

# Restitution d'un effet de salle cohérent avec le lieu d'écoute dans le cadre de Wave Field Synthesis

Terence Caulkins

Rapport de stage de DEA ATIAM

Université de Paris VI - Université Aix Marseille II - Ecole National Supérieure  
des Télécommunications - Institut National Polytechnique de Grenoble

Responsable de stage : Olivier WARUSFEL

Encadrant de stage : Etienne CORTEEL

Laboratoire d'accueil : IRCAM



Année Universitaire 2002 - 2003

# Table des matières

|          |                                                                |           |
|----------|----------------------------------------------------------------|-----------|
| <b>1</b> | <b>Fondements Théoriques de la Wave Field Synthesis</b>        | <b>8</b>  |
| 1.1      | Principe de Huyghens . . . . .                                 | 8         |
| 1.2      | Formulation de Kirchoff-Helmholtz . . . . .                    | 9         |
| 1.3      | Intégrales de Rayleigh . . . . .                               | 11        |
| 1.4      | La source notionnelle . . . . .                                | 12        |
| 1.5      | l'Approximation de Phase Stationnaire . . . . .                | 12        |
| 1.6      | Le Fenêtrage et la diffraction, effets de troncature . . . . . | 14        |
| 1.7      | L'échantillonnage et l'Aliasing Spatial . . . . .              | 16        |
| 1.8      | Le cas particulier des sources focalisées . . . . .            | 18        |
| <b>2</b> | <b>Mise en oeuvre pratique</b>                                 | <b>22</b> |
| 2.1      | Les haut parleurs MAP . . . . .                                | 22        |
| 2.2      | La Multiégalisation . . . . .                                  | 24        |
| 2.2.1    | L'Algorithme de Traitement Multicanal . . . . .                | 25        |
| 2.2.2    | Traitement Hautes Fréquences . . . . .                         | 26        |
| 2.2.3    | Résultats . . . . .                                            | 27        |

|          |                                                                                                          |           |
|----------|----------------------------------------------------------------------------------------------------------|-----------|
| <b>3</b> | <b>Restitution de l'effet de salle</b>                                                                   | <b>29</b> |
| 3.1      | Modélisation de l'effet de salle . . . . .                                                               | 29        |
| 3.2      | Le Spatialisateur . . . . .                                                                              | 31        |
| 3.3      | Reproduction de l'effet de salle et moteur de rendu spatial<br>pour le démonstrateur IRCAM . . . . .     | 32        |
| 3.4      | Intégration d'un dispositif WFS dans une salle d'écoute réelle                                           | 35        |
| <b>4</b> | <b>Directivité des sources</b>                                                                           | <b>37</b> |
| 4.1      | Décomposition d'un champ sonore en harmoniques sphériques<br>ou cylindriques . . . . .                   | 37        |
| 4.2      | Présentation des Résultats WFS . . . . .                                                                 | 39        |
| 4.3      | Amélioration de la reproduction de sources directives par l'ap-<br>proche de multiégalisation . . . . .  | 42        |
| 4.3.1    | Reproduction de sources directives sur le démonstrateur<br>IRCAM . . . . .                               | 44        |
| <b>5</b> | <b>Mise en cohérence du champ sonore reproduit en WFS avec<br/>la salle d'écoute</b>                     | <b>46</b> |
| 5.1      | Conséquences des approximations de la WFS sur le champ<br>effectif rayonné . . . . .                     | 47        |
| 5.2      | Interaction d'une source virtuelle avec la salle d'écoute . . . .                                        | 48        |
| 5.2.1    | Reproduction d'une source non-focalisée . . . . .                                                        | 48        |
| 5.2.2    | Reproduction d'une source focalisée . . . . .                                                            | 50        |
| 5.3      | Importance des régimes de réflexions manquants dans la re-<br>production de sources focalisées . . . . . | 52        |
| 5.4      | Elaboration d'une expérience perceptive . . . . .                                                        | 53        |

|          |                                                                 |           |
|----------|-----------------------------------------------------------------|-----------|
| 5.4.1    | Résultats de l'expérience . . . . .                             | 57        |
| 5.5      | Amélioration de la reproduction des sources intérieures . . . . | 59        |
| <b>6</b> | <b>Conclusion et Prospectives</b>                               | <b>61</b> |

# Remerciements

J'aimerais exprimer mon respect et ma gratitude envers les personnes suivantes :

- ★ Olivier Warusfel, pour son aide et son soutien au cours du stage
- ★ Etienne Corteel, pour le partage de son savoir au sujet de la WFS, pour son assistance technique permanente, pour avoir fait face à des barrages de questions sans s'énerver, etc. etc. etc..
- ★ Cedric Peigné, Mathieu Nogues, Clemens Kuhn pour leur bonne humeur au sein du bureau 105
- ★ Emmanuel Rio et Guillaume Van Der Noot pour leur aide concernant le binaural et le Flock of Birds
- ★ Alexis Baskind pour son assistance sur Linux et LaTeX
- ★ Cyrille Defaye pour ses conseils et son aide tout le long de l'année
- ★ Mes parents pour leur soutien et leur amour depuis toujours

# Introduction

Il existe aujourd'hui des nouvelles techniques de restitution sonore dites "immersives" comme le binaural ou la Wave Field Synthesis. Le binaural permet de restituer un environnement sonore tridimensionnel via des écouteurs. La Wave Field Synthesis permet de restituer par voie électro-acoustique un champ sonore destiné à une large audience, dans lequel on peut se déplacer tout en conservant une sensation de relief entre les objets. Lorsque l'on applique ces techniques dans un contexte lié à la réalité virtuelle, il apparaît la nécessité d'assurer une cohérence, éventuellement visio-auditive, avec le lieu d'écoute. Ceci est d'autant plus important lorsque l'on restitue des sources *à l'intérieur du lieu d'écoute*, ce qui est une possibilité nouvelle offerte par la Wave Field Synthesis. Par souci de réalisme, il peut s'agir également de reproduire des sources dont la directivité n'est pas forcément monopolaire. Au cours de ce mémoire, nous tenterons d'apporter des éléments nouveaux permettant d'avancer sur ces deux problèmes.

La technique de la Wave Field Synthesis (WFS) fut introduite par A.J. Berkhout en 1988 [1]. Elle peut être assimilée à l'équivalent acoustique de l'holographie, ce qui lui vaut l'appellation de "reproduction holophonique". Cette technique vise à reproduire un champ sonore dans une portion de plan délimité par une distribution de haut parleurs.

Les dispositifs classiques de reproduction électro-acoustique comme la stéréophonie ou le 5.1 garantissent un champ sonore spatialement correct à un ou plusieurs endroits précis, dénommés "sweet spots". Ils s'appuient sur l'effet psychoacoustique dit de "sources fantômes", utilisant les différences interaurales de temps (ITD) et les différences interaurales d'intensité (IID). Le problème majeur de ce type de reproduction réside dans le fait qu'un auditeur situé en dehors du "sweet spot" reçoit le front d'onde primaire du haut parleur le plus proche. Ainsi la direction du front d'onde primaire varie avec la position de l'auditeur, ce qui mène à une localisation incorrecte des sources fantômes. Cette erreur sur le positionnement perceptif des sources

sonores peut s'avérer gênante, en particulier dans le cas où on l'associe à une reproduction visuelle.

La reproduction fidèle de la courbure d'un front d'onde constitue le fondement de la théorie WFS. Elle est basée sur le principe de Huyghens qui stipule qu'un front d'onde peut être reproduit par un continuum de sources secondaires. Les dispositifs basés sur la WFS permettent par conséquent de générer un champ d'onde ayant des propriétés spatiales et temporelles correctes dans une zone étendue. Cet élargissement de la zone d'écoute constitue un avantage majeur de l'holophonie par rapport aux techniques classiques.

Les travaux sur l'holophonie ont été initiés par les chercheurs de l'Université technologique de Delft vers la fin des années 80. Depuis, cette technique a suscité l'intérêt de la communauté scientifique européenne, notamment dans le cadre du projet CARROUSO lancé par l'IST (Information Society Technologies)[2]. L'objectif principal du projet CARROUSO (Creating, Assessing and Rendering in Real- Time Of High -quality aUdio-viSual enviroNments) est de fournir une nouvelle technologie permettant le transfert d'un champ sonore généré dans un espace réel ou virtuel vers un espace de restitution. CARROUSO emploie le standard MPEG-4 en matière de codage de l'information, dans lequel on véhicule les données audio séparément de la description de la scène sonore (position des sources, descripteurs perceptifs ou géométriques de la salle). La technique WFS joue un rôle essentiel dans ce projet du fait qu'elle permet de reproduire physiquement un espace sonore virtuel, et non pas uniquement l'illusion perceptive d'un tel espace.

Outre le projet CARROUSO, la Wave Field Synthesis est exploitable dans de nombreuses applications comme le cinéma et le home cinéma, les simulateurs de vol ou les théâtres de réalité virtuelle. Des travaux sont également menés actuellement dans les laboratoires de France Télécom en ce qui concerne l'emploi de la WFS dans des systèmes de visioconférence[3]. Au niveau des applications musicales, la WFS permet d'envisager la reproduction d'enregistrements multipistes. La notion de spatialisation est également étendue par ce système qui donne la possibilité de former du relief dans des scènes sonores par le biais du paramètre de courbure de front d'onde ([4]). Grâce à ce paramètre, la cohérence spatiale peut être maintenue sur une large audience, ce qui permet par la même occasion une navigation sans perte de cohérence. Ceci implique une sensation de présence et d'immersion accrue dans les installations sonores utilisant la WFS.

Le premier chapitre de ce mémoire sera consacré à l'étude théorique de la physique sous-tendant la Wave Field Synthesis. On explicitera les étapes permettant de passer du cadre décrit par le principe de Huyghens

à une situation réelle de reproduction holophonique.

Le deuxième chapitre décrit la mise en oeuvre pratique de l'holophonie sur le système de haut parleurs MAP de l'IRCAM. La particularité de ces haut parleurs implique de procéder à une égalisation multicanale du système, qui sera abordée en détail.

Le troisième chapitre de ce mémoire traite de la restitution d'un effet de salle dans le cadre de la Wave Field Synthesis. On explicitera notamment la structure du moteur de rendu spatial utilisé dans le dispositif WFS de l'IRCAM.

Dans le quatrième chapitre, on examine la possibilité d'introduire une directivité dans les sources reproduites. Pour ce faire, il s'agit d'abord de donner une description succincte des harmoniques sphériques et cylindriques. On exposera ensuite quelques résultats concernant la reproduction des harmoniques de base sur un système WFS. A terme, ce travail permettrait d'envisager la synthèse de figures de directivité complexes (comme celles d'instruments réels) par combinaison linéaire des harmoniques de base.

Dans le cinquième et dernier chapitre on aborde la mise en cohérence d'une reproduction WFS avec la salle d'écoute. Pour ce faire, on étudiera d'abord les conséquences des approximations faites au premier chapitre sur le champ effectif rayonné dans les trois dimensions. Le cas particulier des sources focalisées, qui sont des sources virtuelles reproduites à l'intérieur de l'espace de restitution, est abordé en détail. Cette étude sera étayée par une expérience perceptive pour déterminer le rôle de l'effet de salle manquant lors de la synthèse de ce type de source. On proposera enfin une méthode permettant d'améliorer la reproduction des sources focalisées.



# Chapitre 1

## Fondements Théoriques de la Wave Field Synthesis

*La Wave Field Synthesis est une méthode de reproduction sonore basée sur la reconstruction d'un front d'onde dans le plan à partir d'une distribution de sources secondaires. Cette méthode ressemble au "rideau acoustique" évoqué par Snow en 1953 dans son inventaire des techniques de reproduction stéréophonique [5]. En 1988 Berkhout remplace ce concept intuitif par la WFS en se basant sur une théorie ondulatoire bien fondée [1] [6]. La méthode s'appuie sur le principe de Huyghens, qui date de 1690. Contrairement au procédé "holophonique" décrit par Jessel en 1973 [7], la WFS ne nécessite pas de placer le système de restitution au même endroit que le système de captation lorsqu'on reproduit le rayonnement d'une source sonore. En effet, les chercheurs de la TU Delft montrent qu'il est possible de "propager" le front d'onde capté sur un banc de microphones vers un banc de haut parleurs assurant la restitution du son. Il devient alors possible de dissocier prise de son et restitution, comme dans les systèmes classiques stéréophoniques.*

### 1.1 Principe de Huyghens

Selon le *principe de Huyghens* formulé en 1690, chaque élément d'un front d'onde peut être considéré comme étant le centre d'une onde sphérique individuelle (cf FIG. 1.1). Le front d'onde à un instant ultérieur peut alors être reconstruit en recombinaison ces ondes individuelles. En d'autres termes, le champ induit en aval d'un front d'onde généré par une source (dite

“primaire”) peut être vu comme le champ rayonné par une infinité de sources (dites “secondaires”) réparties le long de ce front d’onde. Dans ce raisonnement, les sources secondaires se substituent de manière parfaitement équivalente à la source primaire.

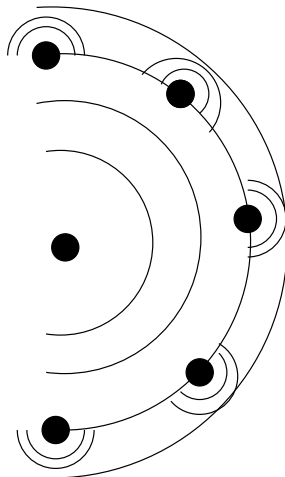


FIG. 1.1 – *Principe de Huyghens*

## 1.2 Formulation de Kirchhoff-Helmholtz

Lors de sa recherche menée sur les phénomènes sismiques [8], Berkhout a tiré parti des théories élaborées par Kirchhoff et Helmholtz pour énoncer une formulation mathématique rigoureuse du principe de Huyghens. Cette formulation permet notamment de généraliser ce principe à n’importe quelle surface en se ramenant à un problème aux limites. Pour ce faire, on définit deux sous-espaces complémentaires  $V$  et  $\bar{V}$ .  $V$  représente le domaine des sources, comprenant toutes les sources primaires dont on cherche à reproduire le rayonnement.  $\bar{V}$  représente le domaine de restitution, libre de sources. On cherche alors à reproduire dans  $\bar{V}$  le champ induit par les sources primaires situées dans  $V$  en utilisant une distribution de sources secondaires sur la frontière  $\partial V$  du domaine des sources(cf FIG. 1.2).

On se ramène donc à un problème aux limites sur  $\partial V$  que l’on peut résoudre à l’aide de la fonction de Green. On peut montrer ([9]) que la pression  $p$  en un point  $r$  situé dans le domaine de restitution  $V$  peut s’écrire :

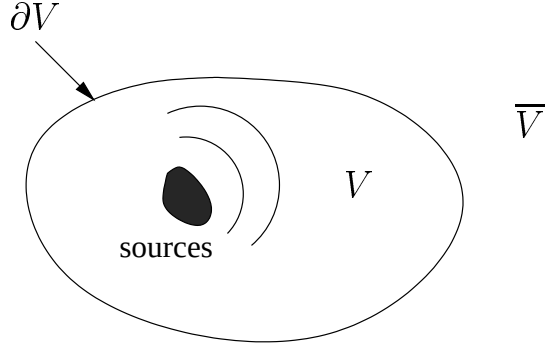


FIG. 1.2 – *Espace des sources/ Espace de restitution*

$$\begin{aligned}
 p(\vec{r}) = & \frac{1}{4\pi} \oint_{\partial V} [j\omega\rho V_n(r_0) \frac{e^{-jk(r-r_0)}}{r-r_0}] dS_0 \\
 & + \frac{1}{4\pi} \oint_{\partial V} [P(r_0) \frac{1+jk(r-r_0)}{(r-r_0)} \cos\phi \frac{e^{-jk(r-r_0)}}{r-r_0}] dS_0 \quad (1.1)
 \end{aligned}$$

avec :

- $r_0$  un point de la surface  $\partial V$
- $V_n(r_0)$  composante de la vitesse particulaire normale à la surface  $\partial V$
- $P(r_0)$  la pression sur la surface  $\partial V$
- $\phi$  l'angle entre les vecteurs  $(\vec{r} - \vec{r}_0)$  et  $\vec{n}$ ,  $\vec{n}$  étant le vecteur unitaire normal à  $\partial V$  orienté vers l'intérieur du domaine de restitution

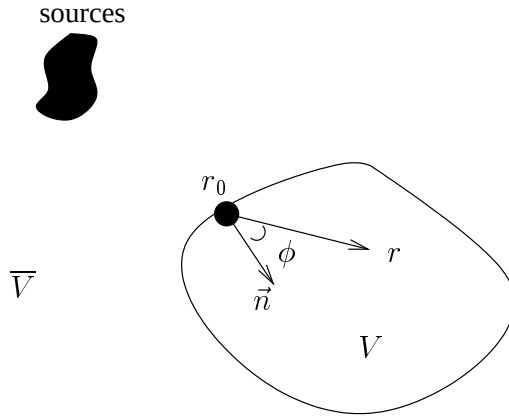


FIG. 1.3 – *Géométrie pour l'intégrale de Kirchhoff-Helmholtz*

L'intégrale de Kirchhoff Helmholtz (eq. 1.1) révèle que tout champ d'onde  $P$  issu de sources extérieures à  $V$  peut être reconstruit à l'intérieur

de  $V$  par le biais d'une distribution continue de monopôles et de dipôles sur la surface  $\partial V$ . Les axes de ces dipôles coïncident avec le vecteur unitaire normal  $\vec{n}$  (cf FIG. 1.3).

Jusqu'ici le cas considéré décrit un volume de reconstruction entouré par la distribution de monopôles et dipôles. Notons qu'il est possible de "retourner" la surface  $\partial V$  de manière à ce qu'elle entoure les sources. La relation (eq. 1.1) demeure alors valable, à condition de prendre le vecteur  $\vec{n}$  orienté vers l'intérieur du domaine de restitution.

