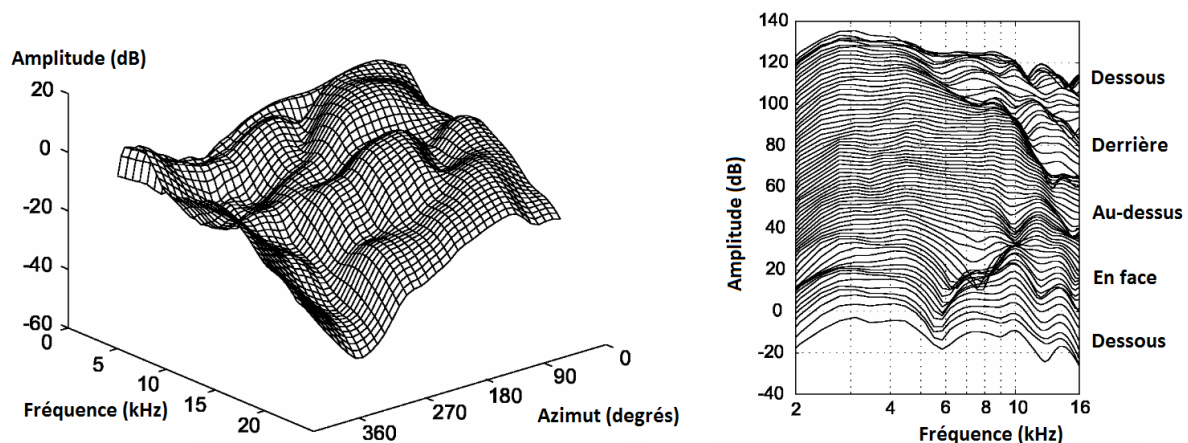


Figure 11 - Exemples de filtrage HRTF en amplitude sur les plans azimutal (à gauche) et zénithal (à droite)

Le recoupage par bande de fréquences des différences de temps et de niveau entre signaux issus des microphones gauche et droite permettrait idéalement de déterminer une position de source, mais dans la pratique on obtient, entre autre à cause de l'ambiguïté à une période près des IPD, plusieurs positions possibles qu'il faut départager, par exemple en cherchant des couples IPD/ILD « naturels » issus de la théorie Duplex. [5]

On se pose de sérieuses questions sur la pertinence de n'utiliser que des HRTF issus de mesures en milieu anéchoïque pour la localisation de sources fantômes dans un champ acoustique typique d'une automobile, et comme on ne sait pas vraiment comment l'implémenter, on ne traitera pas de cette méthode dans notre étude.

⁶ <http://interface.cipic.ucdavis.edu/sound/hrtf.html>

⁷ <http://sound.media.mit.edu/resources/KEMAR.html>

⁸ La littérature anglo-saxonne appelle cela une *HRTF lookup table*

1.4. Effet de précedence – Fusion - Influence du local d'écoute

Dans certains cas, le son qui arrive directement à l'auditeur et ses copies retardées, les réflexions, sont perçus comme un seul évènement sonore. Wallach puis Haas ont montré qu'une fusion perceptive des indices de latéralisation se produit quand deux sources rayonnent des signaux dont certaines composantes diffèrent en temps et en niveau en dessous d'une certaine limite. Un seul évènement auditif est alors perçu, et sa position dépend des différences de niveaux et de temps entre les signaux envoyés aux haut-parleurs.

Pour avoir un ordre de grandeur, on rapporte l'expérience de Haas, en milieu semi-anéchoïque (toit d'immeuble) qui consiste à diffuser avec une même amplitude un son sur deux haut-parleurs placés symétriquement de part et d'autre d'un auditeur en retardant le son du haut-parleur S_T d'un retard τ [4].

Tableau i - Résultats de l'expérience de Haas

Délai sur S_T	Effet perceptif
$0 \text{ ms} < \tau < 0,6 \text{ à } 1 \text{ ms}$	Fusion, une image fantôme flotte entre les haut-parleurs, sa latéralisation dépend des différences d'intensité et de temps entre les canaux alimentant les haut-parleurs
$1 \text{ ms} < \tau < 30 \text{ ms}$	Fusion en une image perçue sur un des haut-parleurs
$\tau > 30 \text{ ms}$	Deux images séparées, chacune sur un haut-parleur, écho

Ce que Blauert résume très clairement dans la figure suivante.

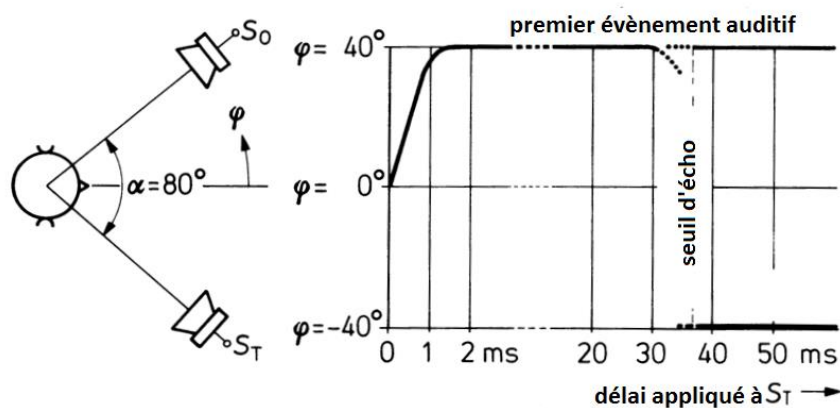


Figure 12 - Résultats de l'expérience de Haas - latéralisation perçue d'un signal de parole en fonction du délai appliqué à un haut-parleur, adapté de [6]

Les expériences citées précédemment ne concernent que des réflexions uniques créées en milieu anéchoïque ou semi-anéchoïque avec deux haut-parleurs ou avec des réflecteurs mobiles. Blauert [4] et Toole [15] ont compilé des résultats d'expériences qui montrent que les nombreuses réflexions issues du local d'écoutes (à gauche dans la figure suivante) vont être plus ou moins sommées perceptivement avec le son direct (si elles sont précoces elles vont être fusionnées) et donner des perceptions temporelle et spatiale qui ressemblent à la figure suivante à droite.

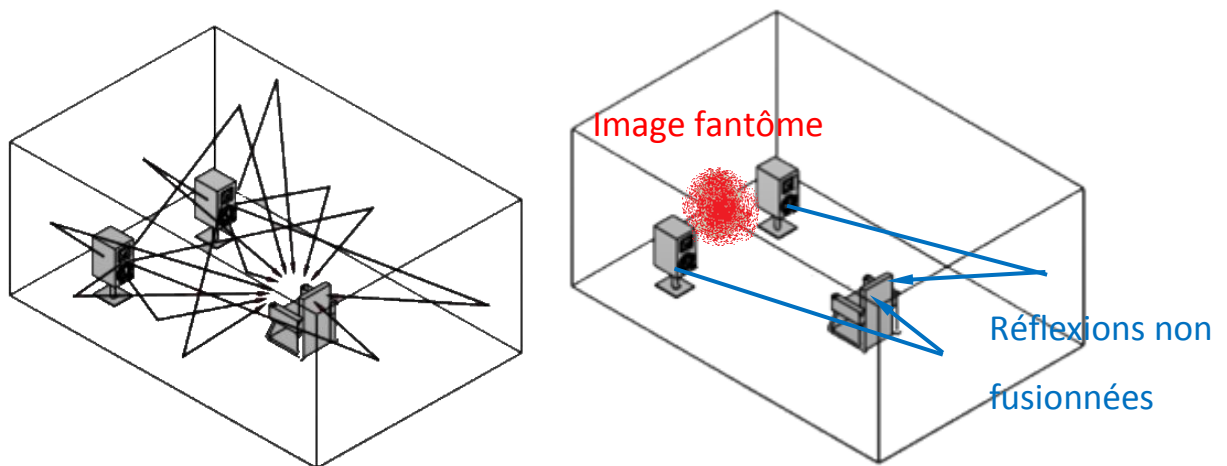


Figure 13 - Représentation des réflexions dans un local et de la perception spatiale résultante, adapté de [15]

Griseinger a montré que les réflexions du local d'écoute peuvent ajouter un caractère plaisant au son qui y est produit, et a proposé les mesures suivantes :

- élargissement de la source perçue par rapport à ses dimensions physiques (*Apparent Source Width ASW*) ;
- enveloppement (*spaciousness/Listener EnVeloppement LEV*).

On trouve dans le standard ISO3382-1997 des mesures qui permettent de qualifier objectivement la perception spatiale du son dans les salles de concerts :

- corrélation croisée inter-aurale (IACC)
- fraction latérale (LF)

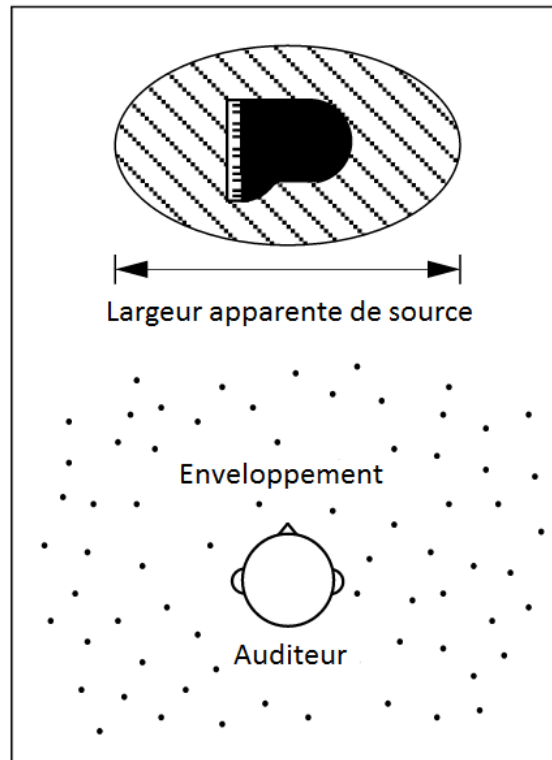


Figure 14 – Perceptions spatiales et réflexions, traduit de [16]

Keet puis Morimoto ont montré une corrélation inverse entre l'IACC calculée sur une durée de 80 ms et l'ASW [17]. Griesinger estime que ce paramètre est peu pertinent dans une petite pièce, et donc dans une automobile. Rumsey suppose lui qu'une source fantôme trop grande serait perçue comme peu naturelle et difficile à localiser.[18]

Soulodre [17] a montré une corrélation entre fraction latérale et enveloppement. Pour calculer la fraction d'énergie latérale précoce, on mesure les réponses impulsionnelles avec un microphone omnidirectionnel $p_o(t)$ et un microphone bidirectionnel $p_f(t)$, placés comme dans un couple microphonique MS, et on calcule

$$LF = \frac{\int_{5ms}^{20ms} p_f^2(t) dt}{\int_0^{20ms} p_o^2(t) dt} \quad [19]$$

Contrairement à l'IACC, la fraction latérale ne donne pas d'indication de position de la source fantôme. On ne dispose pas d'un microphone bidirectionnel de qualité satisfaisante donc cette métrique ne sera pas traitée dans notre étude.

Voici des résultats remarquables : selon Barron et Ando, l'angle d'incidence des réflexions a un impact sur leur perception et leur fusion, qu'on peut relier à l'IACC [15]. Selon Walther la fusion perceptive des réflexions est nettement moins importante si les réflexions viennent d'une autre position que le son direct [20], toujours sur ce point Morimoto a montré que les réflexions tardives venant de l'arrière augmentent plus fortement la sensation d'enveloppement que les autres réflexions [16]. L'utilisation d'une métrique directionnelle est donc à envisager à l'avenir.

La clarté doit être mesurée dans un habitacle afin d'avoir un ordre de grandeur du rapport d'amplitude des réflexions précoces et des réflexions tardives. Si le signal est de la voix, elle est généralement calculée comme :

$$C_{50} = 10 \cdot \log \frac{\int_0^{50ms} p^2(t) dt}{\int_{50ms}^{\infty} p^2(t) dt}$$

Et pour un signal musical [17] :

$$C_{80} = 10 \cdot \log \frac{\int_0^{80ms} p^2(t) dt}{\int_{80ms}^{\infty} p^2(t) dt}$$

Tout ce que l'on vient de voir a été déterminé dans des locaux typiques d'habitation ou voués à la sonorisation, et on se demande si ces métriques sont adaptées à un habitacle automobile, en particulier la fenêtre d'intégration de 80 ms choisie pour l'IACC et la clarté. Käsbach et al estiment que dans des petits locaux, et avec une diaphonie inter-aurale marquée, ce qui est notre cas, le rapport entre IACC et ASW n'est pas si évident, et que de nouvelles métriques s'imposent [21]. On suppose que dans l'habitacle automobile les réflexions sont nombreuses, mais on se pose la question de leur diffusivité, c'est-à-dire si elles viennent principalement de directions identifiables (ordre faible) ou aléatoires (ordre élevé).

Ce qui nous intéresse avant tout, c'est une reproduction perçue comme compacte de la composante corrélée entre les deux canaux de la source audio, car notre expérience des réglages en véhicule a montré que c'est la condition la plus importante pour une reproduction fidèle des caractéristiques spatiales des enregistrements, et donc on cherche à optimiser la corrélation des signaux binauraux quand un signal totalement corrélé est lu.

1.5. Corrélation inter-aurale

L'IACC est un indice de ressemblance entre les signaux issus des deux oreilles qui peut être utilisée pour modéliser la perception spatiale du son. Si on appelle $g(t)$ et $d(t)$ les signaux issus des deux oreilles-microphones, et dont les transformées de Fourier sont notées $G(f)$ et $D(f)$. Alors le produit d'inter-corrélation de ces signaux dans le domaine temporel en fonction du décalage⁹ τ est :

$$\varphi_{gd}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} g^*(t) \cdot d(t + \tau) dt$$

Dans le domaine fréquentiel:

$$\Phi[f] = G^*[f] \cdot D[f], \text{ où } G^*[f] \text{ est le complexe conjugué de } D[f]$$

Avec des signaux discrets cela s'écrit :

$$\varphi_{gd}(n) = \sum_m g^*(m) \cdot d(m + n)$$

$$\Phi[k] = G^*[k] \cdot D[k]$$

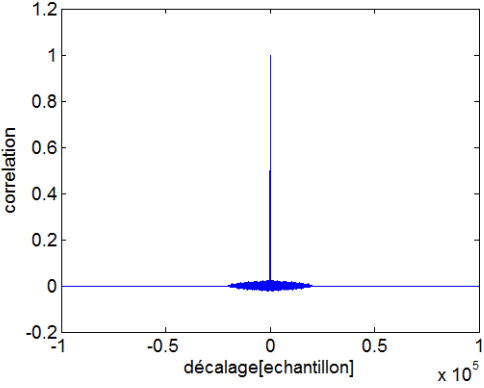
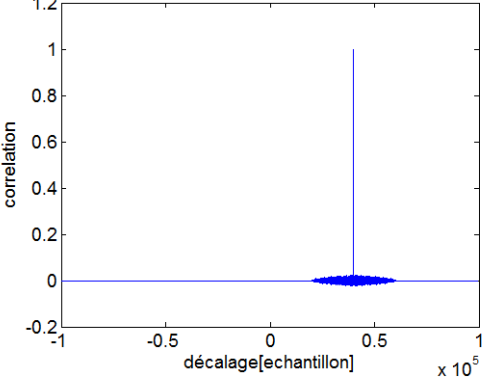
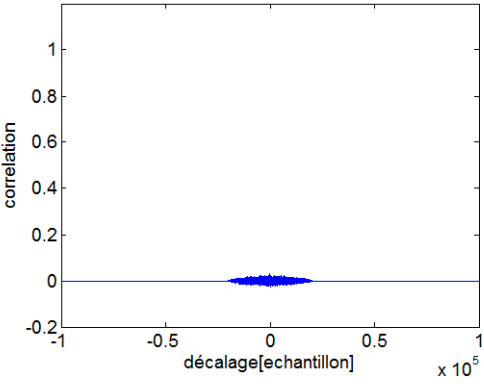
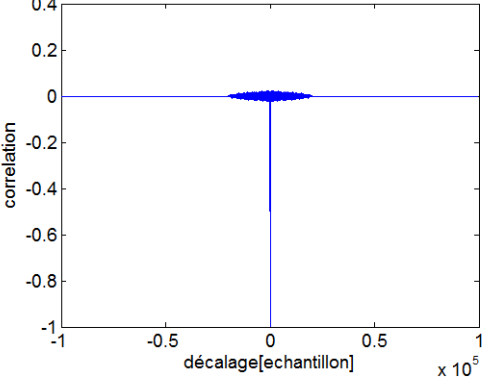
L'inter-corrélation normalisée est aussi utilisée, et en fonction du décalage n elle vaut pour des signaux discrets :

$$\varphi_{gd \text{ norm}}(n) = \frac{\sum_m g^*(m) \cdot d(m + n)}{\sqrt{\sum_p g(p)^2 \sum_q d(q)^2}} = \frac{|\sum_m G^*(m, e^{i\omega}) \cdot D(m + n, e^{i\omega})|}{\sqrt{\sum_p G(p, e^{i\omega})^2 \sum_q D(q, e^{i\omega})^2}}$$

Il est important de noter que l'ITD est le délai interne τ ou n où se trouve le maximum d'inter-corrélation. On calcule l'inter-corrélation normalisée de signaux tests usuels en fonction du décalage n , qu'on présente dans le tableau suivant.

⁹ Parfois appelé « délai interne »

Tableau ii - Inter-corrélation normalisée de signaux tests usuels

Signal test	$\varphi_{gd}(n)$	Phase des signaux
Bruit rose corrélé sur 2 canaux		Nulle
Bruit rose corrélé sur 2 canaux, un canal décalé dans le temps de 0.4×10^5 échantillons		Fonction affine de la fréquence
Bruit rose dé-corrélé sur 2 canaux		Aléatoire
Bruit rose corrélé sur 2 canaux, un canal en opposition de phase		Opposée

De nombreuses études ont montré qu'on pouvait relier la corrélation inter-aurale à la position et la largeur perçues des sources (voir [4]).

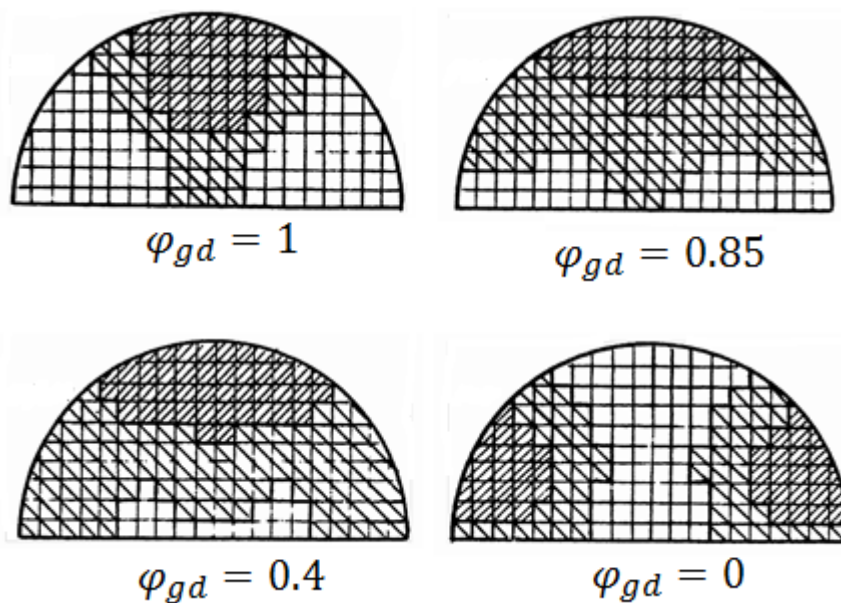


Figure 15 - Perception spatiale en fonction de la corrélation inter-aurale d'un bruit rose au casque, selon Chernyak and Dubrovsky cités dans [22]

La relation entre l'inter-corrélation entre les signaux au niveau des haut-parleurs¹⁰ et la perception spatiale a aussi été étudiée (voir figures suivantes).

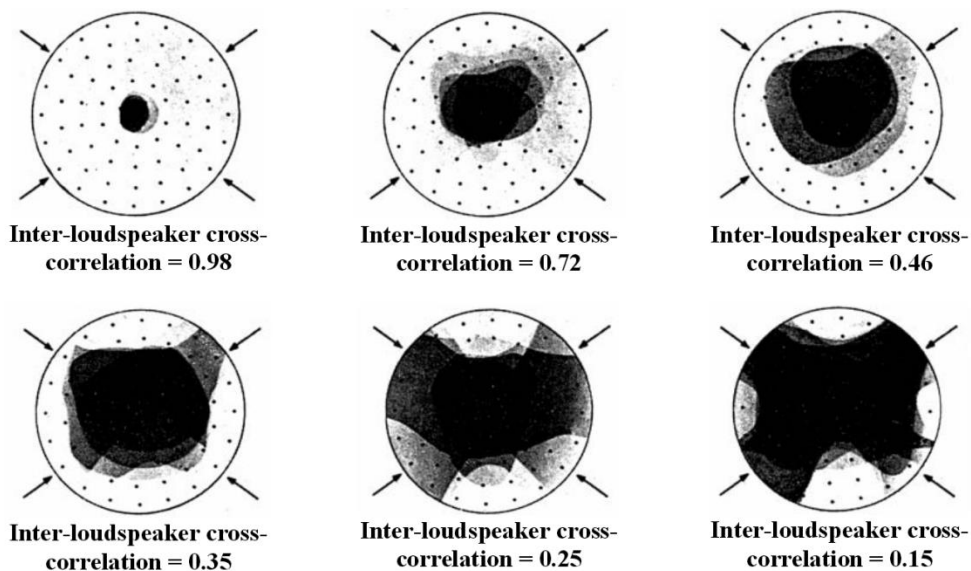


Figure 16 - Perception spatiale de panels d'écoute pour différents niveaux de corrélation de bruits filtrés entre 4 haut-parleurs dont l'orientation est indiquée par les flèches, selon Damaske cité dans [22]

¹⁰ Précédemment on s'intéressait à la corrélation des signaux au niveau des oreilles

Il apparait possible de déterminer la position et la largeur d'une source en mesurant respectivement le décalage du pic qui donne le maximum d'IACC, et la largeur de ce pic.

Sur cette base, la corrélation croisée inter-aurale est définie de façon normative (ISO 3382-1 :2009) par :

$$IACC = \max_{\tau \in [-1ms; 1ms]} \frac{\int_{t_1}^{t_2} g(t)d(t + \tau)dt}{\frac{1}{2T} \sqrt{\int_{t_1}^{t_2} g^2(t)dt \int_{t_1}^{t_2} d^2(t)dt}}$$

On note que cette IACC n'est plus en fonction du délai interne τ . La plage de décalage [-1 ms ; 1 ms] pour le calcul de maximum est à mettre en rapport avec la plage de décalage temporel qui permet l'existence de source fantôme entre les haut-parleurs dans la configuration vue précédemment. Si on observe les réponses impulsionnelles entre $t_1 = 0ms$ et $t_2 = 80ms$; l'IACC est dite « précoce » et on la note $IACC_E$. Si de plus on ne mesure pas l' $IACC_E$ pleine bande mais dans trois bandes de fréquences centrées sur 500 Hz, 1kHz et 2kHz puis que l'on en fait la moyenne, on note le résultat $IACC_{E3}$. Cette mesure est fréquemment utilisée pour qualifier des salles de concerts.

Pour modéliser ce qui se passe dans le système nerveux, de nombreux auteurs ont proposé le concept d'inter-corrélation court terme, aussi appelée inter-corrélation glissante :

$$\varphi_{glissante\ g,d}(t, \tau) = \int_{-\infty}^t g(\alpha)d(t - \alpha)y(\alpha - \tau)p(\tau)d\alpha$$

La fonction $y(t)$ est une pondération temporelle, le plus souvent exponentielle et qui annule le produit dans l'intégrale pour les valeurs « futures », $p(\tau)$ est une exponentielle décroissante qui vise à accentuer la contribution des délais internes τ quand leurs valeurs sont petites. [24] On peut imaginer mesurer cette inter-corrélation glissante tout en modifiant les réglages, et vérifier la pertinence de ces derniers en temps réel.

1.6. Corrélation inter-aurale dans le domaine fréquentiel

La FIACC (*Frequency-dependant Inter Aural Cross Correlation*) peut être utilisée pour qualifier le champ sonore mesuré avec un simulateur tête-torse. Selon Muraoka et Nakazato, les courbes théoriques de FIACC selon le type de champs sonore considéré sont :

- 1 à toutes les fréquences pour deux signaux totalement corrélés
- $\frac{\sin(kr)}{kr}$ pour deux signaux mesurés dans un champ sonore parfaitement diffus avec un écartement des capteurs de r , où $k = \frac{2\pi}{\lambda}$ est le nombre d'onde à la longueur d'onde λ [25].

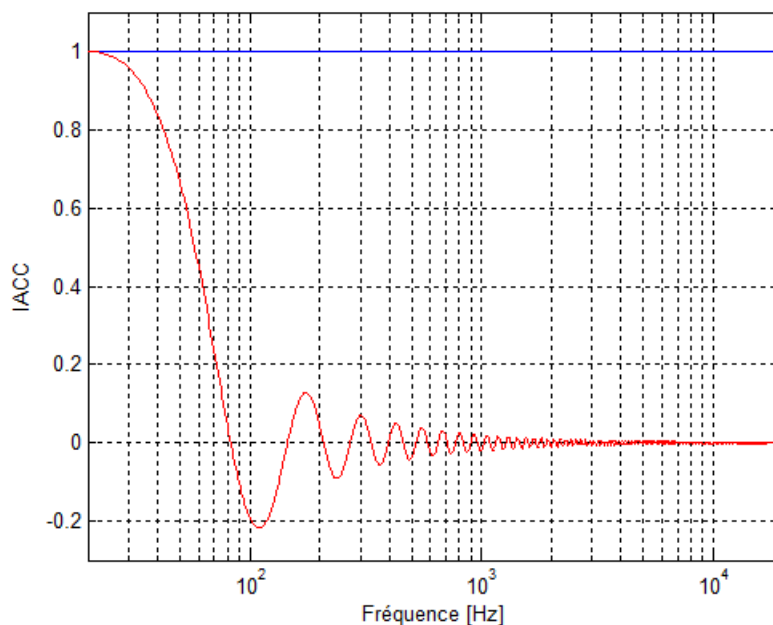


Figure 17 – FIACC théorique de deux signaux totalement corrélés (en bleu) et de deux signaux mesurés dans un champ sonore diffus avec des capteurs distants (en rouge).

Si maximiser la corrélation inter-aurale pour toutes les fréquences quand un signal corrélé est reproduit apparaît comme une priorité, il ne faut pas négliger que la reproduction fidèle de la partie dé-corrélée du signal audio est importante pour transmettre une partie de la spatialité voulue lors de l'enregistrement et du mixage du morceau. A ce sujet, Fajerson s'attend à ce qu'en diffusant un bruit dé-corrélé sur tous les canaux, la FIACC mesurée ressemble au sinus cardinal vu précédemment. [26]

1.7. Modèle de Jeffress et extensions

Jeffress a modélisé l'analyse par le système nerveux de la corrélation entre les signaux des oreilles par un circuit. Il est composé de deux axones qui transmettent chacun les signaux issus d'une oreille avec un retard de propagation dépendant de la position de la source, et des neurones « multiplicateurs » qui jouent le rôle de détecteur de coïncidence.

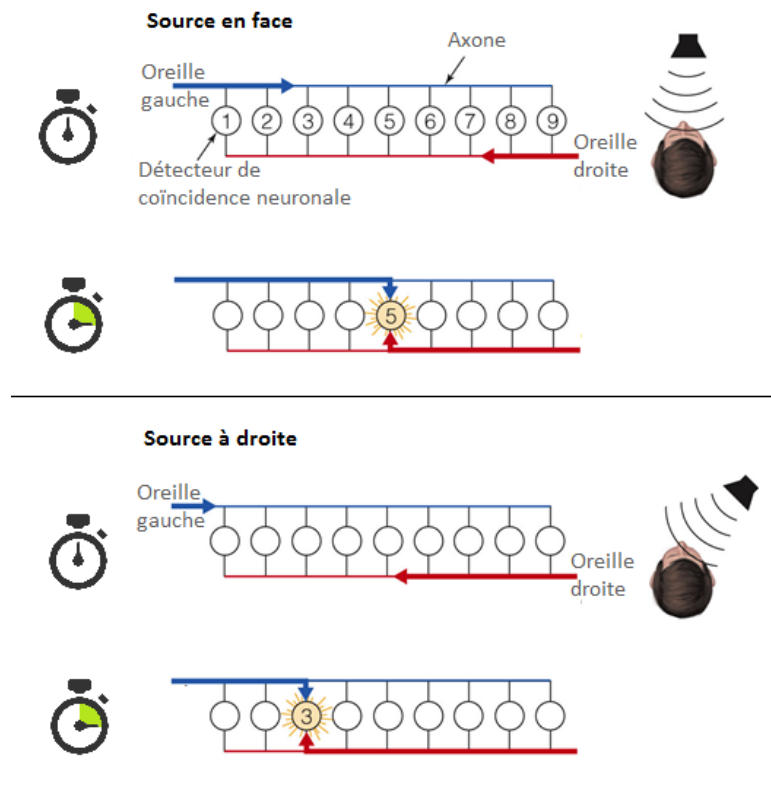


Figure 18 – Explication simplifiée du circuit de Jeffress, traduit de [27]

Des axones transmettent les signaux auditifs gauche (en bleu) et droit (en rouge) aux neurones, représentés par des cercles numérotés. Lorsque la source est en face, le signal part sur les axones gauche et droit en même temps. Plus tard les signaux se rencontrent au neurone 5 et l'activent. Quand la source est à droite, le signal part sur l'axone droit en premier. Plus tard les signaux se rencontrent au neurone 3, ce qui donne une information de latéralisation de la source à droite.

Dans le modèle de Jeffress, l'analyse de corrélation qu'on vient de voir est réalisée après découpage par bandes de fréquences et conversion en influx nerveux des signaux issus des

oreilles, typiquement par redressement simple alternance [9]. Dans les figures suivantes l'inter-corrélation inter-aurale est modélisée par des retards « T » et des multiplicateurs « X ».

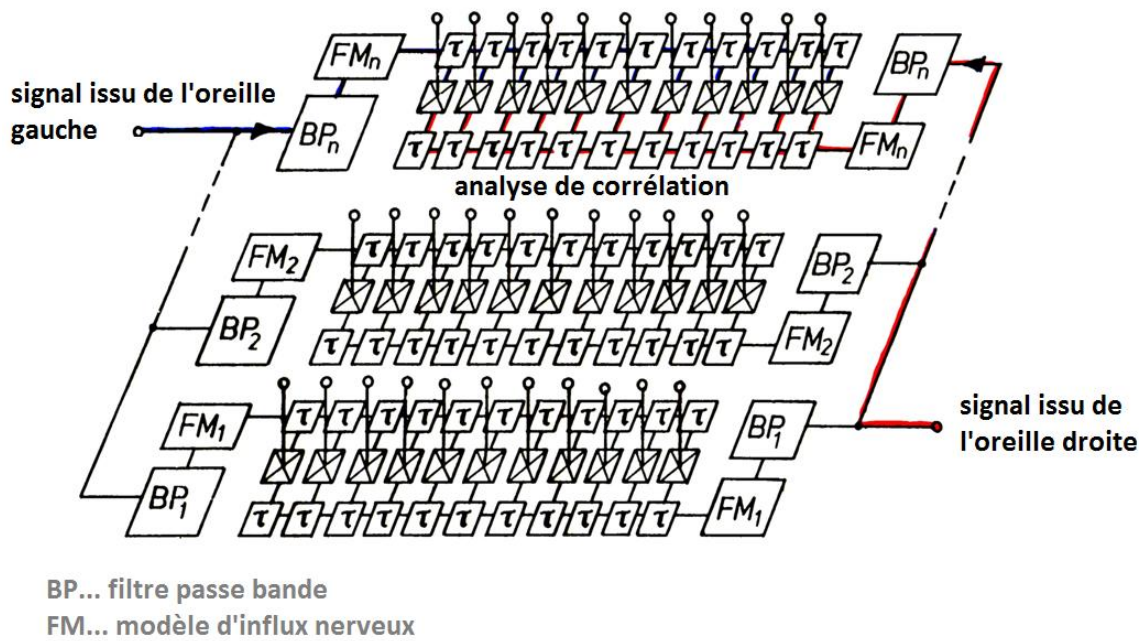


Figure 19 - schéma du modèle de Jeffress, selon [4]

Pour se rapprocher encore plus de la perception spatiale humaine, Lindemann propose de combiner l'inter-corrélation de Jeffress et deux processeurs d'indices monauraux, un par canal, ce qu'on peut schématiser pour chaque bande critique comme dans la figure suivante.

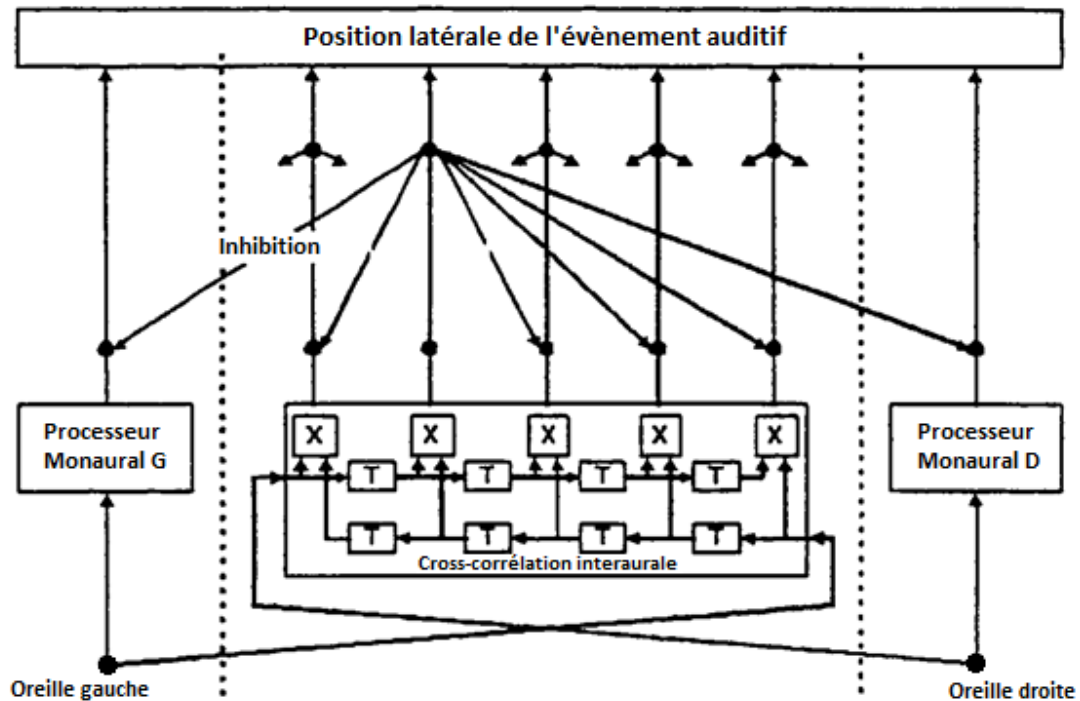


Figure 20 - schéma du modèle de Lindemann, traduit de [4]

Une autre addition de Lindemann au modèle de Jeffress est que les signaux issus du calcul d'inter-corrélation peuvent s'inhiber entre eux, c'est-à-dire amoindrir ou annuler leur prise en compte dans la localisation (le signal partant en premier inhibant les signaux plus tardifs liés au même évènement jusqu'à une durée paramétrable), ou inhiber les indices monauraux. Comme l'a montré Blauert, le modèle de Jeffress-Lindemann permet de plutôt bien modéliser les phénomènes de fusion des réflexions, d'affiner les pics d'inter-corrélation et donc augmenter la précision de latéralisation. Gaik a adapté le modèle de Lindemann afin de le rendre utilisable avec des signaux issus d'un simulateur tête-torse : l'inhibition contralatérale est pondérée de façon asymétrique entre les axones gauche et droite lors de phases d'apprentissage afin de retenir les couples (ITD, ILD) « naturels ». Les résultats expérimentaux cités par Blauert montrent que le modèle de Jeffress – Lindemann – Gaik devrait être testé dans le cadre qui nous intéresse [4].

1.8. Boite à outils de Faller et Merimaa

Cette boîte à outil décrite dans [28], et disponible en ligne, propose un autre modèle de localisation de source sonore en milieu réverbérant. La capacité de l'audition auditive à fusionner le trajet direct et les premières réflexions pour s'affranchir de ces dernières est modélisé par l'analyse conjointe de la corrélation interaurale normalisée, de la différence de temps binaurale et de la différence de niveau binaurale en fonction de la fréquence centrale d'un banc de filtres approximant les bandes critiques de l'audition humaine.

L'idée directrice de ce logiciel est qu'en sélectionnant les différences de temps et les différences de niveau inter-aurales qui coïncident avec une corrélation inter-aurale au-dessus d'un certain seuil, on obtient un ensemble d'indices de localisation proche de ceux obtenus avec les mêmes sources placées au même endroit mais en champ libre, avec une certaine probabilité.

Des résultats des modèles de Jeffress-Lindemann et de Faller-Merimaa seront présentés dans le dernier chapitre.

1.9. Précision de localisation

De nombreuses études ont montré que la résolution de l'audition spatiale sur l'axe horizontal est de l'ordre de quelques degrés ([4] p40), ce qui encore une fois est inférieur à la résolution spatiale de la vue. La localisation par IPD est ambiguë à une période près mais l'analyse de couples (ITD, ILD) à plusieurs fréquences permet de mieux lever les ambiguïtés [29], donc les caractéristiques suivantes rendent les sons durs à localiser par l'audition humaine [4] :

- son continu plutôt que transitoire ou impulsif ([4] et Hartman cité par [9]),
- basses fréquences en milieu réverbérant,
- fréquences pures et spectres restreints, par rapport aux contenus à spectres larges.

On note que :

- un son diffusé avec une opposition de phase entre les deux oreilles apparaît très large et diffus ([4] p148),
- un bruit large bande est plus difficile à localiser dans la réverbération que dans un milieu anéchoïque (Hartmann et Giguère & Abel cités par [9]),
- les mouvements de tête permettent de lever certaines ambiguïtés,
- la connaissance de la source et de l'environnement facilitent la localisation et la perception de la distance de la source.

1.10. Représentation de la scène sonore - Ségrégation

La perception spatiale du son naturel ou reproduit peut être décrite comme une scène sonore, c'est-à-dire la perception d'un ensemble de sources sonores virtuelles ou réelles placées dans l'espace perceptif (paradigme explicité par Rumsey dans [30]). La scène sonore peut être décrite avec un vocabulaire adapté et plutôt standard (la formation LiT en fournit un exemple).