### 1.3 Intégrales de Rayleigh

L'équation (1.1) est valable pour une surface  $\partial V$  quelconque, et non pas seulement pour un front d'onde (comme le suggérait le principe de Huyghens). Cette remarque peut être mise à profit pour dériver les *intégrales de Rayleigh* de la formulation de Kirchhoff-Helmholtz. Pour ce faire, on considère une surface plane infinie  $\partial V$  séparant l'espace de restitution  $V$  et l'espace des sources  $\bar{V}$ . Comme précédemment,  $\partial V$  représente une distribution continue de monopôles et de dipôles. On montre que dans ces conditions, le champ émis par la distribution monopolaire est égal au champ dipolaire dans l'espace de restitution et en opposition de phase avec le champ dipolaire dans l'espace des sources. Par conséquent, il devient possible de reproduire parfaitement le champ émis par les sources dans l'espace de restitution en utilisant un seul type de sources sur la surface  $\partial V$ . Ceci se traduit par l'intégrale de Rayleigh I (eq. 1.2) et l'intégrale de Rayleigh II (eq. 1.3) :

$$p(r_R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} j\omega \rho V_z(r_0) \frac{e^{-jk(r-r_0)}}{r-r_0} dx_0 dy_0 \quad (1.2)$$

$$p(r_R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(r_0) \frac{1 + jk(r-r_0)}{(r-r_0)} \cos \phi \frac{e^{-jk(r-r_0)}}{r-r_0} dx_0 dy_0 \quad (1.3)$$

avec :

- $p(r_R)$  pression en un point de l'espace de restitution
- $V_z(r_0)$  composante de la vitesse particulaire normale à la surface  $\partial V$
- $\phi$  l'angle entre les vecteurs  $(\vec{r} - \vec{r}_0)$  et  $\vec{n}$

Les intégrales de Rayleigh offrent donc une solution exacte au problème de la reconstruction d'un champ sonore. Cette solution est cependant valable

uniquement dans les cas idéal où l'on dispose d'une distribution plane continue et infinie de monopôles ou de dipôles. En réalité, nous disposerons d'une distribution *linéaire*, *discontinue* et *finie* de transducteurs ayant des directivités différentes de celle d'un monopôle ou d'un dipôle. Les sections suivantes seront consacrées à décrire l'effet de ces limitations sur le champ effectif rayonné par un dispositif WFS réel.

## 1.4 La source notionnelle

Nous désignerons par le terme *source notionnelle* toute source primaire dont on cherche à reproduire le rayonnement. La pression rayonnée par une telle source à l'origine est donnée par la formule suivante :

$$P(r, \omega) = S(\omega)D(\theta, \phi, \omega) \frac{\exp -jkr}{r} \quad (1.4)$$

avec :

- $S(\omega)$  fonction d'alimentation de la source (signal qui serait capté par un microphone de proximité)
- $D(\theta, \phi, \omega)$  fonction de directivité de la source

Connaissant les propriétés de rayonnement de la source primaire, on pourra en déduire la pression et la vitesse particulière sur la surface  $\partial V$ . Il devient alors envisageable de reproduire le rayonnement de sources à directivité variable. On se restreindra cependant dans un premier temps à la reproduction de sources monopôlaires, par souci de simplicité.

## 1.5 l'Approximation de Phase Stationnaire

Le dispositif proposé pour la WFS utilise une distribution linéaire de haut parleurs au lieu d'une distribution plane. L'espace de restitution est donc réduit à un plan qui sera désormais désigné comme étant le *plan d'écoute*. Les sources primaires sont également restreintes à ce même plan (cf FIG. 1.4). La difficulté consiste à quantifier la contribution des sources secondaires situées hors du plan d'écoute dans la reproduction d'une source primaire située dans le plan d'écoute. En effet, le dispositif de reproduction étant restreint à une ligne, il va falloir compenser l'effet des sources secondaires qui sont absentes sur le reste de la surface  $\partial V$ . Considerons une

colonne  $C$  traversant le plan d'écoute en  $r_{C,0}$ . En regardant la contribution de toutes les sources secondaires situées sur  $C$ , il est clair que les points situés au voisinage de  $r_{C,0}$  contribuent de manière prépondérante au champ d'onde généré au point de réception  $R$ . On peut montrer que la contribution de la colonne  $C$  peut se réduire à la contribution d'une seule source secondaire située en  $r_{C,0}$ .

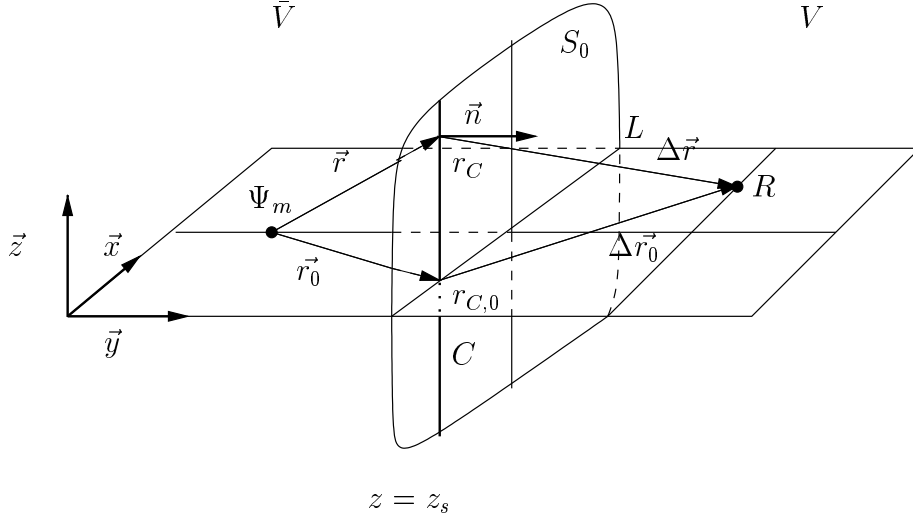


FIG. 1.4 – Schéma d'une source notionnelle  $\phi_m$  à la position  $r = r_\phi$  dans le plan  $z = 0$  et un plan infini de sources secondaires en  $y = y_C$ . La contribution en  $R$  d'une colonne  $C$  de sources secondaires perpendiculaire au plan  $z = 0$  peut être approximée par une contribution de  $r_{C,0}$  uniquement

Après calcul ([9]), on trouve l'expression de la pression acoustique en  $r_R$  :

$$P(r_R) = \int_{-\infty}^{+\infty} g(r_L, r_R) \sqrt{\frac{jk}{2\pi}} S(\omega) \cos \phi_{inc} \frac{e^{-jkr}}{r} \frac{e^{-jk\Delta r}}{\Delta r} dx_L \quad (1.5)$$

où le facteur d'amplitude  $g$  est défini par :

$$g(r_L, r_R) = \sqrt{\frac{r \Delta r}{r + \Delta r}} \quad (1.6)$$

avec  $r = \|\vec{r}\| = \|\vec{r}_L - \vec{r}_\Psi\|$ ,  $\Delta r = \|\Delta \vec{r}\| = \|\vec{r}_R - \vec{r}_L\|$ , et  $r_L = (x_L, 0, z_L)$ . L'angle  $\phi_{inc}$  est celui entre les vecteurs  $\vec{r}$  et  $\vec{n}$ .

## 1.6 Le Fenêtrage et la diffraction, effets de troncature

Il faut cependant réaliser que la distribution de sources secondaires utilisée dans un système de reproduction réel n'est pas une ligne infinie comme on l'a supposé dans cette approximation, mais un segment. Nous verrons dans la partie suivante les effets de cette troncature sur le champ sonore effectivement rayonné. La troncature de la ligne infinie de sources secondaires est un problème qu'il faut cerner lors de l'implémentation pratique du système WFS. Cette restriction implique deux conséquences sur le champ effectif :

- ★ La réduction de la “visibilité” de la source notionnelle vis à vis de la zone de restitution. En effet une source notionnelle peut être correctement reproduite uniquement dans le cas où elle est “vue” à travers la fenêtre que constitue le banc de haut parleurs (cf FIG. 1.5). Partout ailleurs, la reproduction du rayonnement de la source ne peut être assurée correctement.

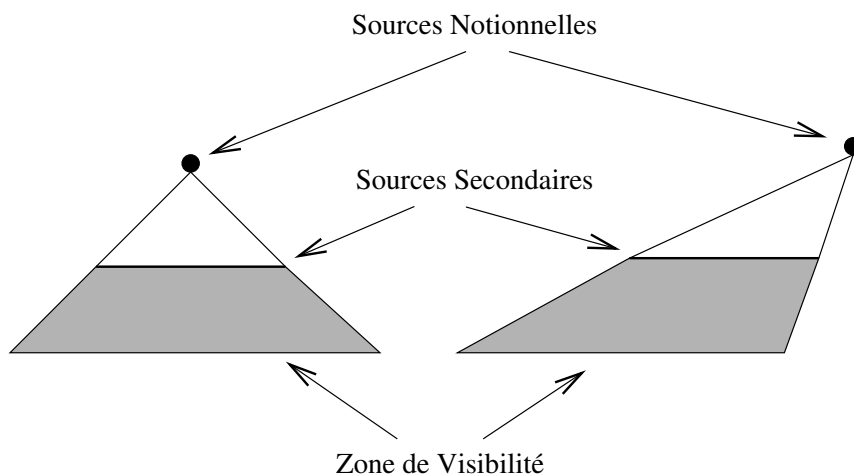


FIG. 1.5 – *Visibilité d'une source notionnelle*

- ★ La création d'ondes diffractées au niveau des bords du banc de haut parleurs. Pour mettre en évidence ce phénomène, on reproduit une onde plane avec des sources secondaires monopôlaire. On enregistre ensuite le résultat de cette synthèse sur un banc de microphones omnidirectionnels situé dans la zone de restitution (cf FIG. 1.6). Sur cette figure on peut observer deux fronts d'ondes parasites émis par les sources secondaires du bord de la distribu-

tion. Ceci s'explique par le fait que contrairement aux sources du centre, ces sources ne sont pas compensées par des sources adjacentes. E. W. Start a étudié la diffraction de manière exhaustive lors de sa thèse conduite à la TU Delft ([10]). Il montre que pour compenser l'erreur de synthèse induite par la troncature de l'intégrale de Rayleigh (eq. 1.2) il faudrait théoriquement modifier les caractéristiques de rayonnement et les fonctions d'alimentation des sources situées aux extrémités de la distribution secondaire. Il existe cependant une méthode plus simple qui consiste à appliquer une fonction de pondération sur le banc de sources secondaires. Cette fonction vaut 1 partout et tend vers 0 de manière sinusoïdale au voisinage des extrémités du banc([11]). En appliquant cette fonction de pondération lors de la reproduction d'une onde plane en utilisant le même dispositif que précédemment, on réduit notablement l'amplitude des ondes diffractées.

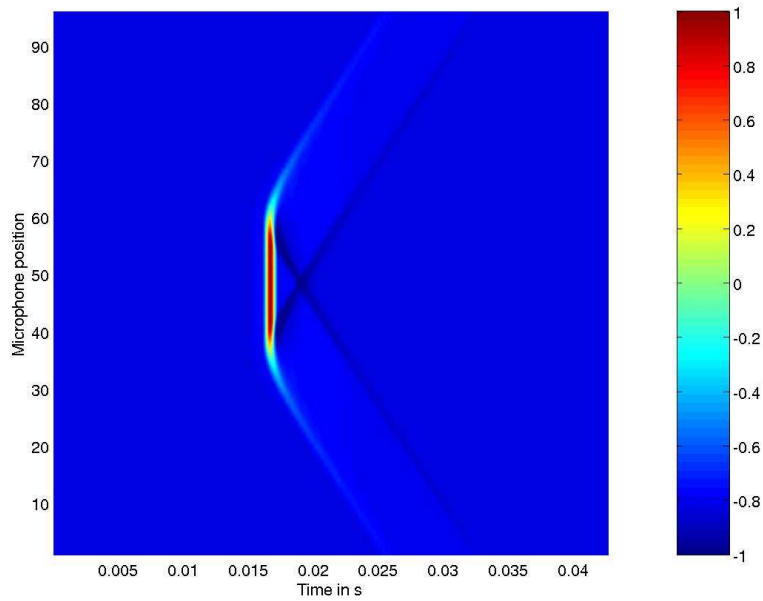


FIG. 1.6 – *Mise en évidence du phénomène de diffraction*



## 1.7 L'échantillonnage et l'Aliasing Spatial

Outre la troncature de la ligne de sources secondaires, une implémentation pratique de la WFS implique l'échantillonnage du continuum de sources secondaires. Il est en effet impossible de fabriquer la distribution continue de monopôles permettant d'obtenir une solution exacte au problème de la restitution d'un champ sonore dans le cadre du principe de Huyghens. Il faut par conséquent opérer une discrétisation de l'intégrale de Rayleigh I (eq. 1.2), ce qui implique un effet d'aliasing spatial dû à l'échantillonnage du banc de monopôles. Une étude menée par les chercheurs de l'université de Delft([12]) montre que la fréquence de coupure est donnée par la formule :

$$f_{aliasing} = \frac{c}{\Delta x (\sin \alpha_{max} + 1)} \quad (1.7)$$

avec  $\alpha_{max}$  l'angle d'incidence maximale du front d'onde reproduit relativement à la distribution de haut parleurs(cf FIG 1.7). En supposant que l'espacement entre les haut parleurs est de  $\Delta x = 10cm$ , la fréquence d'aliasing minimale est de  $f_{aliasing} = 1700Hz$ .

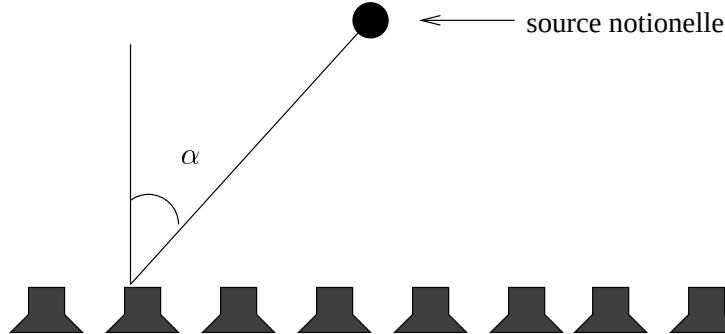


FIG. 1.7 – Angle d'incidence du front d'onde sur le banc de haut-parleurs

La figure 1.8 montre les effets de l'aliasing dans le domaine temporel et dans le domaine fréquentiel. L'ensemble des réponses fréquentielles mesurées au niveau du banc de micros révèle un comportement de champ diffus à partir de la fréquence d'aliasing. Les réponses impulsionnelles mesurées (figure du haut) montrent quant à elles la création d'un champ d'ondes diffus en aval du front d'onde principal. On a appliqué un filtre passe-haut sur ces réponses impulsionnelles dont la fréquence de coupure basse correspond à la fréquence d'aliasing pour mettre en valeur le phénomène. Notons que le premier front d'onde demeure cohérent avec le front d'onde idéal. Ceci implique par effet de précedence que les différences interaurales de temps

seront respectées dans l'ensemble de la zone de visibilité.

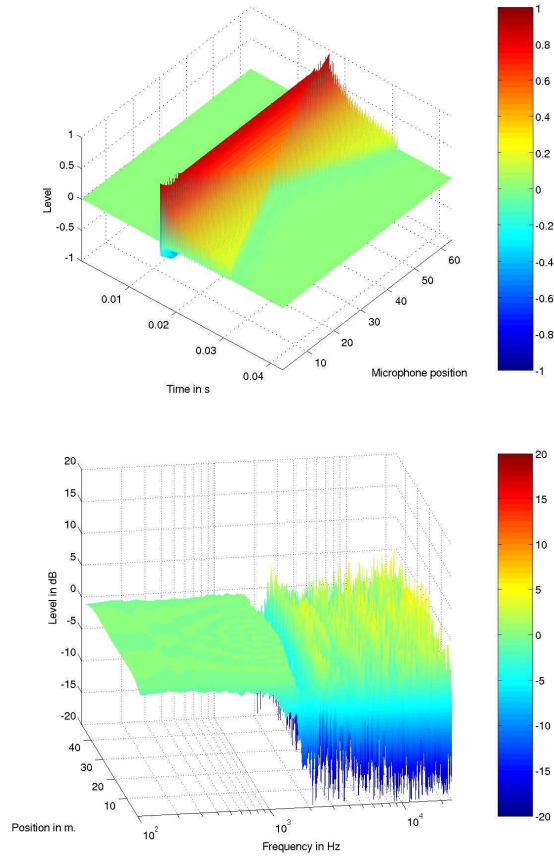


FIG. 1.8 – Mise en évidence du phénomène d'aliasing spatial dans le domaine temporel(haut) et dans le domaine fréquentiel (bas). Réponses mesurées sur un banc de microphones linéaire situé en face du système de diffusion. L'ensemble des réponses impulsionnelles (haut) est filtré passe-haut au dessus de la fréquence d'aliasing (2000Hz).

## 1.8 Le cas particulier des sources focalisées

Le concept de sources virtuelles focalisées est très important en WFS car il représente un type de reproduction sonore qui ne peut pas être réalisé avec des systèmes conventionnels. On peut en effet réaliser la synthèse d'une source située à *l'intérieur de l'espace de reproduction* par réciprocité entre la source et le receveur. Cependant ce type de reproduction sonore sort du cadre théorique prévu par le principe de Huyghens de par le fait que la source synthétisée se situe dans l'espace de restitution.

Néanmoins, l'expérience montre que l'on peut générer des sources focalisées avec un dispositif classique WFS. Il suffit de fabriquer un front d'onde convexe qui converge en le lieu de la source virtuelle que l'on cherche à reproduire. Alors un auditeur situé dans la zone en aval de cette source recevra le front d'onde concave qui aurait été émis par une source réelle située au même endroit (cf FIG. 1.9).