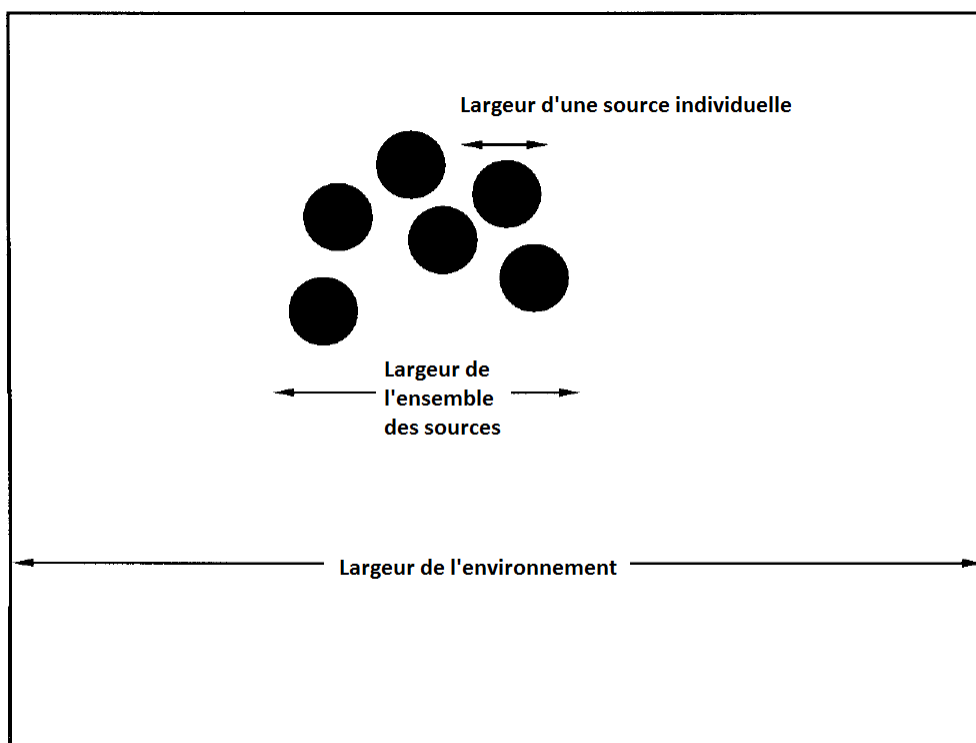


Figure 21 - Exemple de description de la scène sonore (traduit de [30])

Dans [31], Usher et Woszczyk proposent un système de représentation de la scène sonore par les panels d'écoute qui donne ce type de figure :

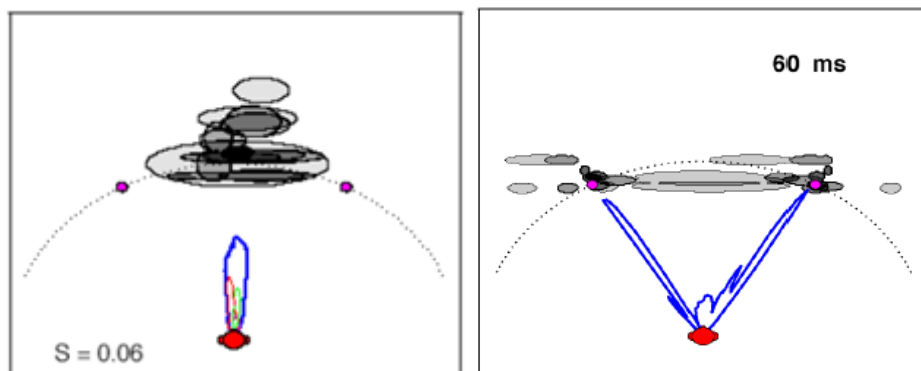


Figure 22 – Représentations de la perception spatiale par un panel d'auditeurs d'une voix centrée, puis avec un décalage temporel entre les deux haut-parleurs.

En 2007, Lecoecur a proposé ce type de représentation pour décrire la spatialisation perçue du son :



Figure 23 - Représentations de la perception spatiale d'une scène sonore par un panel d'auditeurs

On attend d'un algorithme d'analyse informatisée de scène sonore¹¹ tel que celui décrit dans [9] et [32] qu'il puisse représenter d'une façon comparable les sources fantômes, leurs latéralisations, leurs compacités, la largeur, la régularité angulaire de la scène sonore et son

¹¹ en anglais CASA pour *Computer Assisted Auditory Scene Analysis*

côté enveloppant quand on diffuse un signal stéréophonique complexe comme de la musique, et que cela corresponde avec les écoutes critiques.

La source fantôme la plus importante dans le cadre de notre étude est le centre de la scène sonore stéréophonique, le lieu où sont perçus les signaux fortement corrélés dans des conditions d'écoute proche des recommandations. C'est là que se situe le maximum d'énergie et des signaux significatifs, comme la voix parlée et chantée, et des contenus monophoniques. Les cahiers de charges des clients d'Arkamys et les rapports de réglages montrent qu'une voix précisément localisée, frontale et centrée est préférée dans la plupart des cas. On souhaite mesurer en priorité la perception spatiale de cette source fantôme.

La thèse de May [32] traite des algorithmes d'analyse informatisée de scène sonore¹² en général, alors que celle de Hummersome [9] étudie leurs performances en milieu réverbérant. On y apprend que la validité de l'analyse des indices binauraux diminue à mesure que le temps de réverbération augmente et que le rapport direct/réverbéré diminue. Les algorithmes *CASA* doivent donc s'accommoder de la réverbération pour améliorer leur performance en localisation et séparation de sources. Selon ces auteurs, plusieurs logiques existent :

- chaîner un algorithme de dé-réverbération et un algorithme *CASA* conçu pour fonctionner en milieu anéchoïque,
- filtrer spatialement pour séparer le son direct de la source des sons venant d'autres directions,
- modéliser l'effet de précedence,
- combiner plusieurs indices.

Nous avons cherché des modèles d'audition spatiale conçus pour tenir compte de la réverbération, et pour des raisons pratiques nous avons retenu le modèle de Lindemann, qui modélise l'effet de précedence par inhibition, et la boîte à outils conçue par Fallert et Merimaa.

¹² En anglais *CASA*

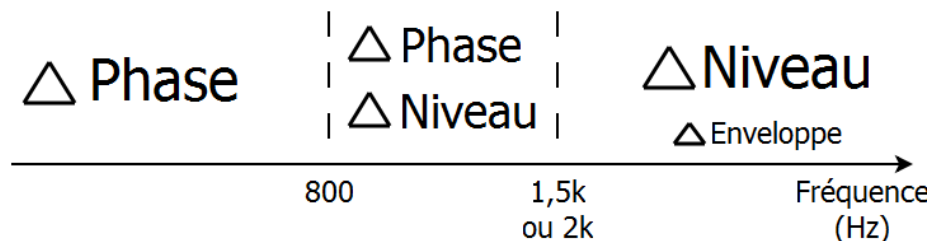
A ce stade on peut se demander :

- quelles sont la vitesse de rafraîchissement et la latence des programmes informatiques embarquant ces modèles ? Est-ce compatible avec un affichage en « temps réel » ?
- quelle précision d'analyse du filtrage morphologique est suffisante ?

La littérature à notre disposition communique malheureusement peu d'éléments sur ces performances, il faudra les vérifier nous-mêmes.

1.11. Résumé - Conclusion partielle - Discussion

Différents modèles permettant de relier le filtrage du son par le corps de l'auditeur et la position latérale perçue d'évènements sonores ont été présentés. Le modèle de base est appelé Duplex, et repose sur l'analyse les changements d'indices binauraux de la façon suivante:



Des équations permettent de relier la latéralisation de la source aux différences binaurales de phase et de niveau mesurées. Ce modèle est limité au plan azimutal et souffre beaucoup d'ambigüités de localisation.

Des bases de données des filtrages en intensité et en temps en fonction de l'incidence sur tout l'espace autour de l'auditeur (HRTF) sont disponibles mais on ne sait pas comment les utiliser de façon efficace dans le cadre de notre étude. Comparer le filtrage mesuré avec une base de HRTF ne permet pas toujours de déterminer précisément une position de source, surtout en milieu réverbérant, mais permet parfois de lever des ambigüités de latéralisation.

L'influence du local sur la perception spatiale est avérée et ne peut être négligée. Selon leur retard, les réflexions peuvent être fusionnées avec le trajet direct, et contribuer à la perception de la taille de la source ou de l'enveloppement sonore. Pour modéliser la fusion et l'effet de précedence, et ainsi augmenter fortement leur performance en milieu réverbéré, des modèles plus évolués se basent sur l'analyse conjointe et conditionnelle de différents indices binauraux (modèle de Jeffress-Lindemann-Gaik, modèle de Faller et Merimaa).

L'expérience des réglages de système de diffusion pousse en priorité à rendre la perception d'un signal corrélé la plus frontale, la plus compacte et la plus ponctuelle possible. Dans ce qui suit, on étudiera principalement les différences de temps d'arrivée en dessous de 2kHz.

2. Reproduction stéréophonique en véhicule

2.1. Stéréophonie

2.1.1. Principe

Un objet qui produit du son peut être qualifié dans notre étude de source réelle (par exemple un haut-parleur). Une source virtuelle ou fantôme, elle, n'existe pas physiquement. Elle est perçue dans l'espace par l'auditeur après l'analyse des sons reçus à ses oreilles, par exemple après que ces sons aient été émis par une ou plusieurs sources réelles dans un environnement donné. Les sources virtuelles peuvent être localisées à des positions autres que celle des sources réelles.

Le format audio le plus répandu en automobile est la stéréophonie, qui repose sur la création de sources sonores virtuelles à partir de deux canaux envoyés à des haut-parleurs ou des écouteurs. La scène sonore stéréophonique est l'ensemble des sources virtuelles, des illusions de sources réelles créées en diffusant des sons plus ou moins semblables. Les liens entre la similitude de signaux reçus et la localisation de sources ont été vus au chapitre 1.

2.1.2. Recommandations et système de référence

L'écoute au casque permet de maîtriser quel signal est envoyé à quelle oreille de l'auditeur. Ce n'est pas le cas quand on écoute avec des haut-parleurs car il se produit un phénomène de diaphonie inter-aurale : chaque oreille reçoit un mélange des signaux issus des chaque haut-parleur et de leurs réflexions dans la pièce.

La configuration d'écoute optimale en stéréophonie avec des haut-parleurs est définie par les recommandations ITU-R BS.775-3[33] comme un triangle où les haut-parleurs sont placés à 30° de part et d'autre de l'axe médian.

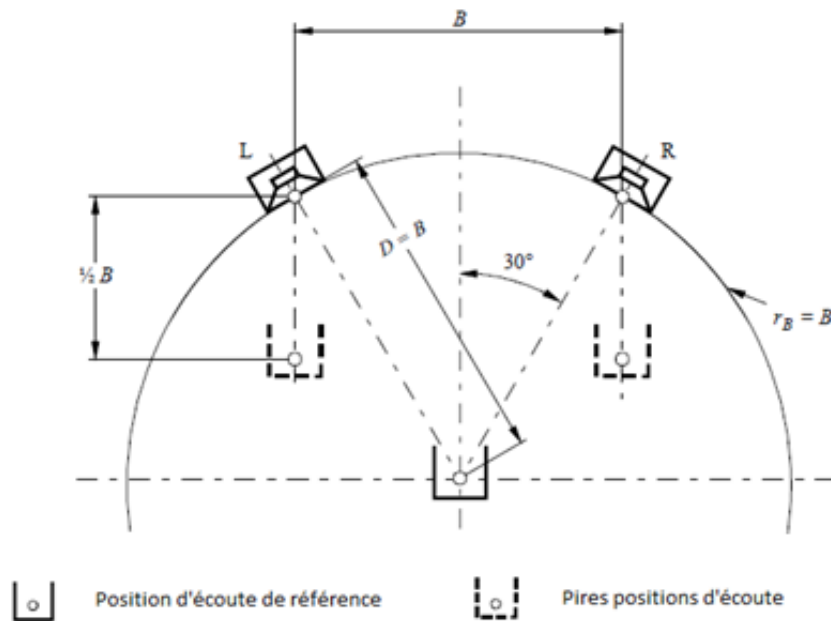


Figure 24 - Extrait de la recommandation ITU-R BS.1116.1

Dans la recommandation ITU-R BS.1116.1 [1], la base B du triangle formé par les haut-parleurs frontaux et l'auditeur est recommandée être une distance entre 2 et 3 mètres, la distance de l'auditeur aux haut-parleurs D est entre 1 et 2 fois la base B . Les « centres acoustiques » des haut-parleurs doivent se trouver à la hauteur des oreilles des auditeurs. Si ces conditions sont respectées, l'auditeur est placé au « *sweet spot* », le lieu optimal pour l'écoute stéréophonique.

La définition d'une pièce d'écoute optimale est un vaste sujet, Toole présente plusieurs aspects d'appréciation dans [15]. Une pièce d'écoute stéréophonique de référence pour une écoute analytique est décrite par la recommandation ITU-R BS.1116.1.

La stéréophonie est capable de reproduire des enregistrements monophoniques, mais Toole¹³ a montré qu'utiliser deux sources réelles de part et d'autre de l'auditeur, au lieu d'une source réelle en face, crée un filtrage en peigne néfaste compensé dans une certaine mesure par les réflexions dans le local d'écoute.

¹³ p152

2.2. Caractérisation acoustique et électro-acoustique des habitacles automobiles

Etudions les facteurs qui limitent la fidélité du système de reproduction sonore dans une voiture tels qu'ils sont appréhendés lors des réglages.

2.2.1. Local d'écoute

L'habitacle est un très petit local d'écoute de forme complexe bien que plutôt parallélépipédique, d'un volume de l'ordre de quelques mètres cubes. A l'intérieur de ses parois composites, on trouve une distribution complexe, mais globalement symétrique par rapport à l'axe médian du véhicule, de surfaces plutôt réfléchissantes (fenêtres et instruments de bord), diffusantes (plastiques grainés) et absorbantes (sièges, habillages en tissus, **passagers**) d'un point de vue acoustique. La présence du volant et de la « casquette » des instruments de bord limite la symétrie entre les deux places avant, et l'expérience montre que les conséquences de cela sur le timbre et la localisation sont importantes, principalement aux fréquences médiums et aigues.

Le local est habituellement clos, l'audio réglée et évaluée avec tous les ouvrants fermés mais des compromis sont parfois demandés à ce sujet, par exemple dans le cas d'une décapotable. Les dimensions précises de l'habitacle et son volume sont rarement connues ou prises en compte avec précision lors des réglages.

L'impédance acoustique des matériaux présents est parfois mesurée ([34] et [35]) mais malheureusement ces données et les plans 3D sont très rarement communiqués pour les études d'implantation de haut-parleurs et les réglages d'égalisation. Une meilleure connaissance des propriétés acoustiques des matériaux constitutifs et des parois de l'habitacle permettrait un meilleur placement des haut-parleurs, une meilleure compréhension des phénomènes complexes qui régissent le champ sonore dans le véhicule. Dans [35], une modélisation par éléments finis tire profit de ces données pour simuler la propagation du son dans un intérieur d'automobile Hyundai.

Dans ce qui suit, on présente comment est considérée la propagation du son lors des réglages.

2.2.2. Position des passagers

Dans une automobile de série typique, la position de la plupart des passagers est fixe, et hors de l'axe médian entre les haut-parleurs sauf pour le passager à l'arrière au centre. Par conséquent à la plupart des positions d'écoute, les haut-parleurs ne sont pas positionnés symétriquement à gauche et à droite de l'auditeur comme dans les recommandations.

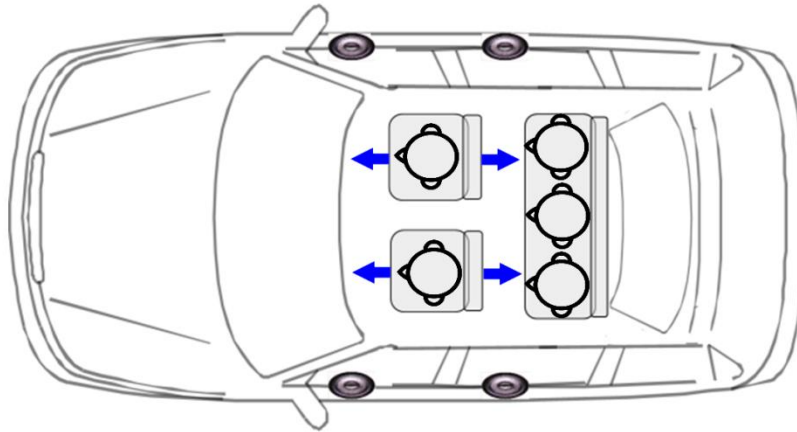


Figure 25 - Positions possibles des auditeurs dans une voiture

Les passagers plus ou moins nombreux et plus ou moins grands sont libres de bouger la tête, certains sièges peuvent coulisser en avant ou en arrière, ce qui donne de nombreuses positions d'écoute possibles. Lorsqu'on mesure la réponse acoustique d'un véhicule, on prend parfois cela en compte en mesurant pour chaque siège à différentes positions longitudinales, et en faisant une moyenne des réponses pondérée par l'importance de la position. Un conducteur est généralement présent et privilégié lors des réglages.

2.2.3. Caractéristiques et implantation des haut-parleurs

En voiture les enregistrements stéréophoniques sont reproduits par des systèmes comprenant à partir de deux haut-parleurs pour des systèmes d'entrée de gamme (typiquement des véhicules utilitaires) et jusqu'à une trentaine de haut-parleurs pour des systèmes de haut de gamme. Le placement des haut-parleurs, le choix de leurs supports et des charges acoustiques (portières, tableau de bord, plage arrière, dessous des sièges pour les caissons de basses) répond à des contraintes parfois étrangères à la qualité acoustique (design, praticité, présence de rangements). Dans une configuration d'entrée de gamme typique à 4 canaux d'amplification et 4 haut-parleurs, sans traitement audio, le routage des signaux revient sensiblement à la représentation suivante.

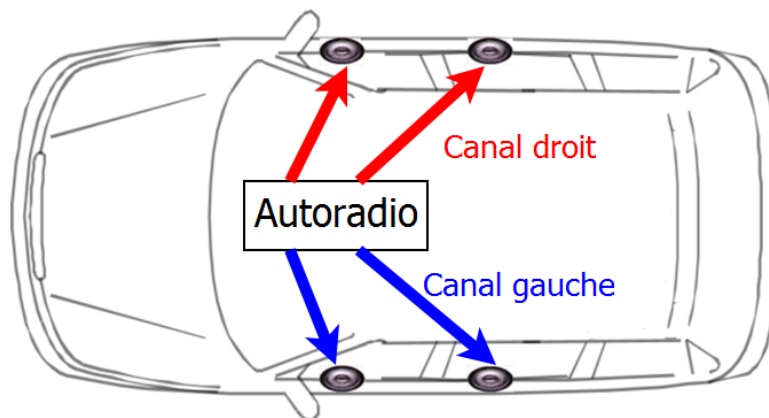


Figure 26 - Signaux envoyés aux haut-parleurs sans traitement

On constate donc qu'indépendamment du traitement d'égalisation :

- les signaux envoyés aux haut-parleurs avant gauche et arrière gauche sont toujours fortement corrélés, puisqu'issues du même canal gauche de la source,
- les signaux envoyés aux haut-parleurs avant droit et arrière droit sont toujours fortement corrélés, puisqu'issus du même canal droit de la source,
- les signaux envoyés aux haut-parleurs avant gauche et avant droit sont plus ou moins corrélés, en fonction de la corrélation des canaux gauche et droit de la source,

- les signaux envoyés aux haut-parleurs arrière gauche et arrière droit sont plus ou moins corrélés, en fonction de la corrélation des canaux gauche et droit de la source.

Les réponses en amplitude en fonction de la fréquence de deux haut-parleurs peuvent varier dans le même lot dans les tolérances fixées contractuellement entre le constructeur automobile et son fournisseur de haut-parleurs. Les haut-parleurs embarqués dans les automobiles ne sont pas des sources omnidirectionnelles, et on imagine que leur directivité renforce les ILD. Les haut-parleurs d'aigus et de médium-graves sont plus éloignés dans une automobile que dans une enceinte Hifi typique et le raccord fréquentiel (*cross over*) entre ces haut-parleurs n'est pas toujours très bien maîtrisé. Pourtant, il est admis qu'un bon ajustement de ce raccord permet de minimiser les problèmes de fidélité fréquentielle et de directivité [36].

Pour les protéger, des grilles sont placées devant les haut-parleurs. Selon la distance entre haut-parleur et grille, l'orientation de la grille, la taille des perforations et le taux de perforation, ces grilles ont une influence sur la réponse des haut-parleurs qui peut être positive (réduction de la directivité de tweeters) ou négative (diminution des fréquences très aiguës reproduites, filtrage en peigne suite au mélange du signal issu du HP et de réflexions sur la grille).

Les panneaux de portes rayonnent également et compliquent davantage l'émission du son. La directivité de l'assemblage haut-parleur-porte-grille n'est jamais connue lors des réglages, mais ce phénomène est simulé dans [35], on voit sur la figure suivante que l'influence du panneau de porte sur le rayonnement est importante.

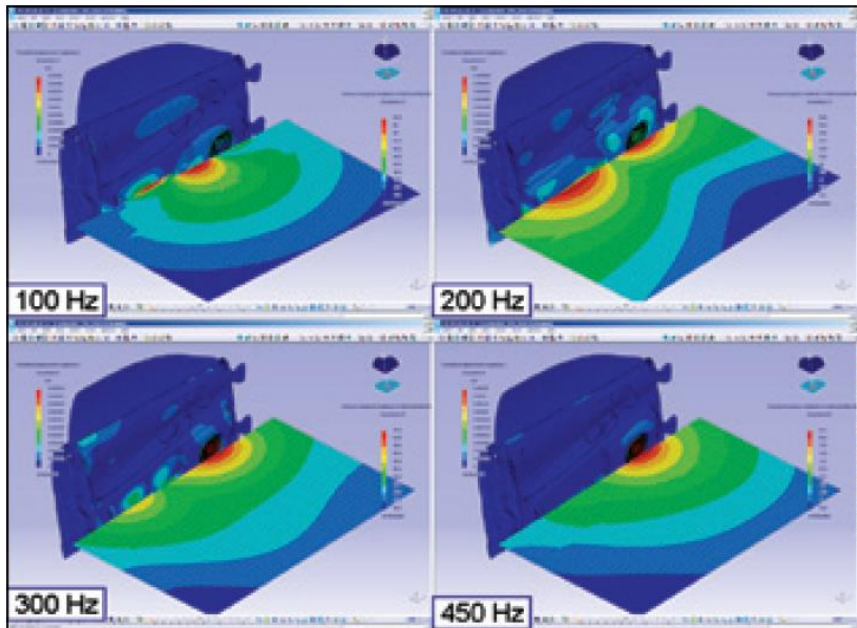


Figure 27 – Modélisation de la pression sonore rayonnée par un haut-parleur placé dans un panneau de porte à différentes fréquences [35]

La conséquence de toutes ces imperfections est que la réponse en amplitude en fonction de la fréquence de chaque haut-parleur pris séparément diffère fortement pour la même position d'écoute.

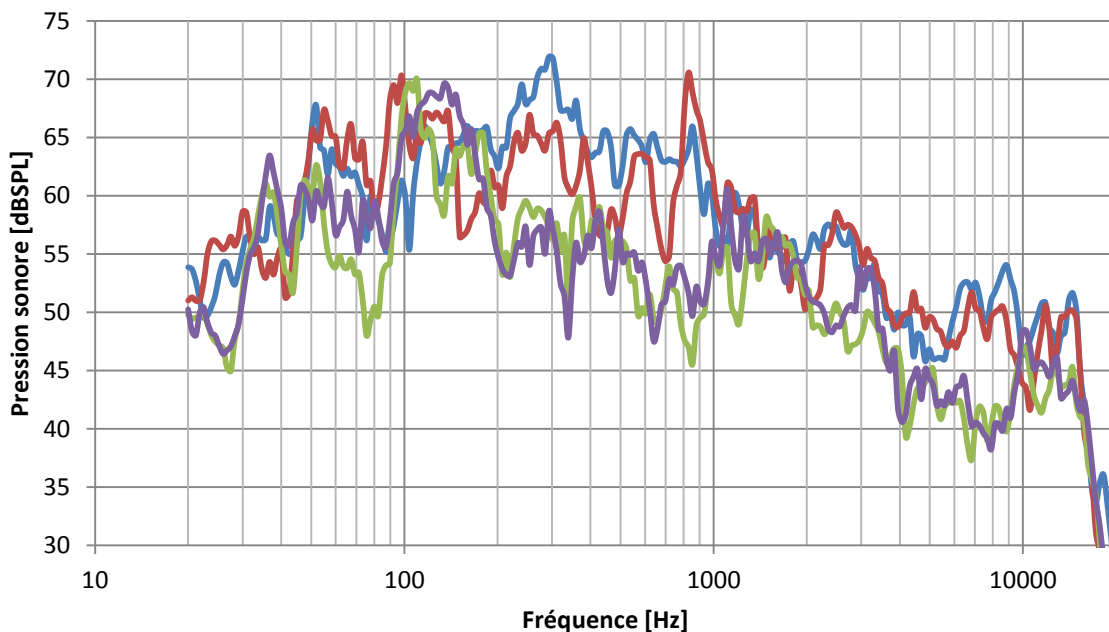


Figure 28 - Comparaison pour une même position de mesure de la réponse amplitude (fréquence) de haut-parleurs du même modèle, placés à différentes positions dans un habitacle

Tout ce qui vient d'être vu limite fortement la fidélité de restitution. Cependant, l'égalisation de systèmes électroacoustique a une efficacité d'autant plus importante que la zone d'écoute

est limitée, et à défaut d'être optimales, les positions des sources et des auditeurs sont plutôt bien définies, donc il est possible d'affiner énormément le compromis lors du réglage.

On propose maintenant d'utiliser des outils classiques d'analyse de l'acoustique d'une pièce pour tenter de qualifier l'acoustique de l'habitacle.

2.2.4. Réverbération, clarté et distance critique

Comme on l'a vu au chapitre 1, la réverbération a un effet sur la localisation perçue [4][15], mais peu d'auteurs partagent des chiffres précis sur la réverbération en véhicule. La durée de réverbération mesurée dans un local dépend des points d'émission et de mesure utilisés. Elle varie aussi selon qu'elle est calculée sur toutes les fréquences audibles ou par bande de fréquence, et selon la fenêtre de décroissance de niveau après arrêt de la source utilisée pour la calculer.

Voici des mesures de TR60 mesurée par bandes de fréquences dans plusieurs berlines par Bennetts et Morris-Kirby [37], en considérant une fenêtre de décroissance de -5dB à -35dB :

Tableau iii - TR60 dans 3 véhicules

Fréquence centrale (Hz)	160	200	250	315	400	500	630	800	1000	1250	1600	2000
TR60(ms) véhicule 1			81.1	--	48.9	54.7	55.5	61.4	50.0	66.5	54.9	52.0
TR60(ms) véhicule 2	128.6	173.3	101.2	178.1	107.6	109.9	102	75.2	96.1	110.2	100.6	110.7
TR60(ms) véhicule 3	111	68.8	51.4	67.8	90.5	70.6	68.9	52.6	43.5	76.4	60.3	56.4

Fréquence centrale (Hz)	2500	3150	4000	5000	6000	8000	10000	12500	16000
TR60(ms) véhicule 1	54.8	60.8	64.8	58.6	64.4	65.1			
TR60(ms) véhicule 2	76.6	91.4	65.4	58.7	54.1	55.3	50.9	44.4	41.2
TR60(ms) véhicule 3	55.3	54.7	58.5	61.9	62.6	63.6	62.8	67.1	50.3

Ces auteurs précisent que leurs résultats suivent la norme ISO 3382 qui préconise au moins trois positions de sources et six points de mesure pour minimiser les effets locaux.

Comparons ces résultats à ceux de TR60 que donne le logiciel REW beta 6¹⁴ à partir de mesures au simulateur tête-torse faites dans le véhicule « G », sans traitement audio, en lisant un balayage fréquentiel sur tous les haut-parleurs du véhicule, présenté dans le tableau et les figures suivants.

Tableau iv - TR60 dans le véhicule "G"

	EDT (0 à -10dB)	T20 (-5 à -25dB)	T30 (-5 à -35dB)
Microphone à gauche	64ms	98ms	326ms
Microphone à droite	40ms	82ms	235ms

On constate que les différentes méthodes de calcul de durée de réverbération globale diffèrent entre elles dans un rapport de 1 à 5 fois pour une même oreille, et diffèrent entre les deux oreilles (TR plus petit d'environ 1,3 fois à droite qu'à gauche). Le calcul pour une décroissance de 30dB apparaît nettement trop élevé par rapport aux résultats de [37] qui utilisaient la même méthode de calcul. Est-ce attribuable au trop faible nombre de points de mesures ? Calculons maintenant le TR selon différentes méthodes par bandes de fréquences.



Figure 29 – Durée de réverbération en fonction de la fréquence selon différentes méthodes dans le véhicule "G", microphone de l'oreille gauche du simulateur tête-torse

¹⁴ <http://www.roomeqwizard.com/>

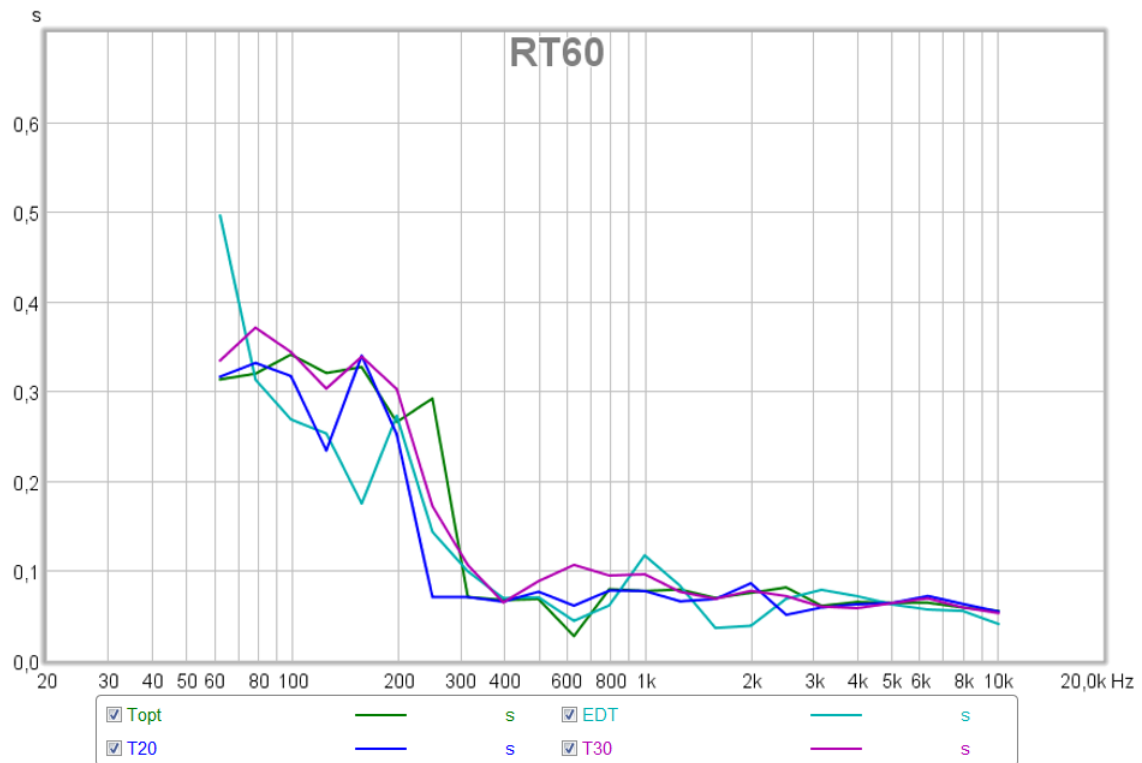


Figure 30 - Durée de réverbération en fonction de la fréquence selon différentes méthodes dans le véhicule "G", microphone de l'oreille droite du simulateur tête-torse

On constate sur les figures précédentes qu'en basses fréquences les méthodes de calcul de temps de réverbération diffèrent fortement, alors qu'elles sont dans un gabarit bien plus serré en hautes fréquences, où l'on trouve des résultats comparables à ceux de [37], et un TR60 d'environ 70ms.

On calcule maintenant la clarté telle que définie au chapitre 1 (code Matlab « clarte.m » en annexe C), sur tout le spectre, et on note chaque résultat comme un couple oreille gauche/oreille droite.

Tableau v - Clartés C50 et C80 dans les véhicules "D", "E" et "F"

Véhicule	« D »	« E »	« F »
C50 sans traitement [dB]	0 / -4	7 / 10	7 / 7
C50 avec traitement [dB]	-6 / -3	15 / 6	7 / 6
C80 sans traitement [dB]	5 / 2	17 / 11	9 / 9
C80 avec traitement [dB]	3 / 7	17 / 9	13 / 10

Sans traitement, la clarté C80 varie donc entre +2 dB et +17 dB dans ces configurations. Par comparaison les locaux d'écoute typiques ont un C80 qui varie de -7 dB à +10dB, et il est habituellement recommandé de rester dans un intervalle -4 dB à +5 dB [38]. De ce point de vue, les voitures apparaissent comme des locaux plutôt clairs pour écouter de la musique. Le réglage n'améliore pas toujours ces mesures de clarté.

2.2.5. Analyse de la propagation

On définit la distance critique comme la distance où le champ direct est égal au champ réverbéré. Pour avoir un ordre de grandeur, en imaginant un champ sonore homogène et diffus, on pose $TR_{60} \cong 70 \text{ ms}$ et on peut approximer grossièrement cette distance par :

$$d_c = \frac{1}{4} \sqrt{\frac{\gamma A}{\pi}}$$

où γ est le coefficient de directivité de la source ($\gamma = 1$ pour une source omnidirectionnelle, comme la directivité des haut-parleurs dans leur charge n'est pas connue, on prend 0.5 pour avoir un ordre de grandeur), A est la surface d'absorption équivalente et TR_{60} la durée de réverbération en s. Par manque de données sur le coefficient d'absorption des parois, on utilise une approximation basée sur la formule de Sabine $TR_{60} = \frac{0.161V}{A}$, où V est le volume du local d'écoute en m^3 . On obtient :

$$d_c \approx 0.057 \sqrt{\frac{\gamma V}{TR_{60}}}$$

Pour avoir un ordre de grandeur, on considère un habitacle parallélépipédique vide de *largeur* = 2,5 m ; *longueur* = 1,5 m et *hauteur* = 1,2 m , alors $V_{\text{intérieur}} = 4,5 \text{ m}^3$ et $S_{\text{intérieure}} = 17 \text{ m}^2$, on obtient entre 30 et 40 cm de distance critique selon les fréquences.

Hänsler et Schmidt mesurent dans un habitacle de voiture une distance critique de 90 cm. Leur protocole expérimental est décrit dans [39].

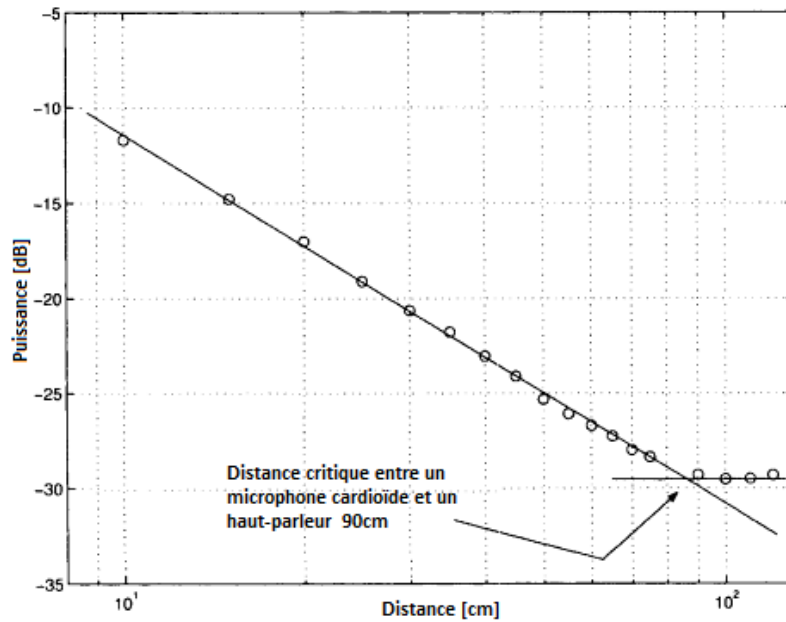


Figure 31 - distance critique mesurée dans un habitacle, traduit de [39]

Cependant, selon [40], il est peu pertinent de parler de distance critique dans un habitacle automobile. Toole quant à lui estime qu'à cause de l'absence de réflexions tardives du fait de la présence d'obstacles, la distance critique a peu de sens dans un petit local absorbant rempli d'objets, et que les auditeurs sont plutôt placés dans un champ particulier, intermédiaire entre champ direct et diffus [15]. On ne s'attarde donc pas sur cet aspect.

Schroeder propose un modèle de propagation du son dans une pièce en découpant deux zones fréquentielles, séparées par une zone de transition caractérisée par la fréquence

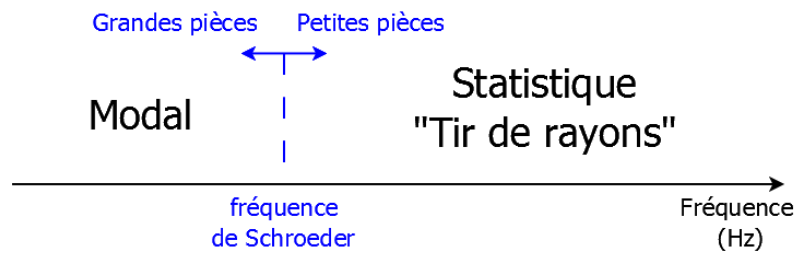
$$f_{Schroeder} = 2000 \sqrt{\frac{TR_{60}}{V}}$$

où $TR_{60} = TR_{60}$ Sabine en s, V le volume du local en mètres cubes

En dessous de cette fréquence¹⁵, la propagation peut être vue comme principalement modale. Au-delà de la fréquence de transition, la densité des modes propres est telle que la

¹⁵ Ce n'est pas une limite précise, de plus le calcul de la fréquence de Schroeder est le résultat d'expérimentations, et sa constante multiplicatrice a été réévaluée dans le temps.

propagation pourrait être considérée de façon statistique par l'acoustique géométrique (« tir de rayons »), sans oublier de considérer les effets de diffraction¹⁶.



Morris-Kirby et Bennetts estiment la fréquence de Schroeder de leur habitacle à 400 Hz, Paik et al à 300 Hz [35]. Les dimensions intérieures du véhicule de notre ordre de grandeur, qui est sans doute trop grand, donnent 250 Hz. Par comparaison les locaux d'écoute typiques ont plutôt une fréquence de Schroeder de l'ordre de 100 à 200 Hz. La propagation doit donc être étudiée de façon modale jusqu'à une fréquence plus élevée que dans une pièce d'écoute typique. Les modes propres créent de fortes différences de rendu sonore entre les différentes positions d'écoute, et cela complique l'atteinte du réglage optimal comme on ne peut plus changer la position des haut-parleurs [15]. On peut cependant se demander si les formules de Sabine et Schroeder sont utilisables dans un local aussi particulier.

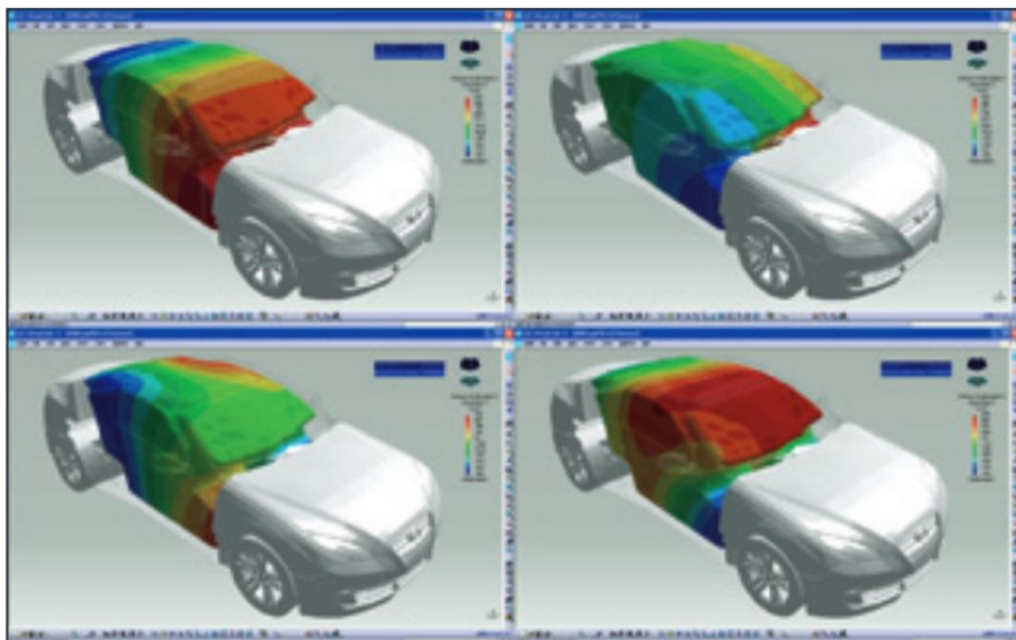


Figure 32 – Simulation des premiers modes longitudinaux et latéraux de pression sonore dans une voiture, aux fréquences 64, 73, 108 et 115 Hz [35]

¹⁶ Des précisions peuvent être trouvées dans Manuel d'acoustique fondamentale, Bruneau p456.

Dès 1991, Grimari proposait pour améliorer la reproduction stéréophonique en véhicule d'adjoindre un haut-parleur central, alimenté par la somme des canaux gauche et droite issus des sources.[41]. On trouve un tel haut-parleur sur la plupart des systèmes haut de gamme, cependant les systèmes d'entrée et de milieu de gamme n'ont que de 2 à 8 haut-parleurs latéraux. La solution est alors d'employer un traitement du signal audio.

2.3. Traitement du signal numérique embarqué

Il est aussi possible de s'accommoder des défauts du système de reproduction en appliquant un traitement au signal envoyé aux haut-parleurs. Dans cette section, on présente succinctement la boîte à outils de filtrage usuelle pour les réglages de système audio embarqué.

2.3.1. Généralités et définitions

La fonction de transfert entre la source audio et l'auditeur a un impact sur les perceptions de la réponse amplitude (fréquence) et de la phase binaurale. L'idée générale est donc de changer la réponse du système de reproduction sonore pour que sa combinaison avec la réponse de l'habitacle limite la dégradation de la fidélité aux points d'écoute. Le réglage est l'opération qui consiste à déterminer les paramètres de filtrages optimaux, en fonction d'un cahier des charges de prestation et des capacités du système de traitement du signal.

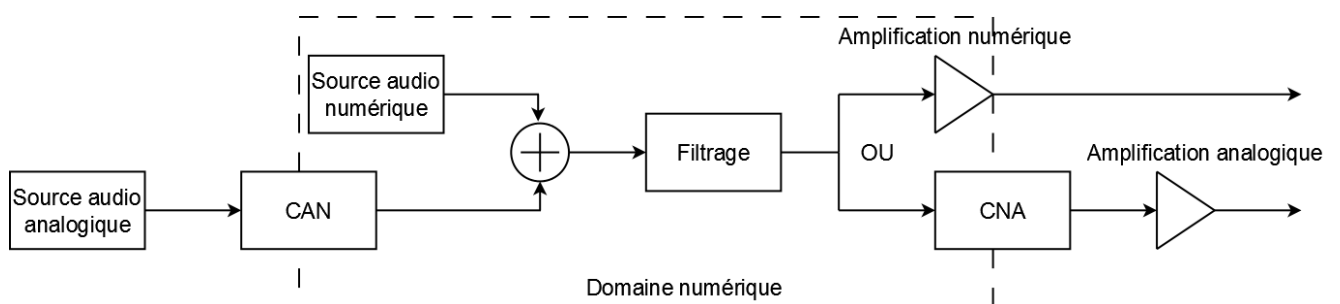


Figure 33 - Position du filtrage dans une chaîne audio typique

Le filtrage peut se faire dans le domaine fréquentiel ou temporel. Chaque canal peut être traité indépendamment. Les signaux issus du DSP et amplifiés sont habituellement routés au niveau de la liaison entre l'amplification et les haut-parleurs comme dans la figure suivante.

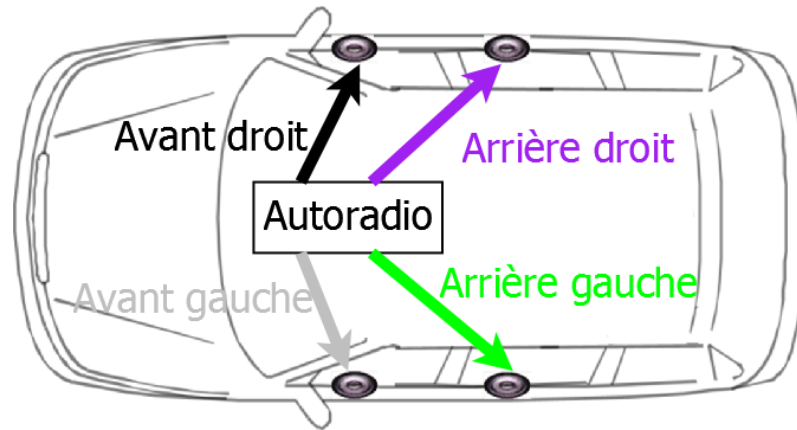


Figure 34 - Traitement audio sur 4 canaux

Les filtres numériques peuvent être de deux types, à réponse impulsionnelle finie (RIF) ou à réponse impulsionnelle infinie (RII). Un atout majeur du filtrage RII est qu'il nécessite significativement moins de coefficients et de lignes à retards qu'un filtrage RIF. Cela explique sans doute pourquoi le filtrage RII est majoritairement présent dans les configurations audio d'entrée et milieu de gamme dans les automobiles. La brique de base du filtrage RII est la cellule biquadratique. L'Interface Homme-Machine de réglage calcule les coefficients à envoyer aux cellules biquadratiques du DSP, et s'assure de leur stabilité. En particulier si ses coefficients sont choisis convenablement, la cellule biquadratique peut être mise à profit pour corriger la phase entre les canaux de sortie.

2.3.2. Alignement temporel

Voici un rapport de distance typique pour un auditeur placé à la place du conducteur :

Tableau vi - Distances typiques entre les haut-parleurs et le conducteur

Haut parleur	Distance (cm)	Différence de temps de parcours par rapport au haut-parleur le plus distant (ms)
Avant gauche	120	-0,88
Avant droit	150	0
Arrière gauche	90	-1,76
Arrière droit	130	-0,59

Le haut-parleur le plus proche est à l'arrière à gauche, le plus distant est à l'avant à droite. Dans un habitacle de voiture sans traitement audio adapté, la composante corrélée entre les canaux du signal à la source audio est souvent perçue comme une source virtuelle peu précise, dissociée voire hors de la base stéréophonique traditionnelle. Typiquement pour le conducteur elle est latéralisée à gauche et un vide est perçu au milieu de la planche de bord. Par rapport à un système stéréophonique de référence, à l'écoute et à la mesure, la différence de temps de parcours entre les haut-parleurs et les auditeurs a des conséquences néfastes :

- distorsion de la scène sonore « ΔT » à cause de l'altération de la phase binaurale en dessous de 2 kHz,
- filtrage en peigne complexe de la composante corrélée entre les haut-parleurs gauches et droits,
- filtrage en peigne entre zone avant et zone arrière. Comme on l'a vu au chapitre 2.2.3, dans un système audio embarqué typique à 4 haut-parleurs, avant gauche est corrélé avec arrière gauche, avant droit est corrélé avec arrière droit, tout le temps, sur tout le spectre.

Les filtrages en peigne réduisent la fidélité de timbre, Smithers et al estiment que pour les distances mises en jeu cela est particulièrement audible entre 1 kHz et 5 kHz [42], où les ILD sont l'indice privilégié de latéralisation.

Il semble qu'un des plus gros problèmes de cette configuration d'écoute est donc que pour chaque auditeur les haut-parleurs sont situés à des distances différentes. Les différences de

temps de parcours entre chaque haut-parleur et chaque oreille, et la recombinaison complexe des signaux plus ou moins corrélés, résultent en une distorsion de la scène sonore.

Il est possible d'optimiser la restitution en appliquant des retards aux signaux destinés aux différents haut-parleurs, pour compenser les différences de temps de parcours. Cela peut s'illustrer par rapport au centre de la tête de l'auditeur considéré, comme ci-dessous :

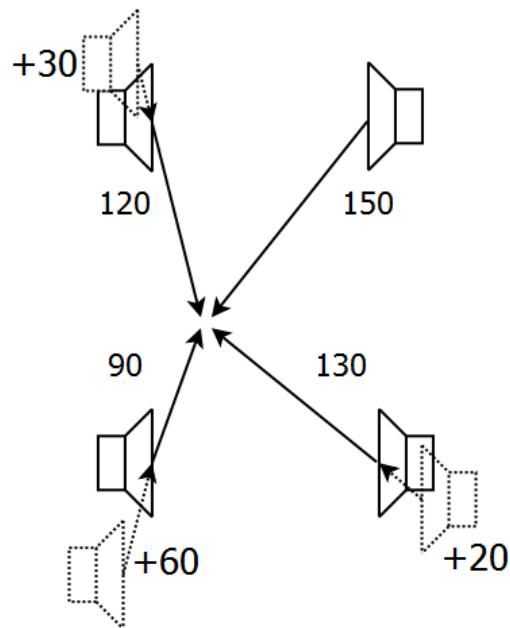


Figure 35 - ajustement temporel mono point

Cette technique permet d'ébaucher l'alignement temporel, mais est simpliste en ce qui concerne la perception spatiale, qui est en partie binaurale.

Il faut donc chercher à égaliser les temps de parcours entre les haut-parleurs et les deux oreilles, comme dans la figure suivante.

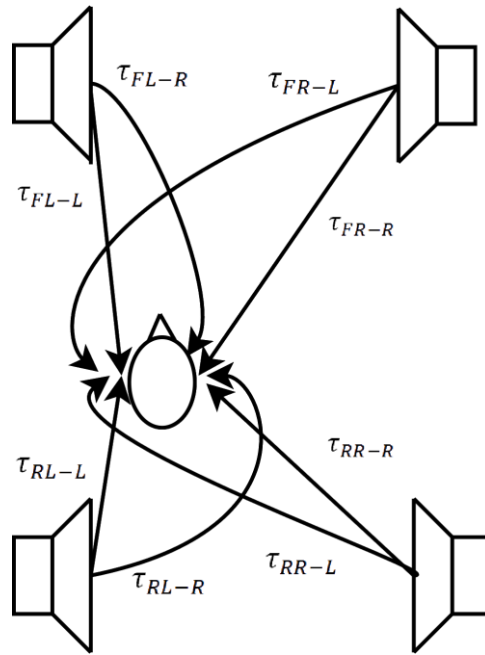


Figure 36 - distances entre les haut-parleurs et les deux oreilles de l'auditeur

On cherche alors à minimiser, à une ou plusieurs positions, les différences de temps de parcours illustrées ci-dessus :

$$\tau_{FR-R} - \tau_{FL-L}$$

$$\tau_{FL-R} - \tau_{FR-L}$$

$$\tau_{RR-R} - \tau_{RL-L}$$

$$\tau_{RL-R} - \tau_{RR-L}$$

L'expérience habituelle lors des réglages de systèmes embarqués est que les effets perçus d'un bon alignement temporel sont les améliorations:

- de la régularité du timbre dans les basses fréquences par réduction du filtrage en peigne (voir la mesure ci-dessous);
- de l'impact perçu du registre grave, par opposition à un son trainant (on le perçoit à l'écoute et on peut le voir sur les réponses impulsionnelles);

- de la scène sonore qui devient plus « cohérente »¹⁷ (la source fantôme correspondant à la composante corrélée est perçue plus compacte et plus frontale) et plus régulière (les sources fantômes sont réparties régulièrement)

Les deux premières améliorations sont dues à une meilleure recombinaison des sons émis par les haut-parleurs, la troisième est due à une optimisation de la corrélation binaurale. L'idéal est de trouver un réglage de retards qui donne un bon compromis entre une amélioration du timbre et une amélioration de la scène sonore, parfois l'un des aspects est favorisé. Dans l'exemple suivant, le seul paramètre qui change entre la courbe verte un peu chahutée dans les graves et la courbe violette plus lisse est l'alignement temporel des haut-parleurs.

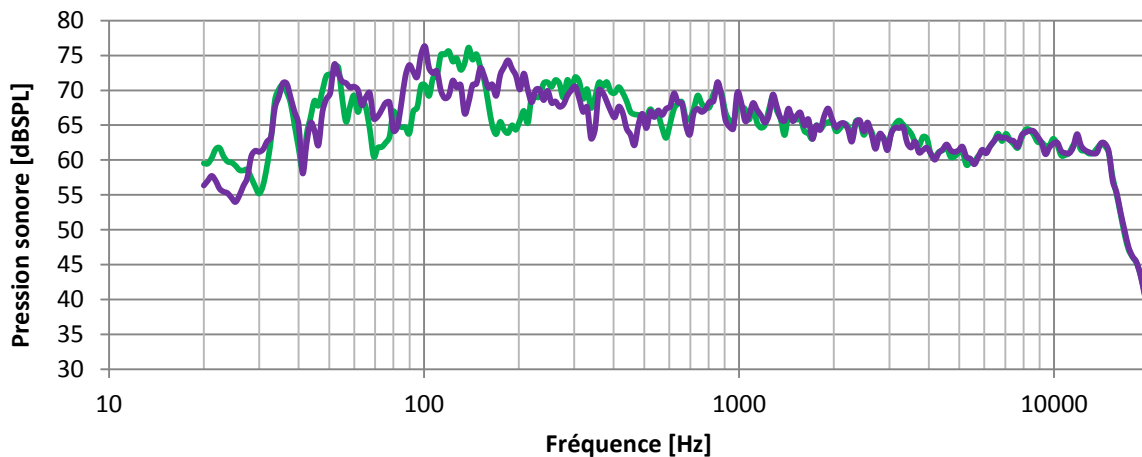


Figure 37 - Exemple d'optimisation de la réponse amplitude (fréquence) par alignement temporel des haut-parleurs, vert = avant alignement temporel, violet = après alignement temporel

L'optimisation du retard est un compromis qui favorise généralement les places avant, sauf en cas de défaut inacceptable aux places à l'arrière. Certaines valeurs de retards améliorent le grave, d'autres valeurs améliorent la ponctualité et la position perçues de la composante corrélée, d'autres valeurs améliorent la scène sonore globale, mais par expérience les valeurs optimales perceptivement ne sont pas obtenues simplement en rendant égales toutes les distances apparentes.

Un attendu principal de la méthode de mesure de scène sonore est de trouver plus rapidement les valeurs de retards optimaux.

¹⁷ Ne pas confondre ce qualificatif utilisé par les ingénieurs audio en automobile avec la définition de la cohérence utilisée en traitement du signal.

2.3.3. Représentation et égalisation phase(fréquence)

On cherche maintenant à analyser dans le domaine fréquentiel l'effet du décalage temporel sur la phase inter-aurale. Commençons par imaginer, en conditions anéchoïques, le signal reçu par un microphone (courbes rouges dans ce qui suit) collé à un haut-parleur qui émet une « impulsion », et les signaux issus d'autres microphones placés de plus en plus loin du haut-parleur (courbe verte et bleue).



Figure 38 – Expérience de pensée pour le calcul de phase – plusieurs microphones

Une simulation de l'effet du décalage temporel sur la phase relative (code en annexe) donne successivement :

- la même réponse impulsionnelle décalée de 0 ms (en rouge), 1 ms soit 34 cm à la vitesse du son (en vert) et 2 ms soit 68 cm (en bleu) ;
- la phase relative dépliée en fonction de la fréquence de chaque point de mesure par rapport au point de mesure le plus proche
- la phase relative repliée en fonction de la fréquence de chaque point de mesure par rapport au point de mesure le plus proche

Avec une échelle fréquentielle linéaire, la différence des temps de propagation résulte en une pente de la phase relative dépliée, d'autant plus importante que la différence de temps de parcours est importante. Avec une échelle fréquentielle logarithmique, plus la différence de temps de parcours est importante et plus la courbure de la phase relative est accentuée.

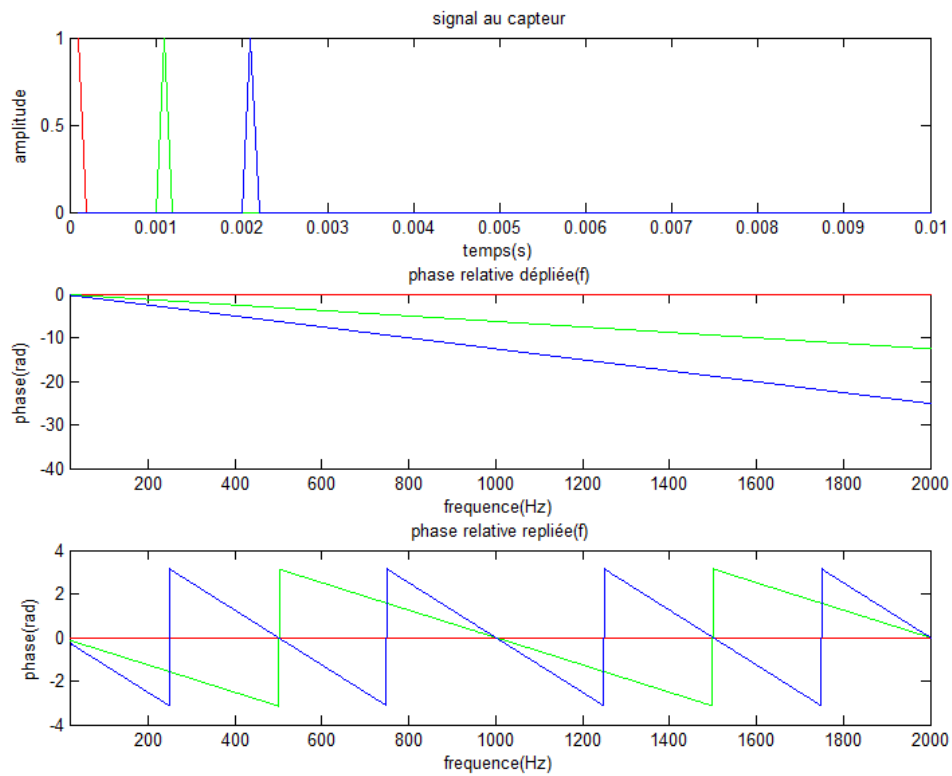


Figure 39- résultats de la simulation - échelle fréquentielle linéaire

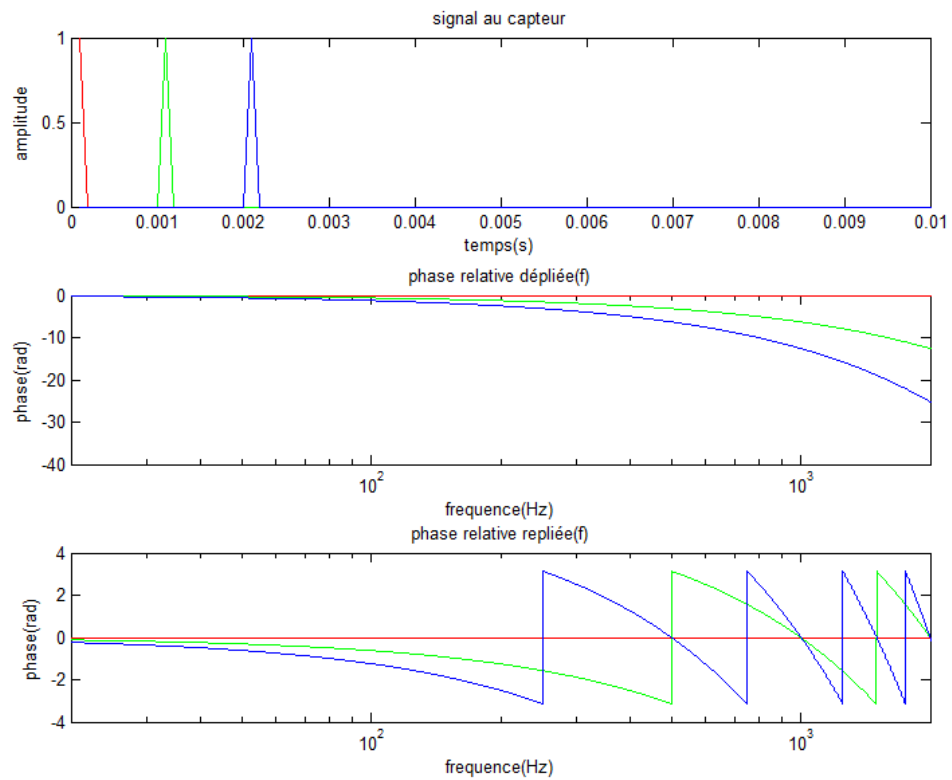


Figure 40 - résultats de la simulation - échelle fréquentielle logarithmique

On peut expliquer cela en considérant la transformation de Fourier :

$$x(t) \stackrel{TF}{\rightleftharpoons} X(f)$$

Si le signal est retardé, le facteur $e^{-j2\pi f t_0}$ apparaît dans le domaine fréquentiel et dans le diagramme de phase relative en fonction de la fréquence :

$$x(t - t_0) \stackrel{TF}{\rightleftharpoons} X(f)e^{-j2\pi f t_0}$$

On propose de relier la perception spatiale à cette analyse des courbes de phase binaurale en fonction de la fréquence :

- la pente locale donnerait une information de position angulaire perçue sur la zone de fréquences considérée,
- si elles sont assez larges, les zones fréquentielles où la phase binaurale est en opposition, ou suffisamment proche de l'opposition, correspondraient à des perceptions spatiales floues « extra larges ».

Carpentier[43] puis Crockett et al [42] ont décrit l'effet d'une écoute hors axe médian entre les haut-parleurs sur la phase binaurale et les limites d'un ajustement par retard simple, et bien que leurs explications sont imprécises et ne semblent pas tenir compte de la diaphonie inter-aurale, les filtrages de la phase entre les canaux du système de sonorisation qu'ils proposent donnent en réalité de bons résultats. Pour tenter d'affiner leur analyse, on se place dans la situation de la figure 36, mais avec seulement deux haut-parleurs devant l'auditeur, on simule le signal arrivant à chaque oreille après émission d'une impulsion par deux haut-parleurs frontaux, en milieu anéchoïque, en prenant en compte la diaphonie inter-aurale et en considérant que le son s'atténue de façon géométrique (code Matlab « ecoutehorsaxe.m » en annexe A). On obtient successivement :

- la même réponse impulsionnelle envoyée par les haut-parleurs avant, mesurée à l'oreille gauche (en vert) et droite (en rouge) pour un auditeur à la place du conducteur,
- la phase binaurale en fonction de la fréquence à la place du conducteur.

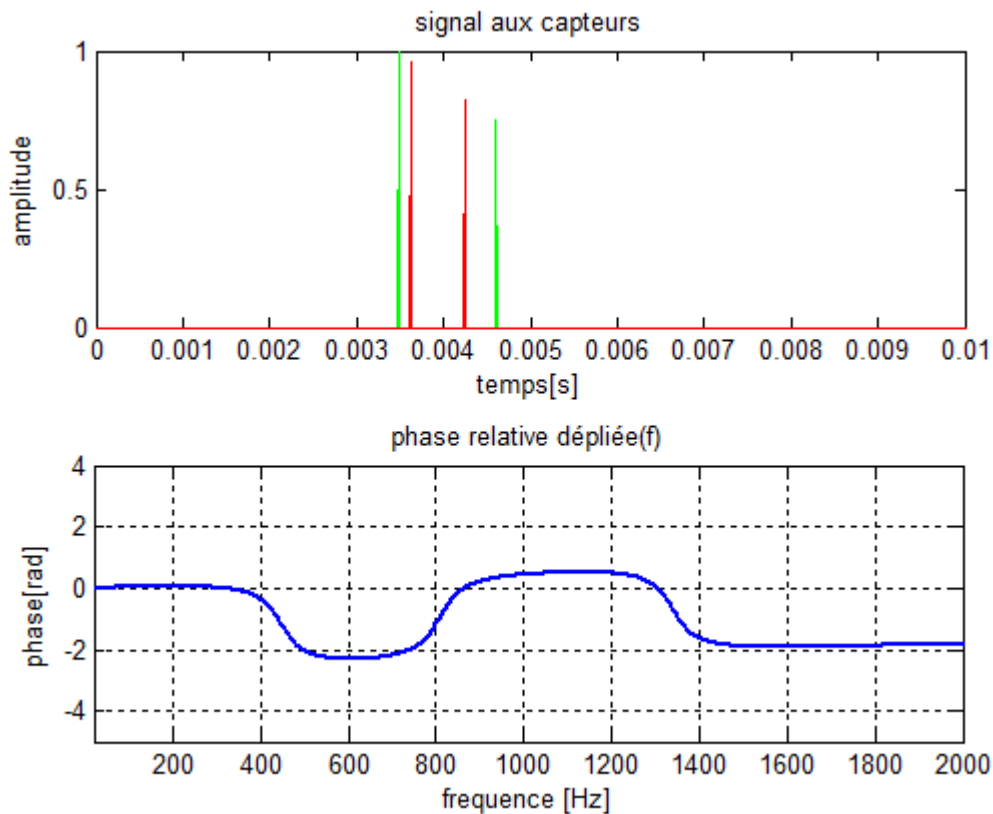


Figure 41 - Résultat de la simulation d'écoute hors *sweet spot* avec diaphonie

On constate que la phase binaurale présente deux irrégularités de phase entre 400 et 800 Hz et entre 1300 et 2000 Hz.

Par symétrie on suppose qu'en refaisant la même simulation pour le passager à l'avant, les courbes vertes et rouge sont permutées dans le premier graphique de la figure 41, et que le signe des phases est inversé.

On attend d'un filtrage adapté qu'il lisse la phase binaurale en fonction de la fréquence et qu'il évite qu'elle atteigne l'opposition sur des bandes de fréquences larges. Le réglage de la phase entre les haut-parleurs se fait empiriquement en mesurant et en cherchant à limiter les filtrages en peigne dus à l'alignement temporel imparfait, et surtout en écoutant de séquences dont les deux canaux sont très corrélés voire totalement corrélés. Si on voulait pousser cette simulation, il faudrait dans un premier temps calculer ce qui se passe si l'on applique une égalisation de phase en fonction de la fréquence à l'impulsion envoyée par le haut-parleur de droite, et déterminer le filtrage qui régularise la phase interaurale.

La réalité est beaucoup plus complexe car l'habitacle est fermé, réfléchissant et rempli d'objets de dimensions proche des longueurs d'ondes mises en jeu, ce qui va modifier encore les ITD et ILD. Les mesures réelles donnent des courbes nettement plus chahutées que les simulations en anéchoïque, et il serait intéressant de comparer des mesures dans une voiture réelle et la simulation dans son modèle acoustique à trois dimensions.

Une technique qui semble bien adaptée pour égaliser la phase est décrite dans le brevet "Audio System Phase Equalization" de M. Christoph et L. Scholz [44]. Elle consiste à mesurer la phase binaurale à différents sièges de l'habitacle pour différentes valeurs de retard de chaque haut-parleur, et à trouver un filtrage qui minimise cette phase pour un maximum d'auditeurs.

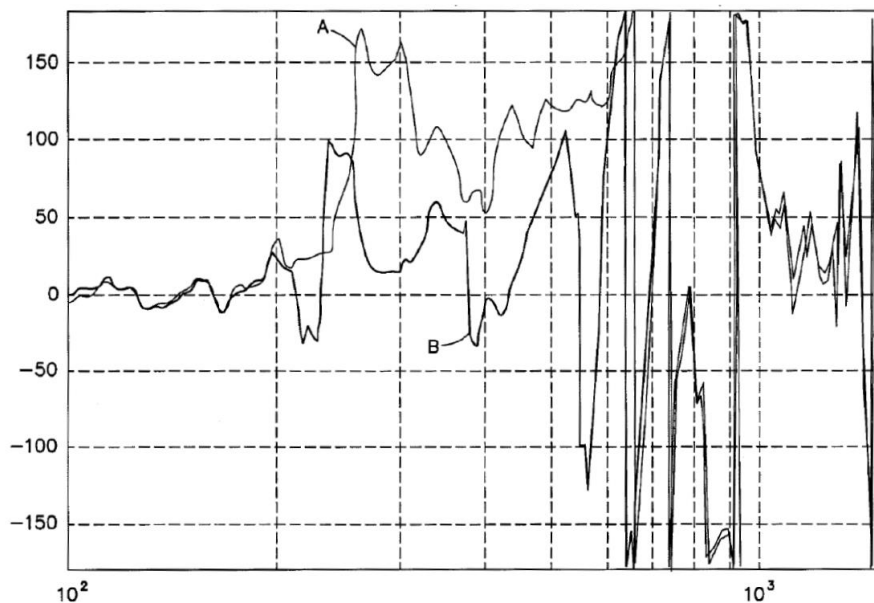


Figure 42 - Exemple d'amélioration de la phase binaurale mesurée, A sans filtrage, B avec filtrage [44]

On voit dans la figure précédente une pente globale montante et des accidents de phase entre 250 Hz et 400 Hz puis entre 600 Hz et 800 Hz sans traitement, ce qui ressemble à la simulation précédente.

2.3.4. Filtrage par inversion

Il existe des méthodes de réglage automatique de filtres RII composés de cellules biquadratiques, mais à notre connaissance elles ne prennent en compte que la réponse amplitude (fréquence) globale du système de reproduction audio, et ne sont pas à même de corriger les effets des phases relatives entre les haut-parleurs sur la réponse amplitude (fréquence) globale et sur la scène sonore perçue. Les réglages de retard et de phase, ainsi que l'évaluation de la qualité des réglages se font donc encore à ce jour par mesures et par évaluations subjectives (écoutes).

Des techniques d'égalisation d'amplitude et de phase « par inversion » avec plusieurs points de mesure ont été proposées [45] [46] [47]. L'intérêt est de trouver « automatiquement » le filtrage optimal qui réalise un compromis d'optimisation de fidélité sur plusieurs places d'écoutes. Considérons un haut-parleur et un microphone, appelons $h(n)$ la réponse impulsionnelle des éléments électroniques, électro-acoustiques et acoustiques entre la source de signal audio et le microphone. Il paraît physiquement réaliste d'imposer que $h(n)$ soit finie : les réponses du système audio et de l'habitacle peuvent être modélisées par un filtre à Réponse Impulsionnelle Finie. Si $h(n)$ est inversible, alors le filtre inverse de réponse impulsionnelle $h_{inv}(n)$ est tel que :

$$h(n) \otimes h_{inv}(n) = \delta(n)$$

Où $\delta(n)$ est la distribution de Dirac centrée en $n = 0$, et \otimes est l'opérateur de convolution.

Si on passe cette relation en transformée en z , elle devient :

$$H(z) \cdot H_{inv}(z) = 1$$

où \cdot est la multiplication.

Au premier abord, la technique d'inversion peut être décrite de la façon suivante :

- détermination de la réponse impulsionnelle $h(n)$ du système à inverser,
- calcul de la fonction de transfert $H(z) = \sum_{-\infty}^{\infty} h(n)z^{-n}$,
- inversion de la fonction de transfert $H_{inv}(z) = \frac{1}{H(z)}$ **sous conditions**,
- implémentation du filtre réalisant la fonction de transfert inverse $H_{inv}(z)$.

L'inversion des fonctions de transfert est sous conditions car il convient de s'assurer que le filtrage inverse obtenu est :

- causal, c'est-à-dire que sa réponse impulsionnelle est nulle avant $n=0$, $h_{inv}(n) = 0 \forall n < 0$
- stable, c'est à dire convergent vers 0. Pour cela les pôles de la fonction de transfert en z , $H_{inv}(z)$ doivent être l'intérieur du cercle unité exclu. Cela signifie que son inverse, la fonction de transfert du trajet dans l'habitacle et du système de sonorisation $H(z)$, a ses zéros à l'intérieur du cercle unité exclu, et comme elle est également stable on dit alors qu'elle est à phase minimale.

Dans la réalité, la réponse impulsionnelle du trajet dans l'habitacle et du système de sonorisation $h(n)$ n'est pas à phase minimale mais est dite « à phase mixte ». Elle peut être décomposée en un produit d'une partie à phase minimale, qui est inversible et donc égalisable en amplitude et en phase, et d'un résidu non inversible. Une technique de filtrage par inversion utilisée en dé-réverbération consiste par exemple à ne prendre en compte qu'une partie de la réponse impulsionnelle, par exemple le trajet direct et les premières réflexions.

Le raisonnement peut être étendu aux trajets entre plusieurs haut-parleurs et plusieurs microphones [48]. Cette technique est fréquemment utilisée en optimisation de captation microphonique pour conversations téléphoniques mains libres, mais encore peu utilisée dans le cadre de réglages de systèmes audio « de confort », c'est-à-dire destinés à écouter la radio ou des enregistrements musicaux, ce qui est le cadre de cette étude. Il serait intéressant de faire comparer à des panels d'écoute :

- les réglages « manuels » obtenus par l'écoute seule,
- les réglages « manuels » obtenus par les écoutes et les procédés de mesure de scène sonore jugés pertinent à l'issue de cette étude,
- le filtrage obtenu par inversion des fonctions de transfert entre les haut-parleurs et les oreilles des auditeurs,
- une approximation de ces filtrages inverses avec les filtres RII disponibles dans un autoradio.

2.4. Conclusion partielle

Le placement des haut-parleurs dans l'habitacle automobile doit s'accommoder de nombreuses contraintes, et ne correspond pas aux recommandations de la diffusion stéréophonique. Lors des réglages audio chez Arkamys, les ingénieurs disposent de peu d'information sur les caractéristiques des haut-parleurs et des constituants de l'habitacle, ce qui ne permet pas de simuler le rayonnement, ni de calculer a priori le filtrage adapté, ce qui renforce la nécessité d'une mesure de scène sonore *in situ*.

L'habitacle a un impact trop important pour ne considérer que le chemin direct du son dans les modèles d'audition. Les mesures acoustiques que l'on a faites ne correspondent pas toujours aux recommandations, mais concordent relativement bien avec la littérature. Elles permettent de déterminer les caractéristiques acoustiques suivantes pour un habitacle automobile typique :

- TR_{60} entre 50 ms et 110 ms, contre environ 500 ms recommandés dans un auditorium,
- $f_{\text{Schroeder}}$ proche de 300 ou 400 Hz, contre 100 Hz dans un auditorium ou 200 Hz dans une pièce d'habitation moyenne,
- clarté C80 entre +2 dB et +17 dB, donc un rendu clair voire très clair de la musique par rapport aux recommandations.

On peut sérieusement s'interroger sur la pertinence d'aspects habituels d'analyse de propagation du son dans une pièce (champ direct/réverbéré, distance critique...) quand on se trouve dans un habitacle automobile, et sur la pertinence des modèles d'audition précités dans ces conditions.

Pour compenser les défauts inhérents à l'écoute stéréophonique en voiture, différentes techniques de filtrage possibles ont été identifiées. Lors des réglages, une chaîne audio composée de filtres RII, de gains et de délais est habituellement utilisée. Une simulation a montré l'importance de la phase binaurale, qui n'est pas ou rarement mesurée lors des mises au point alors qu'elle permettrait d'optimiser le réglage du traitement audio, voire de l'automatiser. Des pistes ont été proposées pour analyser la phase binaurale en fonction de la fréquence et en déduire la perception spatiale associée. Un programme de réglage

automatique perfectionné mesurerait les fonctions de transfert entre chaque oreille à chaque position d'écoute et chaque haut-parleur, puis calculerait le filtrage à appliquer sur chaque zone (avant, arrière, centrale), selon les capacités du DSP et le chaînage audio retenu, ce qui on l'espère, permettrait d'obtenir un compromis optimal et rapide.

On va maintenant vérifier si la mesure de phase binaurale et les résultats des modèles d'audition peuvent assister le réglage du filtrage embarqué.

3. Mesures et scène sonore

3.1. Buts

Jusqu'ici le manque de représentation pratique ne permet pas à un opérateur unique de régler rapidement un filtrage RII pour optimiser la perception spatiale, il lui faut changer de place régulièrement et écouter pour évaluer les conséquences d'un réglage.

Afin de limiter la distorsion angulaire, les vides ou compactages de sources fantômes par rapport au système de référence, et donc pour améliorer la régularité de la scène sonore, les informations attendues d'une mesure de scène sonore sont la perception moyenne de:

- la position angulaire des sources virtuelles,
- la largeur des sources virtuelles,
- une éventuelle opposition de phase.

3.2. Résultats de mesure dans la littérature

3.2.1. Evaluation conjointe des ITD et ILD

Dans [49], Kleiner et Lindgren proposent pour caractériser la scène sonore d'analyser les ITD et ILD mesurées en utilisant des réseaux de courbes de Franssen. Selon eux cette mesure et comparaison avec un système proche des recommandations d'écoute stéréophonique fonctionnerait bien. Cependant, quand plusieurs haut-parleurs reproduisent le même canal de la stéréophonie, ils se heurtent aux ambiguïtés de localisation sur les cônes de confusion. On ne sait pas si ces conclusions se basent sur des expériences.

Les brevets "System for determining the position of a sound source and method therefore" [50] et "Objective quantification of auditory source width of a loudspeakers-room system"[51] déposés par l'entreprise Harman-Becker comprennent un modèle de latéralisation qui donne des diagrammes d'activité binaurale en fonction du temps à partir de mesures dans l'habitacle, en analysant à la fois ITD et ILD, et en recourant à une base de HRTF.

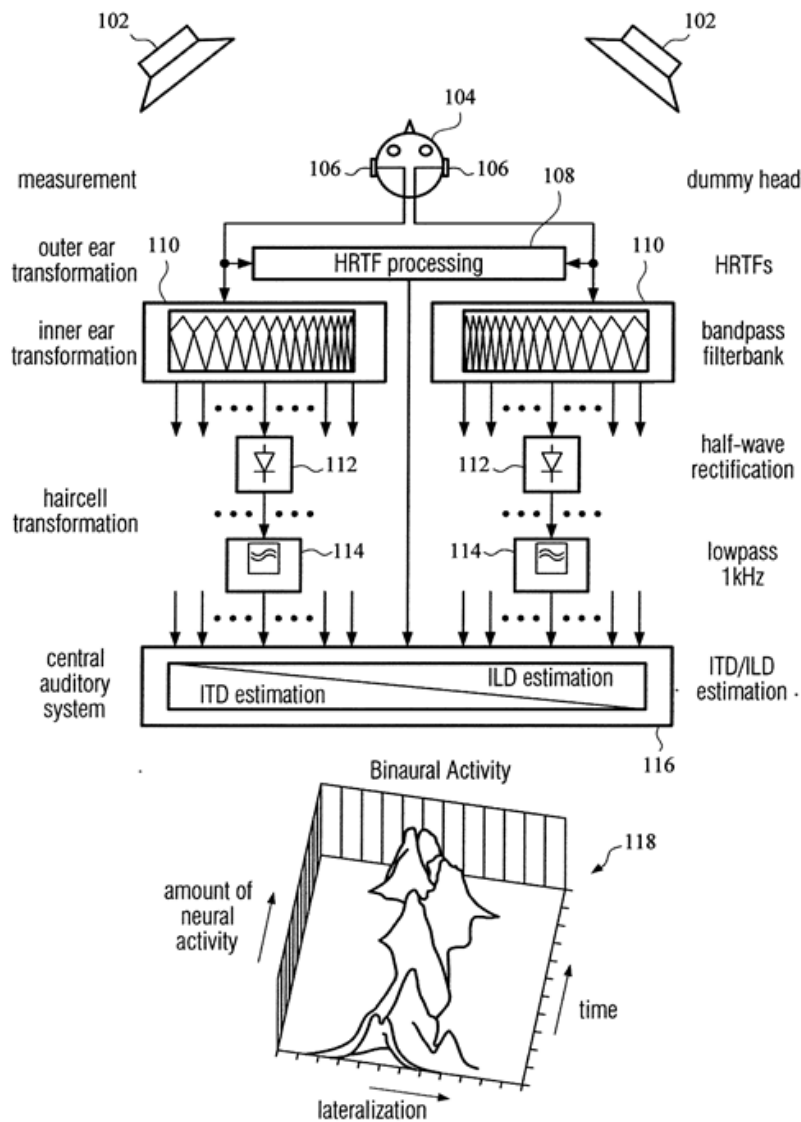


Figure 43 - Modèle embarqué dans les brevets Harman-Becker

3.2.2. Mesure d'IACC

3.2.2.1. IACC globale

Pour déterminer la latéralisation de sources réelles, Marchand et al [5] utilisent en plus des ITD et ILD un calcul de corrélation binaurale qui leur permet de lever les ambiguïtés. Ils donnent des résultats expérimentaux plutôt positifs de cette méthode, mais ils ne précisent pas les conditions acoustiques de ces mesures. On constate un tassement des angles mesurés quand l'angle d'incidence se rapproche de 60°.