Au niveau de la fonction d'alimentation du banc de sources secondaires (eq. 1.5), deux différences avec le cas de sources situées hors zone de restitution apparaissent :

- Le signe du retard s'inverse
- La dépendance fréquentielle en  $\sqrt{jk}$  devient une dépendance en  $\sqrt{-jk}$

L'expression de la pression au point de reception telle qu'elle a été développée dans la partie 1.5 devient alors :

$$P(r_R) = \int_{-\infty}^{+\infty} g(r_L, r_R) \sqrt{\frac{-jk}{2\pi}} S(\omega) \cos \phi_{inc} \frac{e^{jkr}}{r} \frac{e^{-jk\Delta r}}{\Delta r} dx_L \quad (1.8)$$

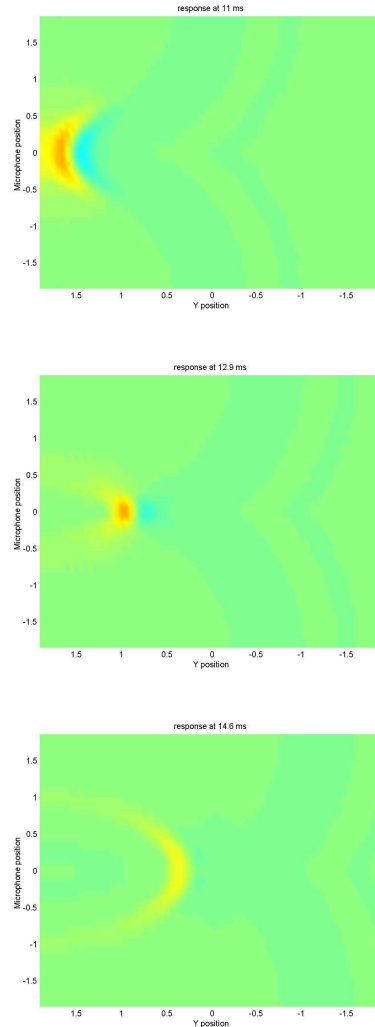


FIG. 1.9 – *Synthèse d'une source focalisée : Mesures sur un échantillonnage spatial de l'espace de restitution à trois temps donnés*

Notons que l'aire située entre le banc de sources secondaires et la source focalisée est à exclure de la zone d'écoute car le front d'onde qui s'y propage est à courbure inversée par rapport à celui de la source que l'on cherche à synthétiser. La figure 1.10 illustre le problème de la visibilité des sources secondaires. En comparant cette figure avec la figure 1.5, on voit que la zone de visibilité est fortement réduite dans le cas d'une source focalisée :

En ce qui concerne l'aliasing, si on compare la réponse impulsionnelle d'une source focalisée avec la réponse impulsionnelle d'une source non focalisée, on se rend compte que les effets de l'aliasing précèdent le front d'onde principal dans le cas de la source focalisée (cf FIG. 1.11). En vertu

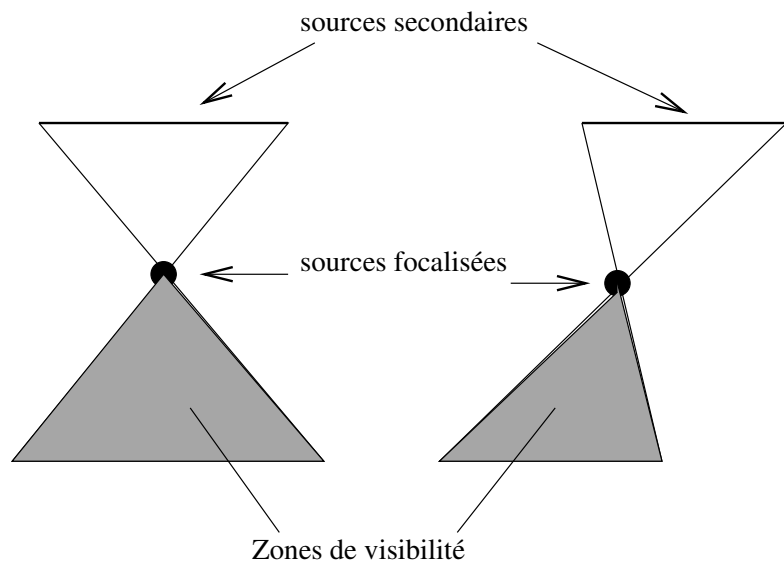


FIG. 1.10 – *Visibilité d'une source notionnelle focalisée*

de l'effet de précedence, ceci peut avoir comme effet de causer une mauvaise localisation de la source virtuelle, car un auditeur situé dans la zone d'écoute recevra un premier front d'onde venant de la mauvaise direction.

Il existe une méthode pour réduire les effets d'aliasing qui sont particulièrement gênants dans le cas des sources focalisées. Cette méthode consiste à appliquer une fenêtre de pondération spatiale sur le banc de haut-parleurs à partir de la fréquence d'aliasing, de manière à limiter l'émission de hautes fréquences par les haut-parleurs situés aux extrémités de la distribution. On observe sur la figure 1.12 les effets de cette pondération.

Nous reviendrons sur les problèmes de visibilité et d'aliasing spatial pour les sources focalisées dans la partie consacrée à l'interaction du système WFS avec la salle d'écoute. Etudions à présent l'implémentation pratique du système Wave Field Synthesis telle qu'elle est menée par l'équipe Acoustique des Salles de l'IRCAM.

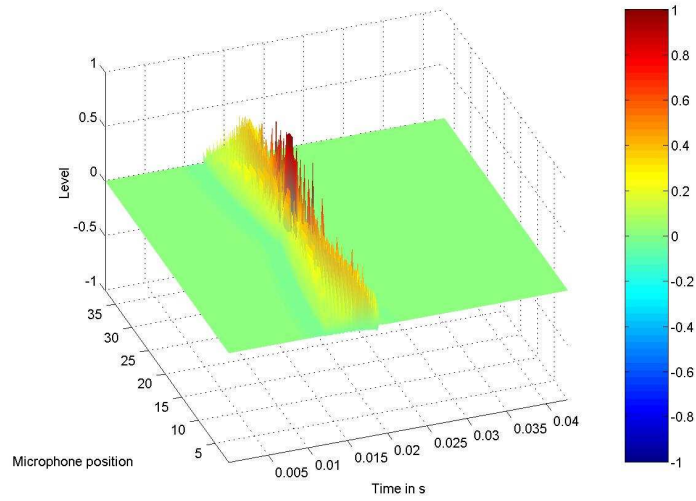


FIG. 1.11 – *Mise en evidence du phénomène d'aliasing spatial sur la réponse impulsionnelle d'une source focalisée mesurée sur un banc de microphones linéaire*

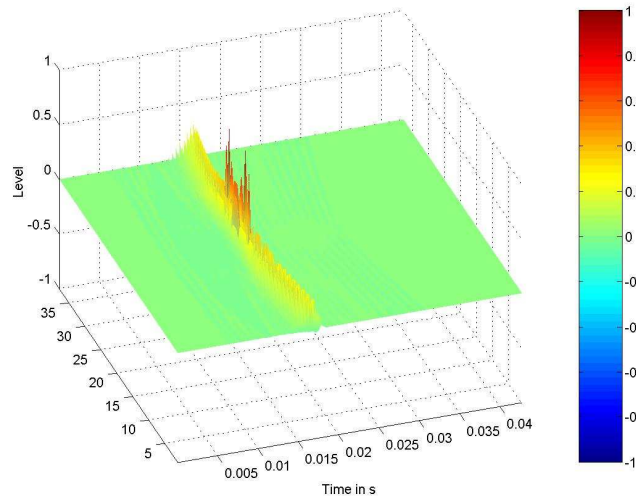


FIG. 1.12 – *Compensation de l'aliasing spatiale dans la reproduction d'une source focalisée*

## Chapitre 2

# Mise en oeuvre pratique

*La reproduction fidèle de scènes sonores tridimensionnelles utilisant la WFS requiert un grand nombre de canaux individuels de haut-parleurs. Cette partie décrit une méthode pratique pour réaliser un tel système en utilisant des haut-parleurs MAP (Multi Actuator Panels) avec du filtrage numérique. En théorie, les sources secondaires employées pour reproduire les champs sonores sont de type monopolaire. Ceci n'est pas le cas lorsqu'on utilise des haut-parleurs réels, qui présentent des figures de directivité variant avec la fréquence. On propose donc un algorithme permettant d'assurer l'égalisation multicanale des haut-parleurs le long d'une ligne de micros située dans l'espace de restitution. L'algorithme approche itérativement la solution idéale au niveau de la ligne visée. Un traitement séparé est envisagé pour assurer une cohérence perceptive du champ sonore au-delà de la fréquence d'aliasing en utilisant cette fois une égalisation individuelle des haut-parleurs.*

### 2.1 Les haut parleurs MAP

Les haut parleurs MAP sont constitués d'une plaque de mousse rigide collée à deux fines plaques de carton. Huit transducteurs, semblables à des haut-parleurs dénués de membrane, sont fixés sur la plaque avant du haut-parleur dans une configuration en "W" (cf FIG. 2.1).

L'écartement vertical entre les transducteurs est suffisamment faible pour que l'on puisse considérer qu'ils forment une ligne de sources secondaires. Parallèlement, cette configuration maximise la distance entre deux

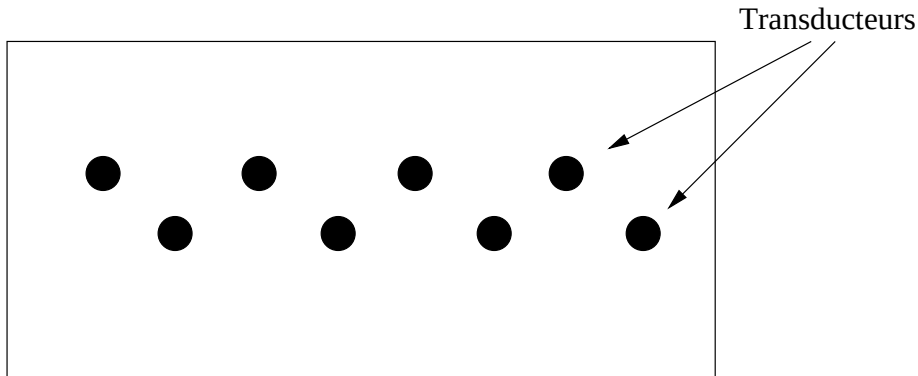


FIG. 2.1 – *Configuration des transducteurs sur un haut-parleur MAP*

haut-parleurs successifs, ce qui permet d'assurer leur indépendance respective.

Une étude menée par STUDER ([13]) a montré que les MAP peuvent être considérés comme des systèmes faiblement non-linéaires à entrées multiples et à sorties distribuées. Le niveau de distorsion dans la bande de fréquences 150-16000 Hz est suffisamment faible (moins de 1%) pour que l'on assimile les MAP à des systèmes multilinéaires. En d'autres termes, il sera possible de calculer les signaux de sortie par superposition, comme dans le cas de systèmes linéaires, tout en conservant une marge d'erreur acceptable.

L'inconvénient majeur des haut-parleurs MAP réside dans la reproduction des basses fréquences. En effet, le fonctionnement de ce type de transducteur dépend des modes de vibration de la plaque avant. Pour espérer reproduire de basses fréquences il faut donc agrandir la taille de cette plaque pour pouvoir baisser la fréquence de son premier mode de résonance. Cet agrandissement de la taille de la plaque implique alors d'utiliser plusieurs excitateurs par panneau. Ceci est dû au fait qu'il faut pouvoir fournir assez d'énergie mécanique pour produire un niveau sonore convenable en basses fréquences. D'autre part, la fréquence d'aliasing est inversement proportionnelle à l'écartement entre les haut-parleurs, ce qui implique de garder un écartement faible si l'on souhaite assurer une fréquence d'aliasing convenable (cf partie 1.7).

Les haut-parleurs MAP présentent des caractéristiques de directivité complexes et variant d'un excitateur à un autre. Ceci apparaît clairement sur la figure 2.2 qui présente les diagrammes de directivité de deux excitateurs adjacents sur un MAP.



D'autre part, étant donné que la fabrication des MAP demeure artisanale, on peut s'attendre à observer des différences entre deux MAP différents. Pour ces raisons, il faut envisager une méthode de synthèse de filtres qui ne présuppose rien sur la directivité du système de diffusion utilisé.

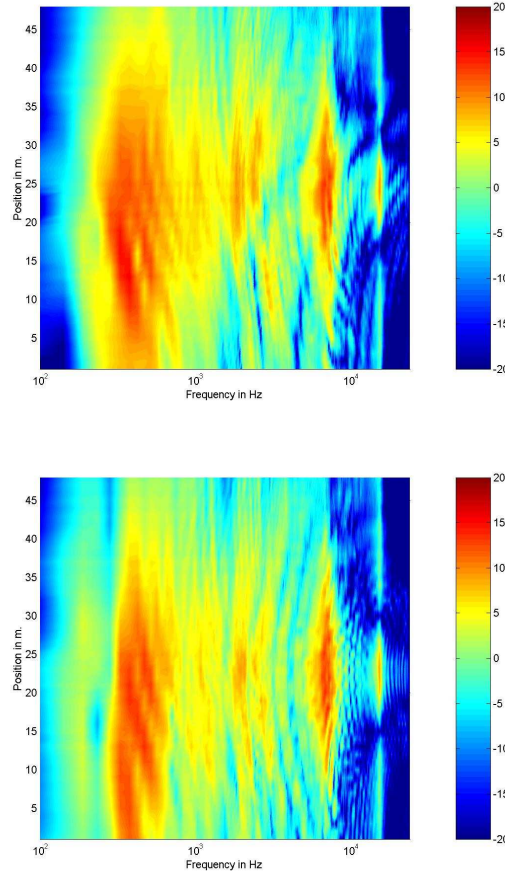


FIG. 2.2 – *Mesures de la réponse fréquentielle de deux excitateurs adjacents d'un haut-parleur MAP sur un banc linéaire de microphones omnidirectionnels situé en face du haut-parleur*

## 2.2 La Multiégalisation

Nous avons vu aux parties précédentes que la théorie sous-tendant la Wave Field Synthesis faisait appel à des réseaux de sources secondaires monopôlaires. En réalité, les transducteurs dont nous disposons pour implémenter un système WFS possèdent des figures de directivité complexes et variant

avec la fréquence. Il va donc s'agir de trouver un moyen de compenser les directivités des haut-parleurs utilisés.

Lorsque l'on dispose d'une distribution linéaire de haut-parleurs identiques pour mettre en oeuvre la WFS, on peut envisager de leur appliquer des filtres d'égalisation individuels choisis de manière appropriée ([14]). Il convient de noter que les filtres à appliquer dépendent de la position de la source virtuelle.

Dans [15], Corteel propose une autre méthode qui permet d'effectuer l'égalisation globale du dispositif. Cette méthode vise à assurer la reproduction d'un champ sonore donné sur une ligne située dans l'espace de restitution. Cela revient à effectuer la WFS au niveau de cette ligne et non plus au niveau des haut-parleurs. On est ainsi assuré du fait que le champ synthétisé se propagera correctement en aval de cette ligne. Un intérêt de cette technique réside dans le fait que l'on n'impose pas de condition a priori sur la directivité des haut-parleurs ou sur la forme du dispositif de restitution. Voyons à présent l'algorithme proposé pour mettre en oeuvre cette approche dite de "multiégalisation".

### 2.2.1 L'Algorithme de Traitement Multicanal

On considère un système de  $M$  haut-parleurs alimentés par des signaux filtrés. On mesure le champ sonore résultant avec une distribution de  $L$  microphones disposés le long d'une ligne dans l'espace de restitution. Le problème consiste à trouver les coefficients des  $M$  filtres appliqués aux haut-parleurs de manière à ce que les  $L$  réponses impulsionnelles mesurées par les microphones approchent au mieux les  $L$  réponses impulsionnelles "idéales" de la source virtuelle aux positions des microphones. On utilise pour cela un algorithme multicanal itératif qui vise à minimiser l'erreur quadratique des réponses impulsionnelles au niveau des microphones.

La méthode proposée possède un certain nombre de limitations en pratique à cause de la troncature et la discrétisation spatiale du banc de haut-parleurs et du banc de microphones (cf parties 1.6, 1.7). En dehors de la zone de visibilité, la procédure d'optimisation décrite ci-dessus ne peut pas produire des résultats valables du fait que cela est *physiquement impossible*. De plus, le nombre fini de haut-parleurs introduit un sous-échantillonnage du champ sonore entraînant un phénomène d'aliasing spatial au-delà d'une certaine fréquence (fréquence d'aliasing). Au delà de la fréquence d'aliasing, la méthode de filtrage inverse proposée ne produit que des solutions *locales*. Le champ sonore devient totalement imprévisible en dehors des points de

contrôle (positions des micros). Il devient donc important d'envisager un traitement séparé dans les hautes fréquences, qui sera décrite dans la partie suivante.

## 2.2.2 Traitement Hautes Fréquences

La figure 2.3 montre des exemples de réponses fréquentielles obtenues en deux points de mesure proches pour une distribution simulée de monopôles idéaux en WFS([13]). Ces réponses présentent des caractéristiques de filtre en peigne à cause d'interférences entre les ondes retardées. Ces interférences dépendent fortement de la position, ce qui explique la grande différence enregistrée entre les deux points de mesure. Ceci implique que le champ sonore au dessus de la fréquence d'aliasing n'est pas prédictible. La procédure d'égalisation s'attache donc à assurer une énergie acoustique reproduite par bandes de fréquence cohérente avec la source virtuelle cible.

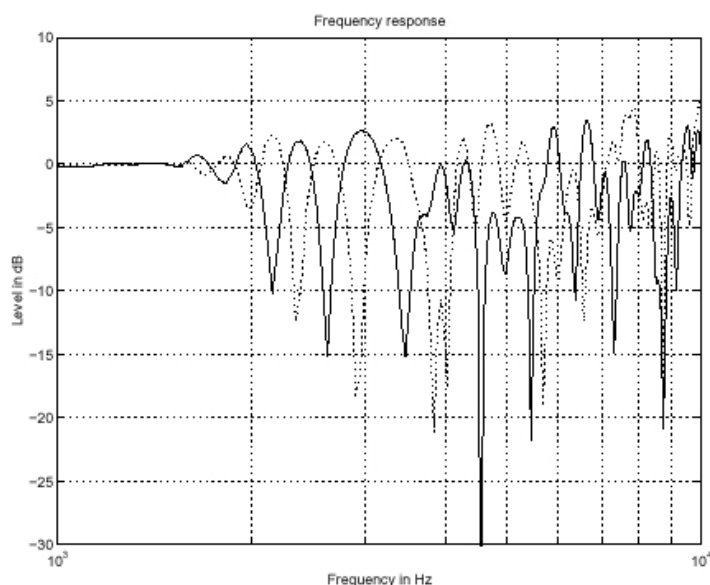


FIG. 2.3 – Réponse fréquentielles (1000-10000 Hz) d'un champ sonore produit en WFS mesurées à 10cm l'une de l'autre

On propose donc une procédure de synthèse de filtre différente de celle vue à la section 2.2.1 pour traiter la partie haute fréquence du spectre. Cette procédure implique une égalisation individuelle des haut-parleurs suivie d'un contrôle énergétique du champ sonore reproduit. Cette méthode

visé à récupérer un champ sonore cohérent en termes de perception et non en termes d'exactitude physique, ce qui serait impossible.

### 2.2.3 Résultats

La méthode proposée est mise en oeuvre sur un dispositif de trois MAP munis respectivement de huit haut-parleurs. On mesure les résultats sur une distribution de 24 microphones espacés de 10cm chacun et placés à 1.5m des haut parleurs (cf FIG. 2.4).

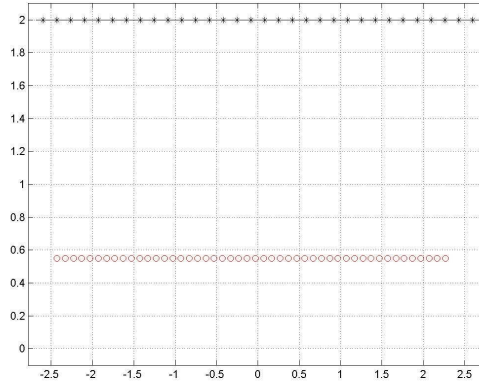


FIG. 2.4 – *Configuration de mesure de quatre haut-parleurs MAP sur 48 positions de microphones. Les croix représentent les haut-parleurs, les cercles représentent les microphones*

Pour les besoins de la démonstration, on simule une onde plane se propageant perpendiculairement au banc de haut parleurs. Les figures 2.5 et 2.6 montrent les réponses fréquentielles des panneaux aux positions des microphones pour deux approches possibles. La figure 2.5 correspond à une approche utilisant des filtres issus directement de la théorie WFS. Ces filtres modélisent les retards et atténuations dus à la propagation, la pondération du banc de haut-parleurs visant à limiter la diffraction ainsi que la compensation en  $\sqrt{j\omega}$  en dessous de la fréquence d'aliasing pour tenir compte de l'approximation de phase stationnaire. La figure 2.6 correspond à l'approche dite de multiégalisation décrite aux sections 2.2.1 et 2.2.2. On observe sur ces figures une fréquence d'aliasing spatiale aux alentours de 2000Hz. En dessous de cette fréquence, l'approche multiégalisation affiche une réponse fréquentielle plate, ce qui n'est pas le cas pour l'approche WFS sans égalisation. Au dessus de la fréquence d'aliasing, le champ sonore revêt une allure de champ diffus dans les deux cas. Cependant la simulation non-

égalisée présente des fluctuations plus importantes que la simulation égalisée, ce qui se traduit au niveau perceptif par une coloration de la source virtuelle synthétisée. Cette démonstration révèle donc l'intérêt d'adopter une approche d'égalisation multicanale en basses fréquences couplée à l'égalisation individuelle des excitateurs en haute fréquences. Elle correspond d'ailleurs au standard proposé dans le cadre du projet CARROUSO ([2]).

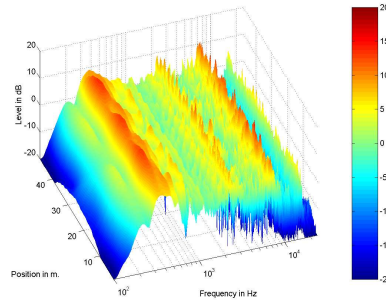


FIG. 2.5 – Réponse fréquentielle des haut-parleurs MAP simulée sur une distribution linéaire de micros utilisant des filtres issus directement de la théorie WFS pour une onde plane se propageant perpendiculairement au banc de microphones

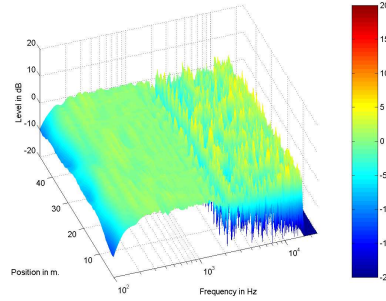


FIG. 2.6 – Réponse fréquentielle des haut-parleurs MAP simulée sur une distribution linéaire de micros utilisant des filtres issus de la méthode dite de “multiégalisation” pour une onde plane se propageant perpendiculairement au banc de microphones

## Chapitre 3

# Restitution de l'effet de salle

*Lorsqu'une source rayonne dans un espace fermé, le champ acoustique est constitué d'un son direct et d'un ensemble de composantes réfléchies par les parois de la salle considérée. On nomme "effet de salle" l'ensemble de ces composantes réfléchies. De manière générale, l'effet de salle renseigne l'auditeur sur l'endroit dans lequel il se trouve, que ce soit en termes de la taille d'une pièce, la réverbérance ou non de ses murs ou sa forme géométrique. Il le renseigne également sur la distance à laquelle il se trouve de la source. Il est donc indispensable de s'intéresser à l'effet de salle lorsque l'on vise à recréer des scènes sonores réalistes, et ce quel que soit le format de rendu choisi (binaural, stéréophonique, transaural, WFS...). Dans ce chapitre, on décrira trois modélisations possibles de l'effet de salle (architecturale, physique ou perceptive). Ces descriptions seront suivies d'une présentation du Spatialisateur qui est un processeur d'acoustique virtuelle conçu pour l'environnement MAX/MSP. On abordera alors les méthodes de reproduction d'effet de salle existantes en Wave Field Synthesis. Une attention particulière sera portée sur la méthode adoptée par l'IRCAM dans le cadre du projet CARROUSO. Pour finir, le problème de l'interaction du dispositif WFS avec la salle de restitution sera introduit à travers un exemple illustrant une installation sonore WFS dans une salle de concert.*

### 3.1 Modélisation de l'effet de salle

De manière générale, l'effet de salle correspond à l'ensembles des composantes du champ sonore réfléchi lors de l'interaction d'une source avec un espace clos. Ces composantes sont classiquement séparées en deux parties :

- Les réflexions précoces (ou primaires), dont la direction et le temps d'arrivée renseignent l'auditeur sur la position de la source ainsi que sur la taille de la pièce ([16])
- Les réflexions diffuses et la réverbération qui renseignent l'auditeur sur la taille de la pièce ainsi que ses caractéristiques d'absorption.

Un effet de salle peut être décrit selon diverses méthodes qui sont classables en trois catégories : modèles architecturaux, modèles physiques et modèles perceptifs([17]).

Les *modèles architecturaux* s'attachent à décrire l'effet de salle partir des données géométriques et des caractéristiques d'absorption de la pièce. A partir de la position et de l'orientation d'une source directive donnée, il est en effet possible de calculer l'effet de salle perçu par un auditeur ayant une position et une orientation de tête connue. Diverses méthodes sont utilisables à cet effet (sources images, rayons, cônes, éléments finis). Notons que le modèle entier doit être recalculé lors d'un déplacement de source, ce qui constitue une forte contrainte dans le cadre d'une éventuelle interaction en temps-réel.

Les *modèles physiques* ont pour objectif de décrire le champ acoustique à partir de la mesure de réponses impulsionnelles. Selon les dispositifs mis en place, la caractérisation du champ sera plus ou moins fine et la zone décrite plus ou moins étendue. L'inconvénient majeur de ce type de méthode réside dans le fait que la description fine d'une pièce nécessite de mesurer puis de stocker des quantités importantes de données sous la forme de réponses impulsionnelles. Dans [18], Edo Hulsebos suggère l'utilisation d'antennes de microphones linéaires ou circulaires pour mesurer la réponse impulsionnelle d'une salle. En décomposant le champ acoustique ainsi mesuré sur une base géométrique donnée (harmoniques cylindriques ou sphériques, ondes planes), il montre qu'il est possible par extrapolation de décrire ce champ sur une zone étendue. On convolve ensuite un signal enregistré en conditions quasi-anéchoïques (microphone de proximité) avec les réponses impulsionnelles issues de la phase d'extrapolation. Enfin, on restitue le résultat sur un système de haut-parleurs WFS. Il convient de noter que l'utilisation d'une distribution planaire de microphones pour la captation d'une scène sonore tridimensionnelle peut induire le repliement de l'espace 3D dans le plan horizontal. Ceci se traduira notamment par une modification du rapport son direct/ son réverbéré du champ sonore reproduit.

Les *modèles perceptifs* permettent de décrire l'effet de salle à l'aide d'un ensemble de paramètres perceptifs. Des paramètres classiquement utilisés à cet effet sont l'enveloppement, le temps de réverbération d'une salle,

ou la présence d'une source ([19]). On estime les paramètres perceptifs à partir de l'analyse d'une réponse impulsionnelle en termes de distribution d'énergie temps/fréquence. Cette description à l'avantage d'être moins coûteuse en calculs que les méthodes précédentes au niveau de la restitution. En revanche, l'effet de salle restitué par ce type de méthode n'est valable qu'au point de mesure. Ce type de méthode permet de reproduire aisément des effets de salles dont on aurait préalablement mesuré les paramètres perceptifs. Elle permettent également de générer l'acoustique d'une salle virtuelle, ce qui peut être intéressant pour des applications de post-production.

Remarquons enfin qu'un modèle physique peut être dérivé d'un modèle architectural; de même un modèle perceptif peut être déduit d'un modèle physique.

## 3.2 Le Spatialisateur

Le *Spatialisateur* est un processeur d'acoustique virtuelle disponible sur l'environnement MAX/MSP. Il permet au travers d'une interface graphique d'effectuer la spatialisation d'une source utilisant diverses techniques (binaurale, ambisonics, stéréophonie, 5.1,...). Ce processeur donne également accès à un contrôle direct de l'effet de salle par l'intermédiaire de paramètres perceptifs réglables (cf figure 3.1).

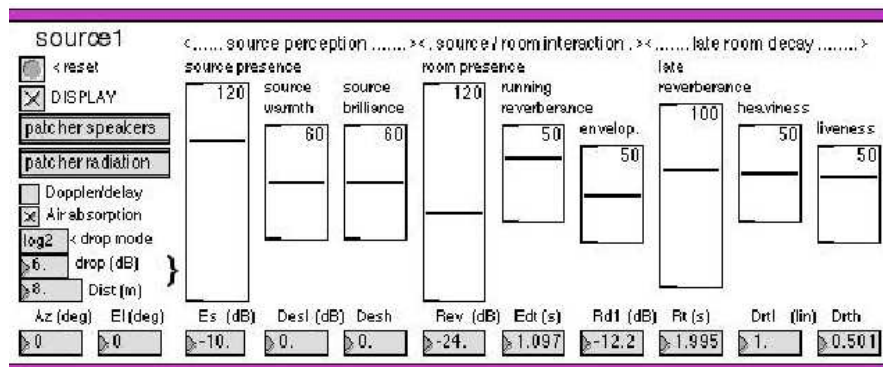


FIG. 3.1 – Interface de contrôle des paramètres perceptifs intégré au *Spatialisateur*

Le calcul de l'effet de salle dans le *Spatialisateur* utilise directement des modèles perceptifs au lieu de passer par une description physique ou géométrique de l'environnement. Trois modules distincts sont chargés de générer l'effet de salle :

- Le module **early** génère les premières réflexions



- Le module **cluster** synthétise les réflexions diffuses
- Le module **reverb** génère la réverbération tardive

La figure 3.2 représente une réponse impulsionnelle découpée en quatre sections temporelles. La section  $R_0$  correspond au son direct. La section  $R_1$  correspond aux premières réflexions qui sont générées par le module **early**. La section  $R_2$  correspond aux réflexions diffuses, générés par le module **cluster**. La section  $R_3$  correspond à la réverbération tardive, qui est générée par le module **reverb**.

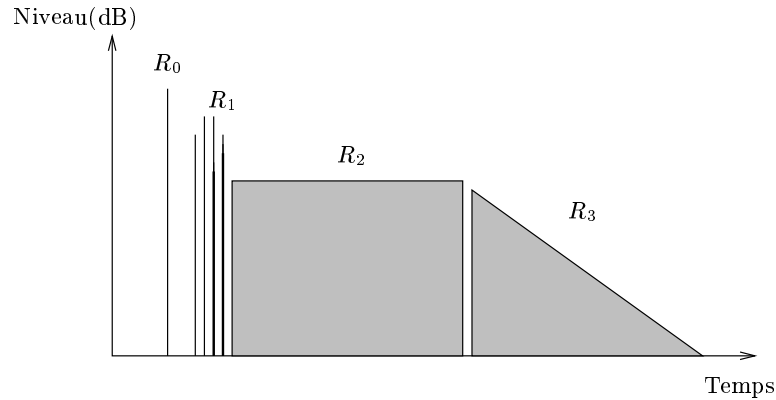


FIG. 3.2 – Représentation de l'effet de salle

Le *Spatialisateur* permet également de générer des trajectoires de sources en temps réel. Grâce au module **SpatCirc** représenté à la figure 3.3 on visualise jusqu'à quatre sources d'une scène sonore. On peut alors composer des déplacements des sources de la scène et s'assurer graphiquement de leur cohérence.

Nous verrons dans la section suivante comment le *Spatialisateur* a été intégré au système WFS de l'IRCAM.

### 3.3 Reproduction de l'effet de salle et moteur de rendu spatial pour le démonstrateur IRCAM

Un démonstrateur WFS constitué de 48 canaux a été développé à l'IRCAM. Dans ce dispositif, l'effet de salle est géré par le *Spatialisateur*. On s'affranchit donc des problèmes liés à la mesure physique, préférant décrire l'effet de salle selon des critères perceptifs. La figure 3.4 présente la structure du moteur de rendu spatial utilisé. Un son anéchoïque présenté à l'entrée du système est envoyé vers le module *Effet de Salle* qui se charge de la

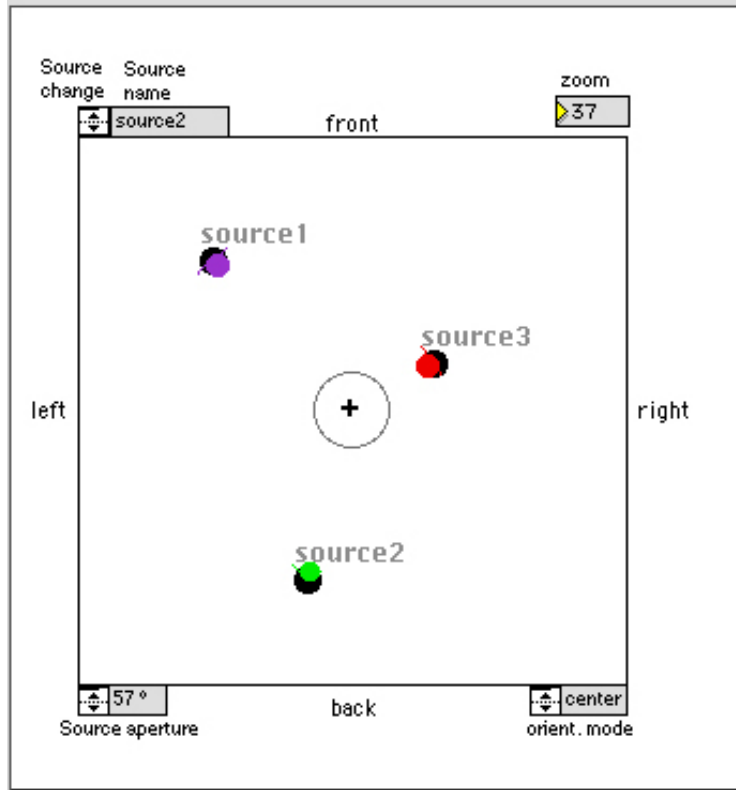


FIG. 3.3 – Interface de contrôle des déplacements de sources intégré au *Spatialisateur*

répartition de l'effet de salle sur  $N$  canaux. Il est également envoyé vers le second module *Positionnement* en vue de pouvoir générer le son direct.

Le module *Positionnement* génère d'une part  $N$  haut-parleurs virtuels sous la forme d'ondes planes pour restituer les  $N$  canaux d'effets de salle. D'autre part, il rajoute une information sur la position de la source permettant ainsi de générer le son direct sous la forme d'une source monopolaire. Il réalise ces deux traitements grâce à des batteries de filtres calculés à l'avance et stockés dans une base de données. Ainsi les  $N$  ondes planes qui restituent l'effet de salle et la source monopolaire qui restitue le son direct possèdent chacun un jeu de  $M$  filtres. Ces jeux de filtres contiennent les informations sur l'angle d'incidence de chaque onde plane et la position de la source monopolaire. Ils réalisent également l'égalisation du système de diffusion par la méthode décrite à la partie 2.2. La figure 3.5 représente la reproduction d'un son direct via une source virtuelle ponctuelle ainsi que la reproduction d'un l'effet de salle via un ensemble d'ondes planes de direction fixée.