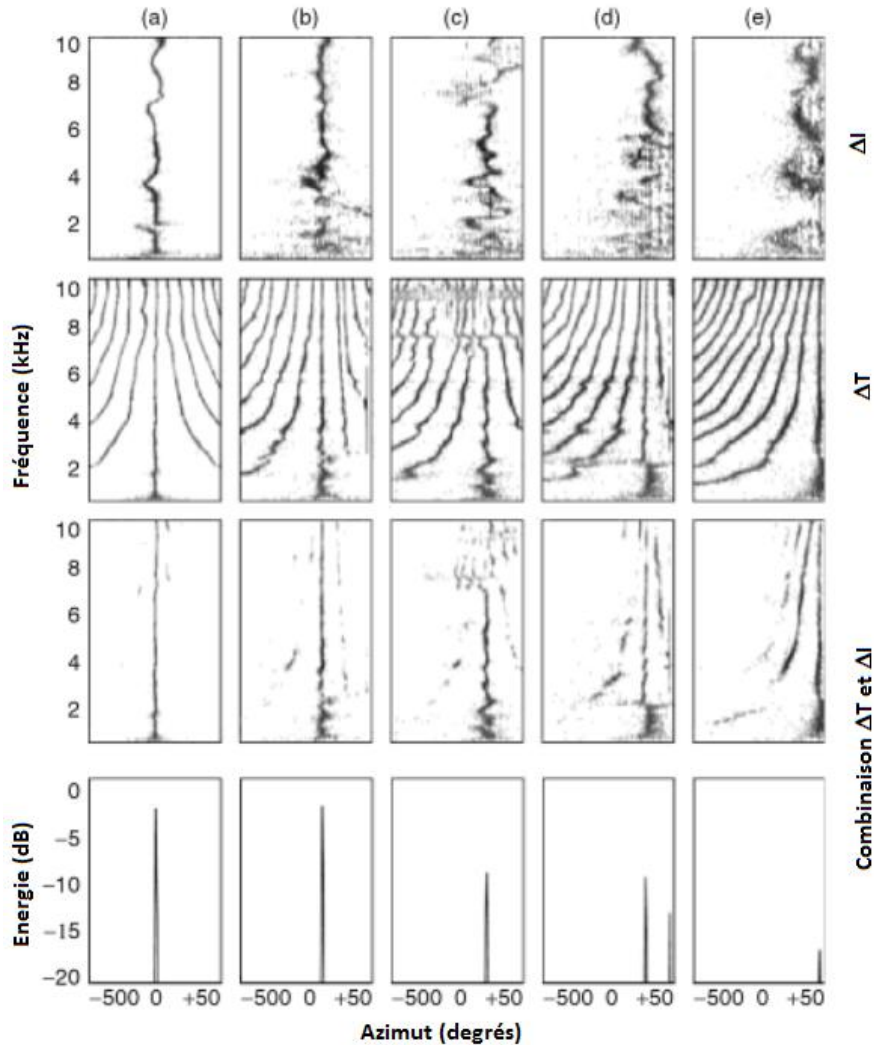


Figure 44 – Estimation d’azimut pour une source latéralisée à 0° ; 15° ; 30° ; 45° et 65°. Première ligne, ILD uniquement, en fonction de la fréquence. Deuxième ligne ITD uniquement, en fonction de la fréquence. Troisième ligne : combinaison ILF et ILD en fonction de la fréquence. Dernière ligne : somme des combinaisons sur toutes les fréquences. Traduit de [5]

Dans [52], Pulkki concluait déjà que sans réverbération, un tel modèle donnait des résultats qualitativement proches de panels d’écoute. En 2003, Farina et Ayalon [53] enregistrent des réponses impulsionnelles de plusieurs locaux avec un microphone Soundfield et avec une tête artificielle, puis calculent respectivement la fraction latérale (LF) et l’IACC¹⁸ à partir de ces mesures. On retiendra de cette étude que les calculs de LF et IACC sur un seul axe peuvent conduire à des erreurs d’appréciation sur l’impression spatiale d’une salle, et qu’il est plus pertinent et discriminant d’enregistrer les réponses impulsionnelles sur tout le plan azimuthal, en faisant tourner le système de mesure, et de représenter LF et IACC en diagramme polaire, comme dans la figure suivante.

¹⁸ Dont ils analysent le maximum, le délai et la largeur de ce pic principal.

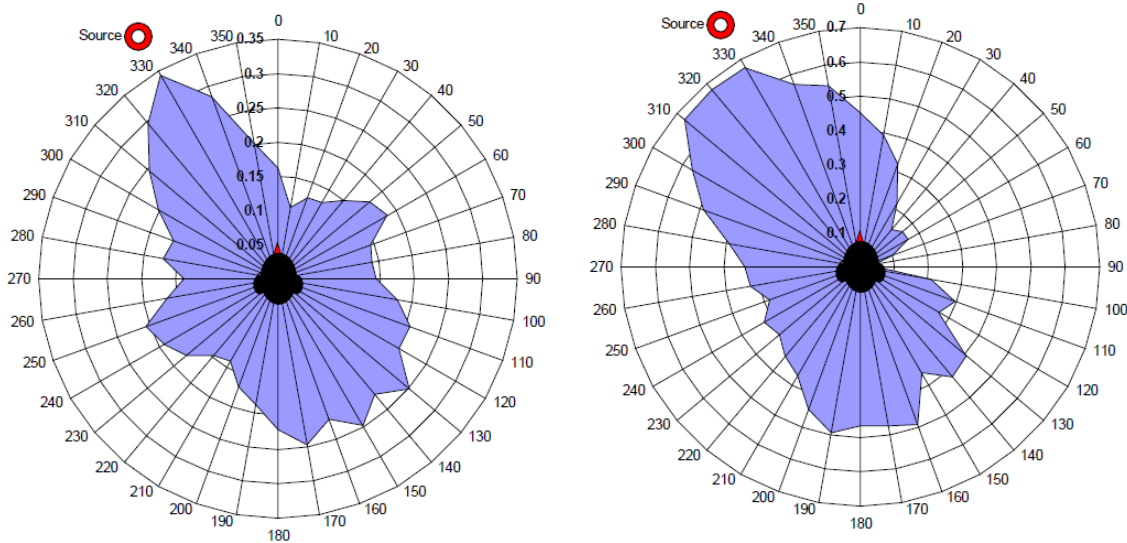


Figure 45 - Mesure d'IACC et 1-LF sur le plan azimutal, dans un même local, extrait de [53]

En 2004, Silverman et Sachar proposent une méthode de visualisation en fonction du temps de la différence de temps de parcours qu'ils appellent le « delayogram » [54], une telle visualisation pourrait être mise à profit pour les réglages. La figure suivante en montre deux exemples.

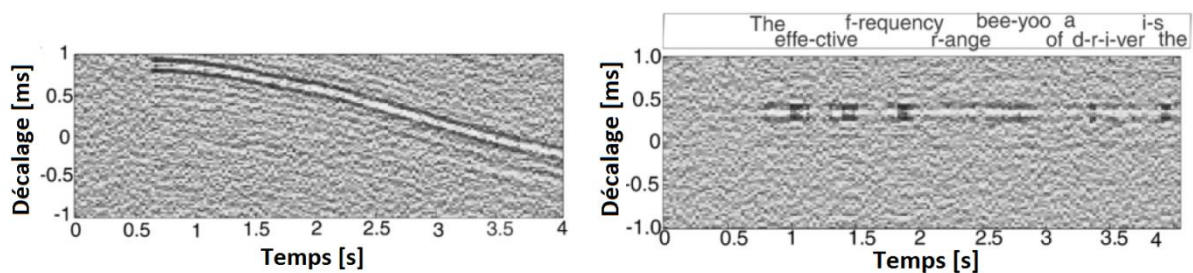


Figure 46 – Delayogram pour un orateur face à deux microphones à 1 m 50, faisant le son « shhhh » en mouvement (gauche) puis fixe et lisant une phrase (droite)

Ces auteurs admettent que leur technique souffre d'un fort taux d'erreur quand il y a des réflexions¹⁹ et un fort bruit de fond [55]. Il faut ajouter des modèles statistiques conditionnels [56], ou filtrer les signaux avant le calcul de l'inter-corrélation, comme dans [57] et [58], pour améliorer la mesure. Paik et al ont simulé le champ acoustique dans un habitacle automobile

¹⁹ Nous avons pu le vérifier, en cherchant à étalonner ce calcul en le mettant face à des distances haut-parleur-microphone connues : ce n'est pas toujours le premier pic ou le plus important qu'il faudrait détecter.

par la modélisation par des éléments finis, et utilisé comme indicateur de la qualité de restitution l'IACC, mesurée en lisant un signal corrélé.

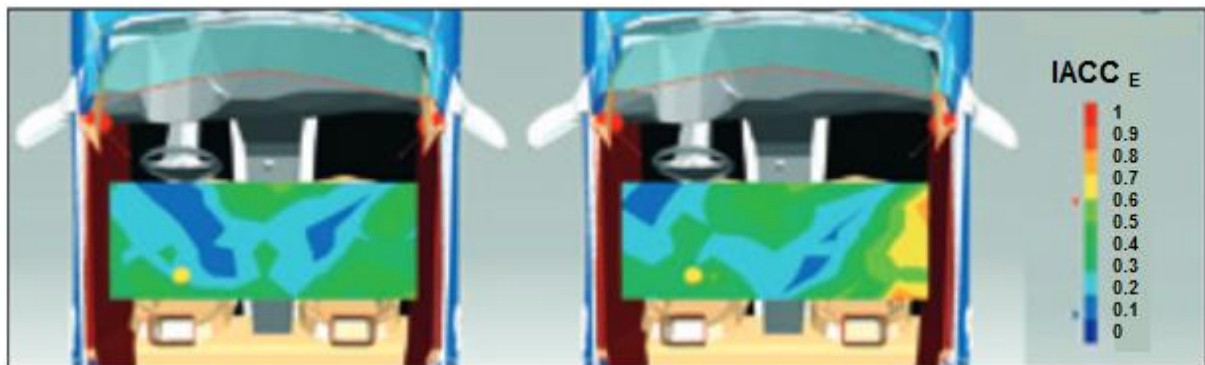


Figure 47 - Cartographie d'IACC précoce, à gauche pour un réglage symétrique, à droite pour un délai de 2 ms sur le tweeter avant gauche et un gain de 5dB sur le tweeter avant droit, adapté de [35]

On voit dans cette cartographie que dans ce véhicule, un réglage du filtrage jugé bon perceptivement permet de faire passer l'IACC de 0.2 à 0.5 pour le conducteur, sans dégradation notable pour le passager avant [35].

3.2.2.2. FIACC

La mesure d'IACC en fonction de la fréquence est mise à profit par Kihara dans son brevet de 1989 afin de déterminer les zones fréquentielles à égaliser en phase sur les canaux d'un système audio dans une automobile [59].

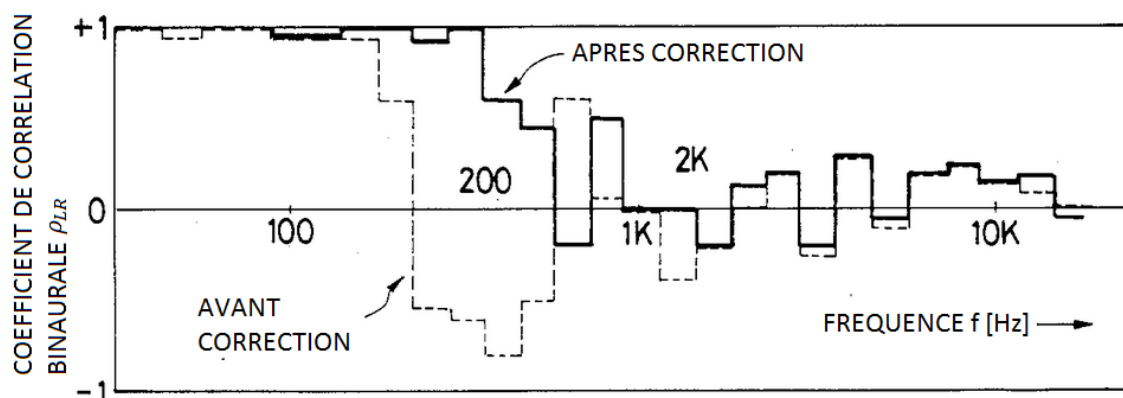


Figure 48 - mesure de FIACC avant et après égalisation, traduit de [59]

La méthodologie consiste à lire un signal test corrélé et rapprocher le plus possible l'IACC de 1 pour toutes les bandes critiques, afin d'optimiser la cohérence et le centrage. Plus tard Wikner et Fajerson [26] analysent cet indice lors d'un réglage. Le calcul de FIACC utilisé consiste à découper le signal en bandes de tiers d'octave, puis à calculer l'expression suivante pour chaque bande de fréquence n.

$$FIACC(n) = \frac{\text{moyenne}(\text{signal gauche filtré} * \text{signal droite filtré})}{\sqrt{\text{moyenne}(\text{signal gauche filtré}^2) * \text{moyenne}(\text{signal droite filtré}^2)}}$$

Ces auteurs comparent les mesures obtenues pour deux réglages du système de reproduction audio avec un signal corrélé, et mettent en rapport ces mesures avec une amélioration perçue de la fidélité spatiale, comme dans les figures suivantes données en exemple.

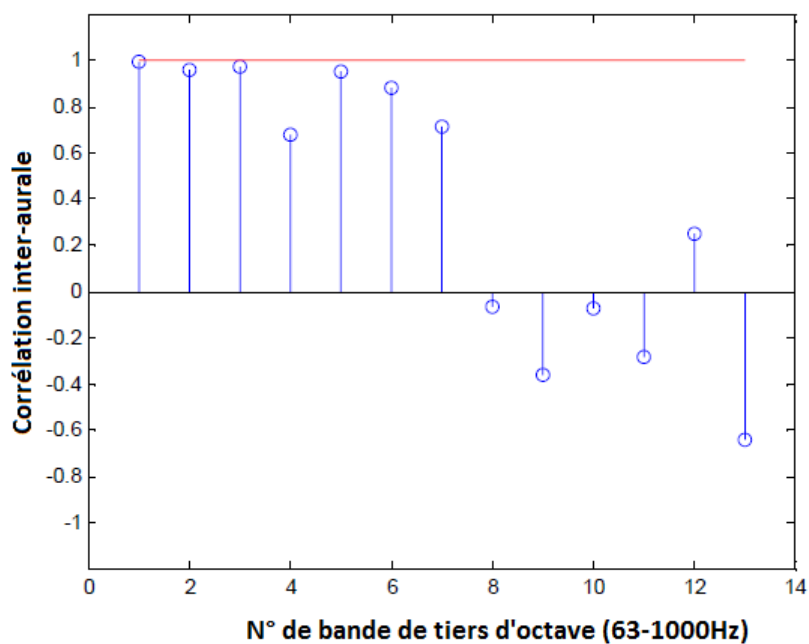


Figure 49 - Inter-corrélation mesurées pour un réglage A – bruit corrélé

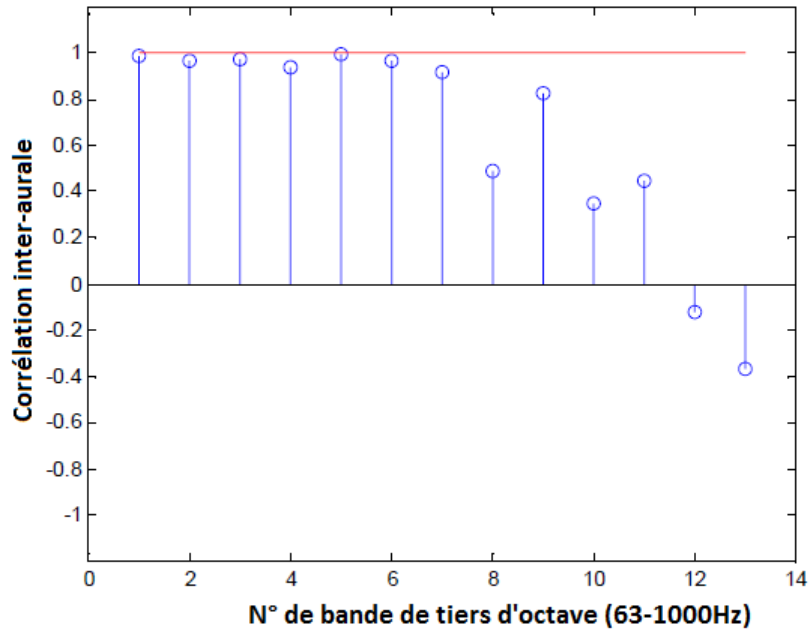


Figure 50 - Inter-corrélation mesurées pour un réglage B – bruit corrélé

Il ressort de leurs écoutes que le meilleur réglage est le B, avec lequel la mesure de FIACC est meilleure pour les bandes de fréquences 8 à 11 et 13, mais pire pour la bande 12. Ces auteurs proposent également de mesurer la FIACC pour un signal test dé-corrélé afin d'assurer l'impression d'espace et l'enveloppement, mais cherchent à ce que la FIACC mesurée se rapproche du sinus cardinal de la fréquence typique d'un champ diffus (voir chap. 1.6), ce qui ne paraît pas si évident.

La limite de cette technique pour notre usage est que la mesure permet de comparer des réglages, de les décomposer en plusieurs étapes et d'analyser l'apport de chaque étape sur la corrélation binaurale, mais ne permet pas encore d'obtenir directement la totalité des réglages d'alignement temporel.

Wittek présente des résultats de mesure de corrélation inter-aurale en fonction de la fréquence²⁰ dans [60] sous la forme intéressante des deux figures suivantes. Cette représentation se rapproche beaucoup de notre cahier des charges. L'azimut n'est pas donné mais on pourrait le calculer par inversion d'une des relations entre IPD et azimut vue au chapitre 1.1.

²⁰ Qu'on pourrait noter $IACC(\tau, f)$.

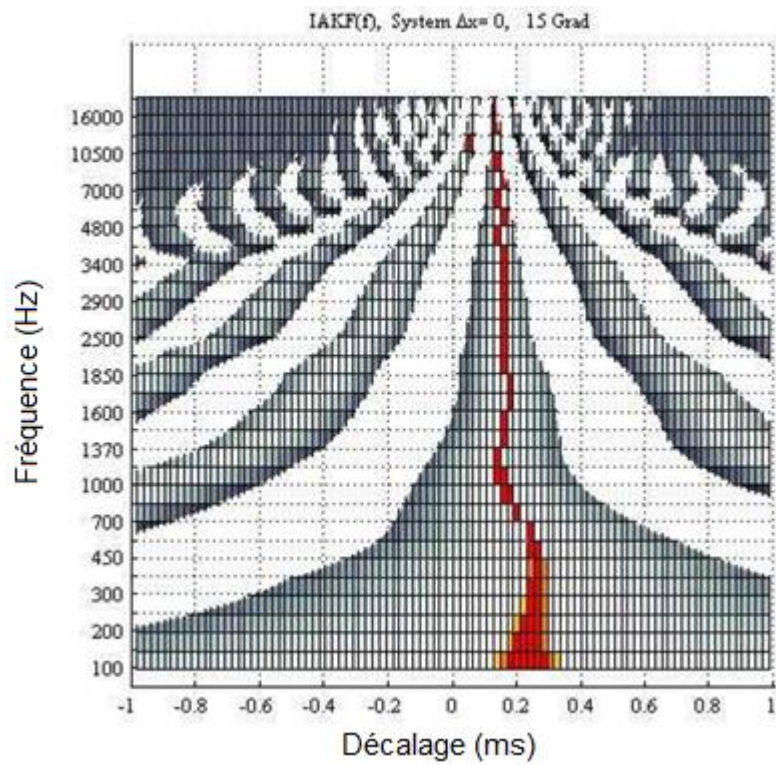


Figure 51 - Analyse de corrélation inter-aurale en fonction de la fréquence, source réelle décalée de 15° à droite[60]

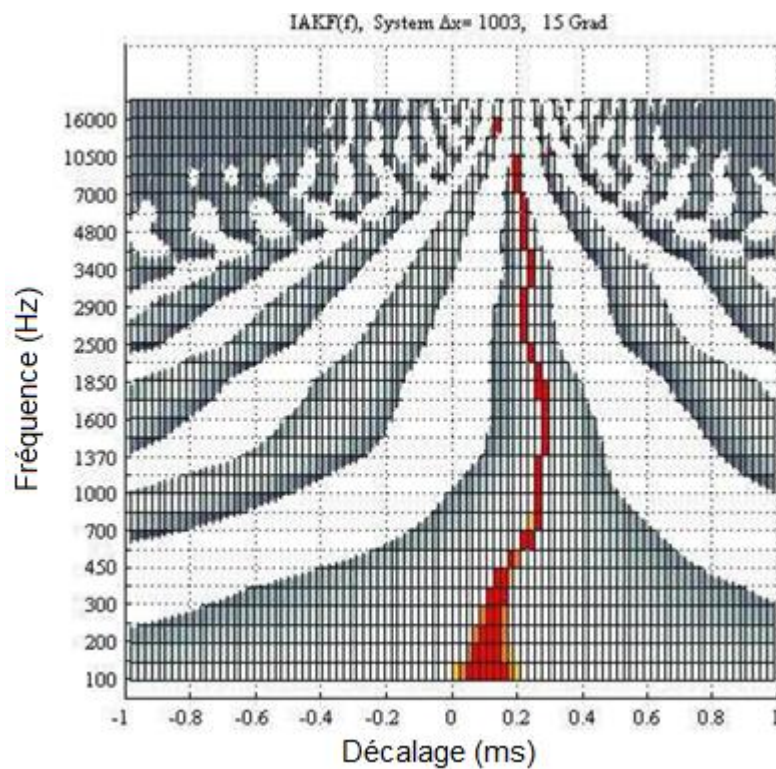


Figure 52 - Analyse de corrélation inter-aurale en fonction de la fréquence, source fantôme créée par un système de sonorisation stéréophonique, décalée perceptivement de 15° à droite par panoramique d'intensité et de temps [60]

3.2.3. Modèle de Mason

Mason [61] propose ce schéma pour un modèle de localisation :

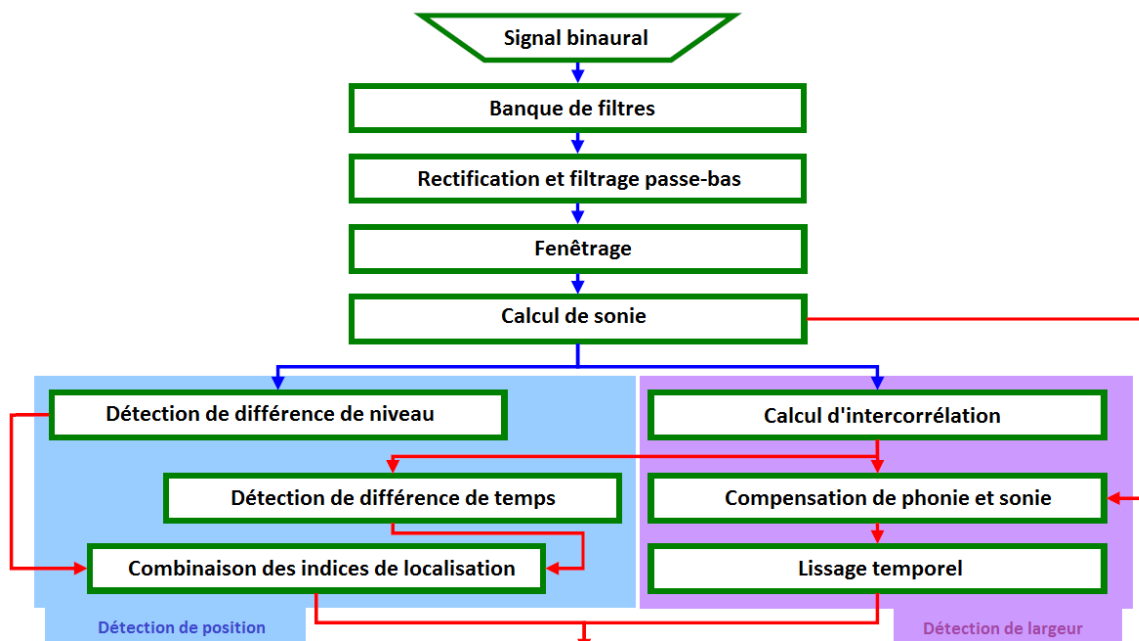


Figure 53 - Modèle proposé par Mason dans [61]

Il obtient grâce à ce modèle des résultats sous la forme d'un diagramme d'angle et de largeur perçus en fonction de la fréquence, tel que dans les figures suivantes.

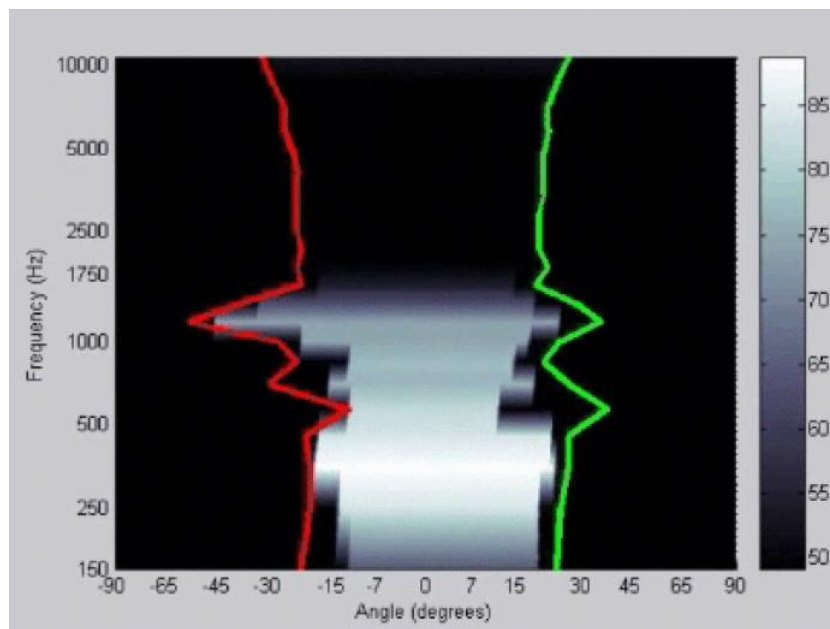


Figure 54 - Diagramme de latéralisation proposé par Mason

Le modèle utilisé par Mason fournit aussi des diagrammes d'angle et de largeur de source perçue en fonction du temps, tels que dans les figures suivantes.

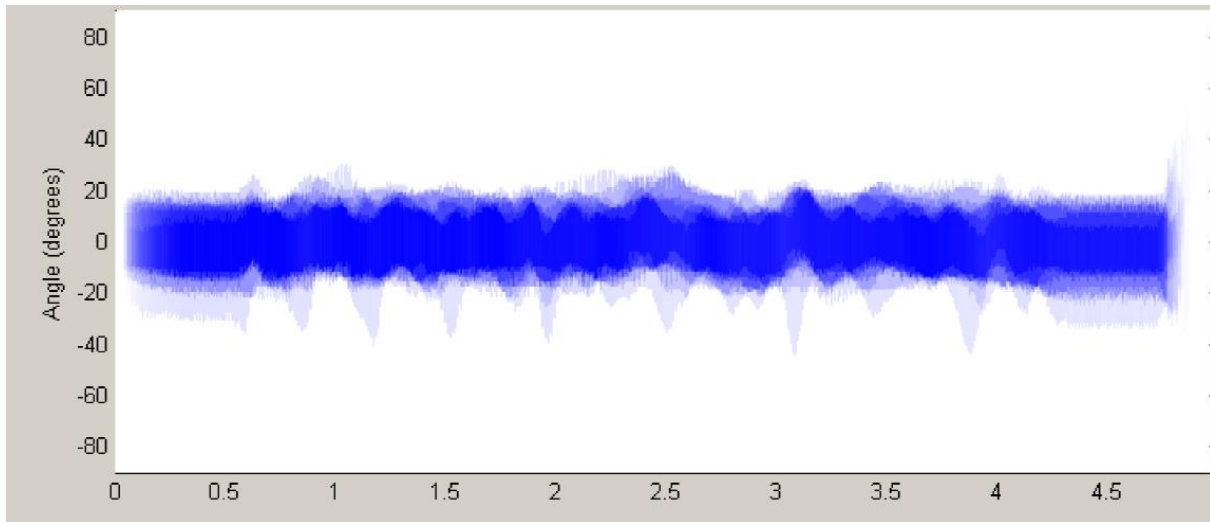


Figure 55 - Diagramme de latéralisation en fonction du temps obtenu dans une pièce au standard ITU-R BS 1116, en lisant un signal corrélé sur un système stéréophonique

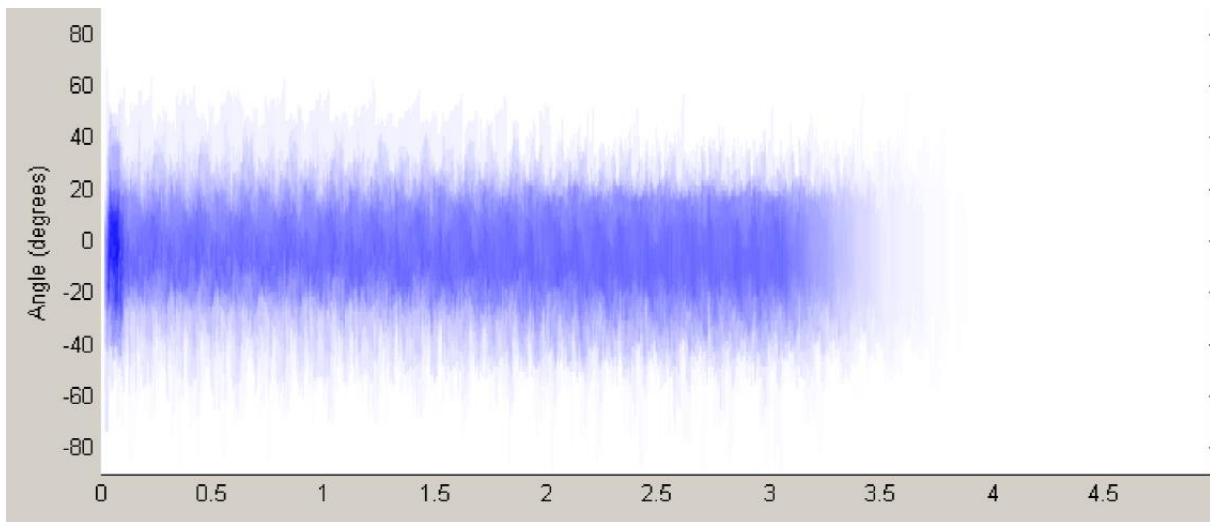


Figure 56 - Diagramme de latéralisation en fonction du temps obtenu dans une automobile équipée d'un système audio typique, en lisant un signal corrélé

On constate que le deuxième système reproduit un signal totalement corrélé de façon moins fidèle car moins ponctuelle, ce à quoi on s'attendait. Malheureusement le logiciel utilisé par Mason n'est pas disponible pour nos tests.

3.2.4. Modèles de type Jeffress - Lindemann – Gaik

Dans [29], Braasch déduit des résultats d'expériences avec une source réelle que le modèle de Gaik lève des ambiguïtés tout en étant plus précis en hautes fréquences que celui de Lindemann. Par exemple, dans les diagrammes d'activités de la figure ci-dessous ont été mesurés pour une même latéralisation d'une source réelle, on constate que la figure en bas à droite ne correspond pas à la bonne latéralisation, ce qui illustre l'imprécision du modèle de Lindemann aux hautes fréquences.

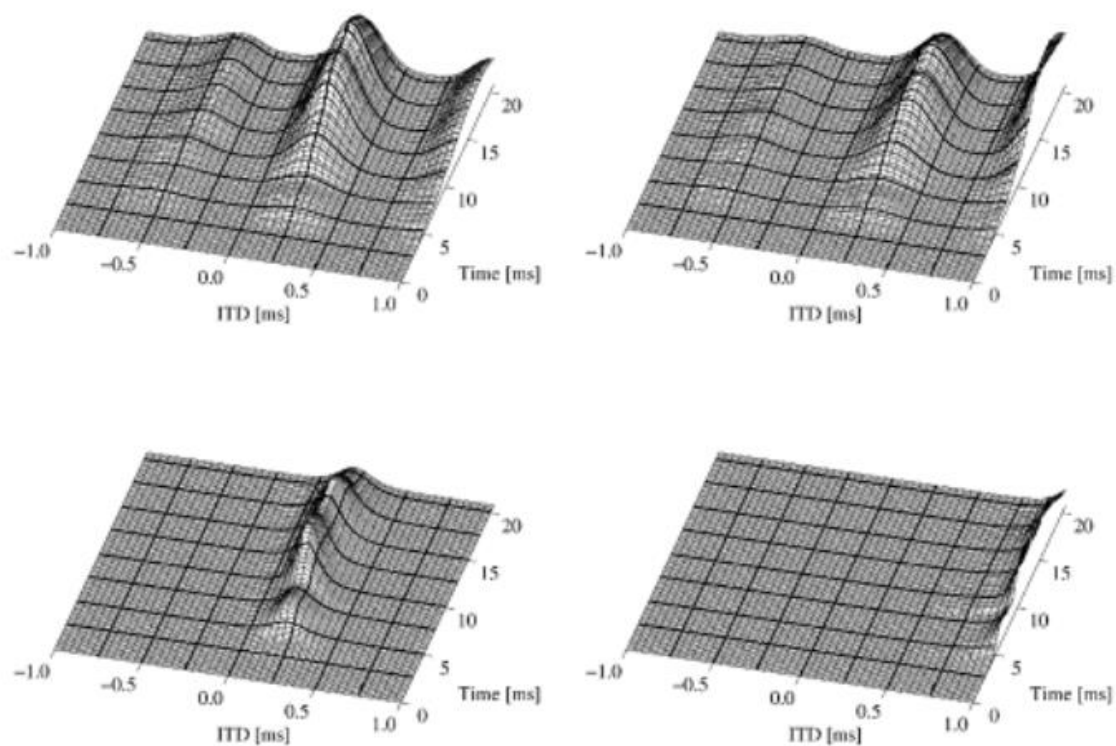


Figure 57 - Diagrammes d'activité binaurale, cadran de gauche modèle de Gaik, cadran de droite modèle de Lindemann, cadran du haut bande de fréquence centrée à 1172Hz, cadran du bas bande de fréquence centrée à 3539Hz [29]

Comme on ne dispose que du code du modèle de Lindemann, on est prévenu de ses limites.

3.2.5. Boîte à outils de Faller et Merimaa

Le code utilisé par ces auteurs:

- génère un signal avec des transitoires, une voix mono ou un bruit rose (le balayage fréquentiel ne donne pas de bons résultats),
- enregistre le signal issu de deux microphones,
- calcule la corrélation inter-aurale dans des bandes de fréquences centrées sur 125, 250, 500, 1000, 2000, 3000, 4000, 6000, 8000 Hz,
- pour chaque bande de fréquences, calcule les ITD et ILD et applique un modèle statistique (on obtient un diagramme d'activité binaurale où les densités de probabilité $p(\Delta L)$ et $p(\tau)$ pour toutes les bandes de fréquences sont représentés)
- pour chaque bande de fréquences, calcule un diagramme d'activité binaurale qui ne prend en compte que les ITD et ILD quand la corrélation inter-aurale est au-dessus d'un seuil paramétrable.

Les auteurs donnent le type de diagramme d'activité binaurale illustré dans la figure suivante, pour une mesure de position azimuthale de sources réelles dans une pièce réverbérante.

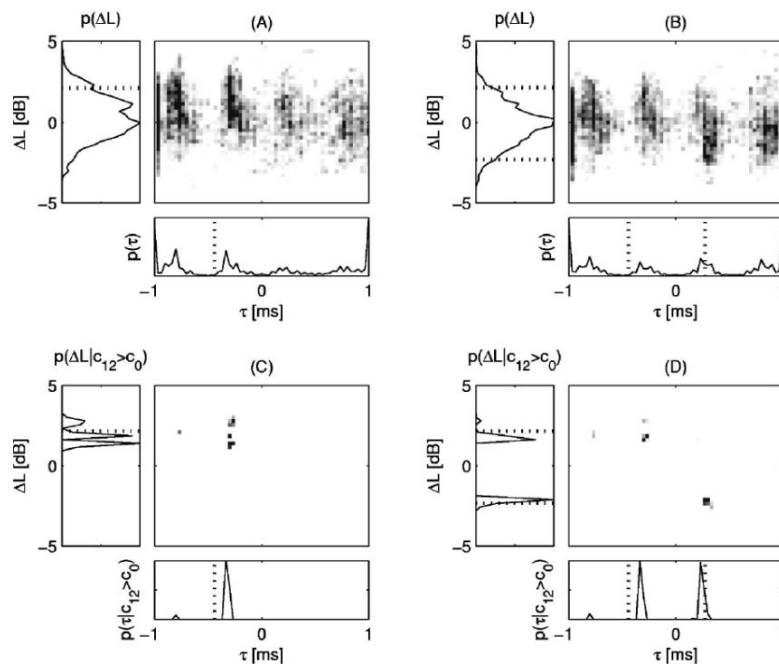


Figure 58 - Résultats de la mesure de Faller et Merimaa [28] pour une source réelle (A) et deux sources réelles (B) sans seuil de corrélation inter-aurale, et respectivement C et D avec seuil

Leurs expériences concernaient des sources réelles, mais l'utilisation d'un seuil de corrélation inter-aurale pour augmenter la précision de localisation apparaît pertinente pour notre usage.

3.2.6. Performance des modèles de latéralisation en milieu réverbéré

La performance des modèles d'audition est habituellement analysée en comparant la latéralisation de sources réelles avec la latéralisation calculée par ces modèles [32] [55], mais la littérature ne regorge pas de comparatif de performance des différents modèles précités. Comme ce sont les sources fantômes qui nous intéressent dans le cadre de cette étude, il serait surtout intéressant de comparer les performances des modèles d'audition aux résultats de panels d'écoute avec des systèmes stéréophonique correspondant aux recommandations vues au chapitre 2, tel que l'a fait Pulki dans [52]. Ce dernier concluait que les résultats de son modèle étaient bons dans un local très peu réverbéré, mais ambigus dans un local réverbérant.

Li & Loizou [62] puis Woodruff & Wang [63], ont montré qu'il est possible d'analyser les performances de séparation des modèles d'audition spatiale en mesurant comment ils extraient le signal utile (dans leur contexte de la voix) du fond (autres signaux concurrents, bruit, réverbération), et en calculant combien cette extraction est proche d'un masque de séparation binaurale idéale (*Ideal Binary Mask Ratio*²¹). Hummersome a comparé différents modèles de cette façon en fonction de la durée de réverbération dans [9], et on en retient la figure suivante.

²¹ On trouvera plus de précisions sur ce concept dans [75] et dans [12] pp88-89.

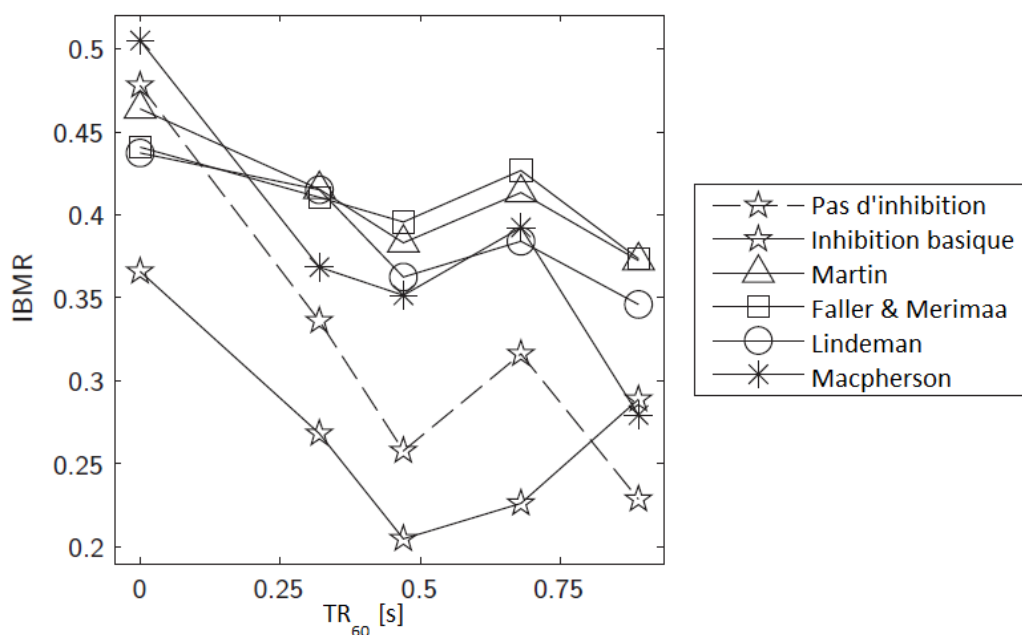


Figure 59 - Performances de séparation spatiale des différents modèles de localisation en fonction du temps de réverbération, traduit de [9]

Avec le recul, on se rend compte que pour la durée de réverbération typique dans l'habitacle automobile (environ 70 ms), utiliser le modèle de Faller et Merimaa serait peut-être moins optimal qu'utiliser le modèle de Macpherson par exemple, ce premier pourrait même être de performance comparable à un modèle sans inhibition dans ces conditions.