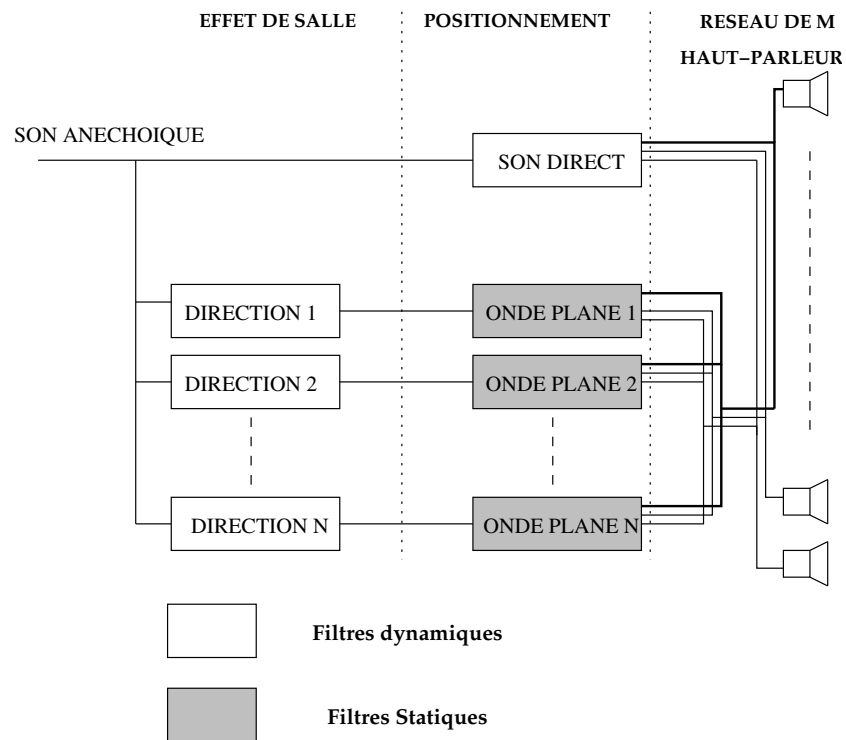


FIG. 3.4 – Architecture du moteur de rendu spatial pour le dispositif WFS de l'IRCAM

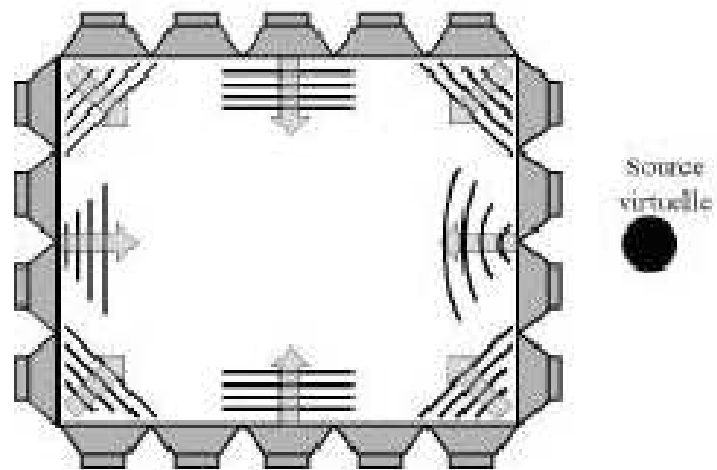


FIG. 3.5 – Reproduction du son direct et de l'effet de salle

L'intérêt majeur de cette structuration de moteur spatial est qu'elle permet le contrôle dynamique d'une scène sonore :

- Le module *Positionnement* fait appel à une banque de filtres précalculés. Il permet un contrôle en temps réel des déplacements d'une source virtuelle dans l'espace de restitution. On utilise à cet effet le module *SpatCirc* du *Spatialisateur* qui envoie des informations sur la position de la source vers la module de *Positionnement*. Celui-ci peut alors charger dynamiquement le jeu de  $M$  filtres correspondant à la position envoyé.
- Le module *Effet de Salle* génère les réponses impulsionnelles des  $N$  filtres correspondant au canaux d'effet de salle avec le *Spatialisateur*. Il permet une modification dynamique de la répartition d'effet de salle lors d'un déplacement de source. On peut également agir directement sur les paramètres perceptifs décrits à la partie 3.2.

### 3.4 Intégration d'un dispositif WFS dans une salle d'écoute réelle

La figure 3.6 présente deux installations possibles d'un dispositif WFS dans une salle de concert. L'implémentation de ce type de dispositif sous-entend la maîtrise de tous les problèmes évoqués dans les chapitres 1 et 2 qui sous-tendent la reproduction de sources virtuelles à l'intérieur ou à l'extérieur d'une la salle de restitution. Elle peut nécessiter également de savoir comment créer (ou recréer) un effet de salle ; les sections précédentes de ce chapitre ont été consacrées à cette question. Ceci étant, dès lors que l'on décide de placer une source virtuelle (focalisée ou non) dans une salle d'écoute réel, il se pose un certain nombre de problèmes nouveaux sur lesquels il convient de se pencher :

- D'une part se pose le problème de la **directivité des sources reproduites** qui n'est pas forcément monopolaire
- D'autre part se pose le problème de la **cohérence de l'effet de salle**. Ceci est particulièrement important dans le contexte des applications de réalité virtuelle, qui nécessitent une cohérence avec le lieu d'écoute qui soit éventuellement visio-auditive.

En ce qui concerne la directivité des sources reproduites, on propose d'appliquer la méthodologie décrite dans [20] qui consiste à recréer une directivité à partir de la combinaison de directivités élémentaires. Cette étude fera l'objet du chapitre 4.

La question de la cohérence de l'effet de salle sous-entend quant à elle d'étudier l'interaction du dispositif WFS avec le lieu d'écoute. Ceci per-

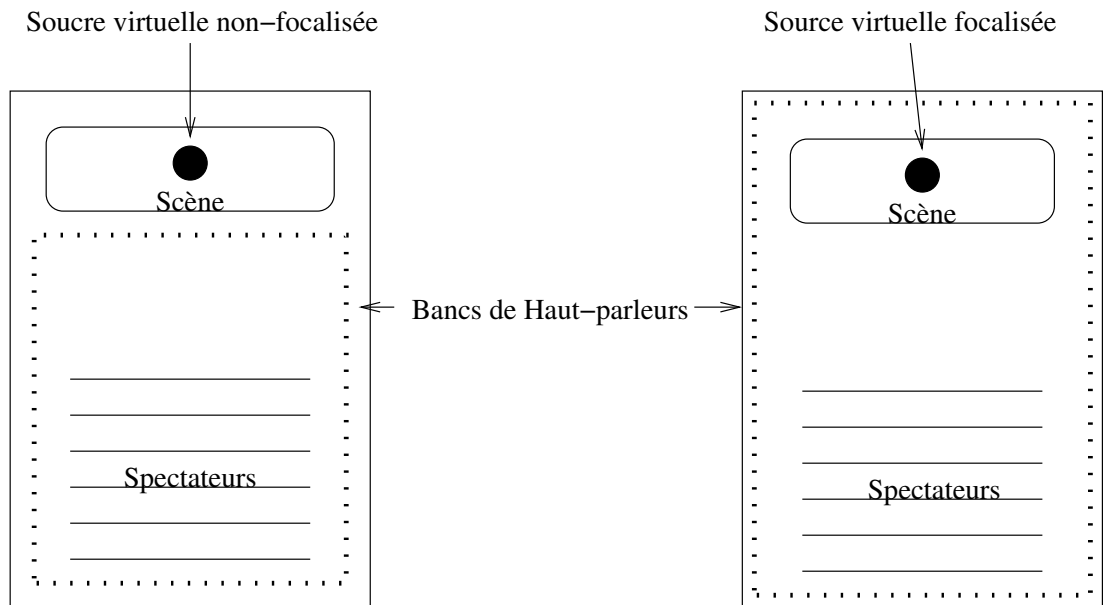


FIG. 3.6 – Deux installations possibles d’un dispositif WFS dans une salle de concert. Le dispositif de gauche est conçu pour la synthèse de sources non-focalisées sur une scène. Le dispositif de droite est conçu pour la synthèse de sources focalisées sur une scène.

mettra de caractériser puis compléter les composantes manquantes de l’effet de salle lors de la synthèse d’une source virtuelle. Cette question fera l’objet du chapitre 5.

## Chapitre 4

# Directivité des sources

*Lors de la reproduction d'une scène sonore, il peut s'agir par souci de réalisme de s'intéresser à la directivité des sources reproduites. Dans ce chapitre, on évoquera de manière théorique une méthode permettant de décomposer un champ sonore en harmoniques sphériques (pour un champ tridimensionnel) ou cylindriques (pour un champ bidimensionnel). Quelques résultats obtenus après l'implémentation d'une directivité variable pour des sources reproduites en WFS seront présentés. Les directivités visées sont celles des harmoniques cylindriques de base (monopôle, dipôle, multipôle d'ordre  $n$ ). Notons que ce sujet a déjà été abordé par Nicolas Misdariis dans son travail sur la "Timée" ([20]). À terme, on vise la reproduction de caractéristiques de directivité d'instruments réels par combinaison linéaire des harmoniques cylindriques de base. Des applications plus libres pourront également être envisagées (effets d'éclairage sonore variables dans une salle, instruments virtuels..).*

### 4.1 Décomposition d'un champ sonore en harmoniques sphériques ou cylindriques

D'après Jérôme Daniel dans [21], le cadre mathématique fourni par la décomposition d'un champ en harmoniques sphériques est à la fois solide et générique. Il fournit à l'approche ambisonique "*un formalisme naturel pour une extension vers des degrés supérieurs de résolution spatiale du champ sonore*". Ceci est possible notamment du fait que l'approche ambisonique est avant tout une "*représentation rationnelle du champ acoustique*". De la même manière que la Wave Field Synthesis, l'approche ambisonique

est basée sur une représentation physique du champ sonore. On retombe d'ailleurs sur la même solution si l'on considère une représentation ambisonique d'ordre infini et une distribution continue de sources secondaires entourant l'espace de reproduction en WFS ([22]).

Les harmoniques sphériques correspondent aux solutions de l'équation des ondes dans le cas particulier où la description du champ sonore est restreinte à la distribution angulaire. Cette restriction conduit à la notion de fonction de directivité. Ces fonctions constituent une base orthogonale qui engendre tout l'espace des fonctions de directivité et autorise une analyse et une synthèse appelées Transformée de Fourier Sphérique ([23]).

Dans le cadre de la Wave Field Synthesis, on dispose d'une distribution linéaire de haut-parleurs. Bien que le champ sonore effectif reproduit soit tridimensionnel (cf 5.1), on vise à restituer une scène sonore de manière exacte uniquement dans le plan d'écoute. Par conséquent, l'implémentation d'une directivité ne pourra être réalisée que de manière bidimensionnelle. On s'intéressera donc aux représentants 2D des harmoniques sphériques, qui se nomment *harmoniques cylindriques*. Les fonctions de directivités de base d'une décomposition en harmoniques cylindriques sont les suivantes ([21]) :

$$\begin{aligned} Y_m^1(\theta) &= \cos(m\theta) \quad \text{pour } m \geq 0 \\ Y_m^{-1}(\theta) &= \sin(m\theta) \quad \text{pour } m \geq 1 \end{aligned} \quad (4.1)$$

On a vu à la partie 1.4 que la pression rayonnée par une source notionnelle dans un espace tridimensionnel était donnée par la formule suivante :

$$P(r, \omega) = S(\omega) D(\theta, \phi, \omega) \frac{\exp -jkr}{r}$$

avec :

- $S(\omega)$  fonction d'alimentation de la source (signal qui serait capté par un microphone de proximité)
- $D(\theta, \phi, \omega)$  fonction de directivité de la source

En se restreignant au plan d'écoute ( $\phi = \frac{\pi}{2}$ ), la pression rayonnée par une source notionnelle possédant la directivité d'une harmonique cylindrique de base s'exprime donc sous la forme suivante (On suppose  $D$  indépendante de la fréquence) :

$$P(r, \omega) = S(\omega) D(\theta) \frac{\exp -jkr}{r}$$

avec :

- $D(\theta) = Y_m^1(\theta) \quad \text{ou} \quad D(\theta) = Y_m^{-1}(\theta)$

En faisant des combinaisons linéaires des figures de directivité de base, on peut réaliser des figures de directivité complexes correspondant à des instruments réels. Voyons à présent les résultats de l'implémentation de la directivité dans des sources reproduites en WFS.

## 4.2 Présentation des Résultats WFS

Pour illustrer l'implémentation d'une directivité en WFS, on choisit d'effectuer une simulation en se plaçant dans la configuration représentée à la figure 4.1.

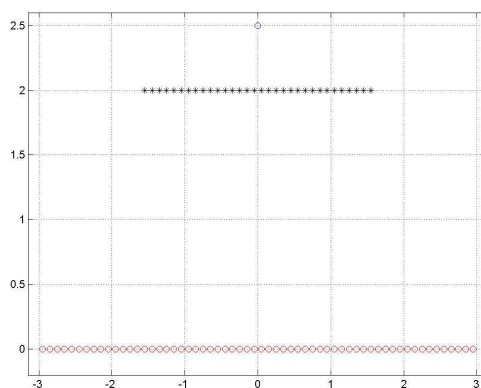


FIG. 4.1 – Configuration choisie pour mettre en évidence la directivité de sources en WFS. Les croix représentent les 24 haut-parleurs espacés de 0.1m, les cercles rouges représentent les 48 microphones espacés de 0.1m, le cercle bleu représente la source reproduite.

La source virtuelle située à 2.5 mètres de l'origine est reproduite par un banc de 24 haut-parleurs monopôlaire situés à 2 mètres de l'origine. Le résultat est mesuré sur un banc de 48 microphones omnidirectionnels situés à l'origine de l'axe perpendiculaire au deux bancs. Les délais et atténuations appliqués aux haut-parleurs sont calculés d'après les formules théoriques de WFS (cf chapitre 1). La reproduction d'un monopôle sur ce dispositif est représentée sur les figures 4.2 et 4.3.

Le monopôle est la figure de directivité la plus simple ; elle correspond à une harmonique cylindrique d'ordre 0. L'harmonique d'ordre 1 correspond au dipôle. Les figures 4.5 et 4.6 représentent la reproduction d'un dipôle orienté à  $0^\circ$  par rapport à l'axe perpendiculaire au banc de haut-parleurs



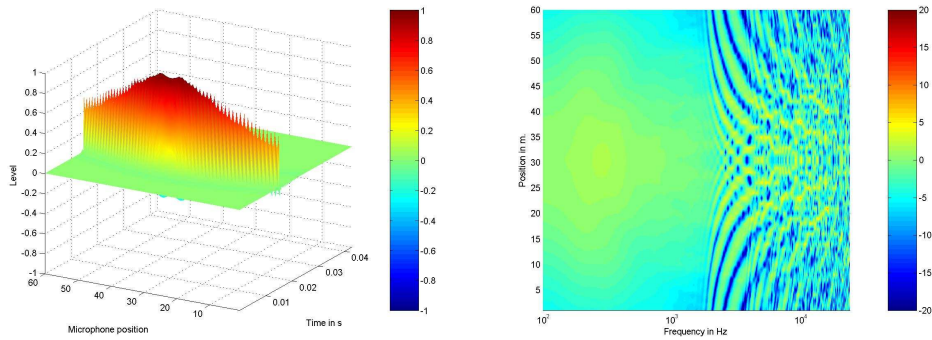


FIG. 4.2 – Réponse impulsionnelle de monopôle

FIG. 4.3 – Réponse fréquentielle de monopôle

(cf figure 4.4.).

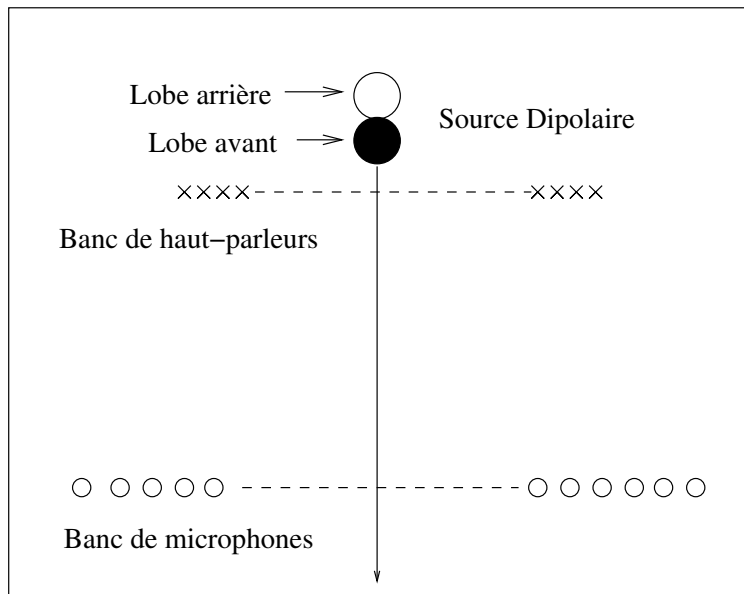


FIG. 4.4 – Dipole orienté à  $0^\circ$  reproduit sur le dispositif WFS proposé

On observe sur la figure 4.5 que le dipôle orienté à  $0^\circ$  ressemble à une source monopolaire au point de vue de sa réponse impulsionnelle. Cependant, la réponse fréquentielle (FIG.4.6) témoigne d'une atténuation entre 600 et 2000 Hz mesurée sur les microphones des bords. Ceci est dû à l'influence d'un lobe arrière en opposition de phase avec le lobe avant qui de ce fait resserre la directivité du dipole sur l'axe.

Un dipôle orienté à  $90^\circ$  par rapport l'axe perpendiculaire au banc de

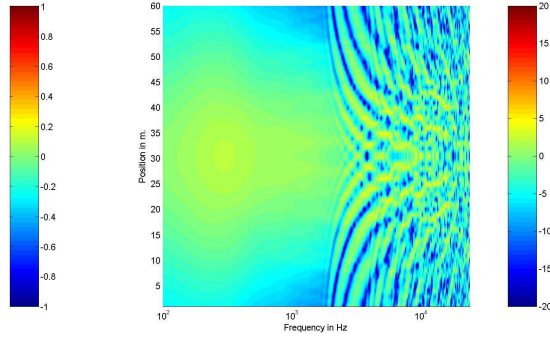
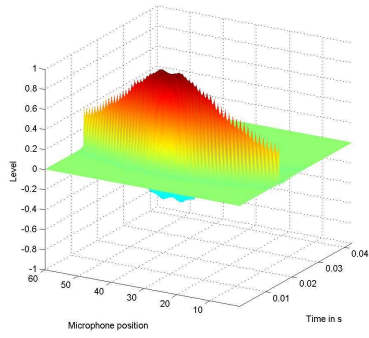


FIG. 4.5 — Réponse impulsionnelle d'une source dipolaire orientée à  $0^\circ$  FIG. 4.6 — Réponse fréquentielle d'une source dipolaire orientée à  $0^\circ$

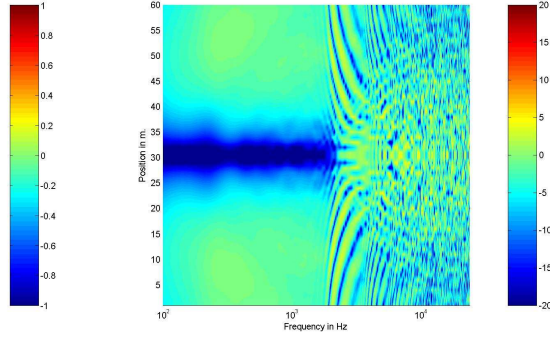
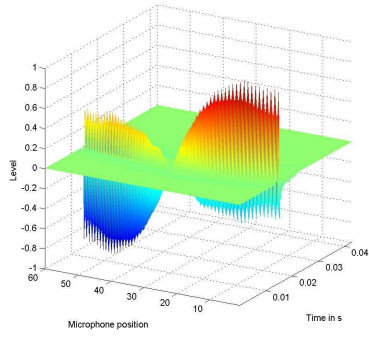


FIG. 4.7 — Réponse impulsionnelle d'une source dipolaire orientée à  $90^\circ$  FIG. 4.8 — Réponse fréquentielle d'une source dipolaire orientée à  $90^\circ$

haut-parleurs est représenté sur les figures 4.7 et 4.8. Sur la réponse impulsionnelle, on observe deux fronts d'onde en opposition de phase. Des artefacts dûs à l'aliasing sont également apparents (front d'onde de niveau faible qui succède au front d'onde principal). La réponse fréquentielle révèle quant à elle deux lobes distincts jusqu'à la fréquence d'aliasing. Une "zone d'ombre" apparaît également très clairement au niveau des microphones centraux. Ce type de figure de directivité être utilisée pour faire varier l'interaction du dispositif de restitution avec la salle. En modifiant l'angle d'orientation du dipole lors de la reproduction, on peut choisir d'illuminer la zone de restitution avec le son direct, ou bien de laisser cette zone dans l'ombre. Lorsque la zone de restitution est dans l'ombre, on fait alors entendre uniquement l'effet de salle. Ce type de variation d'atmosphère sonore peut présenter un réel intérêt au niveau de la composition de musiques spatialisées.