Il nous reste maintenant à tester les différents programmes de mesures disponibles avant et après égalisation dans une voiture, et à analyser la performance de ces outils pour la mise au point en comparant leurs résultats à des écoutes critiques.

3.3. Essais réalisés chez Arkamys

3.3.1. Procédure test

Dans ce qui suit, quand un système audio est dit « réglé » ou « avec réglage », c'est que des panels de quelques auditeurs ont jugé que la fidélité était optimisée par le traitement du signal audio appliqué.

Sauf mention contraire, on lit sur l'autoradio une salve de bruit rose puis un balayage fréquentiel sur tous les haut-parleurs et l'on enregistre le signal des deux microphones d'un simulateur tête torse, placé sur un siège d'une automobile et regardant en face, quand le traitement du signal audio embarqué n'est pas réglé, puis l'on recommence quand il est réglé.

Les différents calculs et modèles utilisés et comparés sont présentés ci-après.

3.3.2. Mesure de la phase binaurale en fonction de la fréquence

3.3.2.1. Mesures de phase binaurale et vérification de sa symétrie entre sièges opposés

Des mesures de phase binaurale, lors de la lecture d'un bruit rose corrélé dans le véhicule B²² sans traitement audio, sont présentées dans les figures suivantes. Elles ont été faites avec deux microphones installés au niveau des oreilles d'un auditeur assis dans le véhicule.

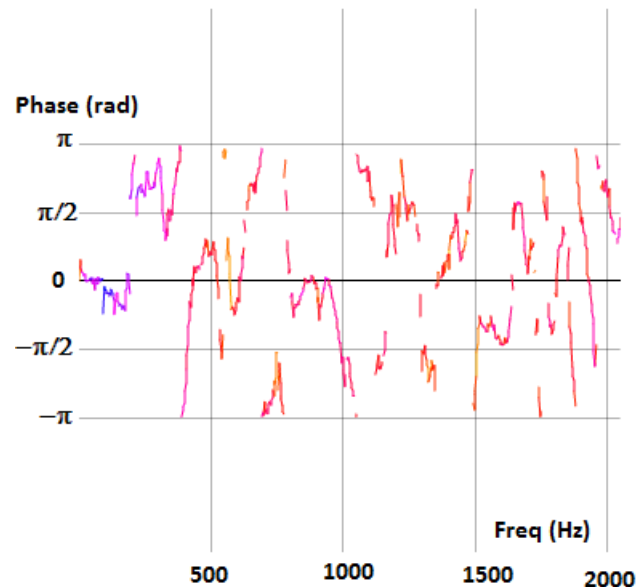


Figure 60 - Phase binaurale sans traitement audio, position conducteur à gauche

²² Une cabine de pick-up à une seule rangée de sièges, équipé de 4 haut-parleurs dont seulement les 2 frontaux sont activés.

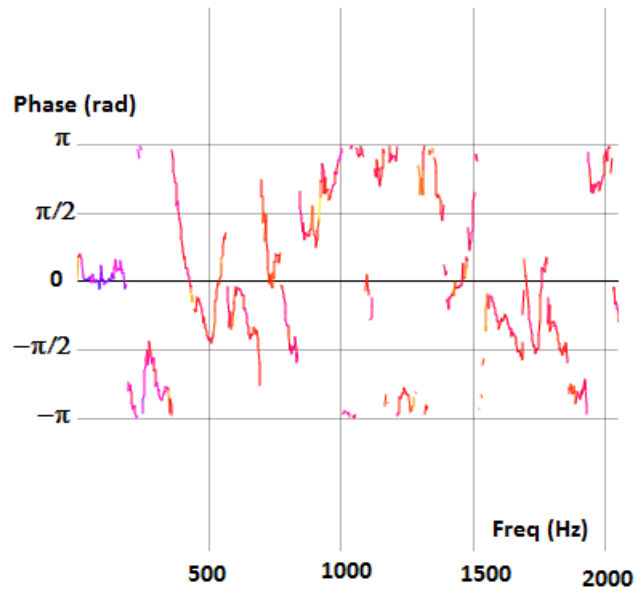


Figure 61 - Phase binaurale sans traitement audio, position passager à droite

Comme l'on pouvait s'y attendre, cela ne ressemble pas totalement à la théorie développée par Carpentier non plus qu'à la simulation prenant compte de la diaphonie inter-aurale à la position du conducteur, un peu plus à la position du passager. On cherche maintenant à vérifier la symétrie de la phase binaurale entre la position du conducteur et du passager, en superposant la phase binaurale à la position du conducteur et le symétrique vertical de la phase binaurale à la position du passager avant dans la figure suivante.

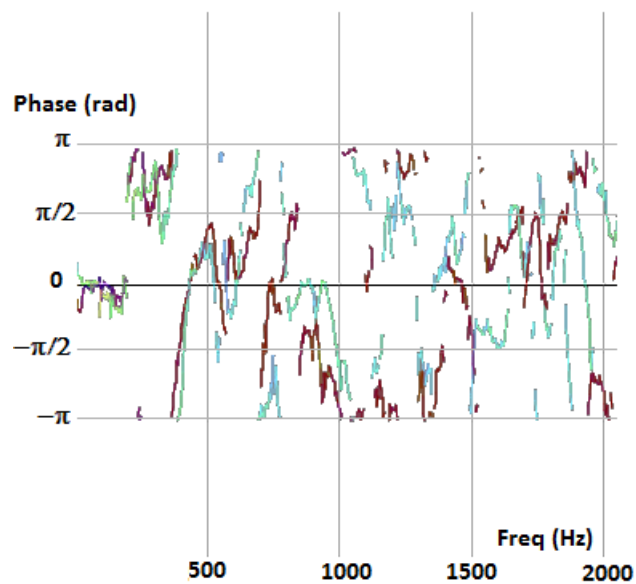


Figure 62 - Phase binaurale sans traitement audio, position conducteur à gauche en vert et cyan, symétrique de la phase binaurale en position passager à droite en bordeaux

On constate que dans ce véhicule, jusqu'à 2 kHz la phase binaurale repliée est globalement symétrique entre conducteur à gauche et passager à l'avant gauche, et particulièrement de 20 à 750 Hz, puis entre 1600 et 1800 Hz. Les oppositions de phase larges et problématiques se situent entre 200 et 400 Hz, et de 900 Hz à 1400 Hz. On peut s'attendre à ce que la première zone d'opposition soit bien corrigée aux deux places avant par rotation de phase entre canal gauche et droite, mais il semble que cette technique soit moins efficace sur la deuxième bande, du fait du peu de symétrie et du côté chahuté des phase relatives entre ces deux places à ces fréquences.

Les mesures faites selon le même procédé dans le véhicule C, un coupé décapotable fermé, avec 4 haut-parleurs avant, sont présentées dans les figures suivantes.

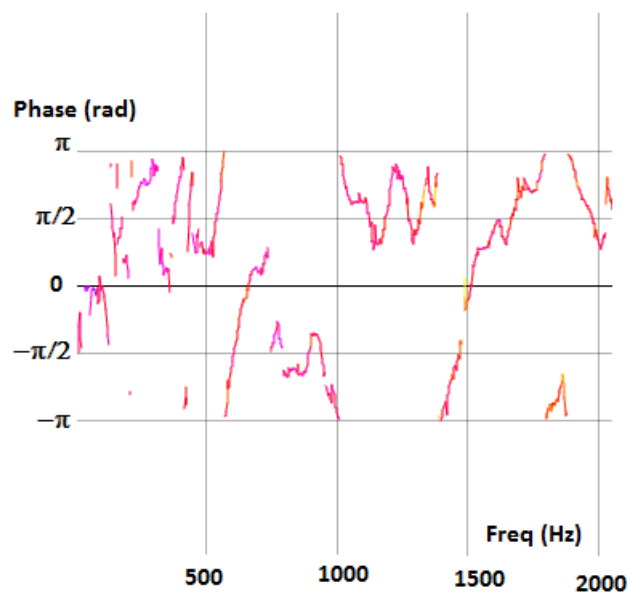


Figure 63 - Phase binaurale sans traitement audio, position conducteur à gauche

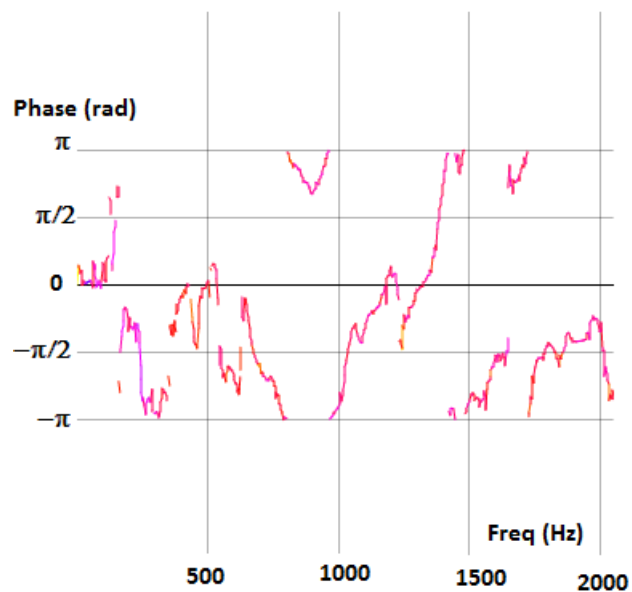


Figure 64 - Phase binaurale sans traitement audio, position passager à droite

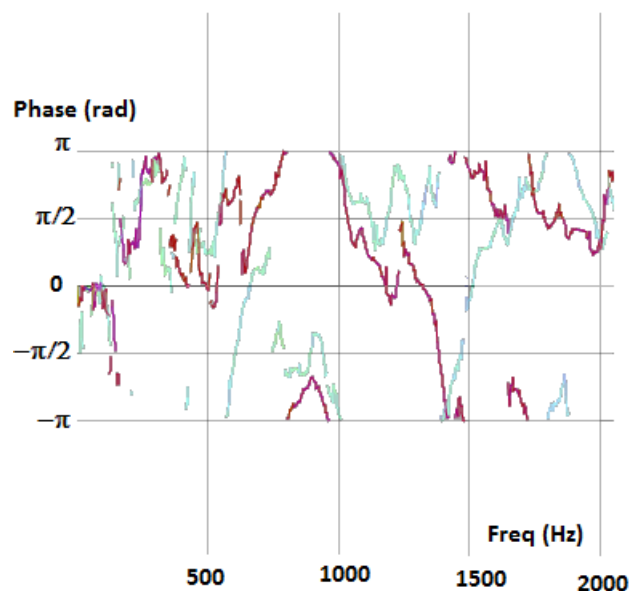


Figure 65 - Phase binaurale sans traitement audio, position conducteur à gauche en vert et cyan, symétrique de la phase binaurale en position passager à droite en bordeaux

La symétrie de la phase binaurale n'est évidente que jusqu'à 400 Hz et autour de 1000 Hz.

Les figures suivantes ont été obtenues avec 6 haut-parleurs, toujours dans le véhicule C.

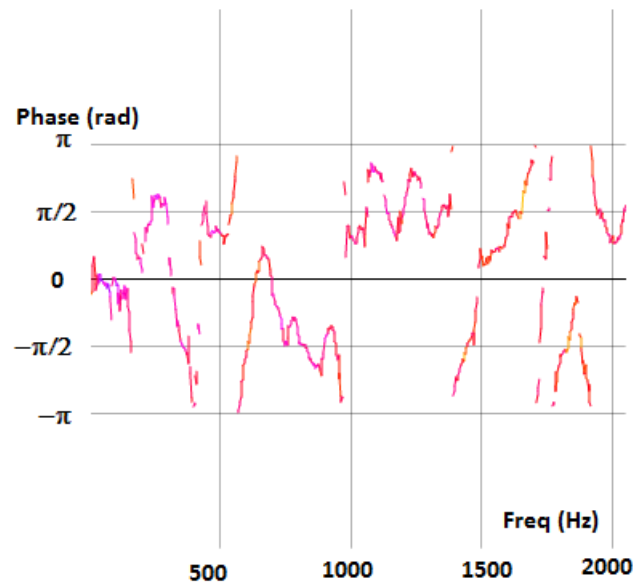


Figure 66 - Phase binaurale sans traitement audio, position conducteur à gauche

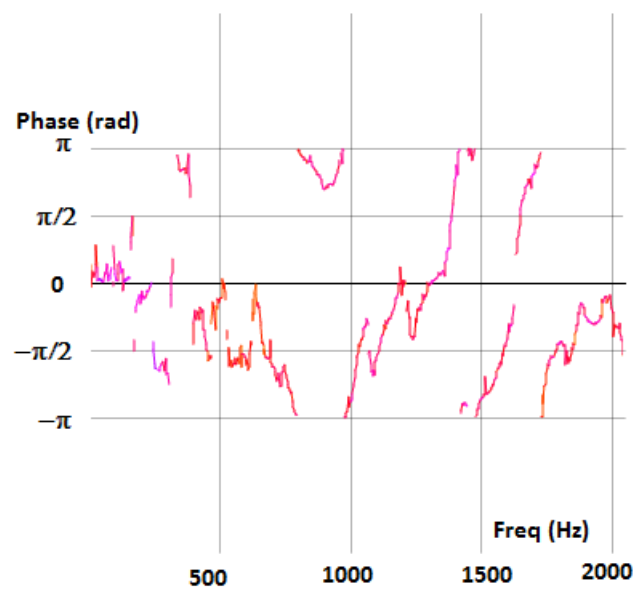


Figure 67 - Phase binaurale sans traitement audio, position passager à droite

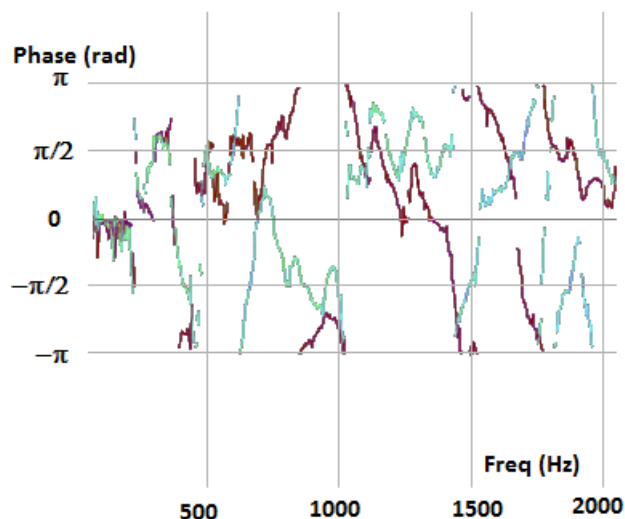


Figure 68 - Phase binaurale sans traitement audio, position conducteur à gauche en vert et cyan, symétrique de la phase binaurale en position passager à droite en bordeaux

La symétrie n'apparaît qu'en dessous de 600 Hz, et un peu entre 1000 et 1200 Hz. Les courbes divergent fortement sur le reste des fréquences.

Conclusion partielle : bien que ces mesures ne correspondent pas totalement avec celles présentées dans Carpentier [43] ni avec la simulation prenant compte de la diaphonie inter-aurale vue au chapitre 3, cette simulation et la méthodologie de réglage habituelle sont en partie validées car l'on constate à cause de la position d'écoute hors axe médian :

- des zones fréquentielles où phase binaurale est proche de π [2π]
- une symétrie locale de la phase binaurale entre les positions d'écoutes symétriques par rapport à l'axe médian, aux fréquences où des filtres d'égalisation de phase sont habituellement appliqués.

3.3.2.2. Comparaison entre réglages

Véhicule D : Les mesures de phase entre 20 Hz et 20kHz sont quasi-inexploitables telles quelles, donc on ne visualise qu'en dessous de 2000 Hz dans la figure suivante.

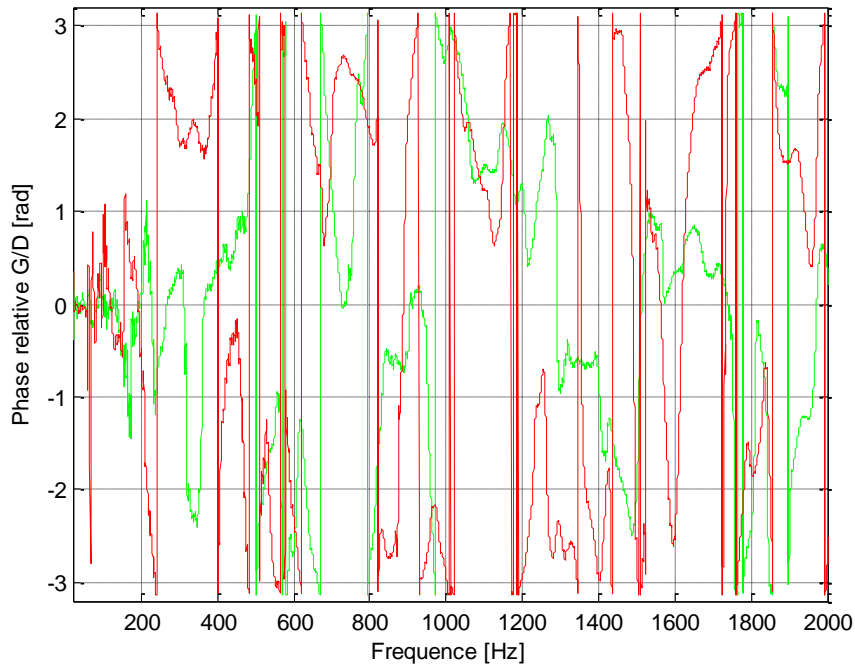


Figure 69 - Phase binaurale, échelle fréquentielle linéaire - rouge sans traitement audio - vert avec réglage

On constate une nette amélioration de la phase binaurale, principalement dans les zones fréquentielles [200 Hz ; 400 Hz], [1200 Hz ; 1400 Hz] et [1600 Hz ; 1800 Hz]. Par contre, un défaut à 170 Hz n'est pas corrigé.

Véhicule E :

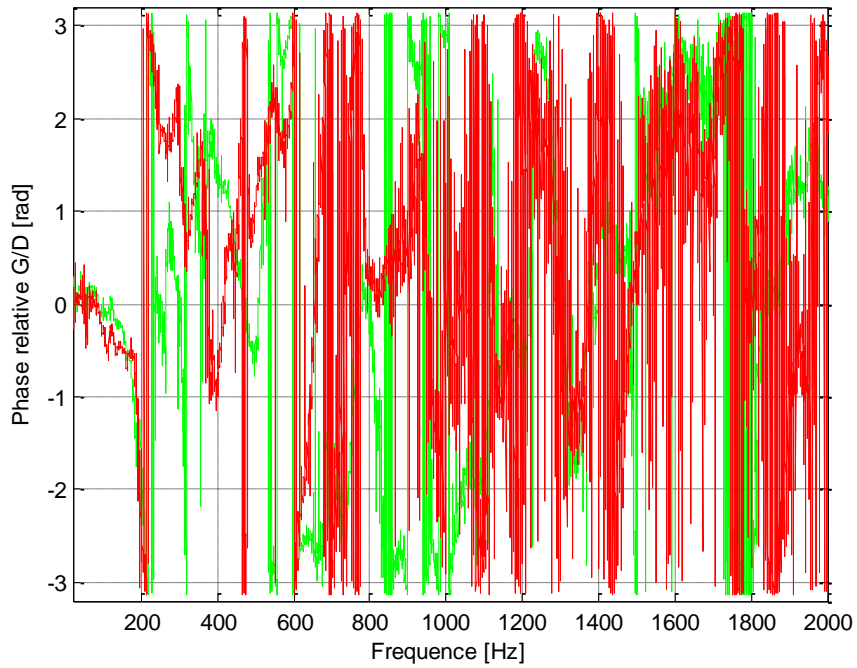


Figure 70 - Phase binaurale, échelle fréquentielle linéaire - rouge sans traitement audio - vert avec réglage

On constate une légère amélioration entre 230 Hz et 320 Hz mais une dégradation entre 800 et 1KHz. Cette figure est très chargée et plus difficile à analyser que la figure 70, il faudrait recalculer avec moins de précision, par exemple en interpolant selon l'axe des fréquences.

Véhicule F :

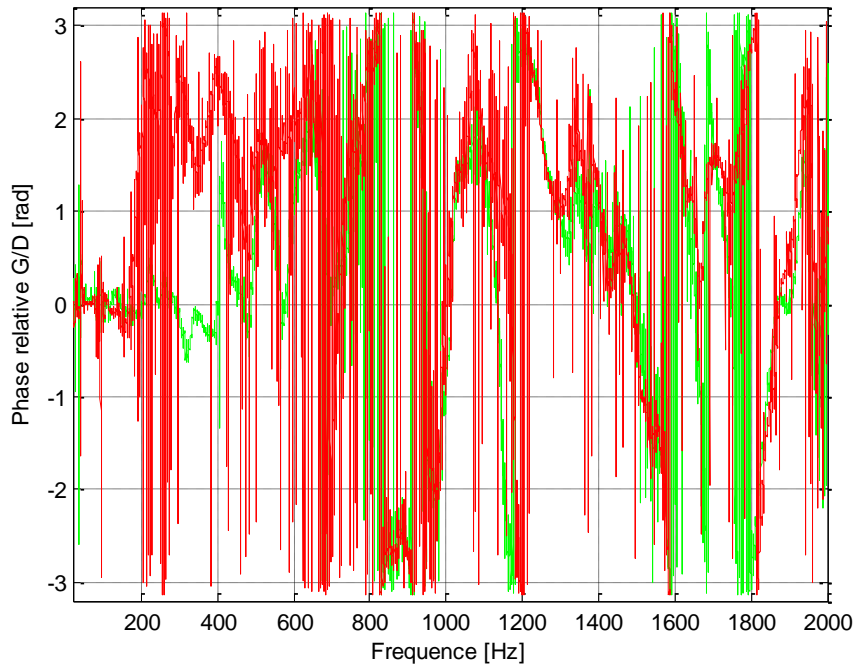


Figure 71 - Phase binaurale, échelle fréquentielle linéaire - rouge sans traitement audio - vert avec réglage

On constate une nette amélioration de la phase binaurale entre 200 Hz et 500 Hz.

Un défaut de cette mesure est qu'elle ne donne pas directement d'indication sur la position ou la largeur des sources fantômes, mais elle informe sur les zones fréquentielles où la phase binaurale est peu lisse ou en opposition.

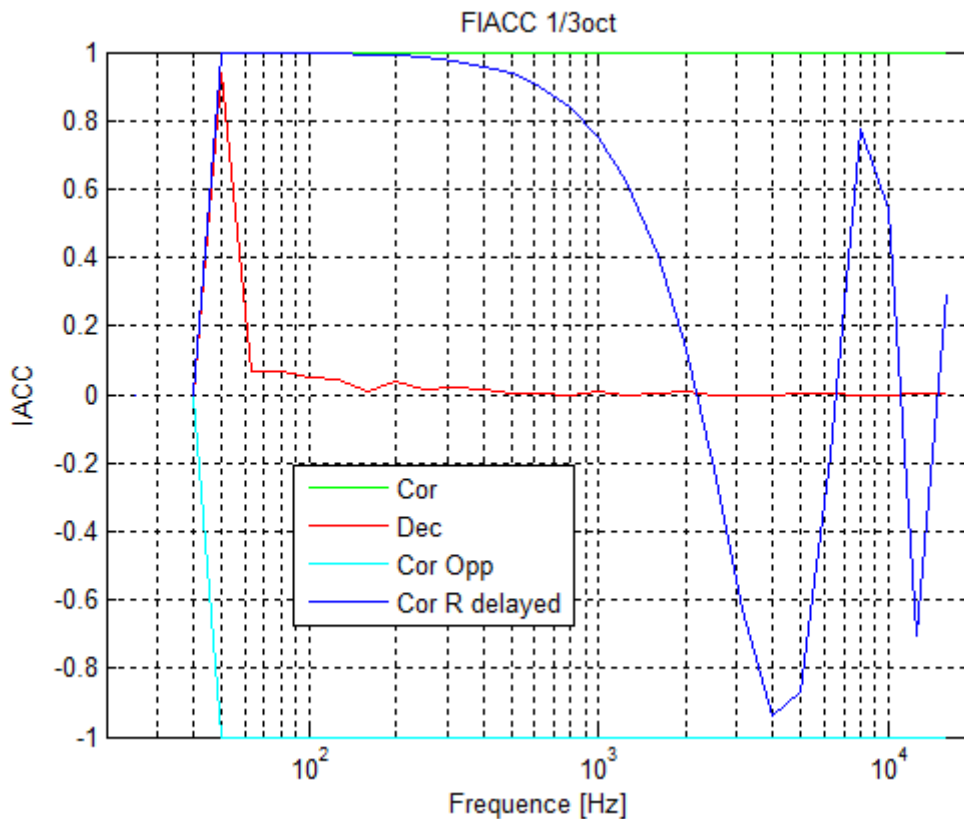
3.3.3. Corrélation inter-aurale dans le domaine fréquentiel

Le code de calcul de FIACC dans [26] est utilisé, la FIACC est normalisée et comprise entre -1 et 1. On modifie ce code pour étendre la plage de fréquences où se fait le calcul jusqu'à 40 Hz, et avoir un axe des fréquences logarithmique, ce qui est plus usuel pour notre usage, plutôt que par numéro de bande critique

Le code résultant est en annexe B.

3.3.3.1. Vérification

Pour vérifier la pertinence du code, on calcule la FIACC pour différents signaux tests à 2 canaux, et on la consigne dans la figure suivante.



Le code donne les résultats auxquels on s'attendait à partir de 60Hz, pour toutes les bandes de fréquences, l'IACC vaut 1 pour des signaux corrélés, -1 pour des signaux corrélés en opposition de phase, 0 pour des signaux totalement dé-corrélés et enfin l'IACC est un sinus cardinal de la fréquence pour un bruit rose corrélé dont un canal a été retardé, dans le cas présent le canal de droite a été retardé de 680µs.

3.3.3.2. Mesures comparatives

On mesure les réponses acoustiques binaurales dans des véhicules sans, puis avec traitement, et on compare les FIACC calculées à partir de ces enregistrements.

Véhicule D :

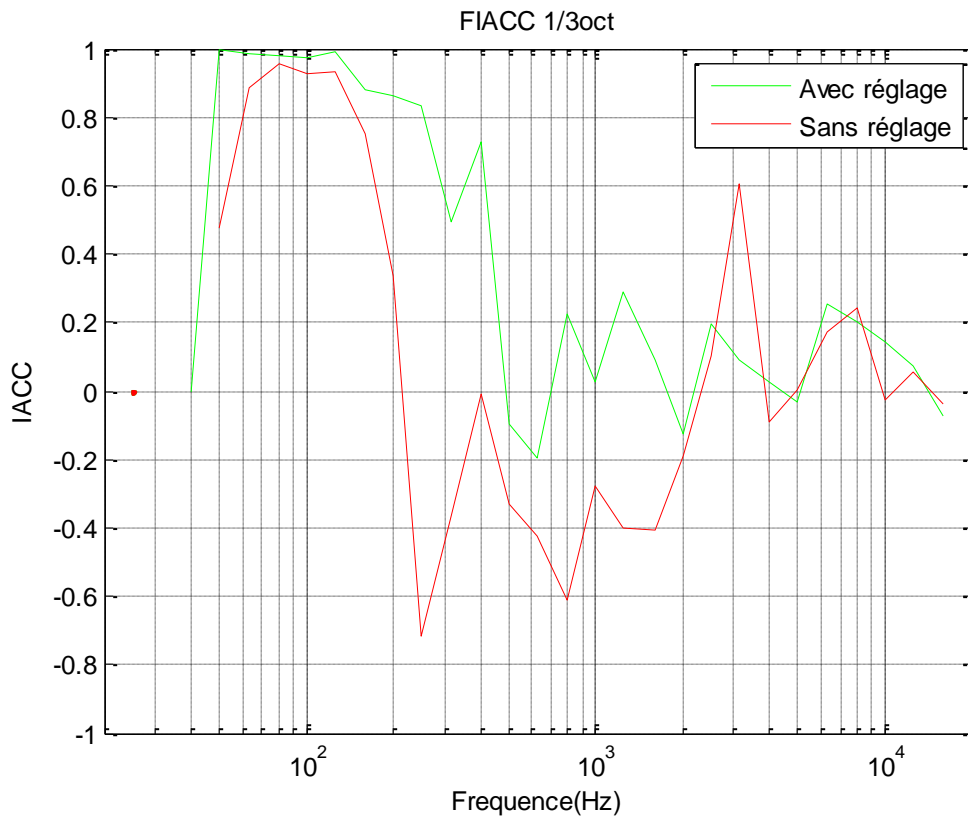


Figure 73 - FIACC avec et sans réglage dans le véhicule D

La FIACC est nettement plus élevée avec le réglage adapté que sans réglage, sauf autour de 3 kHz mais sur une bande de fréquence très fine. Tout cela coïncide avec les résultats d'écoute.

En comparant les FIACC aux phases binaurales mesurées sans et avec réglage adapté, on constate que l'amélioration de cohérence (telle que définie en écoute critique) n'est pas toujours visible sur les mesures de phase relative, les deux mesures apparaissent donc comme complémentaires.

Véhicule E :

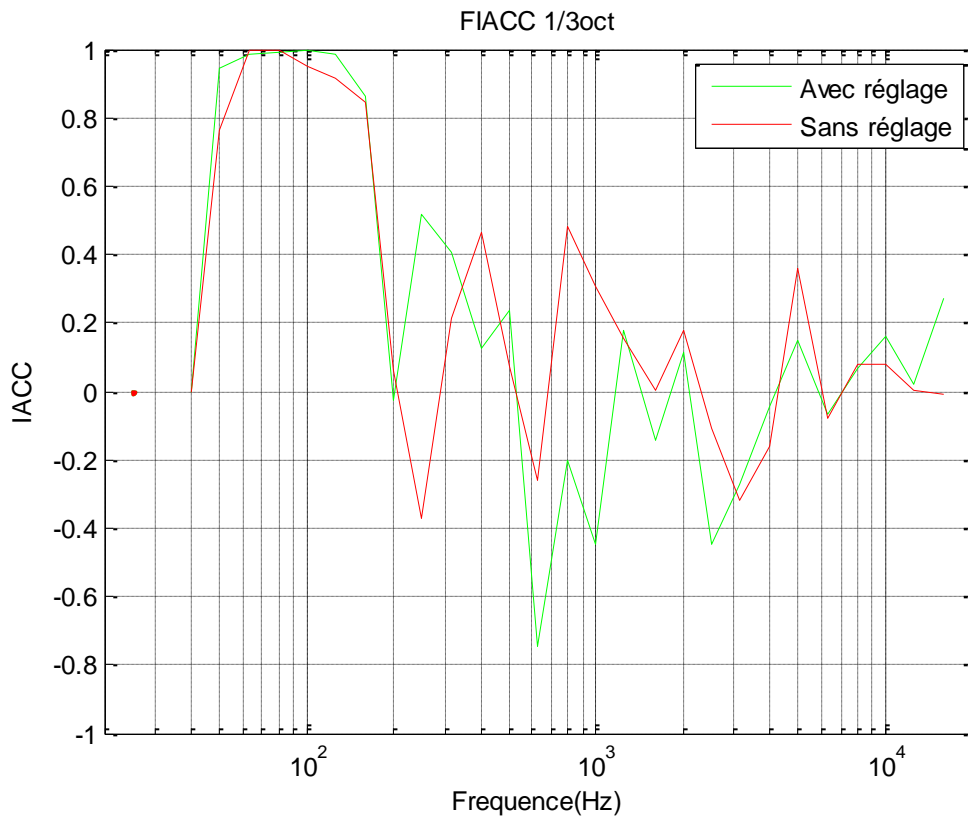


Figure 74 - FIACC avec et sans réglage dans le véhicule E

Le filtrage est moins perfectionné que dans le cas du véhicule D vu précédemment, sur ce chemin audio il n'y a pas de filtrage de phase « passe tout ». Le réglage dégrade nettement la FIACC entre 600 et 1600 Hz, l'améliore autour de 250 Hz, et sinon elle est significativement proche de la FIACC sans traitement. Pourtant les résultats d'écoute indiquent une image plus frontale et plus cohérente avec le traitement. Peut-être aurait-il fallu faire des mesures comparatives avec la tête orientée selon d'autres angles ? Il faudrait calculer la FIACC dans ces deux véhicules en faisant tourner le simulateur tête-torse sur tout le plan azimutal, tel que proposé par Farina et Ayalon dans [53], et comparer à nouveau les résultats.

Véhicule F :

Le filtrage est de complexité comparable à celui du véhicule D. On a appliqué une rotation de phase contrôlée autour de 400Hz et on constate que cette zone fréquentielle voit sa FIACC fortement améliorée.

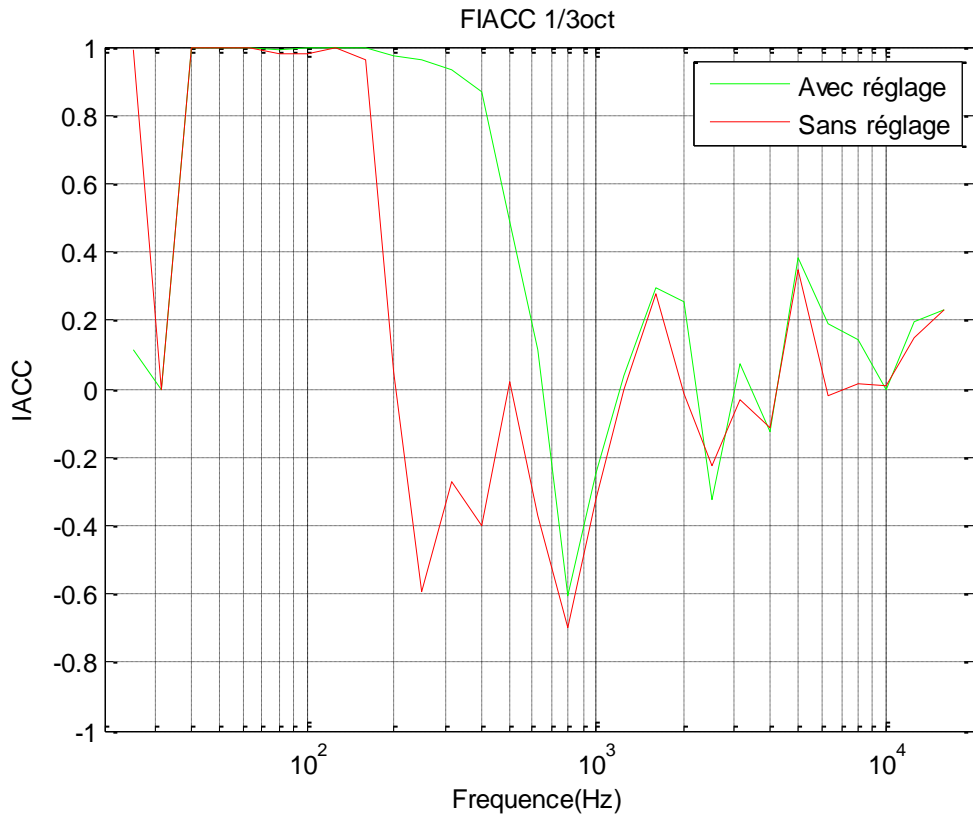


Figure 75 - FIACC avec et sans réglage dans le véhicule F

Le système de mesure était une tête artificielle Schoeps KFM6, avec un mannequin simplifié. L'amélioration globale de la FIACC est cohérente avec les résultats d'écoute.

3.3.4. Mesure de la phase binaurale, de la corrélation binaurale et du diagramme de Lissajous

En plaçant convenablement les fenêtres de mesures fournies avec l'interface audio MOTU 4-Pre, il est possible de visualiser simultanément la phase binaurale, la corrélation binaurale et le diagramme de Lissajous en temps réel, et de comparer différents réglages dans un véhicule. L'idée pour nous est de visualiser la corrélation entre les signaux reçus à chaque place, en même temps que l'on règle le traitement. Le diagramme de Lissajous correspond à ce qui se passe si les signaux des capteurs gauche et droit « alimentent » chacun une des plaques de déviation perpendiculaires d'un oscilloscope analogique : la figure change au fil du temps en fonction des rapports d'amplitude, de fréquence et de phase. La barre verticale cyan (juste à gauche des réponses en phase dans la figure 76) donne des informations en « temps réel » sur la corrélation des signaux binauraux.

Le signal test est un bruit rosé corrélié. Les enregistrements sont faits avec un opérateur assis avec des microphones positionnés à l'embouchure des pavillons des oreilles.

On constate dans la figure 77 que les deux réglages A et B donnent des performances supérieures par rapport au système non réglé :

- phase globalement plus linéaire, nettement plus lisse et plus proche de 0,
- diagrammes de Lissajous plus compacts et plus proches de la diagonale,
- corrélation binaurale supérieure.

Les diagrammes de Lissajous semblent utiles pour définir les réglages en temps réel autrement qu'« à l'oreille ». On cherche les paramètres qui donnaient la figure la plus plate et alignée sur l'axe « en phase », c'est-à-dire la diagonale qui va du coin en bas à gauche à celui en haut à droite. La corrélation binaurale a permis de départager les réglages A et B. Le réglage A créait un défaut en mettant hors phase une bande de fréquence autour de 1kHz, zone très importante pour la perception de la voix parlée, et l'IACC était plus faible qu'avec le réglage B, ce qui correspond aux résultats de tests d'écoutes comparatifs.

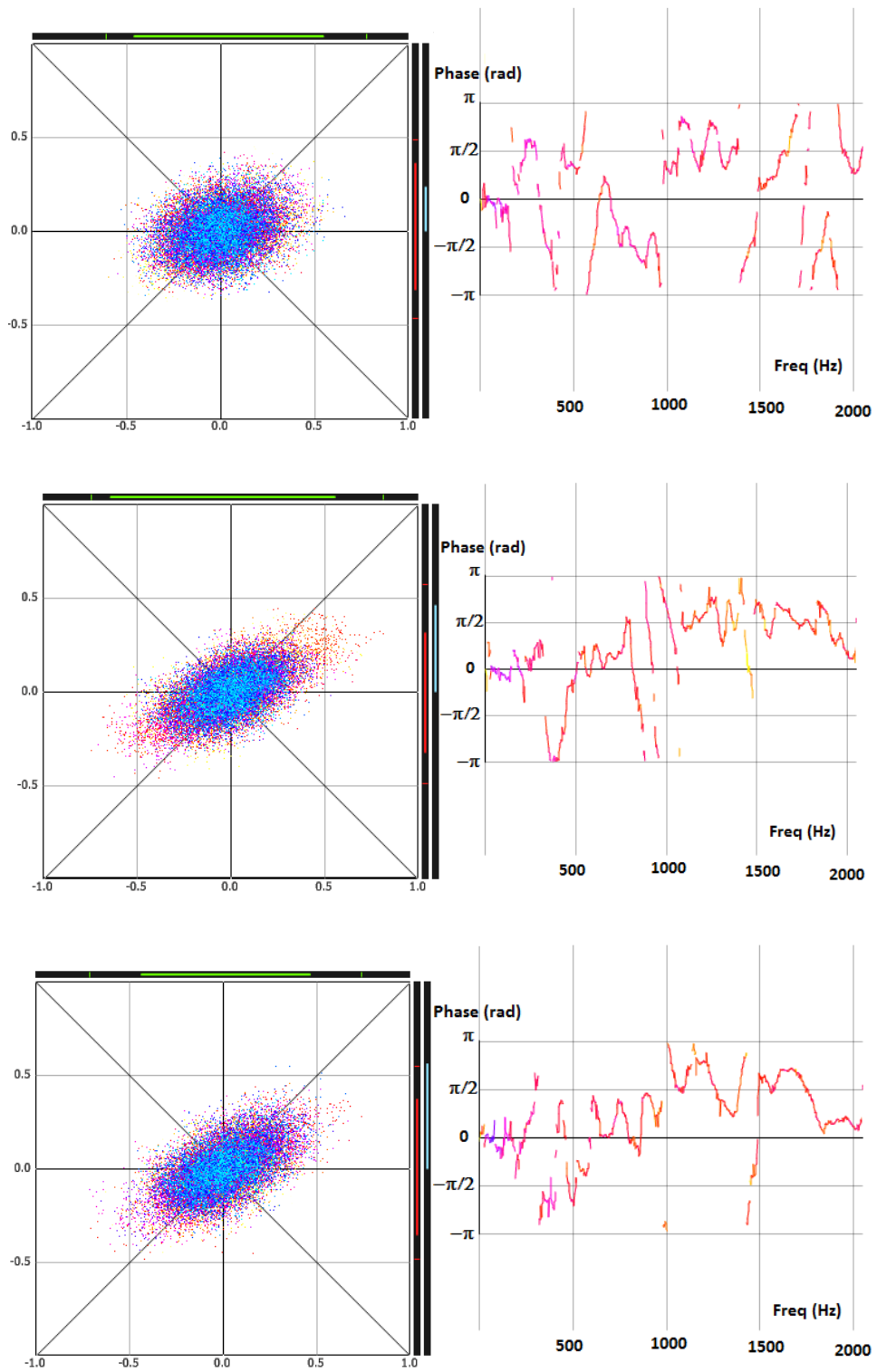


Figure 76 - Résultats à la position conducteur – en haut sans réglage, au milieu avec le réglage A, en bas avec le réglage B

3.3.5. Boite à outils logicielle de Faller et Merimaa

Les raisons principales pour lesquelles nous avons testé ce logiciel est qu'il est destiné aux locaux avec beaucoup de réflexions, et que le code Matlab est disponible en ligne²³, fonctionnel rapidement et aisément modifiable.

3.3.5.1. Test du programme

Quand on « alimente » le modèle avec un bruit rose corrélé, on obtient les diagrammes d'activité binaurale suivants.

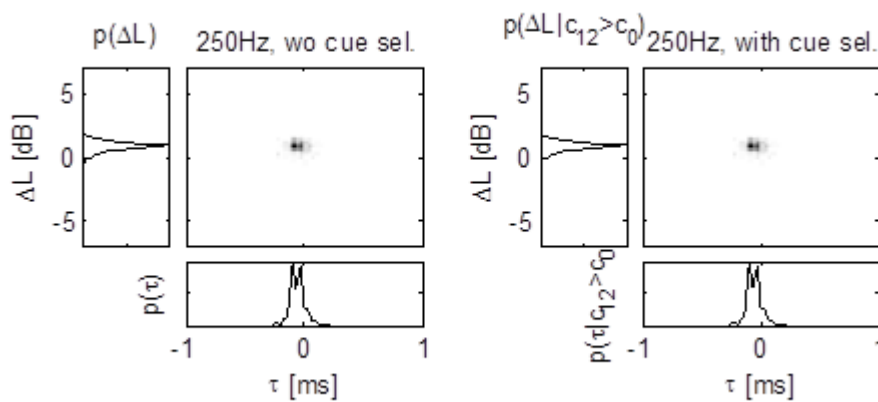


Figure 77 - ITD et ILD dans la bande critique de fréquence centrée sur 250 Hz quand on lit un bruit rose corrélé. A gauche sans seuil de corrélation inter-aurale, à droite avec seuil de corrélation inter-aurale.

Et avec un bruit rose dé-corrélé :

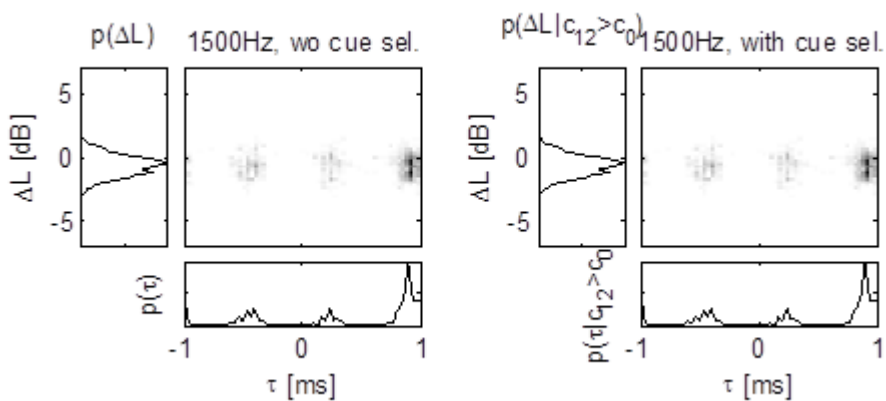


Figure 78 - ITD et ILD dans la bande critique de fréquence centrée sur 250 Hz quand on lit un bruit rose dé-corrélé. A gauche sans seuil de corrélation inter-aurale, à droite avec seuil de corrélation inter-aurale.

Le résultat de la figure précédente est surprenant : on s'attendait à de la symétrie !

²³ <http://legacy.spa.aalto.fi/software/cueselection/>

On propose néanmoins d'analyser les diagrammes d'activité binaurale mesurés en fonction du signal test utilisé (bruit rose corrélé ou dé-corrélé), pour se donner une idée de la pertinence du filtrage. Sur un système bien réglé, un bruit rose corrélé devrait donner une tâche la plus fine possible et idéalement centrée, un bruit rosé dé-corrélé devrait donner trois tâches. Les résultats d'autres tests préliminaires sont surprenants : on ne constate pas de différences significatives selon que l'on utilise deux microphones espacés ou une tête artificielle Schoeps, les ILD ne sont pas strictement 0 quand on place la tête artificielle au milieu de deux haut-parleurs, et on se pose des questions sur la précision fréquentielle du programme.

3.3.5.2. Expérience dans le véhicule A

On utilise deux microphones espacés de 17 cm, on diffuse un balayage fréquentiel avec l'autoradio et on enregistre le signal des deux microphones qu'on fait analyser par le modèle entre 250 Hz et 1.5 kHz.

On compare le résultat sans et avec filtrage. Les probabilités d'ITD et ILD sont représentés sur les axes avec une couleur pour chaque bande de fréquence, et sont fusionnés dans une figure à deux dimensions en niveau de gris au centre du graphique.

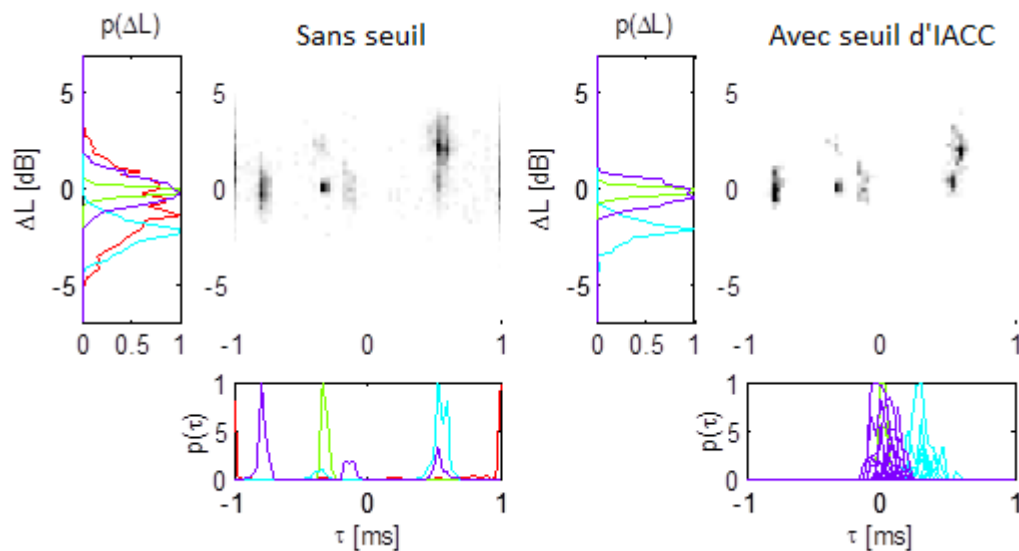


Figure 79 - ITD et ILD, sans filtrage

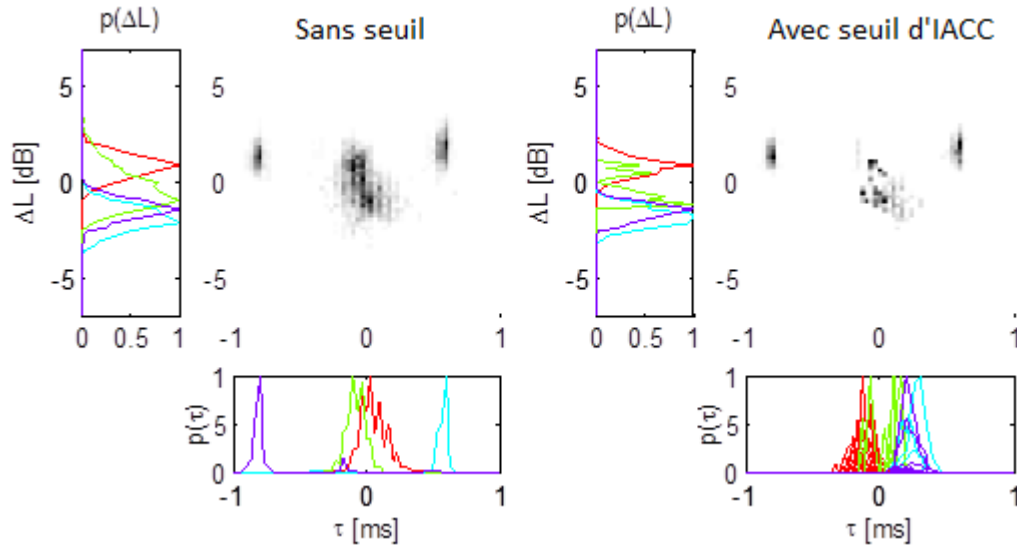


Figure 80 - ITD et ILD, avec filtrage

On voit plutôt bien sur ces figures l'apport du filtrage sur la ponctualité de reproduction d'un signal stéréophonique totalement corrélé : à la lecture d'un tel signal, le diagramme d'activité binaurale se concentre au milieu de la figure quand on active le traitement adapté. On constate également que le réglage améliore l'alignement des tâches latérales sur l'axe des ILD, mais on ne sait pas encore comment interpréter cela. Si on augmente la précision d'analyse en 1/12ème octave, on obtient les figures suivantes.

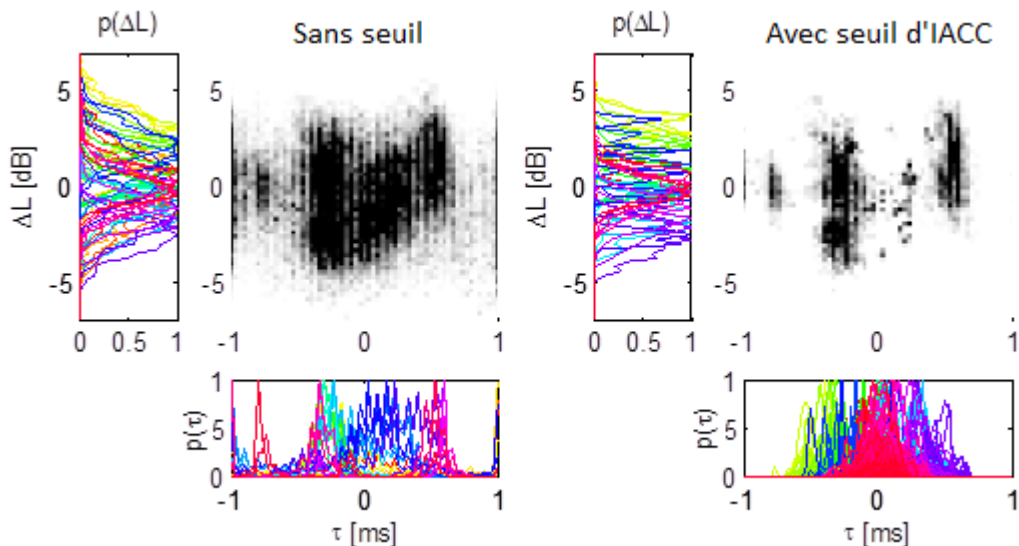


Figure 81 - ITD et ILD, sans filtrage, précision 1/12ème d'octave

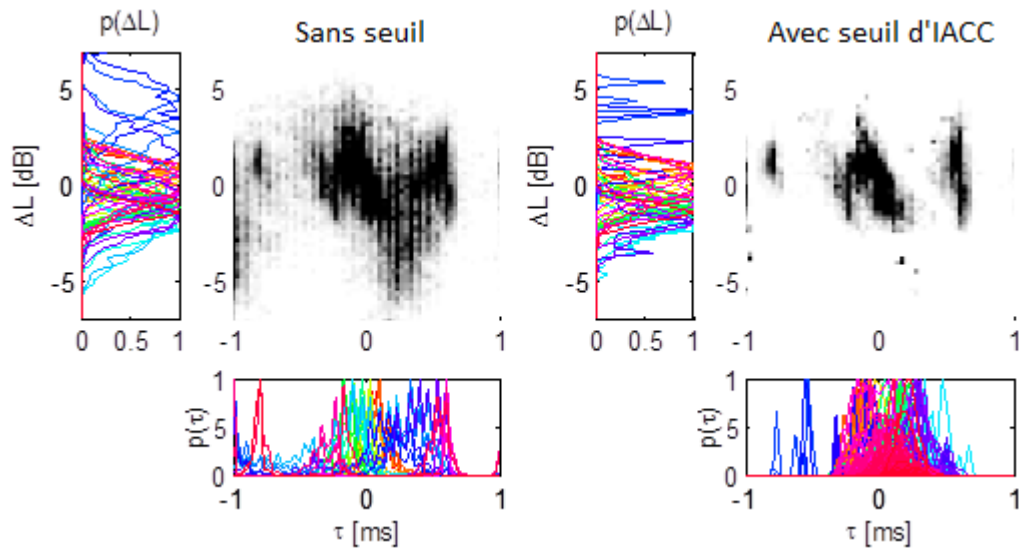


Figure 82 - ITD et ILD, avec filtrage, précision 1/12ème d'octave

L'apport d'une augmentation de la précision fréquentielle reste à prouver. La figure suivante montre les diagrammes d'activité binaurale obtenus quand seulement le haut-parleur avant gauche est activé.

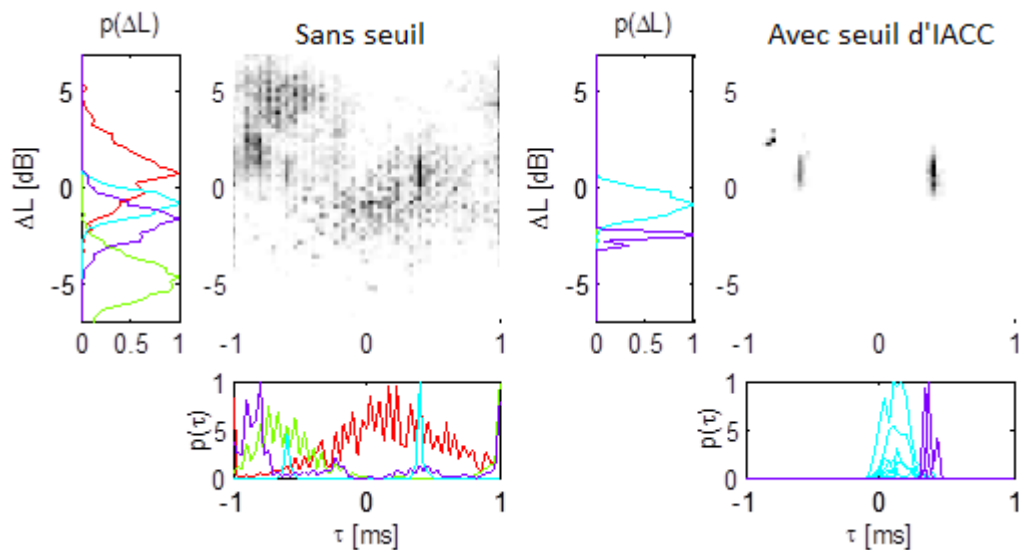


Figure 83 - ITD et ILD, avec seulement le haut-parleur gauche activé

Dans ce cas il n'y a qu'une source réelle et l'apport du seuil d'IACC est net. Il reste à trouver la valeur optimale du seuil d'IACC pour adapter ce procédé à l'étude de sources virtuelles.

On a modifié ce code pour qu'il affiche les ITD en fonction de la fréquence, entre 100 Hz et 1,5 kHz, et constaté une nette amélioration de la corrélation en dessous de 600 Hz, et

particulièrement entre 500 et 600 Hz. Dans ce cadre la définition d'un seuil d'IACC adapté permet de « dé-bruiter » cette mesure d'ITD.

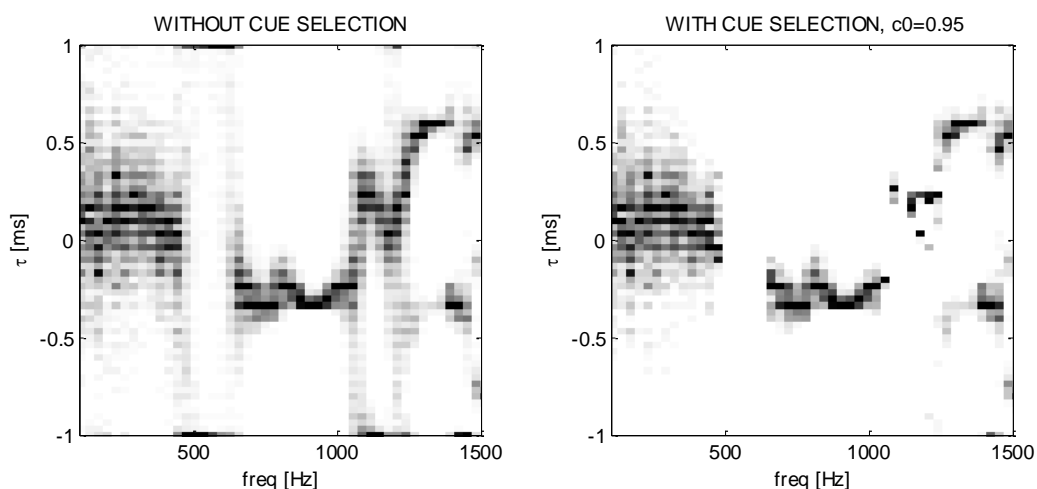


Figure 84 - ITD en fonction de la fréquence, sans filtrage

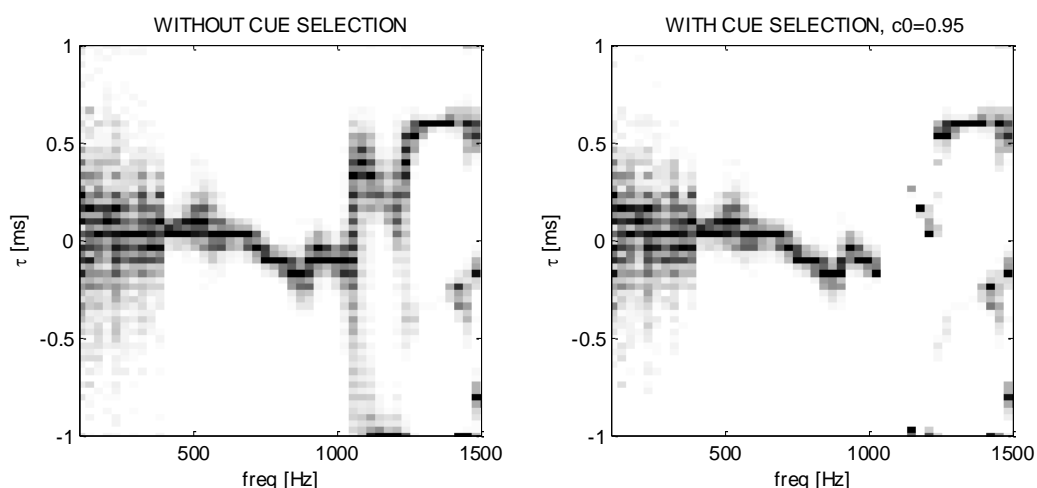


Figure 85 - ITD en fonction de la fréquence, avec filtrage

Des écoutes sont menées dans le véhicule afin de comparer les résultats du modèle de localisation à la perception d'un très petit panel de deux auditeurs formés à l'écoute critique. Les deux auditeurs échangent régulièrement de places entre la place du conducteur et la place du passager avant. Les résultats (voir annexe E) ne correspondent pas toujours à ce qui est mesuré, mais avec le recul les écoutes ne pouvaient pas donner de bons résultats quand le signal test utilisé était un sinus pur, car on a vu qu'ils sont difficiles à localiser. Il est souhaitable de refaire ces écoutes avec des signaux non stationnaires.

En conclusion la boîte à outil de Faller et Merimma :

- est prometteuse pour nos besoins,
- n'est potentiellement pas précise en fréquence,
- mériterait d'analyser plus bas en fréquence,
- nécessite des améliorations et un réglage afin d'être exploitable.

3.3.6. Modèle de Lindemann

On utilise à présent le code « lindemann.m », qui inclut le programme demo_lindemann1986.m fourni dans la boîte à outil amt (<http://amtoolbox.sourceforge.net/>). La boîte à outils Itfat est également nécessaire.²⁴

3.3.6.1. Test du programme

Afin d'être en mesure d'analyser les mesures faites dans les véhicules, on donne à traiter par « lindemann.m » des signaux tests typiques pour avoir des références.

Tableau vii - Signaux utilisés pour le test du code « Lindemann.m »

Signal test	Phases des signaux	Durée
Balayage fréquentiel linéaire mono 20 Hz-20kHz	Identiques	2s
Balayage fréquentiel linéaire 20 Hz-20kHz hors phase	Opposées	2s
Bruit rose dé-corrélé	Aléatoires	2s

²⁴ Les deux boîtes à outils doivent être initialisées avant de lancer lind.m, en exécutant Itfatstart.m et amtstart.m.

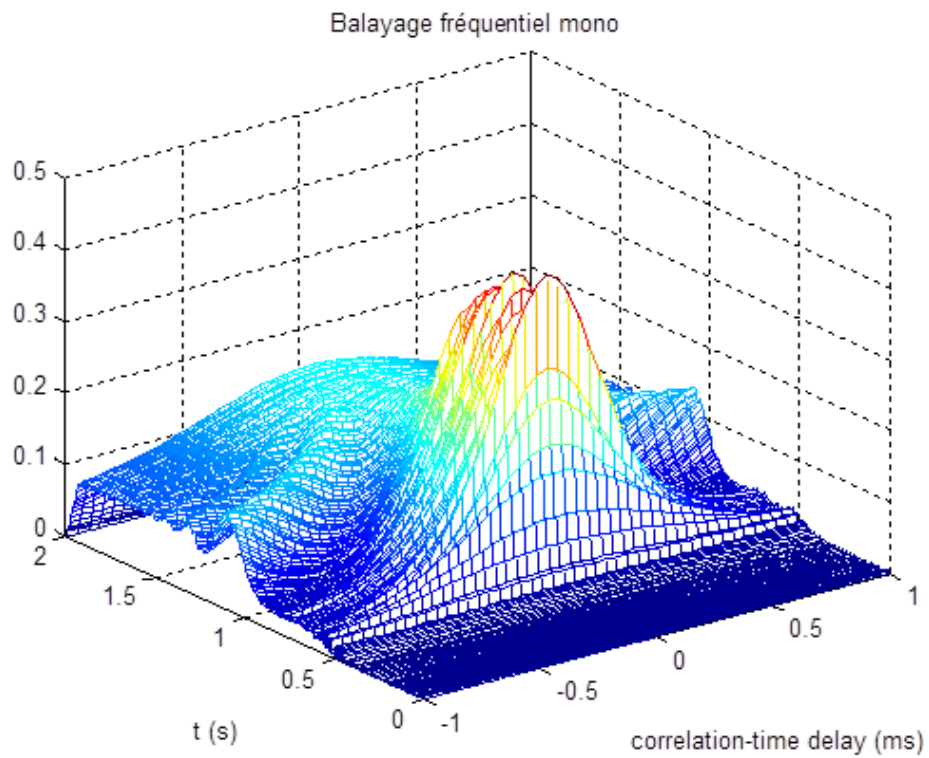


Figure 86 - Résultats pour un balayage fréquentiel corrélé en phase

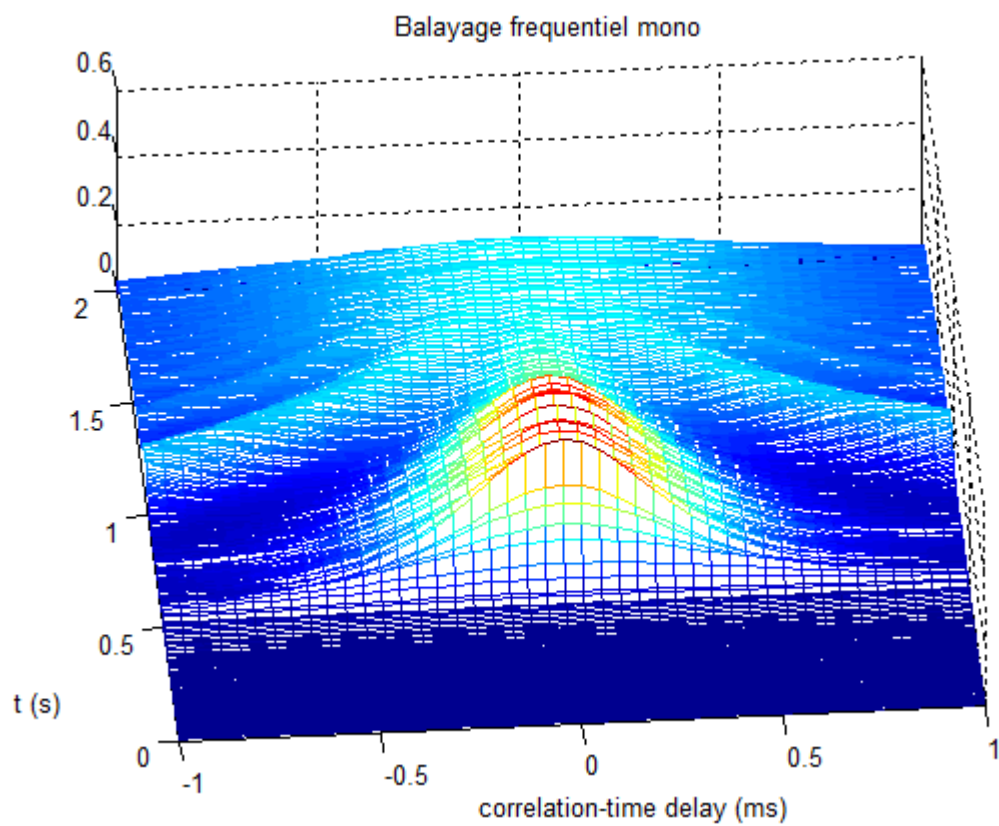


Figure 87 - Figure 86 vue sous un autre angle

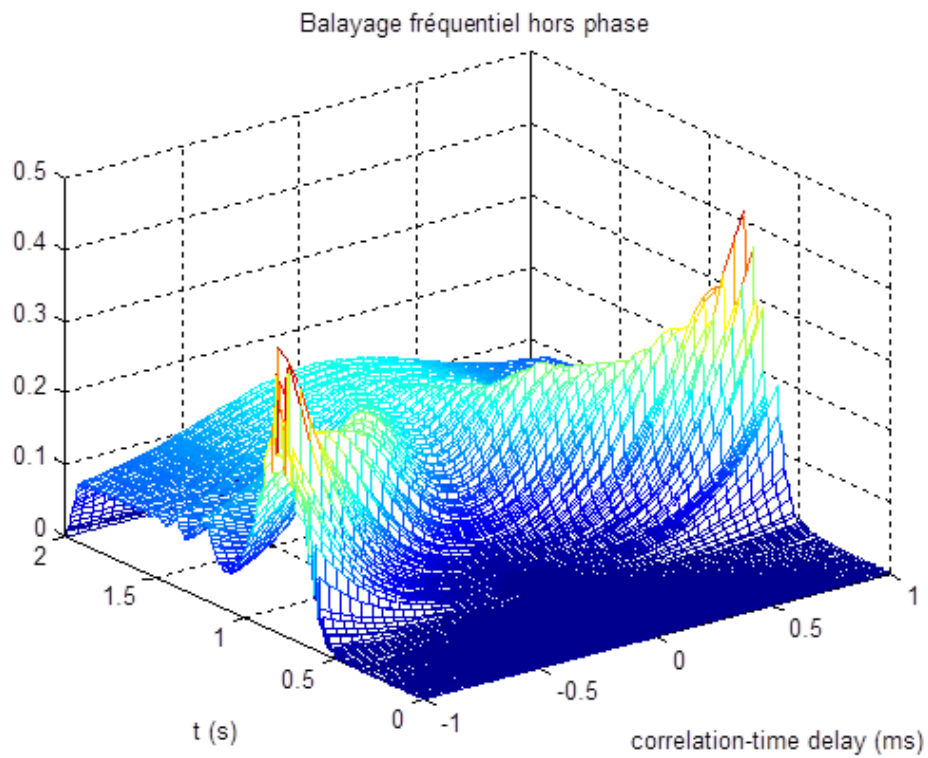


Figure 88 - Résultats pour un balayage fréquentiel, un des canaux étant en opposition de phase

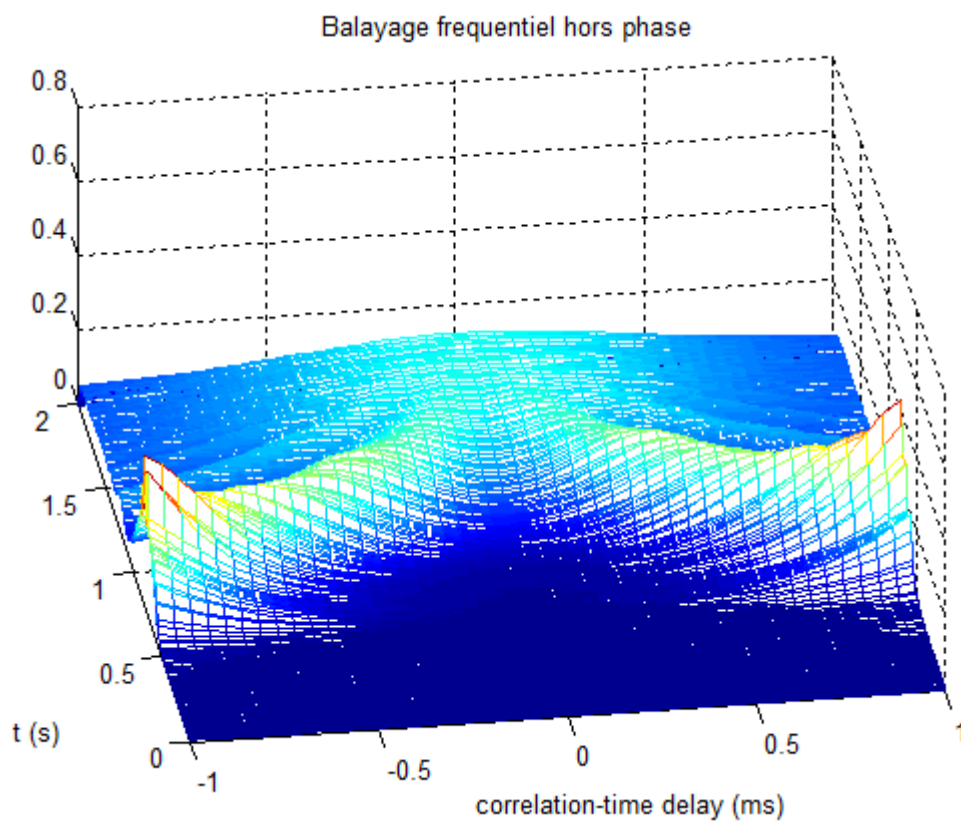


Figure 89- Figure 88 vue sous un autre angle

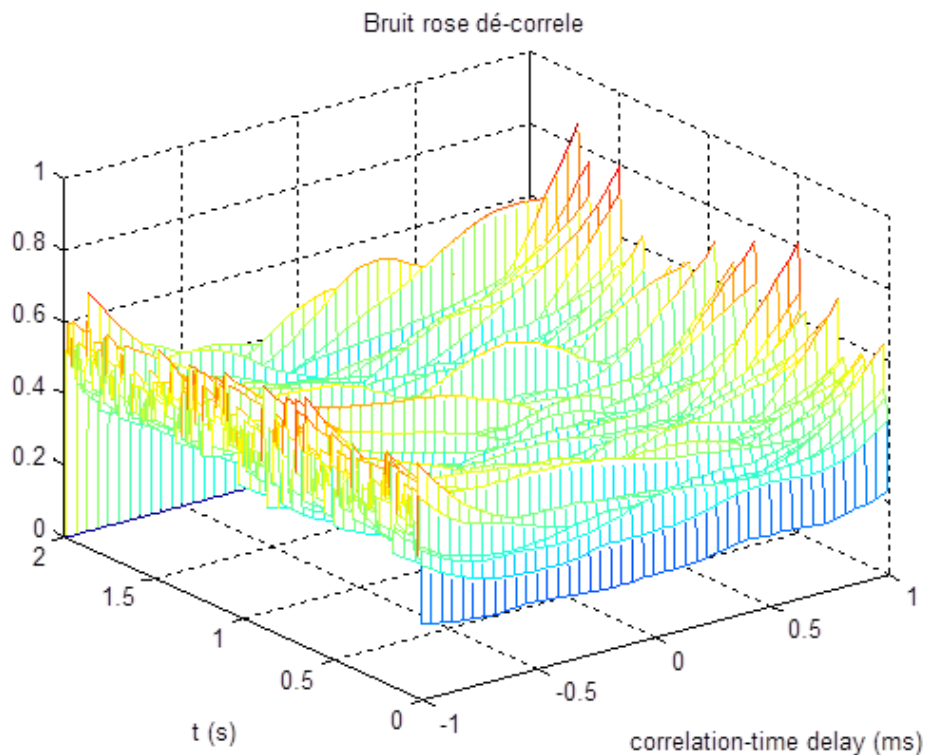


Figure 90 - Résultats pour un bruit rose dé-corrélé

Encore une fois le calcul n'est effectué qu'entre -1 ms et +1 ms de délai d'inter-corrélation, car au-delà de 1 ms de délai inter-aural, certains sons apparaissent totalement latéralisés sur la source réelle en avance.

Le bruit rose dé-corrélé donne une figure entre 0.5 s (100 Hz) et 2 s (20kHz) formant une cuvette relativement plate autours de 0 ms de décalage, et légèrement relevée pour les décalages extrêmes -1 ms et +1 ms.

Le balayage fréquentiel corrélé en phase donne une ligne de pics très nets et alignés sur le délai d'inter-corrélation 0 ms.

Le balayage fréquentiel « hors phase » donne des pics très nets et positionnés sur des délais d'inter-corrélation -1 ms et +1 ms.

Cependant pour ces deux dernier signaux tests le résultat est à 0 en dessous de 0.5s sur l'axe temporel, ce qui pour le balayage fréquentiel utilisé correspond à des fréquences en dessous de 100 Hz environ.

On se dit que c'est peut-être parce que le fichier test est trop court (2 secondes) et on refait le test avec un balayage fréquentiel mono de 10 secondes. Le temps de calcul est nettement rallongé et on obtient le diagramme de la figure suivante.

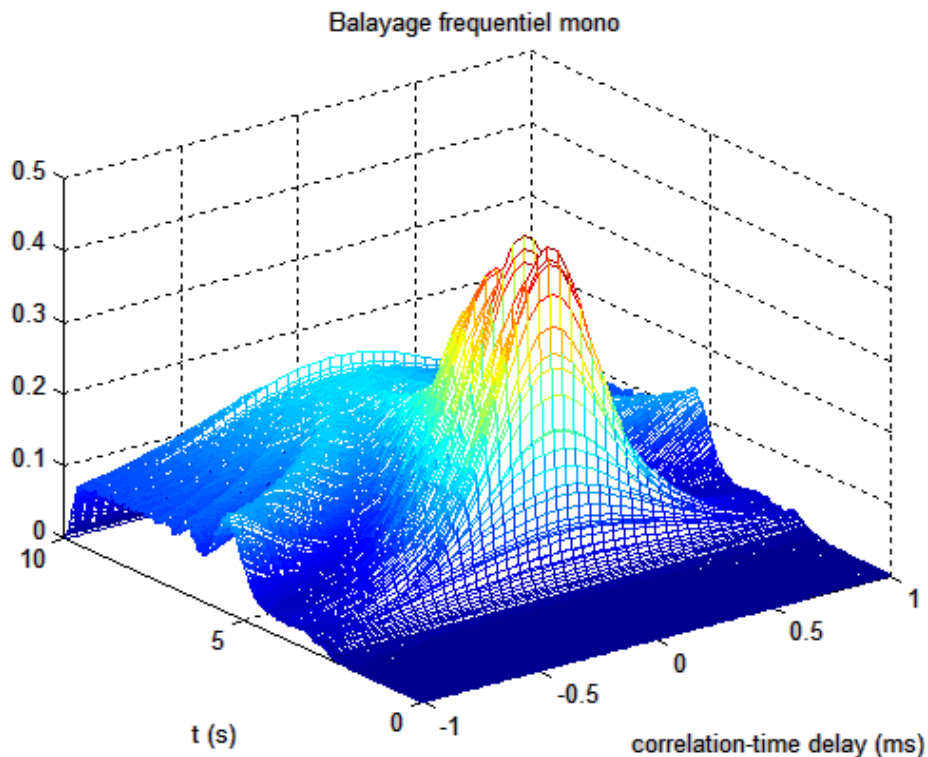


Figure 91 - Résultats pour un balayage fréquentiel mono de 10s

L'explication de ce phénomène de limitation en basses fréquences est donc à chercher dans le code.

Les pics d'inter-corrélation sont nettement plus larges et moins définis au-dessus de 1s sur l'axe de temps, soit pour des fréquences au-dessus de 650 Hz, et identiques que le signal soit en phase ou en opposition de phase. En allongeant la longueur du balayage binaural ce phénomène se produit plus haut en fréquences (900 Hz au lieu de 650 Hz).

EN conséquence, on peut s'attendre à ce que ce programme ne soit utile qu'entre 100 Hz et 650 Hz.

Maintenant on effectue un balayage fréquentiel corrélé mais dont le canal droit est retardé de 0.5 ms, le résultat du modèle est présenté dans la figure suivante.