En augmentant l'ordre des harmoniques cylindriques reproduites, on

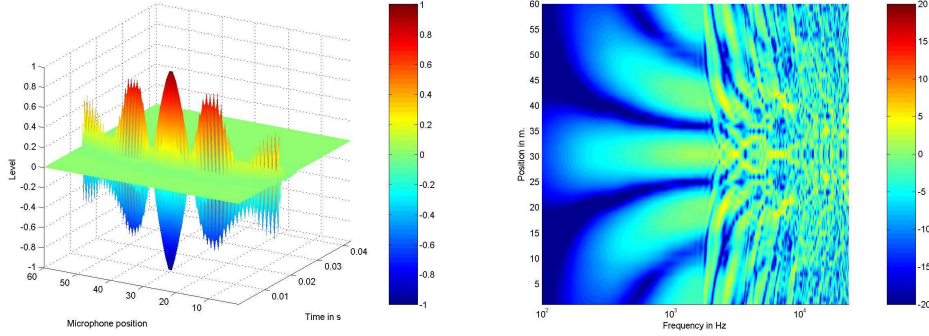


FIG. 4.9 — Réponse impulsionnelle d'une harmonique cylindrique d'ordre 8

FIG. 4.10 — Réponse fréquentielle d'une harmonique cylindrique d'ordre 8

arrive à des figures de directivités plus riches en zones d'ombre et d'éclairage. Les figures 4.9 et 4.10 représentent la reproduction d'une harmonique cylindrique d'ordre 8 sur le dispositif proposé.

### 4.3 Amélioration de la reproduction de sources directives par l'approche de multiégalisation

L'algorithme de traitement multicanal évoqué à la partie 2.2.1 a été mis au point principalement pour pouvoir compenser les directivités variables des haut-parleurs utilisés lors d'une implémentation pratique. Il est cependant possible d'utiliser cet algorithme lors de simulations utilisant des bancs de haut-parleurs monopôlaire. Les figures 4.11, 4.12, 4.13, 4.14 présentent une comparaison entre l'approche de multiégalisation et l'approche WFS sans égalisation. La source reproduite est une harmonique cylindrique d'ordre 5, et le dispositif de restitution est celui représenté à la figure 4.1. On observe principalement une réduction du phénomène d'aliasing lorsqu'on applique la multiégalisation au dispositif de restitution. Cette réduction se manifeste par la limitation du front d'onde succédant au front d'onde principal au niveau de la réponse impulsionnelle (figures 4.11, 4.13), ainsi qu'une amélioration de la directivité au niveau des bords de la zone de restitution (figures 4.12, 4.14).

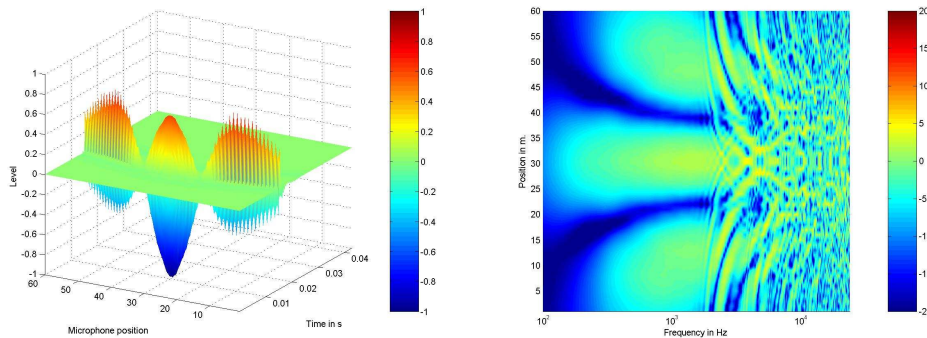


FIG. 4.11 – Réponse impulsionnelle d'une harmonique cylindrique d'ordre 5 en théorie WFS

FIG. 4.12 – Réponse fréquentielle d'une harmonique cylindrique d'ordre 5 en théorie WFS

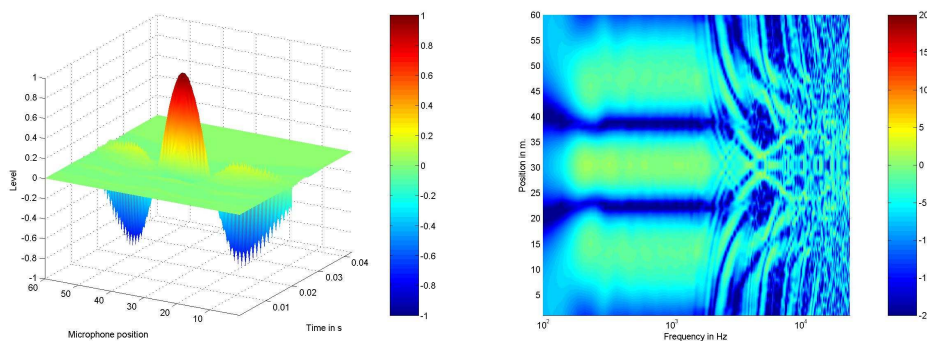


FIG. 4.13 – Réponse impulsionnelle d'une harmonique cylindrique d'ordre 5 après multiégalisation

FIG. 4.14 – Réponse fréquentielle d'une harmonique cylindrique d'ordre 5 après multiégalisation

### 4.3.1 Reproduction de sources directives sur le démonstrateur IRCAM

La dernière étape de ce travail a consisté à effectuer des tests d'écoute sur les haut-parleurs MAP. Pour ce faire, nous avons calculé les filtres à appliquer au dispositif utilisant la technique de multiégalisation exposée au chapitre 2. Le dispositif utilisé dans le démonstrateur est représenté à la figure 4.15.

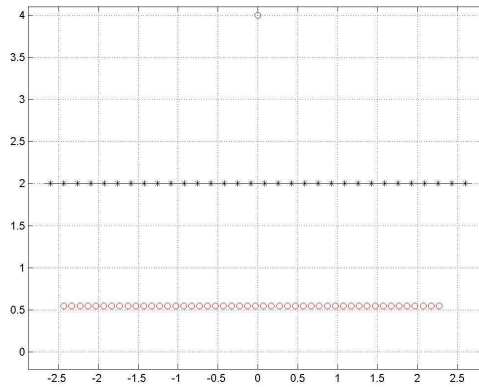


FIG. 4.15 – *Configuration utilisée dans le démonstrateur IRCAM. Les croix représentent les 32 haut-parleurs espacés de 0.14m, les cercles rouges représentent les 48 microphones espacés de 0.1m, le cercle bleu représente la source reproduite.*

Pour les besoins du test d'écoute, on a reproduit un dipole orienté à  $90^\circ$  par rapport à l'axe perpendiculaire au banc de haut parleurs et située à 2 mètres derrière (cf. figures 4.16, 4.16). On a également reproduit une harmonique cylindrique d'ordre 5 située au même endroit (cf. figures 4.18, 4.18). La fréquence d'aliasing est notablement réduite par rapport au simulations du fait d'un espacement supérieur entre les haut-parleurs (13.6 cm au lieu de 10 cm). En dessous de cette fréquence, les fonctions de directivité proposées sont convenablement rendues.

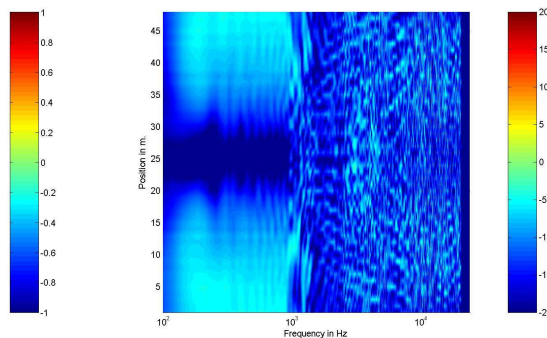
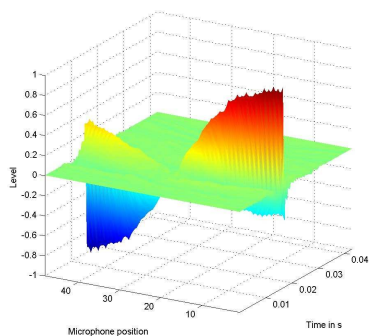


FIG. 4.16 – Réponse impulsionnelle d'un dipôle reproduit sur le démonstrateur IRCAM

FIG. 4.17 – Réponse fréquentielle d'un dipôle reproduit sur le démonstrateur IRCAM

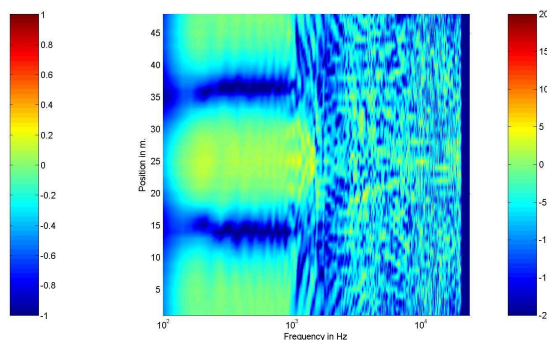
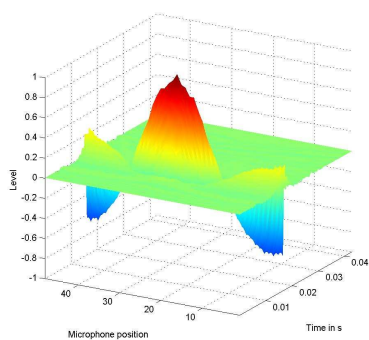


FIG. 4.18 – Réponse impulsionnelle d'une harmonique cylindrique d'ordre 5 sur le démonstrateur IRCAM

FIG. 4.19 – Réponse fréquentielle d'une harmonique cylindrique d'ordre 5 sur le démonstrateur IRCAM

## Chapitre 5

# Mise en cohérence du champ sonore reproduit en WFS avec la salle d'écoute

*Nous avons vu au chapitre précédent une manière de d'introduire une directivité dans les sources reproduites. Nous allons à présent aborder la mise en cohérence du dispositif WFS avec la salle de restitution. On part de l'hypothèse qu'un effet de salle cohérent avec le lieu d'écoute va renforcer le réalisme de la scène sonore reproduite. En d'autres termes, lorsque l'on synthétise une source virtuelle, on vise à reproduire par la même occasion un effet de salle se rapprochant le plus possible de celui qui serait généré par une source réelle. Il s'agit donc d'explicitier la différence entre l'interaction d'une source réelle et l'interaction d'une source reproduite par WFS avec une salle d'écoute. Une fois cette différence explicitée, nous pourrions tenter de vérifier son importance au niveau perceptif.*

*Nous verrons dans un premier temps que les approximations décrites au chapitre 1 ont un certain nombre de conséquences sur le champ effectif tridimensionnel rayonné par un dispositif WFS. On pourra alors se placer dans le cadre d'une salle d'écoute pour étudier l'interaction d'une source virtuelle avec la salle. Il s'avère que la reproduction des sources focalisées présente des particularités quant à l'interaction avec la salle d'écoute. Ces particularités on fait l'objet d'une expérience perceptive conduite lors du stage visant à déterminer l'importance (ou non) des réflexions frontales et latérales dans la reproduction d'une source sonore. Nous terminerons en proposant une méthode pour améliorer la reproduction des sources focalisées.*

## 5.1 Conséquences des approximations de la WFS sur le champ effectif rayonné

Les sections 1.5, 1.6 et 1.7 décrivent les approximations faites par rapport au principe de Huyghens, qui sont nécessaires lors de l'implémentation pratique d'un système WFS. Partant d'une distribution planaire et continue de sources secondaires (suggérée par la formulation de Kirchoff-Helmholtz), on effectue les simplifications suivantes :

1. Réduction du plan infini à une ligne infinie de manière à se restreindre au plan d'écoute horizontal
2. Restriction de la ligne infinie à un segment, limitant par conséquent la zone de restitution fidèle
3. Echantillonnage du segment entraînant des effets d'aliasing spatial au delà d'une certaine fréquence

Ces simplifications influencent directement le champ tridimensionnel effectivement rayonné par le système :

1. La réduction du plan à une ligne entraîne une *symétrie axiale* du champ rayonné par le banc de haut-parleurs. Ce champ s'apparente à un tore ayant pour axe le banc lui-même. La figure 5.1 montre les conséquences de cette symétrie lors de la reproduction d'une source située dans le plan d'écoute. On observe qu'un auditeur situé hors du plan d'écoute percevra la source non plus dans le plan horizontal mais à une position d'élévation inférieure, située sur le tore de rayonnement.

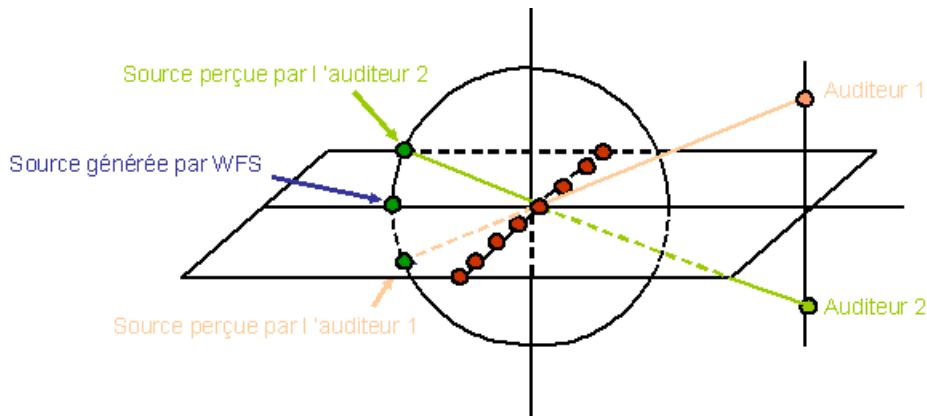


FIG. 5.1 – Rayonnement torique du banc de haut-parleurs



2. La restriction de la ligne infinie à un segment entraîne des effets de fenêtrage et une réduction de la zone de visibilité de la source virtuelle (cf parties 1.6, 1.8). Ceci entraîne l'émission d'ondes diffractées dans la zone d'écoute.
3. L'échantillonnage de la distribution de sources secondaires implique un comportement non-prédictible du champ sonore au-delà de la fréquence d'aliasing. Celui-ci s'apparente alors à un champ diffus, et se propage en dehors de la fenêtre de visibilité.

Outre les approximations faites sur la géométrie du système de reproduction, l'implémentation pratique de la WFS implique une approximation sur la nature même des sources secondaires utilisées. En effet des haut-parleurs réels n'ont pas un comportement monopolaire comme le dicte la théorie. Ils possèdent des caractéristiques de directivité complexes qui modifient le champ 3D effectif rayonné. Nous avons vu au chapitre 2 une méthode permettant de compenser la directivité des haut parleurs à l'intérieur de la zone de visibilité et dans le plan horizontal. Il est possible que le champ restitué hors du plan horizontal dépendent fortement des caractéristique de directivité verticale des haut-parleurs. Ce type de problème entre dans le cadre d'une thèse menée actuellement par Etienne Corteel à l'IRCAM. Cette thèse est dédiée à l'optimisation de la qualité objective et subjective d'un système WFS par l'égalisation du dispositif et la compensation du lieu d'écoute ([13], [24]).

## 5.2 Interaction d'une source virtuelle avec la salle d'écoute

Pour des raisons de simplification, nous considérerons par la suite que le banc de haut-parleurs utilisé est linéaire et qu'il couvre toute la longueur d'un mur de la salle d'écoute. Nous considérerons également qu'il est collé à ce mur (ce qui est le cas dans le démonstrateur IRCAM). Etant donné les problèmes de visibilité évoqués précédemment, il sera intéressant de traiter séparément la reproduction de sources focalisées et la reproduction de sources non-focalisées.

### 5.2.1 Reproduction d'une source non-focalisée

Les sources non-focalisées sont des sources virtuelles situées en dehors de l'espace de restitution. Leur reproduction entre donc dans le cadre

théorique du principe de Huyghens. Lorsque l'on synthétise ce type de source, on donne l'illusion à un auditeur situé dans l'espace de restitution qu'une source sonore est en train d'émettre depuis un point situé derrière le banc de haut-parleurs. La figure 5.2 illustre comment une telle source éclaire le plan horizontal de la salle d'écoute :

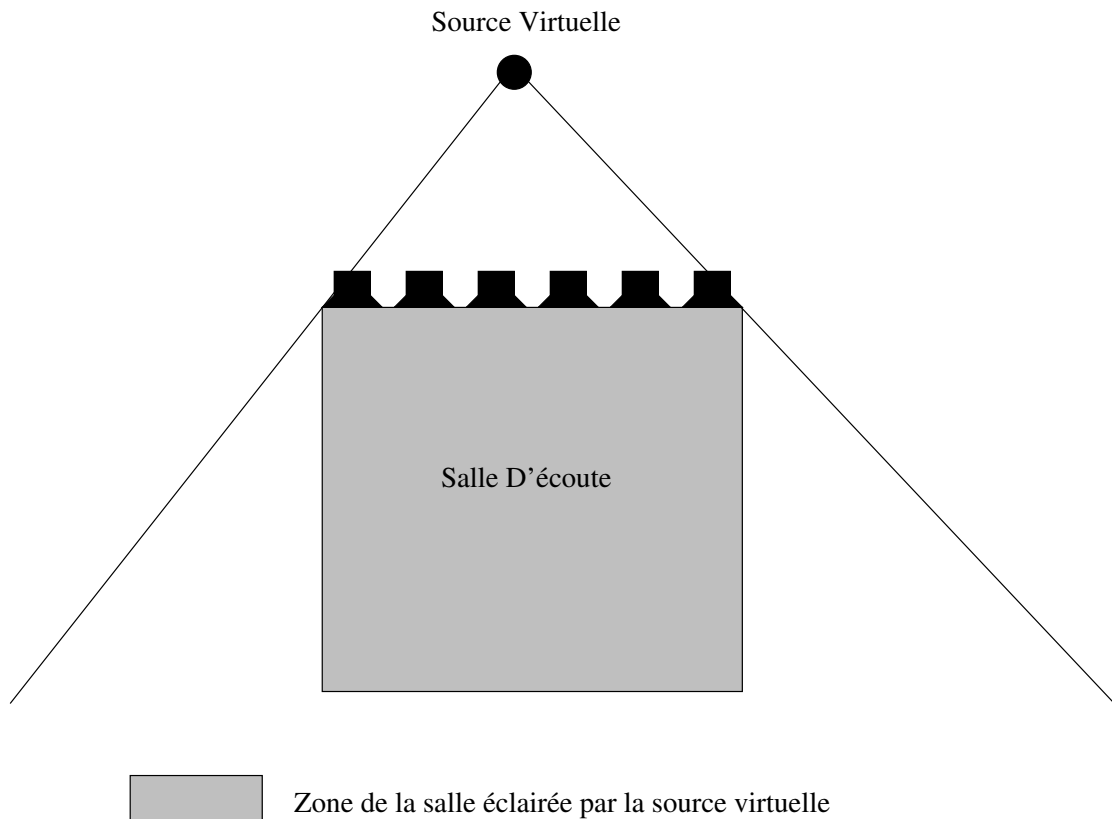


FIG. 5.2 – *Eclairage du plan horizontal de la salle d'écoute lors de la reproduction d'une source virtuelle non-focalisée*

La reproduction de ce type de source ne correspond pas à une situation réaliste. Par conséquent, son interaction avec la salle d'écoute ne peut pas être comparée avec une situation réelle. Ceci n'est pas le cas des sources focalisées, dont l'interaction avec la salle sera décrite dans la partie suivante.

### 5.2.2 Reproduction d'une source focalisée

Les sources focalisées sont des sources virtuelles situées à l'intérieur de l'espace de restitution. La possibilité de reproduire ce type de source est une particularité de la WFS. Ce type de reproduction n'entre pas dans le cadre théorique du principe de Huyghens et constitue un cas limite au regard de la théorie. On a comparé qualitativement le champ émis par une source réelle monopolaire située dans la salle d'écoute avec le champ émis par un système WFS reproduisant la même source. La première constatation concerne les problèmes de fenêtrage évoqués dans les chapitres précédents, qui vont se manifester sous la forme d'un manque d'éclairage de la salle d'écoute par rapport à une source réelle. La figure 5.3 illustre comment une telle source éclaire le plan horizontal de la salle d'écoute :

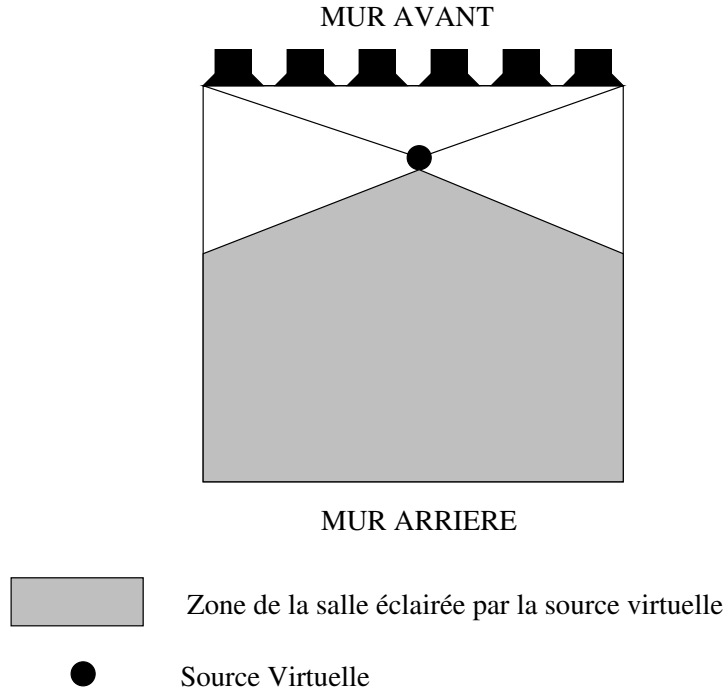


FIG. 5.3 – *Eclairage du plan horizontal de la salle d'écoute lors de la reproduction d'une source virtuelle focalisée*

Sur cette figure, on observe que le mur avant n'est pas éclairé par le son direct de la source synthétisée par WFS. Les murs latéraux sont également affectés partiellement par le fenêtrage ; on voit effectivement que leurs parties avant ne reçoivent pas d'énergie du son direct. Remarquons que ceci est vrai uniquement pour la partie cohérente du front d'onde produite par le banc de haut-parleurs, i.e. en dessous de la fréquence d'aliasing. Au dessus de cette fréquence, le champ sonore reproduit est diffus et les parties

avant des murs latéraux reçoivent de l'énergie supplémentaire venant du son direct. Notons également que des artefacts dûs à la diffraction de l'onde émise par les bords du banc de haut-parleurs vont venir alimenter les murs latéraux. Il apparaît enfin que le mur arrière est éclairé entièrement par le son direct de la source virtuelle.

Dans le plan vertical, la symétrie axiale du champ rayonné par le banc de haut-parleurs implique une légère délocalisation des sources images d'ordre 1 de la source reproduite. En d'autres termes, la direction et le temps d'arrivée des premières réflexions venant du sol et du plafond pour une source reproduite par WFS seront différents de ceux enregistrés pour une source réelle placée au même endroit. La figure 5.4 illustre la délocalisation de la source image par rapport au plafond de la salle de restitution.

Nous avons vu à la partie 5.1 qu'un auditeur situé hors du plan horizontal perçoit une source en décalage avec ce plan, située sur un tore dont l'axe est confondu avec le banc de haut-parleurs. Sur la figure 5.4, on voit que l'auditeur est situé en dessous du plan horizontal contenant l'image du banc de haut-parleurs par rapport au plafond. Par conséquent, il perçoit une position de source virtuelle image décalée par rapport à la position idéale (qu'il percevrait si la source était réelle). Pour la configuration présentée, des considérations géométriques permettent d'évaluer le biais angulaire à  $11^\circ$  et la différence de distance à  $D - D' = 36,5\text{cm}$ . D'après [25], le flou de localisation dans cette zone angulaire du plan médian est de  $\pm 10^\circ$  pour une voix continue connue. Ce flou augmente lorsque l'on présente des sons inconnus à l'auditeur. On en conclut que le biais angulaire de  $11^\circ$  est acceptable en termes d'erreur sur la source image. [25] affirme également que le flou de localisation en distance devient supérieur à 44cm pour des distances auditeur/source supérieures à 2.5 mètres (ce qui est le cas ici). Par conséquent une erreur de 36.5cm est acceptable sur la distance de la source image perçue.

Rajoutons que les résultats obtenus dans [25] concernent des sons directs sans composantes réfléchies. Le son considéré dans cette étude étant une réflexion primaire, il sera par nature moins apte à être localisé précisément.

Etant donné la symétrie de la pièce, ces conclusions demeurent valables pour les sources images par rapport au sol.

Nous considérerons donc dans la suite de cette étude que les réflexions au niveau du mur arrière, du sol et du plafond sont correctement restituées par le dispositif WFS.



Field Synthesis. L'analyse menée à la section précédente a montré que l'on pouvait considérer en première approximation que les sources images d'ordre 1 par rapport au plafond, sol et mur arrière étaient correctement restituées par le dispositif WFS. Celles-ci sont au contraire absentes lorsque l'on considère le mur avant et les murs latéraux. Cette absence engendre par la même occasion la disparition de toutes leurs sources images filles ; il manque donc des régimes entiers de réflexions issus des réflexions frontales et latérales d'ordre 1 dans la reproduction de sources focalisées.

## 5.4 Elaboration d'une expérience perceptive

Pour évaluer l'importance de ces réflexions, et par extension de leur absence, on propose d'effectuer une expérience psychoacoustique basée sur une simulation binaurale. Remarquons qu'une expérience de ce type menée sur un système WFS serait sujette aux caractéristiques de la pièce et du système de haut-parleurs utilisé pour l'expérience. De ce fait, la généralité de l'expérience serait fortement diminuée, d'où le choix de faire une simulation binaurale. Lors de cette expérience, un auditeur dont on a préalablement mesuré les HRTF est placé dans une salle. Il est muni d'un casque auquel on a fixé le récepteur d'un système de headtracking. On synthétise par voie binaurale une source placée à deux mètres devant l'auditeur (figure 5.5). Cette source est placée dans un environnement virtuel dont les dimensions correspondent à la pièce dans laquelle se trouve l'auditeur. On génère quatre scènes sonores pertinentes au regard du problème envisagé, du fait qu'elles permettent de comparer la présence ou d'absence de différents régimes de réflexions absents lors de la reproduction de sources focalisées. Ces situations sont énumérées ci-dessous :

1. *Situation "tout"* : On restitue un son direct et toutes ses réflexions
2. *Situation "ss devant"* : On restitue un son direct et ses réflexions, éliminant la réflexion frontale d'ordre 1 et ses filles
3. *Situation "ss côtés"* : On restitue un son direct et ses réflexions, éliminant les réflexions latérales d'ordre 1 et leurs filles
4. *Situation "min"* : On restitue un son direct et ses réflexions, éliminant les réflexions latérales et frontales d'ordre 1 et leurs filles

Pour préparer ce test, on implémente un programme permettant de générer indépendamment des régimes de sources images jusqu'à l'ordre 4 par rapport aux parois de la pièce. Les parois de la pièce ont des coefficients

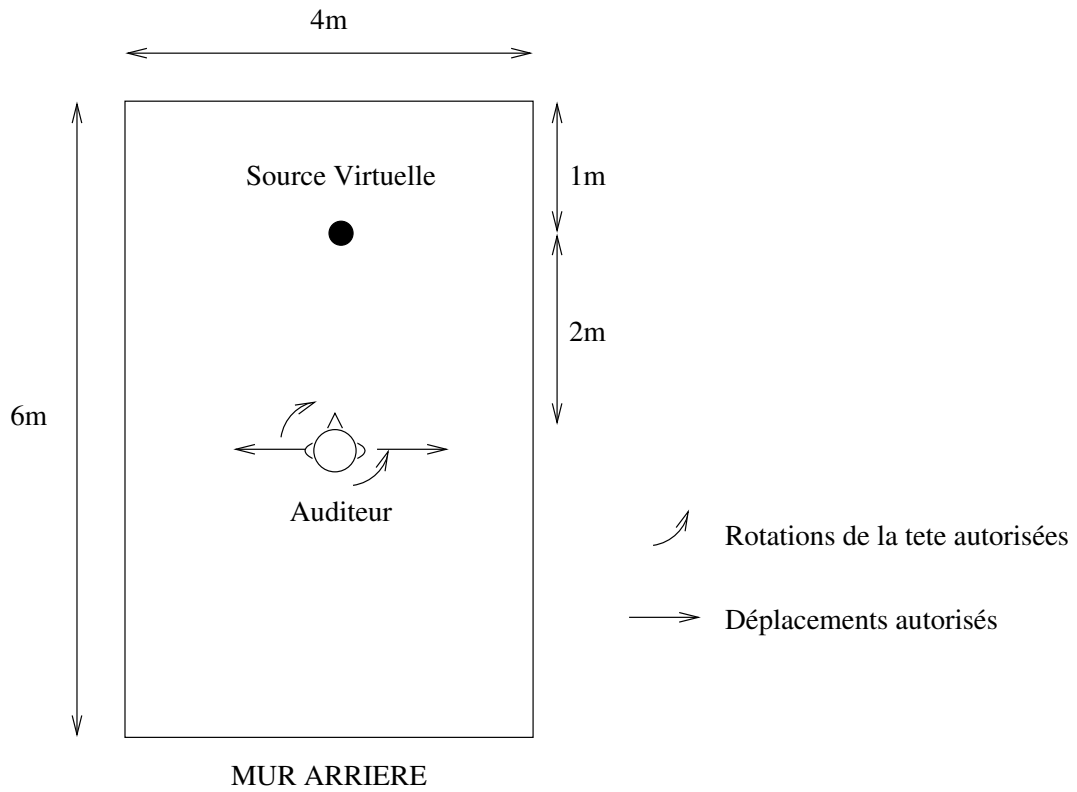


FIG. 5.5 – *Schéma de l'installation utilisée pour l'expérience perceptive*

d'absorption de 0.1 sur la bande des fréquences audibles. Une fois ces ensembles de sources images obtenus, le signal correspondant à chaque source image est filtré par les HRTF correspondantes, pour une position de l'auditeur donnée (angle d'incidence de la source par rapport à l'auditeur + prise en compte du délai de propagation et de l'atténuation de la source). Les signaux “gauche” et “droite” issus du filtrage HRTF sont ensuite sommés et stockés dans une base de données de filtres. Ce calcul est effectué pour des positions d'auditeur tous les dix centimètres sur un segment de deux mètres de largeur au centre de la pièce (cf 5.5). On prend en compte les rotations de la tête de l'auditeur en calculant des filtres tous les  $5^\circ$  sur  $90^\circ$  de chaque côté, et ce pour chaque position sur le segment. Pendant l'expérience, le headtracker envoie les informations sur la position du sujet vers un outil de convolution qui charge en temps réel les filtres correspondants et effectue le filtrage du signal audio désiré. Le résultat du filtrage est ensuite envoyé vers l'oreille correspondante (figure 5.6). Un pied de microphone sert de repère visuel à l'auditeur pour indiquer le lieu de la source.

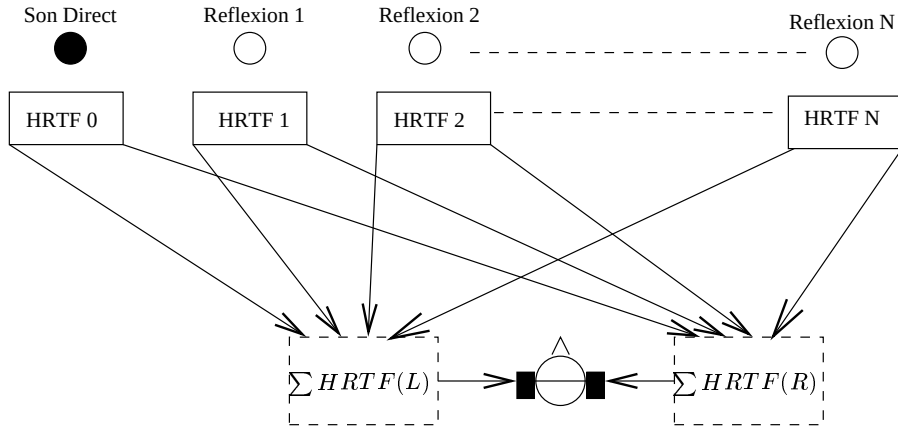


FIG. 5.6 – *Utilisation de filtrage HRTF pour simuler un son direct et ses réflexions dans une pièce*

Le but du test est de déterminer si le sujet arrive à distinguer les quatre situations évoquées précédemment (cf 5.4). On demande à l'auditeur de répondre à un test de type "ABX". Ceci consiste à lui présenter consécutivement une situation "A", une situation "B" et une situation "X", qui est une copie aléatoire de "A" ou "B". Il doit alors déterminer si "X" correspond à "A" ou "B". Les situations sont appariées de manière à pouvoir faire des comparaisons pertinentes au regard du problème considéré. Les paires retenues sont les suivantes :

*Situation "tout" / Situation "min"*  
*Situation "tout" / Situation "ss devant"*  
*Situation "min" / Situation "ss devant"*  
*Situation "tout" / Situation "ss côtés"*  
*Situation "min" / Situation "ss côtés"*

Pour chaque paire retenue, on fabrique 4 triplets pour le test ABX ("ABA", "ABB", "BAA", "BAB") de manière à minimiser l'influence de l'ordre dans la présentation des stimuli. Deux sons différents d'une durée de 3 secondes sont préparés pour le test. On dispose donc au total de 40 triplets "ABX". Lors du test, on présente les 40 triplets de manière aléatoire. 20 triplets utilisent un son de basse ; les 20 autres utilisent une voix masculine. Le volume a été ajusté de manière à ce que les situations soient perçues à volume égal par les auditeurs.

Pour pouvoir caractériser objectivement l'ajustement de niveau effectué, on calcule l'énergie de réponses impulsionnelles captées par un microphone omnidirectionnel pour un ensemble de positions des sujets sur la ligne d'écoute *après ajustement* (cf FIG. 5.7). On remarque que le niveau n'est



pas le même pour toutes les situations : les courbes “tout” et “ss côtés”, qui correspondent à des situations où les réflexions latérales sont présentes, ont été ajustées à un niveau supérieur au deux autres situations. Néanmoins il convient de noter que le niveau perçu par un auditeur qui écoute une source dépend principalement de l’ensemble des contributions dans la direction de la source ; les contributions latérales n’ont qu’une influence limitée ([26]). Un calcul de l’efficacité latérale montre que l’efficacité latérale pour les situations “tout” et “ss côtés” est supérieure à celle calculée pour les deux autres situations (cf 5.8). En d’autres termes, les situations où l’efficacité latérale était importante ont été réajustées à la hausse pour pouvoir les percevoir au même niveau que les autres situations.