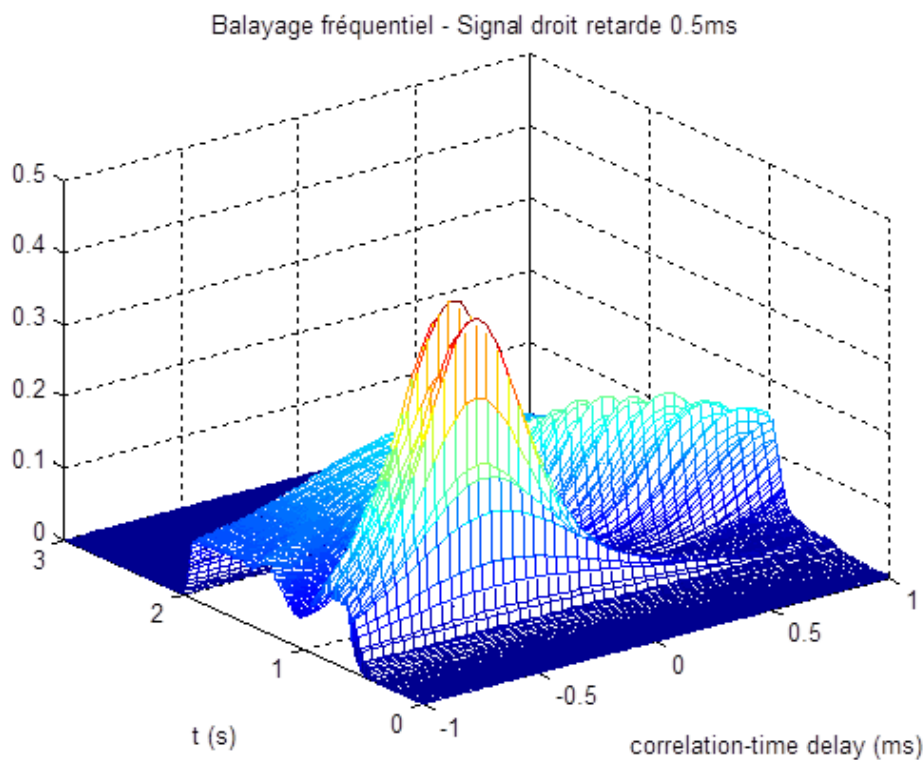


Figure 92 - Résultats pour un balayage fréquentiel avec un retard sur un des canaux

La figure obtenue est décalée vers la gauche et les pics au-dessous de 650 Hz sont tous alignés sur -0.5 ms. On s'attend donc à ce qu'une augmentation de la ponctualité perçue d'une source fantôme fasse apparaître dans cette mesure une série de pics alignés selon l'axe des décalages²⁵, et que plus la source est en face, plus les pics sont proches du décalage 0 ms.

3.3.6.2. Analyse de l'apport d'un réglage

Des enregistrements de réponse binaurale en véhicule sont faits avec un auditeur sur le siège du conducteur, et deux microphones affleurant ses conduits auditifs, dirigés vers l'avant. Les

²⁵ Son étiquette dans les figures est « correlation-time delay (ms) ».

figures suivantes consignent les résultats obtenus en appliquant ce modèle à la réponse à un balayage fréquentiel dans le véhicule D. On compare les résultats du modèle sans traitement audio et avec traitement.

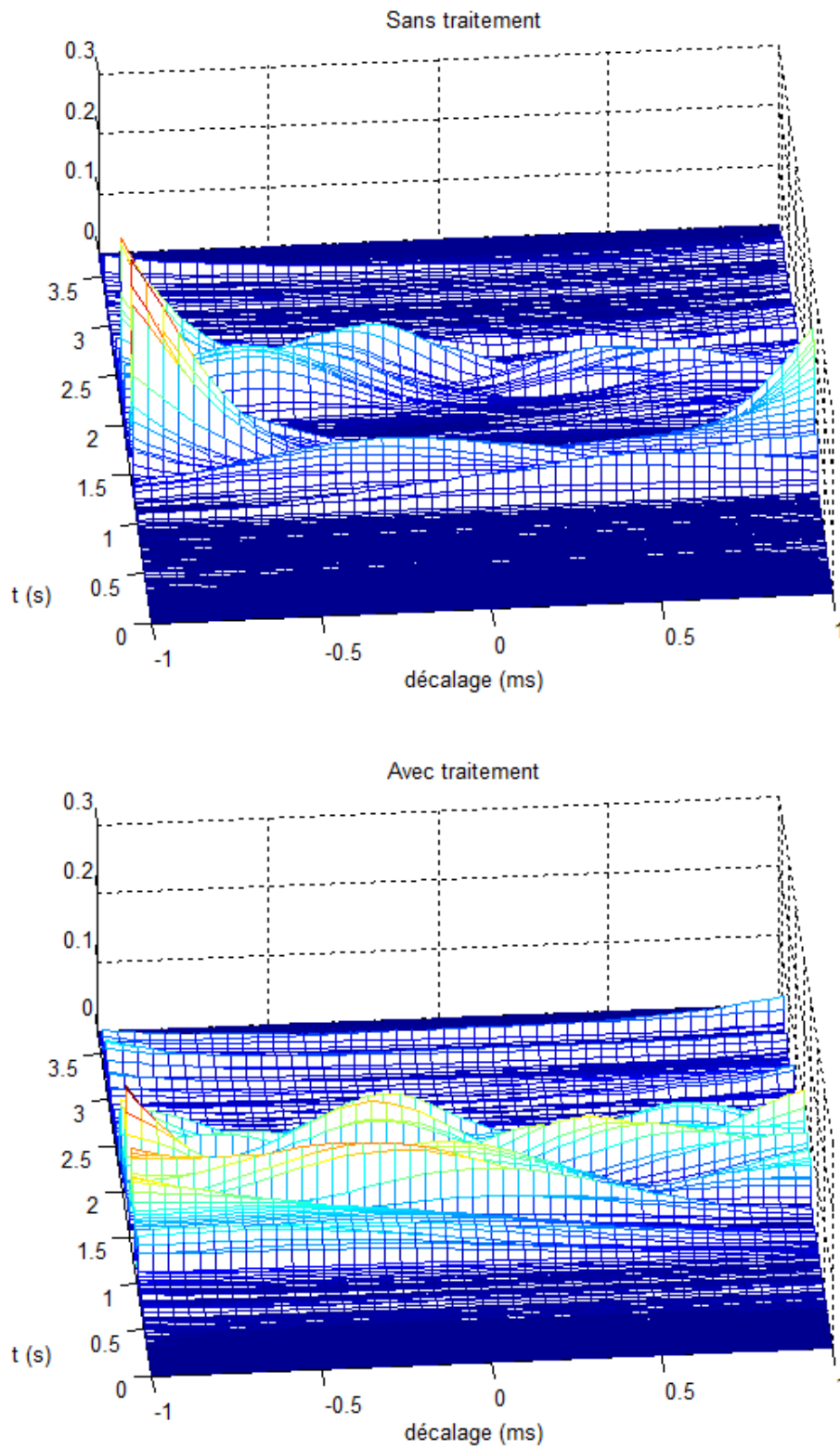


Figure 93 – Résultats du modèle de Lindemann dans le véhicule D

La vue sur le côté est instructive pour déterminer les fréquences où apparaissent les pics :

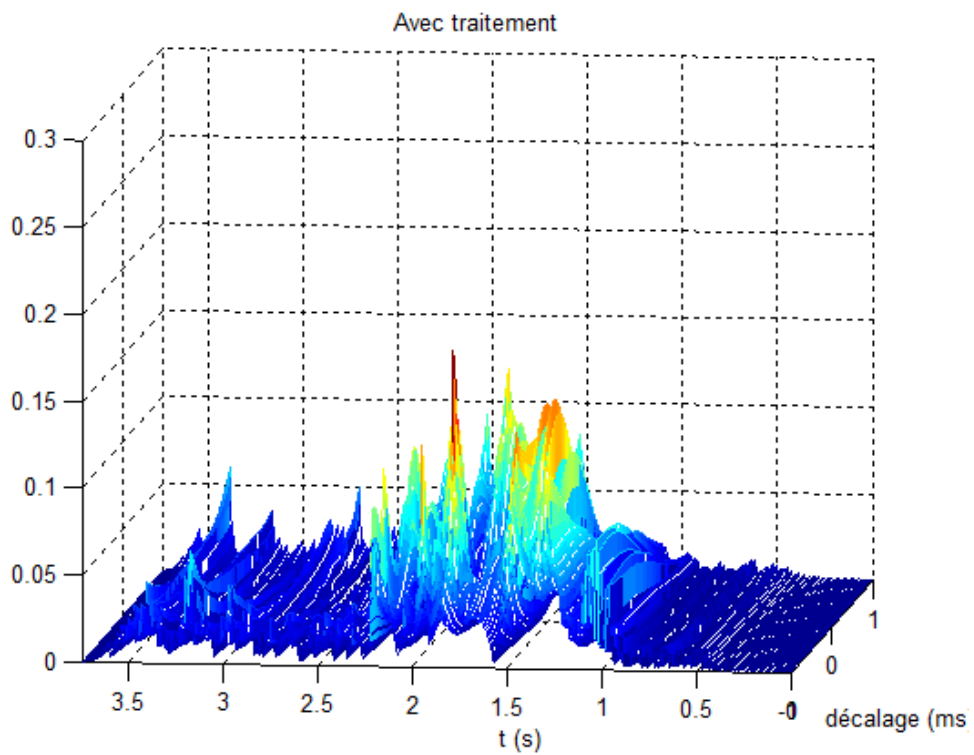
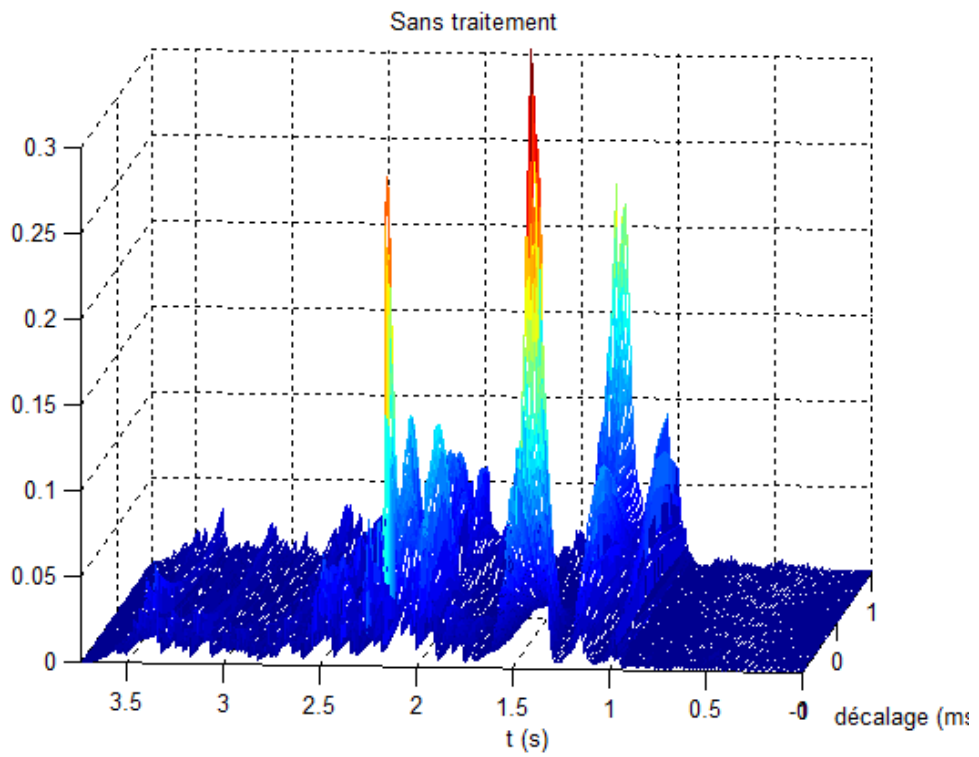


Figure 94 - Figure 93 sous un autre angle

Les pics latéraux à -1 ms de décalage sur la réponse sans traitement correspondent aux fréquences perçues latéralisées, principalement 253 Hz et 1 kHz. L'amélioration avec le traitement est très visible entre 200 Hz et 720 Hz : avec traitement, des pics apparaissent plutôt proches de 0 ms de décalage et les pics latéralisés sont fortement diminués. On compare ces résultats avec la phase binaurale (en degrés, axe vertical) en fonction de la fréquence (axe horizontal)²⁶, la partie du balayage fréquentiel correspondant à 200 Hz-750 Hz est en grisé²⁷ :

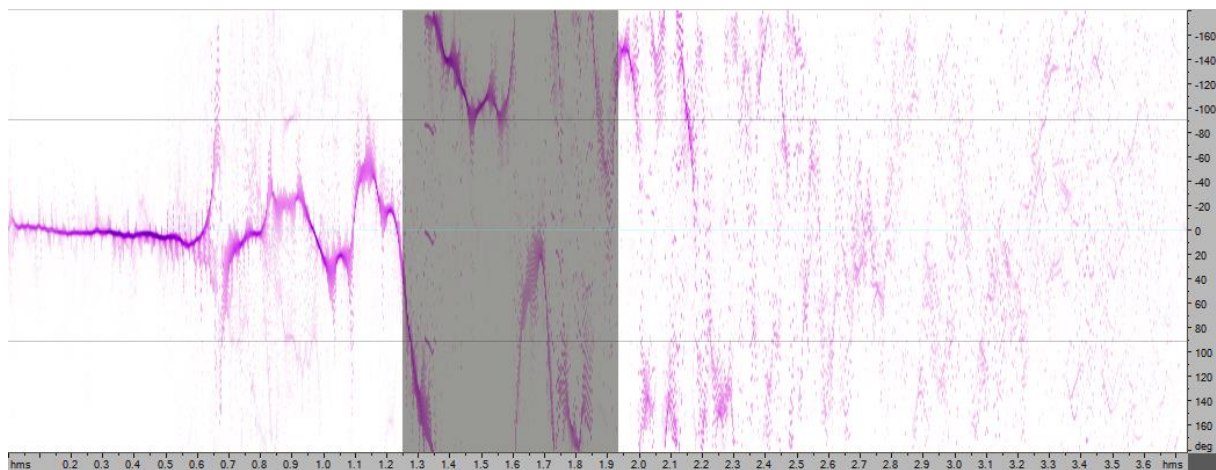


Figure 95 - Phase binaurale dans le véhicule D sans traitement

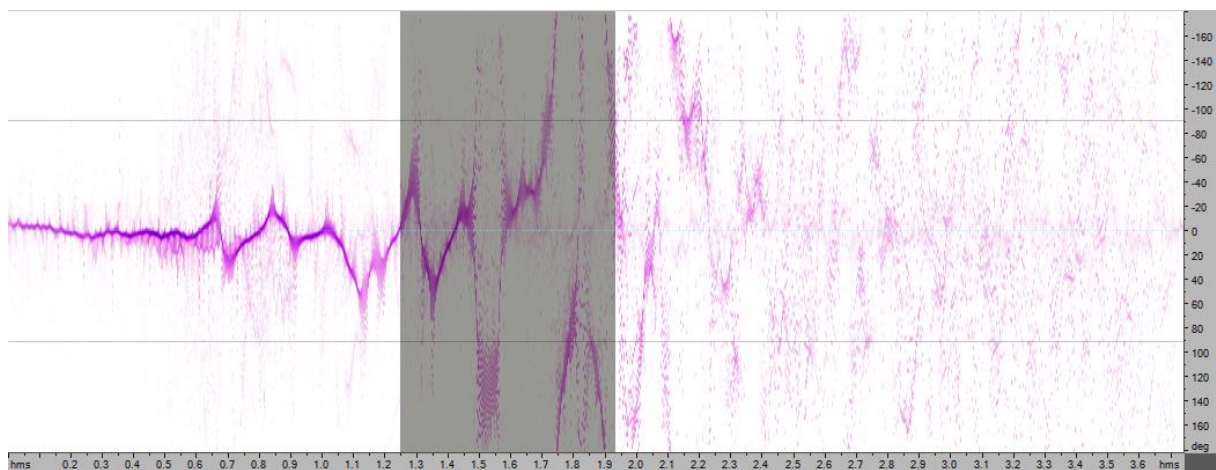


Figure 96 - Phase binaurale dans le véhicule D avec traitement

On constate que la phase binaurale était en opposition sur cette zone sans traitement, et que le traitement a permis de fortement la rapprocher de 0° globalement, ce qui correspond aux résultats d'écoute. On note aussi une nette opposition de phase vers 250 Hz, ce qui

²⁶ Il s'agit en fait d'un axe temporel, mais comme le signal test est un balayage fréquentiel cela est équivalent.

²⁷ Ces mesures sont semblables à la figure 69.

correspond à la fréquence d'un des pics latéralisés dans le motif d'activité binaurale obtenu avec le modèle de Lindemann.

Véhicule E : Dans cette cabine de pick-up à une seule rangée de sièges, où le filtrage est moins perfectionné que dans le cas du véhicule D vu précédemment, la figure suivante présente les résultats du modèle de Lindemann.

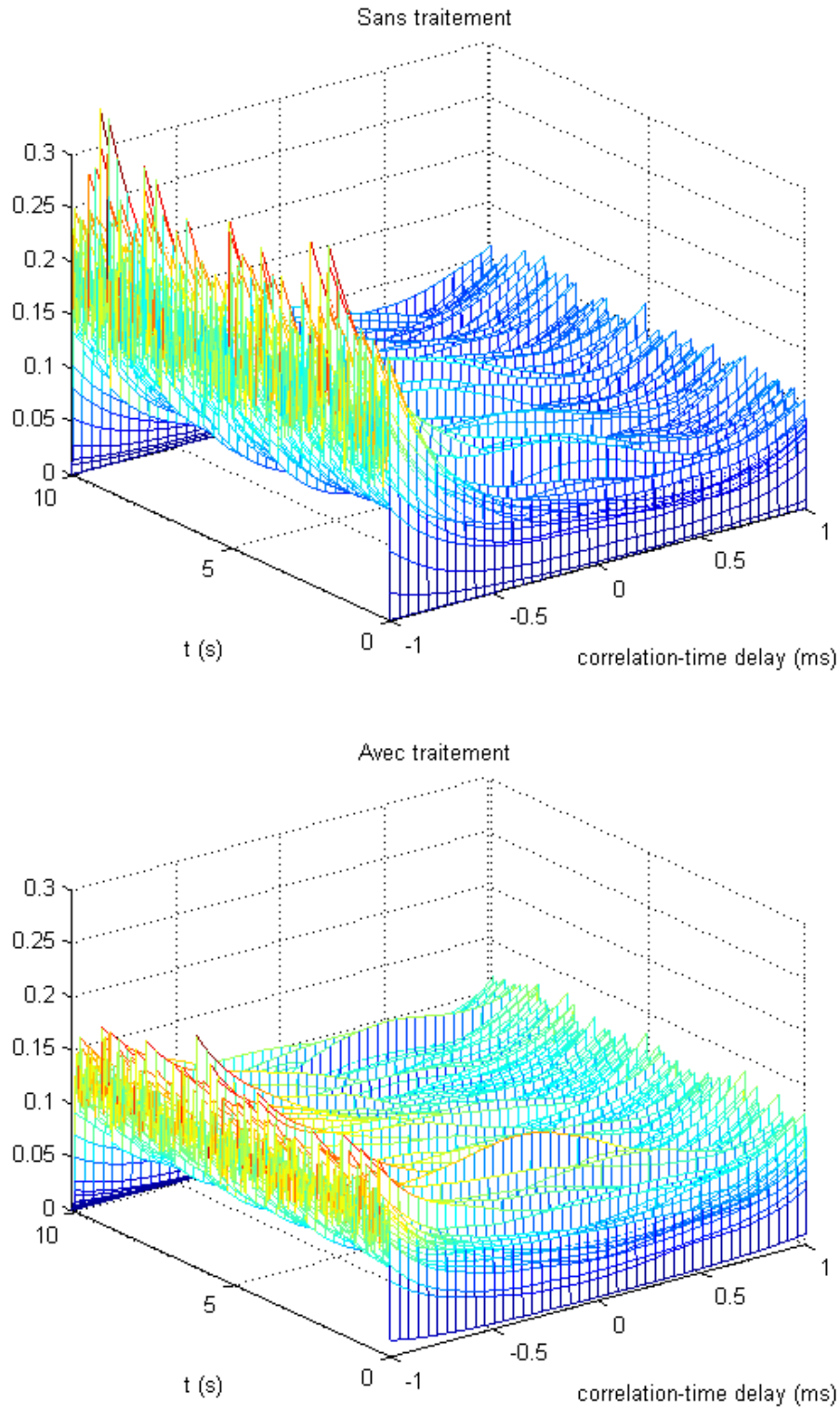


Figure 97 - Résultats du modèle de Lindemann dans le véhicule E

On constate grâce au traitement:

- une diminution de l'amplitude des pics latéraux à gauche (axe -1 ms), qui étaient très marqués sans traitement,
- un renforcement des pics plus centraux, mais beaucoup moins marqués que dans le véhicule précédent.

Véhicule F: Les résultats (fig. 97) sont encore plus nets que précédemment, les enregistrements sont faits avec une tête Schoeps KFM6 et des blocs sont utilisés pour créer une zone d'ombre comparable à un corps humain.

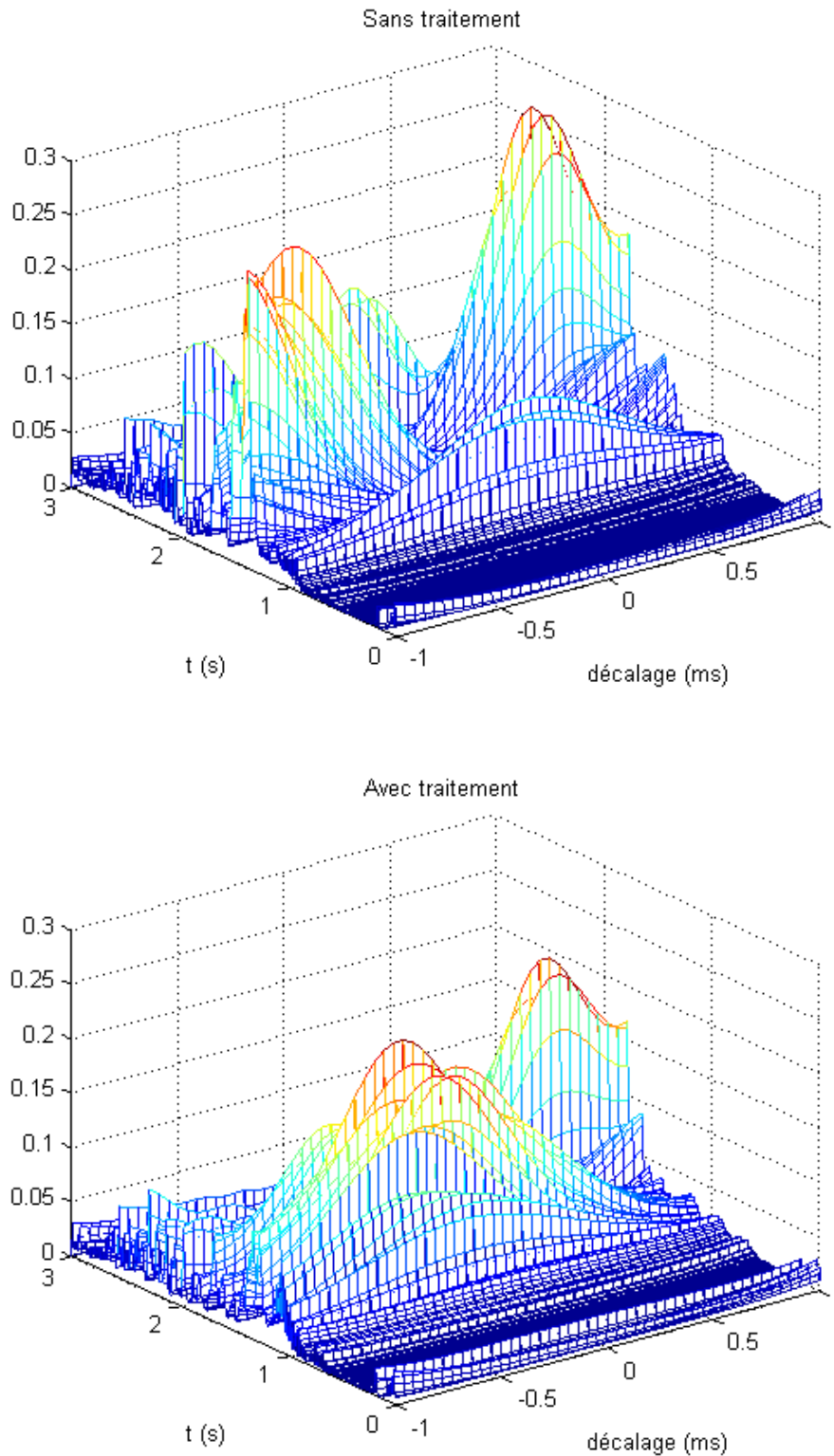


Figure 98 - Résultats du modèle de Lindemann dans le véhicule F

Vu sous d'autres angles :

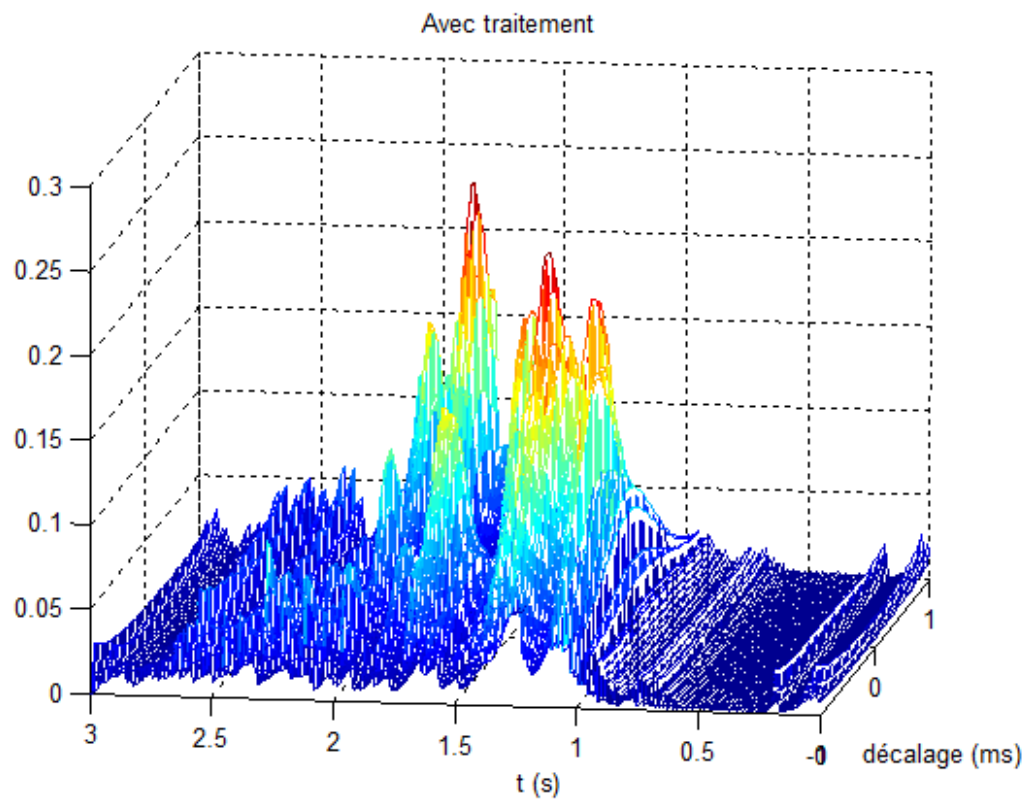
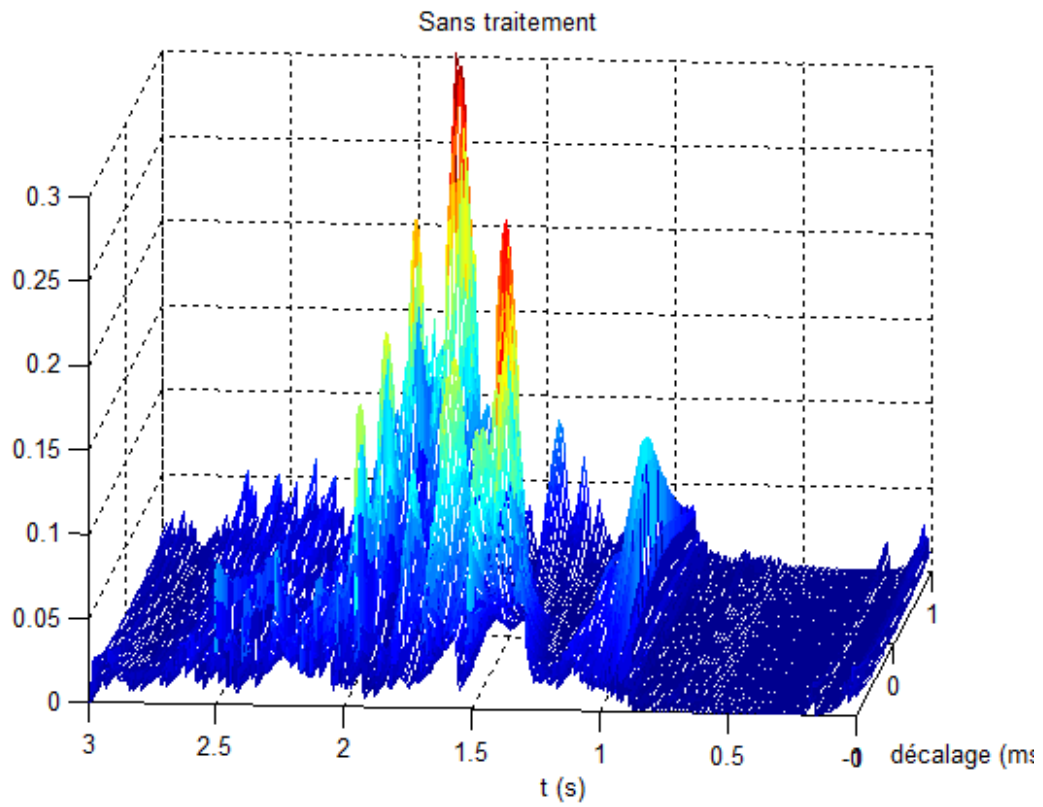


Figure 99 - Figure 98 sous un autre angle

Enfin vu sous un autre angle :

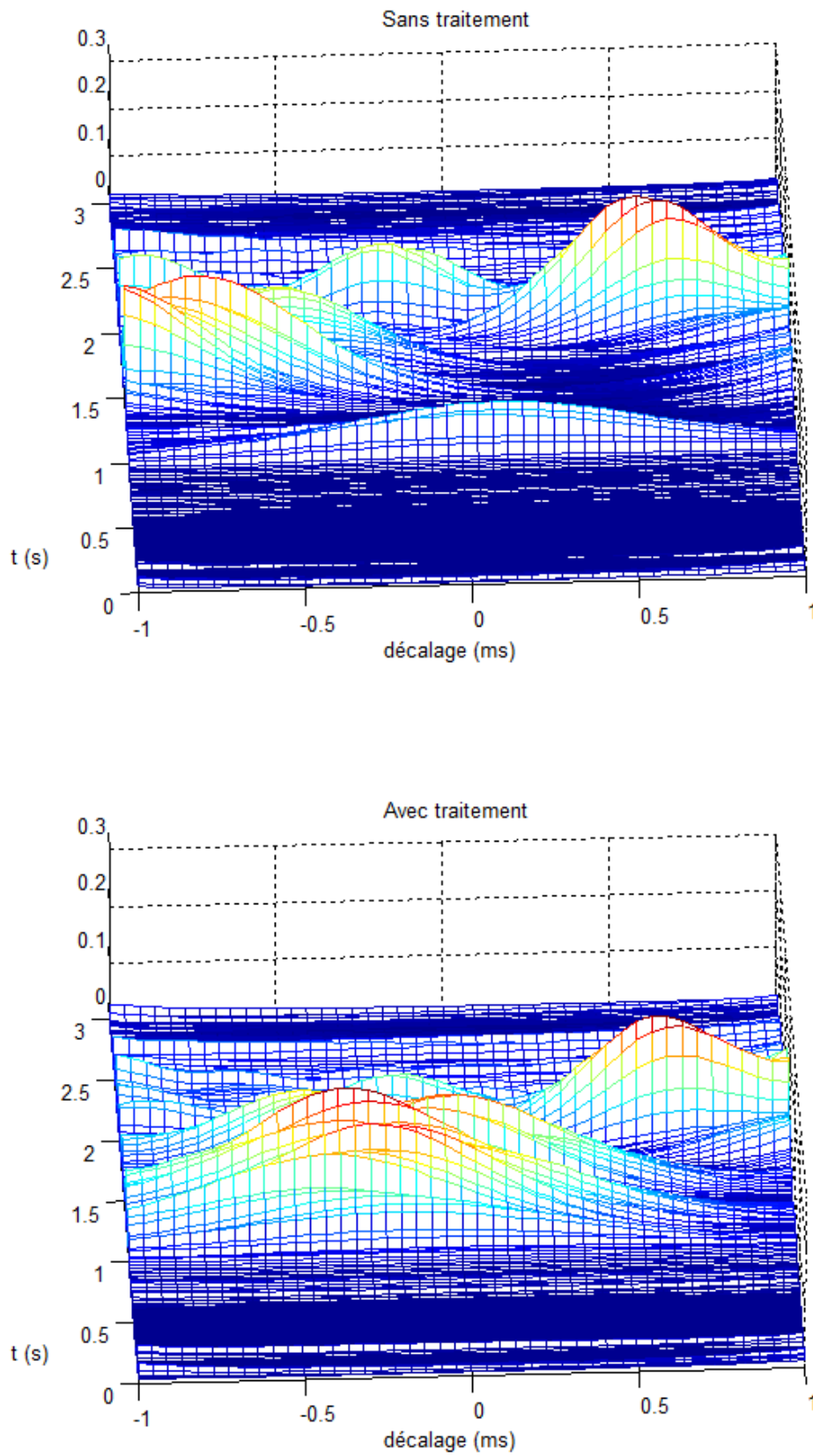


Figure 100 - Figure 98 sous un autre angle

L'amélioration est très nette pour des fréquences entre 130 Hz et 400 Hz (les pics latéralisés à -1ms et +1ms disparaissent et sont remplacés par un pic très net pour $\tau = -0,25$ ms, ce qui est cohérent avec le calcul de FIACC fait sur cette configuration. Par contre, avec le réglage, le pic à 830 Hz vers +0,6 ms de décalage n'est pas recentré vers 0 ms de décalage mais l'amplitude du pic est diminuée de 0,3 à 0,22. Il faut tempérer ces résultats enthousiasmants car le balayage fréquentiel utilisé pour ces mesures n'a pas de transitoires et il apparaît peu optimal pour une mesure proche de la perception (voir [4] p273). Il faudrait refaire ces calculs sur les réponses binaurales à un bruit rose corrélé ou une voix, et pouvoir substituer une échelle fréquentielle à l'échelle temporelle, ce qui impose de modifier le programme de calcul.

3.3.7. Conclusion partielle

Après étude de résultats dans la littérature sur ce sujet, l'amélioration de la fidélité spatiale suite au réglage d'un traitement du signal audio a été analysée avec plusieurs modèles et visualisations, et des conclusions sur leur efficacité ont été esquissées en étudiant la reproduction d'un signal totalement corrélé dans quelques voitures. Une comparaison statistiquement valable de ces modèles reste à établir dans ce contexte.

Si on se contente d'un modèle de type Duplex, la visualisation simultanée d'une figure de Lissajous des signaux binauraux, de la phase binaurale en fonction de la fréquence et de la corrélation binaurale apparaissent comme des indices simples et efficaces d'aide au réglage, exploitable pour en déduire une égalisation adaptée. Ces mesures ont l'énorme avantage d'être réalisables quasiment en temps réel.

Le calcul de la FIACC donne des résultats intéressants mais parfois contradictoires. Il faudrait refaire ces mesures en faisant tourner le système de mesure, car certains auteurs recommandent de calculer l'IACC et la fraction latérale pour plusieurs positions angulaires du système microphonique, ce qui pose des problèmes pratiques. Ne disposant ni d'une table tournante ni de microphones bi-directionnels ou ambisoniques ni d'un automate de mesure, les mesures de fraction latérale et de diffusivité ont été exclues de cette étude, et les mesures de corrélation binaurale ont été faites sur un seul axe. Il reste fortement souhaitable de tester

cette façon de mesurer dans un habitacle automobile, et de comparer avec les résultats obtenus en utilisant un système fixe.

Pour améliorer la prédiction de perception spatiale dans les milieux réverbérant, des modèles plus complexes, analysant les différences de phase et d'intensité et incorporant des modèles conditionnels, des seuils d'inter corrélation ou des filtres ont été testés. La boîte à outils de Faller et Merimaa donne des résultats encourageants, mais nécessite des améliorations pour être vraiment exploitable. Si la mesure en temps réel n'est pas nécessaire, le modèle de Jeffress-Lindemann apparaît comme le plus utile de cette catégorie. Cependant, il est fortement souhaitable de poursuivre l'expérimentation avec ce modèle en diffusant un signal de type bruit rose et de comparer les résultats obtenus avec ceux obtenus en diffusant un balayage fréquentiel, et aussi de lui appliquer les améliorations de Gaik pour améliorer sa précision.

Aucun des modèles disponibles ne donnaient d'information sur l'azimut des sources virtuelles. Pour ce faire il faudrait leur ajouter un calcul supplémentaire, basé sur les relations et/ou les bases de données présentées au chapitre 1. Les résultats obtenus sont prometteurs, certains exploitables, mais ils restent très éloignés d'une cartographie de scène sonore comparable à celle faite par le système nerveux d'un auditeur humain.

4. Conclusion

La sonorisation des véhicules automobiles a une fidélité limitée car les habitacles et les systèmes de sonorisation automobile ne correspondant pas aux recommandations pour la reproduction des enregistrements stéréophoniques. Dans une voiture avec un système audio d'entrée de gamme, la majorité des auditeurs ne sont pas assis sur l'axe médian entre les haut-parleurs, ce qui a un impact très négatif sur leur perception de la scène sonore reproduite. Des contraintes limitent la qualité de restitution des haut-parleurs, leur nombre et leur placement font que la recombinaison des signaux qu'ils émettent est mal maîtrisée. Une analyse succincte de l'acoustique de l'habitacle automobile a montré des différences notables avec les lieux d'écoute habituels de voix et de musique. Cette analyse mérite d'être approfondie pour mieux comprendre l'influence de l'habitacle sur la propagation du son et sur la combinaison des ondes sonores issues des haut-parleurs.

Afin d'optimiser la durée des réglages, en particulier ceux de filtrages à réponse impulsionnelle infinie, des techniques de mesure et des modèles d'audition spatiales ont été étudiés. A défaut de donner une cartographie de scène sonore comme on espérait au départ, la phase binaurale, la corrélation binaurale et les diagrammes d'activité binaurale de Jeffress-Lindemann ont changé de façon plutôt prévisible, explicable et analysable selon qu'un traitement du signal audio idoine était appliqué ou pas. Utiliser ces modèles apparaît pertinent dans l'acoustique d'un habitacle d'automobile et dans un contexte de reproduction stéréophonique. Cela renforce l'intérêt à employer ce type de modèle dans le réglage et la comparaison de systèmes audio. Les performances des différents modèles ont été comparées succinctement, ce qui a donné des pistes sur ceux qui sont les mieux adaptés à l'usage précité, mais une étude plus approfondie est nécessaire.

Afin d'aider à trouver l'alignement temporel qui répond le mieux au cahier des charges, sur la base des outils existants ou développés dans le cadre de cette étude, il faut développer un logiciel unique qui permette à partir d'enregistrements binauraux fait sur une ou plusieurs positions d'écoutes de calculer:

- la phase binaurale, la corrélation binaurale et le diagramme de Lissajous, en « temps réel » lors des réglages et moyennés dans le temps pour les rapports,

- les figures issues du modèle de Jeffress-Lindeman,
- la FIACC.

Pour obtenir des mesures de l'IACC et de la fraction latérale pertinentes, la littérature montre qu'il faut trouver un moyen de faire ces mesures sur plusieurs positions angulaires. Un système pour faire cela dans une automobile reste à trouver. On envisage l'utilisation de simulateurs tête-torse ou de mannequins complets munis d'un système de captation binaural, « assis » aux autres places du véhicule avec l'opérateur de mesures, ou de microphones ambisoniques.

Ces résultats doivent être vérifiés sur plus de véhicules et avec plus d'auditeurs afin de s'assurer de leur validité, ce qui permettrait également d'améliorer la compréhension de l'audition spatiale lors de la diffusion sonore dans des très petits locaux. La boîte à outil de Fallert et Merimaa mérite d'être calibrée puis testée à nouveau. Les améliorations que Gaik a apportées au modèle de Lindemann méritent d'être implémentées et testées, en particulier des tests de plausibilité entre les indices binauraux. Afin d'obtenir une scène sonore précise et régulière, il apparaît nécessaire dans le futur de tester toutes ces mesures en lisant des signaux de différents niveaux de corrélation entre les canaux.