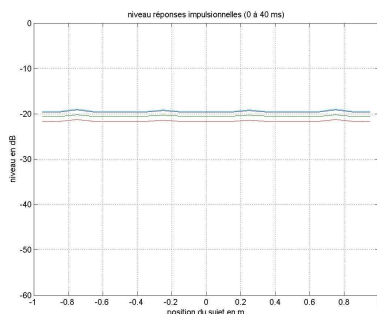


FIG. 5.7 – *Niveaux captés sur un microphone omnidirectionnel après ajustement perceptif sur l’ensemble des positions de l’auditeur ; La couleur bleu foncé correspond à la situation “tout”, bleu clair à “ss devant”, vert à “min”, rouge à “ss côtés”*

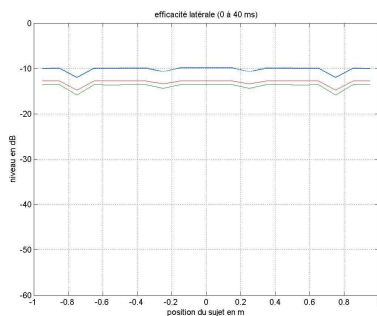


FIG. 5.8 – *Efficacité latérale calculée sur l’ensemble des positions de l’auditeur pour les quatre situations ; la couleur bleu foncé correspond à la situation “tout”, bleu clair à “ss devant”, vert à “min”, rouge à “ss côtés”*

Une interface de test conçue sur MAX MSP permet à l’auditeur de saisir sa réponse et de passer à l’exemple suivant (figure 5.9).

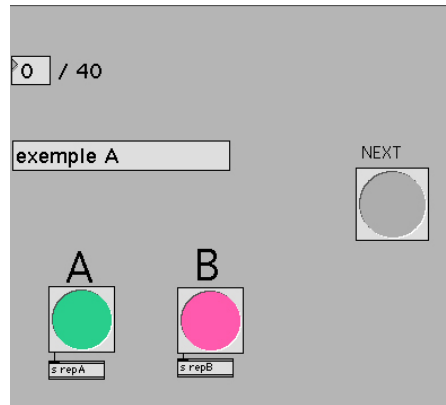


FIG. 5.9 – *Interface du test perceptif*

#### 5.4.1 Résultats de l'expérience

On recueille les résultats de l'expérience pour 9 sujets. Chacune des paires de situations servant à la comparaison est donc représentée 36 fois. On suppose l'indépendance entre chaque tirage. La probabilité  $p$  que le choix entre deux situations s'apparente à un tirage au hasard est calculée selon la loi du  $\chi^2$  à un degré de liberté. Les résultats sont compilés dans le tableau 5.1. Remarquons que le nombre limité de sujets implique qu'il s'agit uniquement de donner une tendance quant à la discriminabilité de certains régimes de réflexions.

| SON DE VOIX       | tout/min | tout/ss devant | min/ss devant | tout/ss côtés | min/ss côtés |
|-------------------|----------|----------------|---------------|---------------|--------------|
| % bonnes réponses | 69.4%    | 69.4%          | 72.2%         | 91.7%         | 66.7%        |
| % choix au hasard | 1.9631%  | 1.9631%        | 0.7661%       | 0.0001%       | 4.5500%      |
| SON DE BASSON     | tout/min | tout/ss devant | min/ss devant | tout/ss côtés | min/ss côtés |
| % bonnes réponses | 63.9%    | 55.6%          | 66.7%         | 66.7%         | 61.1%        |
| % choix au hasard | 9.558%   | 50.498%        | 4.5500%       | 4.5500%       | 18.242%      |

TAB. 5.1 – *Résultats du test perceptif*

On observe tout d'abord que pour le stimulus voix, toutes les situations sont discriminées, et ce malgré l'ajustement en niveau. Il apparaît aussi un écart important entre les résultats obtenus utilisant le stimulus "VOIX" et le stimulus "BASSON". On en déduit une dépendance de la perception des réflexions précoces vis-à-vis du type de son utilisé.

D'autre part les résultats pour le stimulus "BASSON" s'écartent sensiblement du hasard ( $p < 5\%$ ) uniquement dans deux configurations : lorsque

l'on compare "min" (c'est à dire le son direct et ses réflexions privées des réflexions latérales et frontales) avec "ss devant" (c'est à dire le son direct et ses réflexions privées des réflexions frontales) et lorsque l'on compare "tout" (son direct et toutes ses réflexions) avec "ss côtés" (son direct et réflexions privées des réflexions latérales). Ces configurations obtiennent également les maximums de discrimination lorsque l'on utilise le stimulus "VOIX". Il apparaît donc que les sujets discernent de manière privilégiée le rajout de réflexions latérales sur la situation "min" et le retrait des réflexions latérales dans la situation "tout". Ces réflexions, auxquelles on attribue de manière classique le paramètre perceptif d'enveloppement, semblent donc plus discernables que les réflexions frontales.

La comparaison entre la situation "tout" et la situation "min" n'est pas significative pour le son basson ( $p > 5\%$ ). Dans le cas du stimulus "VOIX", on relève 69% de bonnes réponses, soit moins que pour les deux configurations évoquées précédemment. Ce résultat est difficilement interprétable; on s'attend à priori à enregistrer le plus grand discernement étant donné que ces deux situations sont les plus éloignées du point de vue de la présence ou non de certains régimes de réflexions. On est amené à conclure que les réflexions frontales viennent "équilibrer" la scène sonore lorsqu'on les combine avec les réflexions latérales; on discerne plus facilement le rajout de réflexions latérales lorsqu'il s'effectue sans le rajout simultané de réflexions frontales. Cette tendance est toutefois faible (63% contre 67% pour le basson, 69% contre 72% pour la voix).

Les tendances relevées lors du test sont donc les suivantes :

- La perception du rajout ou du retrait des différents régimes de réflexions absents dépend du stimulus sonore utilisé
- Les réflexions issues des réflexions latérales d'ordre 1 se discernent plus facilement que les réflexions issues de la réflexion frontale d'ordre 1
- On discerne plus facilement la présence ou l'absence de réflexions latérales lorsque les réflexions frontales sont absentes

Ces tendances sont à confirmer en effectuant des tests avec plus de sujets. On pourrait cependant envisager des tests plus élaborés à l'avenir. D'une part, la discrimination entre les situations dépend du type de son employé. Un autre test utilisant des sons filtrés passe bande pourrait être mis en oeuvre pour déterminer l'influence du contenu fréquentiel dans la détection des réflexions primaires frontales et latérales. On pourrait également faire varier l'attaque d'un son donné pour étudier l'influence de ce paramètre quant à la discrimination de certaines réflexions. D'autre part, ce test pour-

rait permettre d'étudier l'influence de la position de la source et de la taille de la salle, le cadre de ce test étant restreint à une position de source et une taille de salle fixes.

Une étude plus approfondie consisterait à simuler un dispositif WFS en binaural pour déterminer l'influence des artefacts qui apparaissent lors de la reproduction de sources focalisées (aliasing, diffraction..).

## 5.5 Amélioration de la reproduction des sources intérieures

Suite à l'analyse menée dans la section 5.2.2, on propose une approche pour améliorer la reproduction des sources focalisées en WFS. Ceci peut être réalisé en injectant artificiellement les réflexions manquantes le long des murs avant et latéraux :

- **Mur avant** : Il est possible générer une source image virtuelle de l'autre côté du mur avant de manière à recréer la réflexion primaire qui existerait si la source était réelle. Quelle que soit la fréquence, ce mur est invisible au regard du son direct ; il suffit donc de tenir compte du coefficient d'absorption du mur lorsque l'on rajoute cette source image virtuelle.
- **Murs latéraux** : En ajoutant des extensions au banc de haut-parleurs le long des sections frontales des murs latéraux, on est capable de générer des sources images virtuelles sur ces sections. Les haut-parleurs utilisés pour reproduire ces sources images devront cependant être restreints à la zone non-éclairée par le son direct. De plus, cette restitution ne peut être réalisée qu'en dessous de la fréquence d'aliasing où le front d'onde reproduit demeure cohérent. Au delà de la fréquence d'aliasing, la zone de visibilité décrite n'est plus valable et le haut parleurs situés sur le mur avant rayonnent également sur les murs latéraux. Il s'agirait donc d'évaluer l'énergie rayonnée sur ces murs au-delà de la fréquence d'aliasing. On pourrait alors déterminer s'il manque de l'énergie par rapport au cas d'une source réelle, dans quel cas un apport d'énergie supplémentaire en hautes fréquences serait envisageable.

La figure 5.10 illustre l'injection de sources images virtuelles pour combler le manque de réflexions frontales dans la reproduction de sources focalisées :

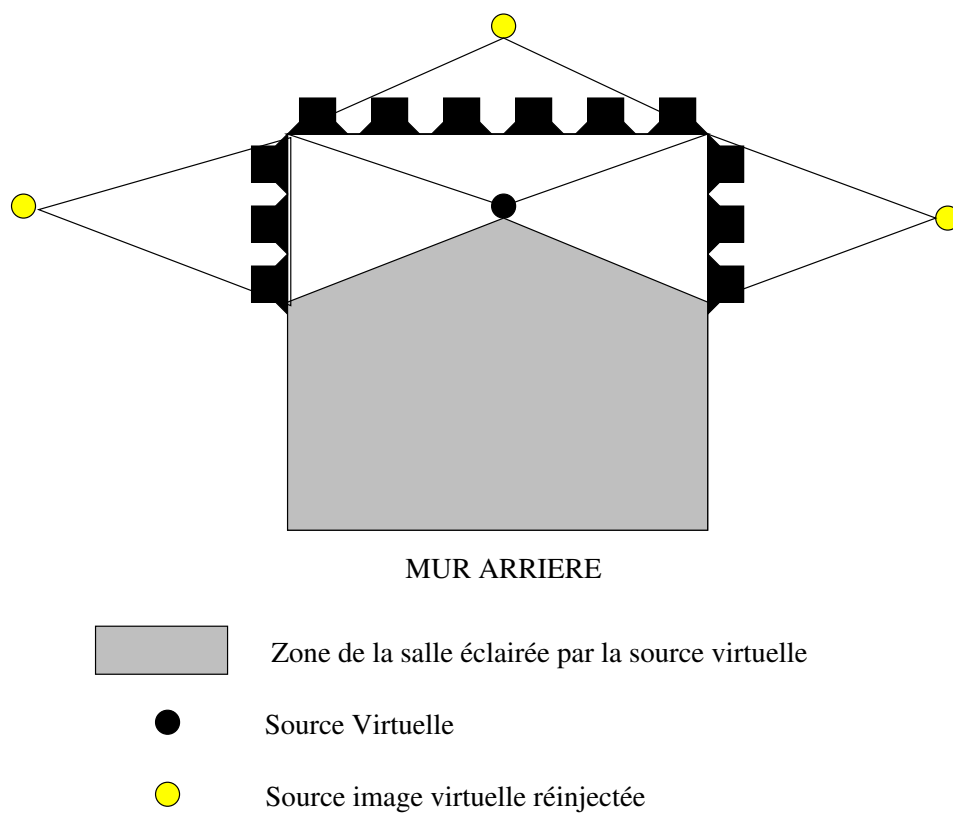


FIG. 5.10 – *Ajout de sources images virtuelles utilisant des bancs de haut-parleurs situés sur les zones frontales de la salle de restitution.*

## Chapitre 6

# Conclusion et Perspectives

Ce stage a commencé par un travail bibliographique permettant une familiarisation avec la théorie sous-tendant la Wave Field Synthesis. Une attention particulière a été portée à la compréhension du processus d'égalisation multicanale utilisé dans le dispositif de reproduction WFS mis en place à l'IRCAM. Parallèlement, un apprentissage pratique a été effectué pour comprendre l'utilisation des outils technologiques et logiciels nécessaires à une implémentation pratique. Une fois cette phase d'apprentissage terminée, un travail sur la directivité des sources a été envisagé. Après avoir conduit une étude théorique sur les harmoniques sphériques et cylindriques, l'implémentation d'un programme permettant d'introduire une directivité dans les sources notionnelles reproduites en WFS a été réalisée. Cette implémentation a donné lieu à une comparaison entre la reproduction de sources directives avec et sans égalisation multicanale du dispositif.

D'autre part, une modélisation de l'interaction du dispositif avec la salle de restitution a été réalisée dans le but de mettre en cohérence les sources virtuelles reproduites avec la salle d'écoute. Cette modélisation a permis notamment de cerner les problèmes liés à la reproduction de sources situées à l'intérieur de l'espace de restitution (sources focalisées). Une expérience perceptive permettant d'étudier l'influence des régimes de réflexions absents lors de la reproduction de sources focalisées a été conduite. Les résultats de cette expérience tendent à montrer l'importance des réflexions latérales dans la perception d'un effet de salle. Une méthode permettant d'améliorer le rendu des sources focalisées a également été proposée.

Le travail réalisé au cours de ce stage permet d'envisager de nombreux développements. Tout d'abord, des expériences perceptives s'inspi-

rant de celle décrite dans ce mémoire pourraient être dérivées permettant de donner plus de précisions sur le rôle des différents régimes de réflexions dans la perception d'un effet de salle. Une simulation binaurale du dispositif WFS pourrait également être réalisée pour pouvoir déterminer l'influence des phénomènes d'aliasing et de diffraction dans la localisation d'une source focalisée. D'autre part, le travail réalisé sur la directivité des sources permet d'envisager de recréer les figures de directivité d'instruments réels en faisant des combinaisons linéaires des harmoniques cylindriques de base. Des effets d'éclairages sonores variables de la salle d'écoute pourront également être imaginés, ce qui pourrait présenter un intérêt réel au niveau de la composition musicale spatialisée.

# Bibliographie

- [1] A.J. BERKHOUT. A holographic approach to acoustic control. *J. Audio Eng. Soc.*, pp.977-995, 1988.
- [2] S. BRIX et al. Carrouso, a european approach to 3d audio. *AES 110th Convention*, 2001.
- [3] R. NICOL. *Etude de la restitution du son spatialisé en zone étendue : Application à la téléprésence*. PhD thesis, Université du Maine, 1999.
- [4] M. NOGUES. Perception et gestion de l'effet de distance en wave field synthesis. Technical report, Conservatoire National Supérieur de Musique et de Danse de Paris, 2003.
- [5] W.B. SNOW. Basic principles of stereophonic sound. *Journal of the SMPTE*, vol.61 pp.567-589, 1953.
- [6] A.J. BERKHOUT, D. DE VRIES, and P. VOGEL. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.*, vol. 93, pp. 2764-2778, 1993.
- [7] M. JESSEL. *Acoustique théorique : propagation et holophonie*. Masson, 1973.
- [8] A.J. BERKHOUT. *Applied Seismic Wave Theory*. Elseiver, 1987.
- [9] E.N.G. VERHEIGEN. *Sound Reproduction by Wave Field Synthesis*. PhD thesis, TU Delft, 1997.
- [10] E.W. START. *Direct Sound Enhancement by Wave Field Synthesis*. PhD thesis, TU Delft, 1997.
- [11] E. CORTEEL. Wave field synthesis (holophonie) :possibilités offertes et limitations. Technical report, ENST, 2000.
- [12] D. DE VRIES, E.W. START, and V.G. VALSTAR. The wave field synthesis concept applied to sound reinforcement : Restrictions and solutions. *AES 96th Convention*, 1994.
- [13] E. CORTEEL, U. HORBACH, and R.S. PELLEGRINI. Multichannel inverse filtering of multiexciter distributed mode loudspeakers for wave field synthesis. *AES 112th Convention*, 2002.



- [14] D. DE VRIES. Sound reinforcement by wave field synthesis : Adaptation of the synthesis operator to the loudspeaker directivity characteristics. *JAES*, 44, 1996.
- [15] E. CORTEEL. Loudspeaker array equalization using multichannel adaptative techniques, application to the reproduction of virtual sources in an extended area. Technical report, STUDER Professional Audio A. G., 2001.
- [16] R. PELLEGRINI. *A Virtual Reference Listening Room as an Application of Auditory Virtual Environments*. PhD thesis, Université de Bochum, Allemagne, 2001.
- [17] S. ROUX. Synthèse d'un effet de salle dans le contexte de diffusion holophonique. Technical report, Stage de DEA ATIAM, 2002.
- [18] E. HULSEBOS, D. DE VRIES, and E. BOURDILLAT. Improved microphone array configurations for the auralization of sound fields by wave field synthesis. *AES 110th Convention, Amsterdam*, 2001.
- [19] J. M. JOT and O. WARUSFEL. Le spatialisateur, le son et l'espace. *Rencontres Musicales Pluridisciplinaires Informatique et Musique, Musiques en Scène, Lyon*, 1995.
- [20] N. MISDARIIS. Radiation control on multi-loudspeaker device : La timée. *XXth International Computer Music Conference (ICMC)*, 2001.
- [21] J. DANIEL. *Représentation de Champs Acoustiques, Application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. PhD thesis, Université Paris VI, 2000.
- [22] R. NICOL. *Restitution sonore spatialisée sur une zone étendue : Application à la téléprésence*. PhD thesis, Université du Maine, 1999.
- [23] A. LABORIE. Capture échantillonnage et manipulations de fonctions de directivité. application à l'enregistrement de scènes sonores 3d avec un micro soundfield. Technical report, Mastère Spécialisé ENST, "Traitement du Signal et Reconnaissance des formes", 2000.
- [24] E. CORTEEL and R. NICOL. Listening room compensation for wave field synthesis : What can be done ? *AES 23rd Conference*, 2003.
- [25] J. BLAUERT. *Spatial Hearing*. MIT Press, 1974.
- [26] E. KAHLE. *Validation d'un modèle objectif de la perception de la qualité acoustique sur un ensemble de salles de concert et d'opéra*. PhD thesis, Université du Maine, 1995.