Pour l'instant ces méthodes ne donnent ni la cartographie perceptive espérée, ni directement les coefficients de réglage. Dans l'optique d'un réglage automatique, on propose plutôt d'évaluer la pertinence d'une inversion de la fonction de transfert, suivie d'une approximation du filtrage inverse par un filtrage RII comme décrit dans [64], puis comparer les résultats par des écoutes critiques et par les modèles d'audition.

Annexes

Annexe A – Simulation d'écoute hors axe (code Matlab « ecoutehorsaxe.m »)

```
clc
close all
clear all
fs=100000;

i=1:fs;
t=i/fs;
f=i/2;

% delais
d1=1.18;
d2=1.23;
d3=1.44;
d4=1.56;
t1=d1/340*fs;
t2=d2/340*fs;
t3=d3/340*fs;
t4=d4/340*fs;

%generation des signaux
a1=[zeros(1,t1),1,zeros(1,t4-t1),0.75,zeros(1,fs-t4-1)];
a2=[zeros(1,t2),0.96,zeros(1,t3-t2),0.82,zeros(1,fs-t3)];

%fft puis calcul phase(f) wrapped et unwrapped
d1=unwrap(angle(fft(a2)./fft(a1)));
e1=(angle(fft(a2)./fft(a1)));

%plot amplitude(t)
subplot(2,1,1)
plot(t,a1,'g');
hold on
plot(t,a2,'r');
title('signal aux capteurs')
xlabel('temps[s]')
ylabel('amplitude')
axis([0;0.01;0;1]);

%phase déroulée
subplot(2,1,2)
plot(f*2,d1,'LineWidth',1.5);
hold on
title('phase relative dépliée(f)')
xlabel('frequence [Hz]')
ylabel('phase[rad]')
axis([20;2000;-5;4]);
grid on
```

```

[x, lag]=xcorr(a2, a1);
figure
plot(lag, x)
xlabel('lag [ms]')
ylabel('IACC')
axis([-200 200 0 1])

```

Annexe B - Calcul de FIACC (code Matlab « FIACC.m »)

```

% This program is based on the work by T. Fajersson H. Wikner
% It computes, compares and plots the FIACC of two binaural acoustic
% responses to a sweep sine
% Alterations to the original code:
%     Removed unused computation
%     Extended the frequency range of the filter bank
%     Replaced stems in frequency bands with plot with %semilog frequency axis
% Alexandre FENIERES

close all

%First binaural response
[signal,fs]=wavread('ALLCAR.wav');%first .wav file to analyze
EL=signal(:,1);
ER=signal(:,2);

%Number of third—octave bands analysed
freq_vector = [25 31.5 40 50 63 80 100 125 160 200 250 315 400 500 630 800 1000 1250 1600 2000
2500 3150 4000 5000 6300 8000 10000 12500 16000];

%Create FIACC and reference vectors
fiacc = zeros(length(freq_vector),1);
sin_kr = fiacc;
one_vec = ones(length(freq_vector),1);
c = 344; % Speed of sound
r = 0.328; % Acoustical distance between the ears

for(n=1:length(freq_vector))
    fc=freq_vector(n);
    N=3;
    [B,A] = oct3filter(freq_vector(n),fs,3);
    filt_signal_L = filter(B,A,signal(:,1));
    filt_signal_R = filter(B,A,signal(:,2));
    filt_signal = [filt_signal_L;filt_signal_R];
    fiacc(n) = mean(filt_signal_L.*filt_signal_R) / sqrt(mean(filt_signal_L.^2) * mean(filt_signal_R.^2));
    k = (2*pi*freq_vector(n))/c;
    sin_kr(n) : sin(k*r)/(k*r);
end

% Choose a reference (sin(kr)/(kr) or oneivec):
ref = one_vec;

```

```

%ref : sinikr;

%Second binaural response
[signal,fs]=wavread('BYPASS.wav');%second.wav file to analyze
EL=signal(:,1);
ER=signal(:,2);

%Create FIACC2
fiacc2 = zeros(length(freq_vector),1);
one_vec = ones(length(freq_vector),1);

for(n=1:length(freq_vector))
    fc=freq_vector(n);
    N=3;
    [B,A] = oct3filter(freq_vector(n),fs,3);
    filt_signal_L = filter(B,A,signal(:,1));
    filt_signal_R = filter(B,A,signal(:,2));
    filt_signal = [filt_signal_L;filt_signal_R];
    fiacc2(n) = mean(filt_signal_L.*filt_signal_R) / sqrt(mean(filt_signal_L.^2) * mean(filt_signal_R.^2));
    k = (2*pi*freq_vector(n))/c;
    sin_kr(n) : sin(k*r)/(k*r);
end

%Comparison of FIACC
figure
semilogx(freq_vector,fiacc,'g')
hold on
semilogx(freq_vector,fiacc2,'r')
grid on
title('FIACC 1/3oct');
xlabel('Frequence( Hz)');
ylabel('IACC')
legend('Avec réglage','Sans réglage')
axis([20 2e4 -1 1])

```

Annexe C - Calculs de clarté C50 et C80 (code Matlab « clarte.m »)

```

close all;
clear all;
[a,fs]=wavread('ALLCAR.wav');
%[a,fs]=wavread('ALLB2.wav');
%[a,fs]=wavread('PRIUSALL.wav');
t=1:length(a)/(2);

a1=a(:,1);
a2=a(:,2);

%Calcul réponse impulsionnelle abs
A1=(abs(iff(a1))).^2;
Am1=max(A1);
A2=(abs(iff(a2))).^2;

```

```

Am2=max(A2);

C1=A1(t)/Am1;
C2=A2(t)/Am2;

plot(t/fs,C1,'b')
hold on
plot(t/fs,C2,'r')
axis([0 0.2 0 1])
x=0.05*ones(1000,1);
w=0.08*ones(1000,1);
y=[0:1:999];
plot(x,y,'g--')
plot(w,y,'m-.')
title('allcar')

disp('*****')
disp('*   ALLCAR   *')
disp('*****')

%Calcul clarté
u=1:0.05*44100;
l=length(u);
v=l:length(a)/(2);
C50left=10*log(sum(C1(u))/sum(C1(v)))
C50right=10*log(sum(C2(u))/sum(C2(v)))

u=1:0.08*44100;
l=length(u);
%v=l:length(a)/(2);
v=l:44100;
C80left=10*log(sum(C1(u))/sum(C1(v)))
C80right=10*log(sum(C2(u))/sum(C2(v)))

[a,fs]=wavread('BYPASS.wav');
%[a,fs]=wavread('BYPB2.wav');
%[a,fs]=wavread('PRIUSBYP.wav');
%[a1,fs]=wavread('P32RG.wav');
%[a2,fs]=wavread('P32RD.wav');
t=1:length(a)/(2);

t=1:length(a)/(2);

a1=a(:,1);
a2=a(:,2);

%Calcul réponse impulsionnelle abs
A1=(abs(iff(a1))).^2;
Am1=max(A1);

```

```

A2=(abs(iff(a2))).^2;
Am2=max(A2);

C1=A1(t)/Am1;
C2=A2(t)/Am2;

figure
plot(t/fs,C1,'b')
hold on
plot(t/fs,C2,'r')
axis([0 0.2 0 1])
x=0.05*ones(1000,1);
w=0.08*ones(1000,1);
y=[0:1:999];
plot(x,y,'g--')
plot(w,y,'m-.')
title('bypass')

disp('*****')
disp('*   BYPASS   *')
disp('*****')

%Calcul clarté
u=1:0.05*44100;
l=length(u);
v=l:length(a)/(2);
C50left=10*log(sum(C1(u))/sum(C1(v)))
C50right=10*log(sum(C2(u))/sum(C2(v)))

u=1:0.08*44100;
l=length(u);
%v=l:length(a)/(2);
v=l:44100;
C80left=10*log(sum(C1(u))/sum(C1(v)))
C80right=10*log(sum(C2(u))/sum(C2(v)))

```

Annexe D - Comparaison entre écoutes critiques et résultats de la boîte à outils de Faller et Merimaa appliquée aux mesures dans le véhicule A

Les résultats des mesures d'ITD étaient :

- l'inversion de phase entre haut-parleurs entre 200 et 500 Hz améliore la cohérence entre 500 et 600 Hz
- l'inversion de phase entre haut-parleurs entre 500 et 2000 Hz améliore la « déviation » d'ITD entre 1kHz et 1,1kHz

Des écoutes sont menées dans le véhicule afin de comparer les résultats du modèle de localisation à la perception d'un très petit panel de deux auditeurs formés à l'écoute critique.

Les deux auditeurs échangent régulièrement de places entre la place du conducteur et la place du passager avant.

Tableau viii - Résultats d'écoute avec et sans filtrage, le 16/05/2013.

Signal	Auditeur 1 siège avant gauche	Auditeur 2 siège avant gauche	Auditeur 1 siège avant droit	Auditeur 2 siège avant droit
Sinus 200 Hz	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage
Sinus 300 Hz	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage
Sinus 400 Hz	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage	Meilleur, plus en face avec filtrage
Sinus 450 Hz	Meilleur avec filtrage			Sans Filtrage latéralisé à droite, avec filtrage en face
Sinus 500 Hz	Meilleur avec filtrage	Aucun changement	Gros changement	Sans filtrage en face, avec Filtrage totalement à gauche et hors phase
Sinus 550 Hz		Amélioration avec filtrage mais niveau moins fort	Amélioration avec filtrage	Peu de differences
Sinus 600 Hz		Sans Filtrage tout dans l'oreille gauche, avec filtrage en face	Sans filtrage tout à gauche, avec filtrage un peu à droite dur à localiser	Pire avec filtrage, part à gauche, bien que sans filtrage le corrélé semble venir de derrière hors phase

Bruit filtré Bessel 550 Hz ordre 6	Plus en face et cohérent avec filtrage	Plus en face et cohérent avec filtrage	Plus en face et cohérent avec filtrage	Plus en face et cohérent avec filtrage
------------------------------------	--	--	--	--

Tableau ix - Résultats d'écoute avec et sans filtrage, le 16/05/2013.

Signal	Auditeur 1 siège avant gauche	Auditeur 2 siège avant droit
Sinus 300 Hz	Meilleur avec Filtrage : curieux car flou sur les figures avec ou sans Filtrage	Meilleur avec Filtrage : curieux car flou sur les figures avec ou sans Filtrage
Sinus 400 Hz	Idem 300 Hz mais un peu moins marqué	Idem 300 Hz mais un peu moins marqué
Sinus 500 Hz	Pire avec Filtrage puis 2 ^{ème} essai : devant mais légèrement hors phase	Pire avec Filtrage puis 2 ^{ème} essai : devant mais peu compact
Sinus 550 Hz	Un peu mieux avec Filtrage mais sensation hors phase	Meilleur avec Filtrage
Sinus >600 Hz	Sans Filtrage, la latéralisation augmente avec la fréquence	Sans Filtrage, la latéralisation augmente avec la fréquence
Sinus 620 Hz	Plus précis avec Filtrage mais légèrement latéralisé	Plus précis avec Filtrage mais légèrement latéralisé
Sinus 700 Hz	Meilleur en face avec Filtrage	La phase change beaucoup avec la position d'écoute longitudinale Peu de différences avec/sans Filtrage

Ces résultats sont donnés en annexe pour montrer le type de panel d'écoute disponibles lors de la réalisation de ce document.

Annexe E - Références des véhicules utilisés, des techniques de mesures et des fichiers de mesure résultants

Véhicule A : Mercedes Class A, 2 microphones Apex seuls

Véhicule B : GM K2RC UQ5

Véhicule C : GM A1AC Camaro

Véhicule D : DS5 (fichiers BYPASS.wav ; ALLCAR.wav ; BYPASSSHORT.wav ; ALLCARSHORT.wav)

Véhicule E : GM K2RC UQ5 (fichiers BYPB2.wav ; ALLB2.wav)

Véhicule F : Prius, tête Schoeps (fichiers PRIUSBYP.wav ; PRIUSALL.wav)

Véhicule G : Nissan Qashqai 2014

Sauf mention contraire, les mesures ont été faites avec 2 microphones Apex appairés, placés aux oreilles de l'auditeur

Bibliographie

- [1] ITU-R, 1997, *Recommendation ITU-R BS.1116-1 - Method for the subjective assessment of small impairments in audio systems including multichannel sound systems.*

http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1116-1-199710-S!!PDF-E.pdf

- [2] ITU-R, 2001, *Recommandation ITU-R BS. 1387 2001 - Method for objective measurements of perceived audio quality.*

<http://www.itu.int/rec/R-REC-BS.1387>

- [3] JAMES B., AMY P., 2014, 2014, Achilles' Ear? Inferior Human Short-Term and Recognition Memory in the Auditory Modality, *PLoS ONE*, vol. 9, issue 2 e89914.

- [4] BLAUERT J., 1977, *Spatial Hearing - The psychophysics of Human Sound Localization*, MIT Press.

- [5] EVANGELISTA G., MARCHAND S., PLUMBLEY M. D., VINCENT E., 2011, Sound Source separation, *Udo Zolzer DAFX: Digital Audio Effects 2nd Edition*, Wiley-Blackwell, pp. 551-586.

- [6] HANNETON S., Fiche ITD, http://sylvain.hanneton.free.fr/fiches/fiche_ITD.pdf. [Accès le 20 01 2014].

- [7] AARONSON N. L., HARTMANN W. M., 2013, Testing, correcting, and extending the Woodworth model for interaural time difference, *Journal of the Acoustical Society of America*, vol. 135, n°12, pp. 817-823.

- [8] LARCHER V., 2001, *Technique de spatialisation des sons pour la réalité virtuelle*, Thèse de doctorat de l'Université de Paris VI.

- [9] HUMMERSONE C., 2011, *A Psychoacoustic Engineering Approach to Machine Sound Source Separation in Reverberant Environments*, Thèse de doctorat, Department of Music & Sound Recording, Faculty of Arts & Human Sciences, University of Surrey.

- [10] ZHANG P. X., HARTMANN W. M., 2006, Lateralization of sine tones—interaural time vs phase, *Journal of the Acoustical Society of America*, vol. 120, n°16, p. 3471–3474.
- [11] MACAULAY E. J., HARTMANN W. M., 2010, The acoustical bright spot and mislocalization of tones by human listeners, *J Acoust Soc Am.*, vol. 127, n°13, p. 1440–1449.
- [12] CAMPBELL D. R., Effect of Ear Separation and the Human Head, Localization, HRTF
<http://media.paisley.ac.uk/~campbell/AASP/Human%20Head.PDF> [Accès le 01 Décembre 2014].
- [13] OAG H. J., 2002, «Localization of phantom sources derived from contradicting signal differences». [En ligne] [Accès le 01 Décembre 2014].
http://www.hauptmikrofon.de/HW/Oag_ContradictingSD.pdf.
- [14] SHINN-CUNNINGHAM B. G., SANTARELLI S., KOPCO N., 2000, Tori of confusion: Binaural localization cues for sources within reach of a listener, *J. Acoust. Soc. Am.*, vol. 107, n° 13, pp. 1627-1636.
- [15] TOOLE F. E., 2008, *Sound Reproduction - Loudspeakers and rooms*, Focal Press.
- [16] MORIMOTO M., 2002, «The relation between spatial impression and the precedence effect», présenté à l'*International Conference on Auditory Display*, Kyoto, Japon.
- [17] SOULODRE G. A., 2006, Can Reproduced Sound be Evaluated using Measures Designed for Concert Halls?, *Spatial audio & sensory evaluation techniques*, Université de Surrey, Royaume Uni.
- [18] RUMSEY F., 2001, *Spatial Audio*, Oxford, Focal Press, pp. 36-39.
- [19] DICK D. A., VIGEANT M. C., 2014, A comparison of late energy (GLL) and lateral fraction (LF) measurements using a spherical microphone array and conventional methods, *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany.
- [20] WALTHER A., 2013, «Perception and Reproduction of Auditory Spatial Impression», Thèse

de l'École polytechnique fédérale de Lausanne EPFL, n° 5913

- [21] JOHANNES K., MARTON M., BASTIAN E., TORSTEN D., 2013, The relation between perceived apparent source width and interaural cross-correlation in sound reproduction spaces with low reverberation, *Deutsche Gesellschaft für Akustik*.
- [22] MASON R. 2002, «Elicitation and measurement of auditory spatial attributes in reproduced sound», thèse de doctorat, Université de Surrey, School of Performing Arts, Royaume Uni.
- [23] KENDALL G., 1995, «The Decorrelation of Audio Signals and Its Impact on Spatial Imagery», *Computer Music Journal*, vol. 19, n°14, pp. 71-87.
- [24] WANG D., BROWN G. J., 2006, *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, Wiley-Blackwell.
- [25] MURAOKA T., NAKAZATO T., 2007, Examination of Multichannel Sound-Field Recomposition Utilizing Frequency-Dependent Interaural Cross Correlation (FIACC), *J. Audio Eng. Soc.*, vol. 55, n°14, pp. 236-256.
- [26] FAJERSSON T., WIKNER H., 2008, *Binaural Measures of Spatial Properties of Audio reproduced in Cars*, Thèse de Master, Lulea University of Technology, Lulea.
- [27] BOYNTON G., 2008, Sound Localization and the Auditory Scene. [En ligne]
http://courses.washington.edu/psy333/lecture_pdfs/Week9_Day2.pdf
- [28] MERIMAA J., FALLER C., 2004, «Source localization in complex listening situations: Selection of binaural cues based on interaural coherence», *Journal of the Acoustical Society of America*, vol. 116, n°15, p. 3075–3089.
- [29] BRAASCH J., 2005, Modelling of binaural hearing, *Communication Acoustics*, edited by J. Blauert, Springer, Berlin, Allemagne, p. 75–108.
- [30] RUMSEY F., 2002, Spatial quality evaluation for reproduced sound: terminology, meaning, and a scene-based paradigm, *Journal of the Audio Engineering Society*, vol. 50, n°19, pp.

651-666.

- [31] USHER J., WOSZCYK W., 2003, Design and testing of a graphical mapping tool for analyzing spatial audio scenes, *AES 24th International Conference on Multichannel Audio*, Banff.
- [32] MAY T., 2012, Binaural Scene Analysis - Localization, Detection and Recognition of Speakers in Complex Acoustic Scene, Thèse de doctorat, Presses Universitaires de la Technische Universiteit Eindhoven.
- [33] ITU-R, 2012, *Recommendation ITU-R BS.775-3 - Multichannel stereophonic sound system with and without accompanying picture*, Genève.
- [34] KLEINER M., LIDGREN C., 1998, Objective Characterization of Audio Sound Fields in Automotive Spaces, *Proceedings of the AES 15th International Conference*.
- [35] PAIK S., DE GEEST M., VANSANT K., 2013, Interior Acoustic Simulation for In Car Audio Design, *Sound and Vibration*.
- [36] GEDDES E., 2010, Directivity in loudspeaker systems.
<http://www.gedlee.com/downloads/directivity.pdf>. [Accès le 18 08 2014].
- [37] BENNETTS A., MORRIS-KIRBY R., 2004, RT60 : its use as an optimiser in automotive cabins.
<http://www.baysystems.ltd.uk/documents/1/RT60%20and%20SPTF%20Its%20use%20as%20an%20optimiser%20in%20automotive%20cabins.pdf>. [Accès le 15 04 2014].
- [38] CAMPANELLA A. J., 1995, Preliminary values of clarity and intelligibility for small auditoria, meeting, and teleconference rooms, *J. Acoust. Soc. Am.* , vol. 97, n° 15, p. 3337.
- [39] HANSLER E., SCHMIDT G., 2004, *Acoustic Echo and Noise Control: A Practical Approach*, Oxford, Wiley-Blackwell, p. 374.
- [40] MIJICAND M., MASOVIC D., 2010, Reverberation Radius in Real Rooms, *Telfor Journal*, vol. 2, n°12, pp. 86-91.

- [41] GRIMARI A., 1991, A Stereophonic Imaging System for Car Audio, *91st AES convention*, New York.
- [42] CROCKETT B., SMITHERS M., BENJAMIN E., 2006, Next Generation Automotive Research and Technologies, *120ème convention de l'AES*, Paris.
- [43] CARPENTIER C., *Les technologies Isophases*, 2003.
- [44] CHRISTOPH M., SCHOLTZ L., 2010, Audio System Phase Equalization Pub. N° US2011/0103590 A1, *United States Patent Application Publication*.
- [45] DYREBY J., CHOISEL S., 2007, Equalization of loudspeaker resonances using second-order filters based on spatially distributed impulse response measurements, *Audio Engineering Society Convention 123*, New York.
- [46] BINELLI M., FARINA A., 2008, Digital equalization of automotive sound systems employing spectral smoothed FIR filters, *AES 125th Convention, San Francisco, CA, USA*.
- [47] CARINI A., CECCHI S., ROMOLI L., 2001, Multipoint Room Response Equalization with Group Delay Compensation, *Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis*, Dubrovnik.
- [48] FARINA A., RIGHINI F., 1997, Software Implementation of an MLS Analyzer with Tools for Convolution, Auralization and Inverse Filtering, *Audio Engineering Society Convention 103*, New York.
- [49] KLEINER M., LIDGREN C., 1998, Objective Characterization of Audio Sound Fields in Automotive Spaces, *Proceedings of the AES 15th International Conference*.
- [50] HESS W. G., NITZPON H.-J., ZELLER M., 2003, System for determining the position of a sound source, Brevet US 7386133 B2.
- [51] HESS W., 2007, Objective quantification of auditory source width of a loudspeakers-room system, Brevets US 8238589 B2, EP1962559A1, US20080247556.

- [52] PULKII V., KARJALAIEN L., HUOPANIEMI J., 1999, Analyzing Virtual Sound Source Attributes Using a Binaural Auditory Model, *J.AudioEng.Soc.*, vol. 47, n°14, pp. 203-217.
- [53] FARINA A., AYALON R., 2003, Recording Concert Hall acoustics for posterity, *AES 24th International Conference on Multichannel Audio*, Banff.
- [54] SILVERMAN H. F., SACHAR J. M., 2005, The Time-Delay Graph and the Delayogram—New Visualizations for Time Delay, *IEEE Signal Processing Letters*, vol. 12, n°. 4, pp. 301 - 304.
- [55] CAVALIERE S., SANTANGELO P., Spectrum Weighted HRTF based sound localization, 2007 , *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, Barcelone.
- [56] STROBEL N., RABENSTEIN R., 1999, Classification of time delay estimates for robust speaker localization, *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3081 - 3084.
- [57] BEN-REUVEN E., SINGER Y., 2002, Discriminative Binaural Sound Localization, *Advances in Neural Information Processing Systems 15*, MIT Press, Cambridge, MA, pp. 1229--1236.
- [58] BHARGAVA M. S., *The Generalized Correlation Method for Estimation of Time Delay*, Rapport de projet de l'Indian Institute of technology Kanpur.
- [59] KIHARA H., 1989, Brevet US4817162 A - Binaural correlation coefficient correcting apparatus.
- [60] WITTEK H., 2007, *Perceptual differences between wavefield synthesis and stereophony*, Thèse de doctorat, Department of Music and Sound Recording - School of Arts, Communication and Humanities - University of Surrey.
- [61] MASON R., 2006, Implementation and application of a binaural hearing model to the objective evaluation of spatial impression, *AES 28th International Conference*, Piteå, Sweden.
- [62] LI N., LOIZOU P. C., 2008, Factors influencing intelligibility of ideal binary-masked speech -

Implications for noise reduction, *J. Acoust. Soc. Am.*, vol. 123, n°13, p. 1673–1682.

- [63] WOODRUFF J., WANG D., 2012, Binaural Localization of Multiple Sources in Noisy and Reverberant Rooms, *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, vol. 20, n° 15, pp. 1503-1512.
- [64] KWAN H. K., JIANG A., 2006, Recent Advances in FIR Approximation by IIR Digital Filters, *2006 International Conference on Communications, Circuits and Systems Proceedings*, Guilin.
- [65] FARINA A., UGOLOTTI E., 1998, Automatic Measurement System for Car Audio Applications, *104th AES convention*, Amsterdam.
- [66] DAVILA-CHACON J., HEINRICH S., LIU J., WERMTER S., 2012, Biomimetic Binaural Sound Source Localisation with Ego-Noise Cancellation, *Artificial Neural Networks and Machine Learning – ICANN 2012 Lecture Notes in Computer Science Volume 7552*, Berlin, Springer Berlin Heidelberg, pp. 239-246.
- [67] FRANCOIS T., 2013, «Le futur du son spatialisé».
- <http://www.lesnumeriques.com/futur-son-spatialise-partie-2-a1725.html>. [Accès le 2014].
- [68] SONDERGAARD P. L., CULLING J. F., DAU T., LE GOFF N., JEPSEN M. L., MAJDAK P., WIERSTORF H., Towards a binaural modelling toolbox.
- <http://amtoolbox.sourceforge.net/notes/amtnote002.pdf>
- [69] KUO S. M., LEE B. H., TIAN W., 2006, Real-Time Digital Signal Processing: Implementations and Applications, *Real-Time Digital Signal Processing: Implementations and Applications, 2nd Edition*, Wiley & Sons Ltd, 2006 (E-book).
- [70] SHIVELY R., BAILEY J., HALLEY J., KURANDT L., MALBOS F., RUIZ G., SVOBODNIK, A., 2010, Considerations for the Optimal Location and Boundary Effects for Loudspeakers in an Automotive Interior, *Processings of the 128th AES Convention*, Londres.

- [71] SHIVELY R., HALLEY J., MALBOS F., RUIZ G., 2010, Optimal Location and Orientation for Midrange and High Frequency Loudspeakers in the Instrument Panel of an Automotive Interior, *Processings of the 129th AES Convention*, San Francisco, USA.
- [72] WATERHOUSE R. V., 1958, Output of a Sound Source in a Reverberation Chamber and Other Reflecting Environments, *The Journal of the Acoustical Society of America*, vol. 30, n°11, pp. 4-13, 1958.
- [73] KUTTRUFF H., 1998, Sound Fields in Small Rooms, *15th International Conference of the AES: Audio, Acoustics & Small Spaces*, Copenhagen, Danemark.
- [74] HARTMANN W. M., MACAULAY E. J., 2014, Anatomical limits on interaural time differences: an ecological perspective, *Frontiers in Neuroscience*, vol. 8, article 34
- [75] FARINA A., UGOLOTTI E., 1997, Subjective comparison of different car audio systems by the auralization technique, *Pre-prints of the 103rd AES Convention*, New York.
- [76] FRANK M., 2013, Source Width of Frontal Phantom Sources: Perception, Measurement, and Modeling, *Archives of Acoustics*, vol. 38, n° 13, pp. 311-319.
- [77] WANG D., From Auditory Masking to Binary Classification: Machine Learning for Speech Separation, 2009. [En ligne].
- http://www.waspaa.com/waspaa13/wordpress/wp-content/uploads/2011/09/Wang_WASPAA13.pdf.

Table des illustrations

Figure 1 - Infériorité de la mémoire de l'audition par comparaison au toucher et à la vue, traduit de [3].....	7
Figure 2 - Calcul de la différence de parcours avec une tête sphérique [5]	10
Figure 3 - ITD selon la formule de Woodworth pour une distance entre les oreilles de 18 cm	11
Figure 4 - Comparaison de la relation de Woodworth avec $r=11$ cm (en bleu) et de la relation de Kuhn avec $r=9,3$ cm en dessous de 500 Hz (rouge), à 1000 Hz (orange) et à 2000 Hz (vert)	12
Figure 5 - ILD mesuré en fonction de l'angle d'incidence à différentes fréquences selon Gulick [12].....	13
Figure 6 - ILD calculée en fonction de l'angle d'incidence à différentes fréquences de sinus pur, traduit de [11].....	14
Figure 7 - Un abaque de Franssen pour une bande critique, traduit de [16]	15
Figure 8 – Un « cône » de confusion considéré pour une oreille, adapté de [4].....	17
Figure 9 - Position du filtrage gammatone dans un modèle d'audition	17
Figure 10 - Réseau de filtrage en bandes critiques, selon [4]	18
Figure 11 - Exemples de filtrage HRTF en amplitude sur les plans azimutal (à gauche) et zénithal (à droite)	19
Figure 12 - Résultats de l'expérience de Haas - latéralisation perçue d'un signal de parole en fonction du délai appliqué à un haut-parleur, adapté de [6].....	20
Figure 13 - Représentation des réflexions dans un local et de la perception spatiale résultante, adapté de [15]	21
Figure 14 – Perceptions spatiales et réflexions, traduit de [16]	22
Figure 15 - Perception spatiale en fonction de la corrélation inter-aurale d'un bruit rose au casque, selon Chernyak and Dubrovsky cités dans [22].....	26
Figure 16 - Perception spatiale de panels d'écoute pour différents niveaux de corrélation de bruits filtrés entre 4 haut-parleurs dont l'orientation est indiquée par les flèches, selon Damaske cité dans [22].....	26
Figure 17 – FIACC théorique de deux signaux totalement corrélés (en bleu) et de deux signaux mesurés dans un champ sonore diffus avec des capteurs distants (en rouge).	28
Figure 18 – Explication simplifiée du circuit de Jeffress, traduit de [27]	29
Figure 19 - schéma du modèle de Jeffress, selon [4]	30

Figure 20 - schéma du modèle de Lindemann, traduit de [4]	31
Figure 21 - Exemple de description de la scène sonore (traduit de [30])	33
Figure 22 – Représentations de la perception spatiale par un panel d’auditeurs d’une voix centrée, puis avec un décalage temporel entre les deux haut-parleurs.....	34
Figure 23 - Représentations de la perception spatiale d’une scène sonore par un panel d’auditeurs.....	34
Figure 24 - Extrait de la recommandation ITU-R BS.1116.1	39
Figure 25 - Positions possibles des auditeurs dans une voiture.....	41
Figure 26 - Signaux envoyés aux haut-parleurs sans traitement	42
Figure 27 – Modélisation de la pression sonore rayonnée par un haut-parleur placé dans un panneau de porte à différentes fréquences [35]	44
Figure 28 - Comparaison pour une même position de mesure de la réponse amplitude (fréquence) de haut-parleurs du même modèle, placés à différentes positions dans un habitacle	44
Figure 29 – Durée de réverbération en fonction de la fréquence selon différentes méthodes dans le véhicule "G", microphone de l’oreille gauche du simulateur tête-torse.....	46
Figure 30 - Durée de réverbération en fonction de la fréquence selon différentes méthodes dans le véhicule "G", microphone de l’oreille droite du simulateur tête-torse.....	47
Figure 31 - distance critique mesurée dans un habitacle, traduit de [39]	49
Figure 32 – Simulation des premiers modes longitudinaux et latéraux de pression sonore dans une voiture, aux fréquences 64, 73, 108 et 115 Hz [35]	50
Figure 33 - Position du filtrage dans une chaîne audio typique.....	51
Figure 34 - Traitement audio sur 4 canaux.....	52
Figure 35 - ajustement temporel mono point.....	54
Figure 36 - distances entre les haut-parleurs et les deux oreilles de l’auditeur	55
Figure 37 - Exemple d'optimisation de la réponse amplitude (fréquence) par alignement temporel des haut-parleurs, vert = avant alignement temporel, violet = après alignement temporel	56
Figure 38 – Expérience de pensée pour le calcul de phase – plusieurs microphones.....	57
Figure 39- résultats de la simulation - échelle fréquentielle linéaire.....	58
Figure 40 - résultats de la simulation - échelle fréquentielle logarithmique	58
Figure 41 - Résultat de la simulation d’écoute hors <i>sweet spot</i> avec diaphonie.....	60

Figure 42 - Exemple d'amélioration de la phase binaurale mesurée, A sans filtrage, B avec filtrage [44]	61
Figure 43 - Modèle embarqué dans les brevets Harman-Becker.....	67
Figure 44 – Estimation d'azimut pour une source latéralisée à 0° ; 15° ; 30° ; 45° et 65°. Première ligne, ILD uniquement, en fonction de la fréquence. Deuxième ligne ITD uniquement, en fonction de la fréquence. Troisième ligne : combinaison ILF et ILD en fonction de la fréquence. Dernière ligne : somme des combinaisons sur toutes les fréquences. Traduit de [5]	68
Figure 45 - Mesure d'IACC et 1-LF sur le plan azimutal, dans un même local, extrait de [53] .	69
Figure 46 – Delayogram pour un orateur face à deux microphones à 1 m 50, faisant le son « shhhh » en mouvement (gauche) puis fixe et lisant une phrase (droite)	69
Figure 47 - Cartographie d'IACC précoce, à gauche pour un réglage symétrique, à droite pour un délai de 2 ms sur le tweeter avant gauche et un gain de 5dB sur le tweeter avant droit, adapté de [35]	70
Figure 48 - mesure de FIACC avant et après égalisation, traduit de [59].....	70
Figure 49 - Inter-corrélation mesurées pour un réglage A – bruit corrélé.....	71
Figure 50 - Inter-corrélation mesurées pour un réglage B – bruit corrélé.....	72
Figure 51 - Analyse de corrélation inter-aurale en fonction de la fréquence, source réelle décalée de 15° à droite[60]	73
Figure 52 - Analyse de corrélation inter-aurale en fonction de la fréquence, source fantôme créée par un système de sonorisation stéréophonique, décalée perceptivement de 15° à droite par panoramique d'intensité et de temps [60]	73
Figure 53 - Modèle proposé par Mason dans [61]	74
Figure 54 - Diagramme de latéralisation proposé par Mason	74
Figure 55 - Diagramme de latéralisation en fonction du temps obtenu dans une pièce au standard ITU-R BS 1116, en lisant un signal corrélé sur un système stéréophonique	75
Figure 56 - Diagramme de latéralisation en fonction du temps obtenu dans une automobile équipée d'un système audio typique, en lisant un signal corrélé.....	75
Figure 57 - Diagrammes d'activité binaurale, cadran de gauche modèle de Gaik, cadran de droite modèle de Lindemann, cadran du haut bande de fréquence centrée à 1172Hz, cadran du bas bande de fréquence centrée à 3539Hz [29]	76

Figure 58 - Résultats de la mesure de Faller et Merimaa [28] pour une source réelle (A) et deux sources réelles (B) sans seuil de corrélation inter-aurale, et respectivement C et D avec seuil	77
Figure 59 - Performances de séparation spatiale des différents modèles de localisation en fonction du temps de réverbération, traduit de [9].....	79
Figure 60 - Phase binaurale sans traitement audio, position conducteur à gauche.....	80
Figure 61 - Phase binaurale sans traitement audio, position passager à droite.....	81
Figure 62 - Phase binaurale sans traitement audio, position conducteur à gauche en vert et cyan, symétrique de la phase binaurale en position passager à droite en bordeaux.....	81
Figure 63 - Phase binaurale sans traitement audio, position conducteur à gauche.....	82
Figure 64 - Phase binaurale sans traitement audio, position passager à droite.....	83
Figure 65 - Phase binaurale sans traitement audio, position conducteur à gauche en vert et cyan, symétrique de la phase binaurale en position passager à droite en bordeaux.....	83
Figure 66 - Phase binaurale sans traitement audio, position conducteur à gauche.....	84
Figure 67 - Phase binaurale sans traitement audio, position passager à droite.....	84
Figure 68 - Phase binaurale sans traitement audio, position conducteur à gauche en vert et cyan, symétrique de la phase binaurale en position passager à droite en bordeaux.....	85
Figure 69 - Phase binaurale, échelle fréquentielle linéaire - rouge sans traitement audio - vert avec réglage	86
Figure 70 - Phase binaurale, échelle fréquentielle linéaire - rouge sans traitement audio - vert avec réglage	87
Figure 71 - Phase binaurale, échelle fréquentielle linéaire - rouge sans traitement audio - vert avec réglage	88
Figure 72 – FIACC de signaux tests usuels; vert - bruit rose corrélé ; bleu – bruit rose corrélé avec un canal retardé ; rouge - bruit rose dé-corrélé ; cyan - bruit rose corrélé avec un canal hors phase.	89
Figure 73 - FIACC avec et sans réglage dans le véhicule D	90
Figure 74 - FIACC avec et sans réglage dans le véhicule E.....	91
Figure 75 - FIACC avec et sans réglage dans le véhicule F.....	92
Figure 76 - Résultats à la position conducteur – en haut sans réglage, au milieu avec le réglage A, en bas avec le réglage B	94

Figure 77 - ITD et ILD dans la bande critique de fréquence centrée sur 250 Hz quand on lit un bruit rose corrélé. A gauche sans seuil de corrélation inter-aurale, à droite avec seuil de corrélation inter-aurale.	95
Figure 78 - ITD et ILD dans la bande critique de fréquence centrée sur 250 Hz quand on lit un bruit rose dé-corrélé. A gauche sans seuil de corrélation inter-aurale, à droite avec seuil de corrélation inter-aurale.	95
Figure 79 - ITD et ILD, sans filtrage	96
Figure 80 - ITD et ILD, avec filtrage.....	97
Figure 81 - ITD et ILD, sans filtrage, précision 1/12ème d’octave.....	97
Figure 82 - ITD et ILD, avec filtrage, précision 1/12ème d’octave	98
Figure 83 - ITD et ILD, avec seulement le haut-parleur gauche activé.....	98
Figure 84 - ITD en fonction de la fréquence, sans filtrage	99
Figure 85 - ITD en fonction de la fréquence, avec filtrage	99
Figure 86 - Résultats pour un balayage fréquentiel corrélé en phase	101
Figure 87 - Figure 86 vue sous un autre angle.....	101
Figure 88 - Résultats pour un balayage fréquentiel, un des canaux étant en opposition de phase.....	102
Figure 89- Figure 88 vue sous un autre angle	102
Figure 90 - Résultats pour un bruit rose dé-corrélé	103
Figure 91 - Résultats pour un balayage fréquentiel mono de 10s	104
Figure 92 - Résultats pour un balayage fréquentiel avec un retard sur un des canaux.....	105
Figure 93 – Résultats du modèle de Lindemann dans le véhicule D	106
Figure 94 - Figure 93 sous un autre angle	107
Figure 95 - Phase binaurale dans le véhicule D sans traitement.....	108
Figure 96 - Phase binaurale dans le véhicule D avec traitement	108
Figure 97 - Résultats du modèle de Lindemann dans le véhicule E	110
Figure 98 - Résultats du modèle de Lindemann dans le véhicule F	112
Figure 99 - Figure 98 sous un autre angle	113
Figure 100 - Figure 98 sous un autre angle	114

Table des tableaux

Tableau i - Résultats de l'expérience de Haas	20
Tableau ii - Inter-corrélation normalisée de signaux tests usuels.....	25
Tableau iii - TR60 dans 3 véhicules	45
Tableau iv - TR60 dans le véhicule "G"	46
Tableau v - Clartés C50 et C80 dans les véhicules "D", "E" et "F"	47
Tableau vi - Distances typiques entre les haut-parleurs et le conducteur.....	53
Tableau vii - Signaux utilisés pour le test du code « Lindemann.m ».....	100
Tableau viii - Résultats d'écoute avec et sans filtrage, le 16/05/2013.....	124
Tableau ix - Résultats d'écoute avec et sans filtrage, le 16/05/2013.	125

Mesures de la spatialisation pour la sonorisation automobile

Alexandre FENIERES

Mémoire d'ingénieur C.N.A.M., Paris 2016

Spécialité mécanique option acoustique

RÉSUMÉ

Après une présentation des modèles usuels d'audition spatiale humaine, une analyse acoustique des systèmes audio typiquement embarqués dans les automobiles montre pourquoi ces conditions diffèrent des recommandations de diffusion stéréophonique. Pour trouver plus rapidement les paramètres de traitement du signal audio qui permettent d'améliorer la fidélité spatiale dans ce contexte, certains modèles de latéralisation sonore sont appliqués à des mesures acoustiques binaurales faites dans des véhicules, et leurs résultats sont analysés et comparés à ceux dans la littérature.

Mots clés : Sonorisation automobile, traitement du signal embarqué, modèles d'audition spatiale

SUMMARY

After a presentation of the usual models of human spatial hearing, an acoustic analysis of typical automotive audio systems shows why the conditions inside a car cabin differ from the recommendations for the stereophonic reproduction of a sound field. To find the parameters for a processing of the audio signal that would increase the spatial fidelity more easily in that context, some models of auditory lateralization are applied to acoustic binaural measurements carried out in vehicles, and their results are analyzed and compared to those in the literature.

Key words: Automotive audio, embarked signal processing, models of spatial hearing